



Universidade do Minho

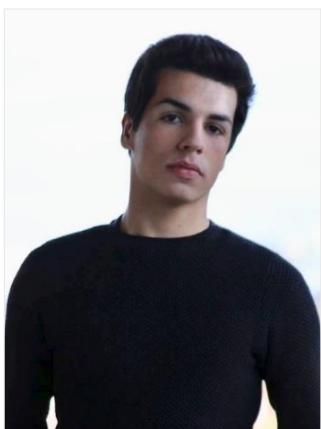
Licenciatura em Engenharia e Gestão de Sistemas de Informação

Criminalidade no Reino Unido no contexto de Igualdade de Género

Relatório Final

Engenharia de Dados para Suporte à Tomada de Decisão | 2023/2024

Grupo 3 | PL 3



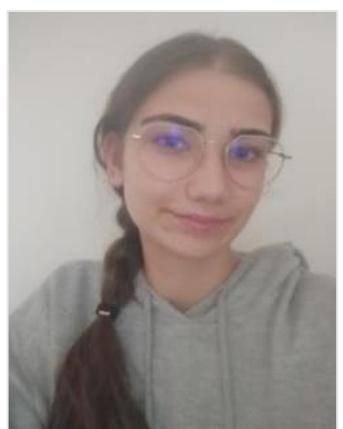
A95549
Pedro Miguel Araújo Pires
a95549@alunos.uminho.pt



A100436
Landro Macedo da Cruz
a100436@alunos.uminho.pt



A101441
Catarina Brandão Fernandes
a101441@alunos.uminho.pt



A93347
Maria Inês da Costa Maia
a93347@alunos.uminho.pt

Índice

Introdução.....	4
Questões Analíticas e KPI's	5
Diagrama da Arquitetura	7
Ficheiros e diretórias Bronze no HDFS.....	8
Ficheiros e diretórias Silver no HDFS	10
Tabelas Silver correspondentes ao Homicide_Dataset.....	11
Tabelas Silver correspondentes ao Crime_Dataset	12
Tabelas Silver correspondentes ao GenderStats_Dataset	13
Tabelas Silver correspondentes ao Child_Protection_Dataset.....	14
Ficheiros e diretórias Gold no HDFS.....	15
Transformações	16
Gender_Stats_Series_Dataset	16
Tabelas Silver resultantes	17
Tabelas Gold resultantes.....	19
Crime_Dataset	20
Tabelas Silver resultantes	21
Homicide_Victim.....	24
Tabelas Silver resultantes	25
Tabelas Gold resultantes.....	26
Dashboards e esquema no Tableau.....	27
Child_Protection_Dataset.....	28
Dashboards	31
Gender_Stats_Dataset	31
Gender Data per Country.....	31
Stop_and_Search	33
Stop and search in UK.....	33
Homicide_Dataset.....	36
Homicide Victims in London	36
Homicide Aggressors in London.....	38
.....	38
Child_Protection_Dataset.....	40
Children Careplan	40
Data Profiling and Data Quality	42

Data Profiling and Dictionary	43
Dataset: Homicide_Victim	43
Datasets: Crime_Dataset	56
Dataset: Child_Protection_Dataset	81
Dataset: Gender_Stats_Dataset.....	93
Auto e Hétero avaliação.....	Erro! Marcador não definido.
Auto avaliação:.....	Erro! Marcador não definido.
Hétero avaliação do trabalho:	Erro! Marcador não definido.

Introdução

No contexto da Unidade Curricular de Engenharia de Dados para Suporte à Tomada de Decisão, foi-nos desafiado a desenvolver um projeto com a finalidade de estabelecer um sistema Data Lakehouse. O tema central fornecido pelos docentes incide sobre a Igualdade de Género e a Criminalidade no Reino Unido. A partir de uma seleção criteriosa, optamos por explorar quatro datasets (6 ficheiros CSV) relevantes que nos permitissem mergulhar no subtema da relação entre criminalidade e questões de género no Reino Unido.

Na primeira fase do projeto, após definir o tema, estabelecemos os nossos KPIs e formulámos as questões analíticas a serem respondidas. Seguidamente, realizámos o Data Profiling e a avaliação da Qualidade dos Dados utilizando o Talend Open Studio For Data Quality. Concluímos esta etapa com a implementação inicial do projeto na estrutura Landing Zone, colocando os dados correspondente à zona de Bronze no HDFS.

Avançando para a segunda fase, começamos o processo ETL, aplicando transformações necessárias baseadas na análise realizada anteriormente.

Na terceira e última fase do projeto, aprofundámos o processo ETL para uma preparação ainda mais refinada dos dados, com o intuito de extrair insights mais precisos e abrangentes acerca do tema. Aplicamos também modelos em estrela que facilitam a análise multidimensional, permitindo uma interpretação mais eficaz e integrada dos dados. Além disso, dedicámo-nos à criação de dashboards interativas e informativas, utilizando ferramentas de visualização de dados avançadas. Este esforço não só facilitou a visualização e compreensão das complexidades associadas à criminalidade e à igualdade de género, mas também forneceu uma base sólida para a tomada de decisões informadas e a formulação de estratégias efetivas para abordar estas questões cruciais.

Em suma, objetivo do nosso projeto é realizar uma análise descritiva aprofundada da criminalidade no Reino Unido, focando especialmente na igualdade de género. Isto será explorado através de comparações temporais, geográficas e percentuais sobre diversas métricas, como os tipos de crime, as taxas de vitimização e o envolvimento diferenciado de géneros em atividades criminais. Um aspecto adicional e significativo do nosso estudo é a inclusão de dados sobre crianças que foram inscritas em planos de proteção, uma circunstância muitas vezes resultante de ambientes criminosos ou abusivos.

Questões Analíticas e KPI's

Questão analítica	Dataset
Quais são as discrepâncias entre faixas etárias na população do UK entre os anos 2010 e 2022? Haverá mais população do género feminino ou do género masculino?	Gender_Stats_Dataset
Alterando o intervalo de anos para entre 2017-2021, qual é o número total de mortes infantis em UK? Como esse número se distribui em relação a raparigas e rapazes?	Gender_Stats_Dataset
Dentro do mesmo intervalo de anos anterior e avaliando em taxas, a mortalidade afeta mais os adultos ou as crianças?	Gender_Stats_Dataset
Serão os homens ou as mulheres mais suscetíveis a cometerem homicídios intencionalmente?	Gender_Stats_Dataset
Qual é a tendência ao longo do tempo das operações de stop and search? Há um aumento ou diminuição nas operações realizadas e isso varia por gênero?	Crime_Dataset
Existe alguma relação entre os picos de Operações Stop and Search com algum pico de homicídios em Londres?	Crime_Dataset
Há diferenças nos tipos de busca realizadas em homens e mulheres? E quanto aos objetos da busca (por exemplo, drogas controladas, armas ofensivas, etc.)?	Crime_Dataset
Existe uma diferença na legislação invocada quando se trata de stop and search entre homens e mulheres? Isso pode sugerir algum viés de gênero na aplicação da lei?	Crime_Dataset
A distribuição geográfica das operações de stop and search difere entre homens e mulheres? Alguma área apresenta maior incidência de Operações entre ambos os géneros?	Crime_Dataset
Como a distribuição por gênero se correlaciona com a faixa etária dos indivíduos revistados? Existem diferenças notáveis entre os gêneros em diferentes faixas etárias?	Crime_Dataset
A etnia de um agente implica que um gênero seja mais sujeito a ser revistado que outro?	Crime_Dataset
Há uma discrepância significativa entre o número de homens e mulheres submetidos a operações de stop and search? Se sim, quais poderiam ser as razões para essa discrepância?	Crime_Dataset
Há uma maior discrepância de géneros nas vítimas ou nos agressores em Londres?	Homicide_Dataset
Os Boroughs onde estão registados mais crimes contra homens são os mesmos onde há mais crimes contra mulheres?	Homicide_Dataset

Qual o método de homicídio mais sofrido por cada género? Qual a evolução desse método ao longo do tempo?	Homicide_Dataset
Qual a relação entre o a faixa etária das vítimas de Homicídio entre homens e mulheres nos últimos anos? E nos agressores? Estarão as autoridades a dar a devida atenção?	Crime_Dataset / Homicide_Dataset
Qual o género mais afetado por crimes de violência doméstica? Como tem sido a evolução ao longo dos anos para esse género?	Homicide_Dataset
O número de agressores numa cidade poderá dar origem a desleixos no que toca a crianças. É possível visualizar, em Barnet, uma relação entre os crimes ocorridos em crianças e o número de homicídios por género?	Homicide_Dataset / Chil_Protection_Dataset
Qual a categoria de abuso com maior percentagem de vítimas por género?	Child_Protection_Dataset
Ao longo dos anos, qual o ano onde foi registado o maior número de crianças que iniciaram um plano de proteção, e qual o género com mais registo?	Child_Protection_Dataset
Qual a variação percentual ao longo do tempo de crianças que sofreram diversos tipos de abuso por género?	Child_Protection_Dataset
Será que há disparidades na duração dos planos em crianças de géneros opostos?	Child_Protection_Dataset

KPI	Cálculo	Dataset
Número total de homicídios/agressores/crianças/paragens	$\sum X$, sendo x a variável em estudo com filtros.	Homicide_Dataset / Crime_Dataset / Child_Protection_Dataset
Percentagem de homicídios/agressores/crianças/paragens	$\frac{\sum X}{T} \times 100$, sendo X a variável em estudo com/sem filtros e T o número total de ocorrências desse objeto sem filtros.	Homicide_Dataset / Crime_Dataset / Child_Protection_Dataset
Relação entre vítimas/agressores do género x pelo género y	$\frac{\sum X}{\sum Y}$	Homicide_Dataset
Duração média para as crianças completarem o plano	$\bar{x}_{\text{ano final}} - \bar{x}_{\text{ano inicial}}$	Child_Protection_Dataset
Razão entre número de mortes em crianças do género masculino pelo género feminino	$\frac{X_{\text{masculino}}}{X_{\text{feminino}}}$, sendo X o número de mortes em crianças	Gender_Stats_Dataset

Diagrama da Arquitetura

Com o término deste projecto, o grupo concluiu a fase Gold. Superámos o que era esperado, organizando e melhorando os dados para a análise. Resolvemos problemas de qualidade dos dados e efectuámos as alterações necessárias para garantir que a informação estava correcta e relevante.

Na Figura 1Figura 1 - Arquitetura, agora atualizada, mostramos as tabelas finais do nosso trabalho até ao final da fase Gold. Agora, além de termos terminado esta fase, também analisámos os dados em detalhe. Recorremos a modelos em estrela para organizar os dados, o que nos ajudou a entender melhor os resultados da nossa análise.

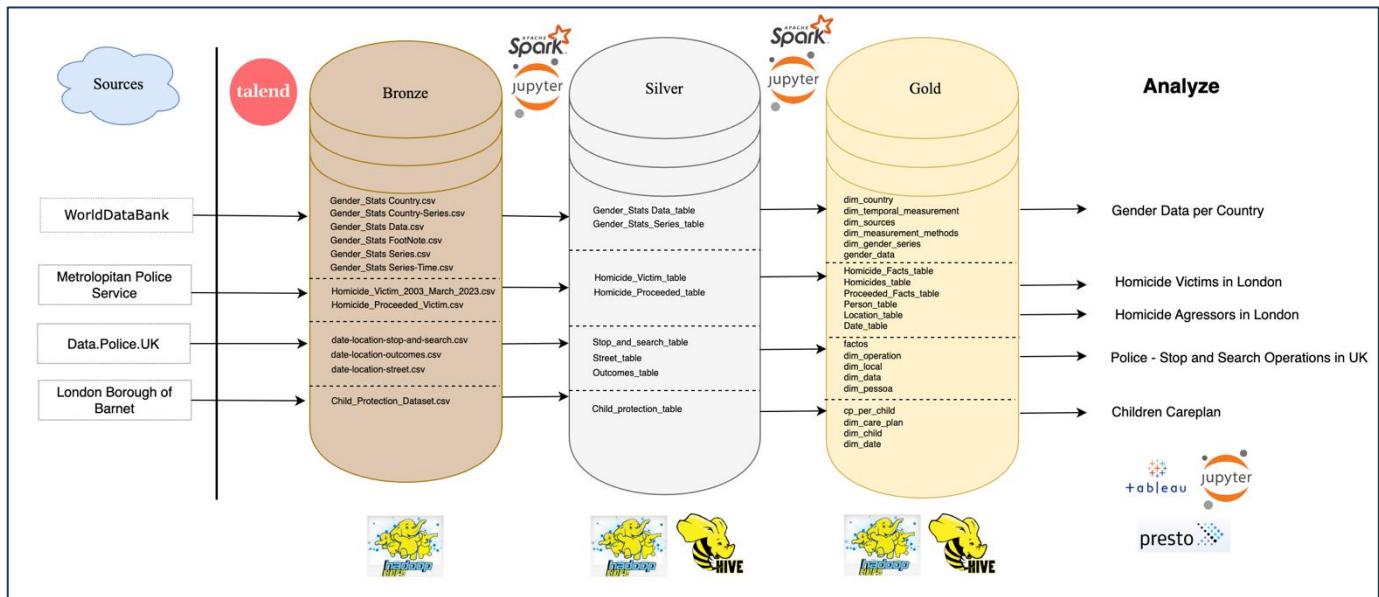


Figura 1 - Arquitetura

Ficheiros e diretórias Bronze no HDFS

Browse Directory										
/Datasets/Bronze										
<input type="text" value="/Datasets/Bronze"/> <input type="button" value="Go!"/> <input type="button" value="New"/> <input type="button" value="Upload"/> <input type="button" value="Download"/> <input type="button" value="Edit"/> <input type="button" value="Delete"/> <input type="button" value="Search..."/>										
Show	25	entries								
□	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name		
□	drwxr-xr-x	anonymous	hdfs	0 B	Oct 12 19:01	0	0 B	Child_Protection_Dataset	<input type="button" value="Edit"/>	<input type="button" value="Delete"/>
□	drwxr-xr-x	anonymous	hdfs	0 B	Oct 12 19:37	0	0 B	Crime_Dataset	<input type="button" value="Edit"/>	<input type="button" value="Delete"/>
□	drwxr-xr-x	anonymous	hdfs	0 B	Oct 12 19:07	0	0 B	Gender_Stats_Dataset	<input type="button" value="Edit"/>	<input type="button" value="Delete"/>
□	drwxr-xr-x	dr.who	hdfs	0 B	Oct 12 18:57	0	0 B	Homicide_Dataset	<input type="button" value="Edit"/>	<input type="button" value="Delete"/>
Showing 1 to 4 of 4 entries										
<input type="button" value="Previous"/> <input type="button" value="1"/> <input type="button" value="Next"/>										
Hadoop, 2023.										

Figura 2 - Diretoria Bronze

Browse Directory										
/Datasets/Bronze/Gender_Stats_Dataset										
<input type="text" value="/Datasets/Bronze/Gender_Stats_Dataset"/> <input type="button" value="Go!"/> <input type="button" value="New"/> <input type="button" value="Upload"/> <input type="button" value="Download"/> <input type="button" value="Edit"/> <input type="button" value="Delete"/> <input type="button" value="Search..."/>										
Show	25	entries								
□	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name		
□	-rw-r--r--	anonymous	hdfs	267.61 KB	Oct 15 14:50	3	128 MB	Gender_Stats Country-Series.csv	<input type="button" value="Edit"/>	<input type="button" value="Delete"/>
□	-rw-r--r--	anonymous	hdfs	153.24 KB	Oct 15 14:51	3	128 MB	Gender_Stats Country.csv	<input type="button" value="Edit"/>	<input type="button" value="Delete"/>
□	-rw-r--r--	anonymous	hdfs	112.84 MB	Oct 15 14:51	3	128 MB	Gender_Stats Data.csv	<input type="button" value="Edit"/>	<input type="button" value="Delete"/>
□	-rw-r--r--	anonymous	hdfs	48.54 MB	Oct 15 14:52	3	128 MB	Gender_Stats FootNote.csv	<input type="button" value="Edit"/>	<input type="button" value="Delete"/>
□	-rw-r--r--	anonymous	hdfs	7.43 KB	Oct 15 14:52	3	128 MB	Gender_Stats Series-Time.csv	<input type="button" value="Edit"/>	<input type="button" value="Delete"/>
□	-rw-r--r--	anonymous	hdfs	2.05 MB	Oct 15 14:53	3	128 MB	Gender_Stats Series.csv	<input type="button" value="Edit"/>	<input type="button" value="Delete"/>
Showing 1 to 6 of 6 entries										
<input type="button" value="Previous"/> <input type="button" value="1"/> <input type="button" value="Next"/>										
Hadoop, 2023.										

Figura 3 - Diretoria /Datasets/Bronze/Gender_Stats_Dataset

Browse Directory

/Datasets/Bronze/Child_Protection_Dataset

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
<input type="checkbox"/>	-rw-r--r--	anonymous	hdfs	553.18 KB	Oct 12 19:01	3	128 MB	Child_Protection_Dataset.csv

Showing 1 to 1 of 1 entries

Previous 1 Next

Hadoop, 2023.

Figura 4 - Diretoria /Datasets/Bronze/Child_Protection_Dataset

Browse Directory

/Datasets/Bronze/Homicide_Dataset

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
<input type="checkbox"/>	-rw-r--r--	anonymous	hdfs	308.56 KB	Oct 15 14:41	3	128 MB	Homicide_Proceeded_Victim.csv
<input type="checkbox"/>	-rw-r--r--	anonymous	hdfs	267.74 KB	Oct 15 14:43	3	128 MB	Homicide_Victim_2003.csv

Showing 1 to 2 of 2 entries

Previous 1 Next

Hadoop, 2023.

Figura 5 - Diretoria /Datasets/Bonze/Homicide_Dataset

Browse Directory

/Datasets/Bronze/Crime_Dataset

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
<input type="checkbox"/>	-rw-r--r--	anonymous	hdfs	267.61 KB	Oct 12 19:36	3	128 MB	2021-07-gwent-outcomes.csv
<input type="checkbox"/>	-rw-r--r--	anonymous	hdfs	267.61 KB	Oct 12 19:37	3	128 MB	2021-07-gwent-stop-and-search.csv
<input type="checkbox"/>	-rw-r--r--	anonymous	hdfs	267.61 KB	Oct 12 19:37	3	128 MB	2021-07-gwent-street.csv

Figura 6 - Diretoria /Datasets/Bronze/Crime_Dataset

Ficheiros e diretórias Silver no HDFS

Browse Directory

/warehouse/

Show 25 entries

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Projeto.db

Showing 1 to 1 of 1 entries

Previous 1 Next

Figura 2 - Arquitetura da warehouse contendo o Projeto da fase 2.

Browse Directory

/warehouse/Projeto.db

Show 25 entries

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Child_Protection_table
drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Gender_Stats_Data_table
drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Gender_Stats_Series_table
drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Homicide_Proceeded_table
drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Homicide_Victim_table
drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:49	0	0 B	Outcomes_table
drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:48	0	0 B	Stop_and_search_table
drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:50	0	0 B	Street_table

Showing 1 to 8 of 8 entries

Previous 1 Next

Figura 3 – Diretoria das tabelas da fase 2.

Tabelas Silver correspondentes ao Homicide_Dataset

Browse Directory

/warehouse/Projeto.db/Homicide_Victim_table

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Barking & Dagenham
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Barnet
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Bexley
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Brent
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Bromley
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Camden
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Croydon
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Ealing
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Ealing
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Enfield

Figura 4 – Partições da Homicide_Victim_table

Browse Directory

/warehouse/Projeto.db/Homicide_Proceeded_table

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Barking & Dagenham
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Barnet
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Bexley
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Brent
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Bromley
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Camden
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Croydon
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Ealing
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Enfield
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Greenwich
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:47	0	0 B	Borough=Hackney

Figura 5 – Partições da Homicide_Proceeded_table

Tabelas Silver correspondentes ao Crime_Dataset

Browse Directory								
/warehouse/Projeto.db/Stop_and_search_table								Go!
Show 25 entries		Search:						
□	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:48	0	0 B	_delta_log
□	-rw-r--r--	jovyan	hdfs	3.68 MB	Nov 15 11:47	3	128 MB	part-00000-2c0949e0-d8c2-4490-815a-53c847292308-c000.snappy.parquet
□	-rw-r--r--	jovyan	hdfs	3.34 MB	Nov 15 11:47	3	128 MB	part-00001-cac7c452-7936-49cf-8a1e-e4779eec70cc-c000.snappy.parquet
□	-rw-r--r--	jovyan	hdfs	3.11 MB	Nov 15 11:48	3	128 MB	part-00002-e1e0ac67-4e66-4ee5-8646-69b2a6a84d66-c000.snappy.parquet
□	-rw-r--r--	jovyan	hdfs	2.67 MB	Nov 15 11:48	3	128 MB	part-00003-b9286a0b-f8e3-4056-a11d-33ae3f378ceb-c000.snappy.parquet
□	-rw-r--r--	jovyan	hdfs	1.97 MB	Nov 15 11:48	3	128 MB	part-00004-2509a225-b884-4fce-a898-b103e61f808f-c000.snappy.parquet
□	-rw-r--r--	jovyan	hdfs	1.42 MB	Nov 15 11:48	3	128 MB	part-00005-0cc0ec19-c14b-4cd1-ae0f-d5d0d3363f1f-c000.snappy.parquet

Figura 6 - Partições da Stop_and_search_table

Browse Directory								
/warehouse/Projeto.db/Outcomes_table								Go!
Show 25 entries		Search:						
□	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:49	0	0 B	_delta_log
□	-rw-r--r--	jovyan	hdfs	9.7 MB	Nov 15 11:48	3	128 MB	part-00000-ef707f88-8997-426f-a7f5-cbd0c486d333-c000.snappy.parquet
□	-rw-r--r--	jovyan	hdfs	6.31 MB	Nov 15 11:48	3	128 MB	part-00001-e0e1433a-8cc9-4518-9bd2-74f3f84b6bff-c000.snappy.parquet
□	-rw-r--r--	jovyan	hdfs	8.4 MB	Nov 15 11:48	3	128 MB	part-00002-30ad8a13-6980-42f8-b4fc-7e594c195172-c000.snappy.parquet
□	-rw-r--r--	jovyan	hdfs	6.39 MB	Nov 15 11:48	3	128 MB	part-00003-a2628ae4-443b-466f-9413-290de29736f9-c000.snappy.parquet
□	-rw-r--r--	jovyan	hdfs	7.84 MB	Nov 15 11:48	3	128 MB	part-00004-f380e444-5d79-4ba4-ad71-df7344eee5bc-c000.snappy.parquet
□	-rw-r--r--	jovyan	hdfs	7.52 MB	Nov 15 11:48	3	128 MB	part-00005-7645253e-75eb-4524-bee1-0f0ed0efdcfe-c000.snappy.parquet

Figura 7- Partições da Outcomes_table

Browse Directory

/warehouse/Projeto.db/Street_table

Show 25 entries

Search:

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
□	drwxr-xr-x	joyyan	hdfs	0 B	Nov 15 11:50	0	0 B	_delta_log
□	-rw-r--r--	joyyan	hdfs	5.79 MB	Nov 15 11:49	3	128 MB	part-00000-a7a7d8ef-76d3-4b54-aef2-705c3aa1972b-c000.snappy.parquet
□	-rw-r--r--	joyyan	hdfs	6.61 MB	Nov 15 11:49	3	128 MB	part-00001-d171730e-965d-4f4e-9103-f94f39a1ee0a-c000.snappy.parquet
□	-rw-r--r--	joyyan	hdfs	6.65 MB	Nov 15 11:50	3	128 MB	part-00002-28b5b6b0-d891-4746-93cd-878d3bdc5bef-c000.snappy.parquet
□	-rw-r--r--	joyyan	hdfs	6.2 MB	Nov 15 11:50	3	128 MB	part-00003-a965792a-35cd-4a0e-8ca3-df257ef74e7a-c000.snappy.parquet
□	-rw-r--r--	joyyan	hdfs	6.52 MB	Nov 15 11:50	3	128 MB	part-00004-ec2ff341-1fa2-481e-bc70-d53e25ab9432-c000.snappy.parquet
□	-rw-r--r--	joyyan	hdfs	5.88 MB	Nov 15 11:50	3	128 MB	part-00005-53df5b61-0220-4577-97f2-f70fba6facac-c000.snappy.parquet

Figura 8 - Partições da Street_table

Tabelas Silver correspondentes ao GenderStats_Dataset

Browse Directory

/warehouse/Projeto.db/Gender_Stats_Series_table

Show 25 entries

Search:

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
□	drwxr-xr-x	joyyan	hdfs	0 B	Nov 15 11:51	0	0 B	_delta_log
□	-rw-r--r--	joyyan	hdfs	259.44 KB	Nov 15 11:51	3	128 MB	part-00000-917c7009-7043-4c47-b069-e36166c8589b-c000.snappy.parquet

Showing 1 to 2 of 2 entries

Previous 1 Next

Figura 9 - Partições da Gender_Stats_Series_table

Browse Directory

/warehouse/Projeto.db/Gender_Stats_Data_table

Show 25 entries

Search:

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Country_Region_Situation_Name=Afghanistan
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Country_Region_Situation_Name=Africa Eastern and Southern
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Country_Region_Situation_Name=Africa Western and Central
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Country_Region_Situation_Name=Albania
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Country_Region_Situation_Name=Algeria
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Country_Region_Situation_Name=American Samoa
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Country_Region_Situation_Name=Andorra
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Country_Region_Situation_Name=Angola
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Country_Region_Situation_Name=Antigua and Barbuda
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Country_Region_Situation_Name=Arab World
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	Country_Region_Situation_Name=Argentina

Figura 8 - Partições da Gender_Stats_Data_table

Tabelas Silver correspondentes ao Child_Protection_Dataset

Browse Directory

/warehouse/Projeto.db/Child_Protection_table

Show 25 entries

Search:

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
□	drwxr-xr-x	jovyan	hdfs	0 B	Nov 15 11:51	0	0 B	_delta_log
□	-rw-r--r--	jovyan	hdfs	99.01 KB	Nov 15 11:51	3	128 MB	part-00000-aae2da28-6201-4bf0-8c04-fc56f43a3e6c-c000.snappy.parquet

Showing 1 to 2 of 2 entries

Previous 1 Next

Figura 9 - Partições da Child_Protection_table

Ficheiros e diretórias Gold no HDFS

Na fase Gold foi utilizado a mesma diretória para guardar as tabelas e não foi feita nenhuma partição em nenhuma tabela. Assim, a diretória final ficou a seguinte:

/warehouse/Projeto.db										Go!	File	Up	Download	Print	Help
Show 50 entries										Search:					
	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name							
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	Child_Protection_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	Date_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	Gender_Stats_Data_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	Gender_Stats_Series_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	Homicide_Facts_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	Homicide_Proceeded_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:31	0	0 B	Homicide_Victim_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	Homicides_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	Location_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	Outcomes_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	Person_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	Proceeded_Facts_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	Stop_and_search_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	Street_table								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	cp_per_child								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	dim_care_plan								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	dim_child								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	dim_country								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	dim_data								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	dim_date								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	dim_gender_series								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	dim_local								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	dim_measurement_methods								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	dim_operation								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	dim_pessoa								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	dim_sources								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	dim_temporal_measurement								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	factos								
drwxr-xr-x	jovyan	hdfs	0 B	Jan 03 14:32	0	0 B	gender_data								
Showing 1 to 29 of 29 entries										Previous	1	Next			

Figura 7 - Tabelas Gold

Transformações

Gender_Stats_Series_Dataset

Transformações Bronze-Silver

Renomeação de Colunas e Correção de tipos de dados:

- ▶ As colunas em **Gender_Stats_Series_table** foram renomeadas para substituir espaços por sublinhados (_).
- ▶ Nas duas tabelas foram renomeadas colunas para nomes mais claros.
- ▶ Foi também necessária uma alteração de tipo de dados para o tipo mais adequado na tabela **Gender_Stats_Data_table**.

Filtragem de Dados:

- ▶ Foi removida uma coluna sem informação em **Gender_Stats_Series_table**.

Adição de Chaves Artificiais:

- ▶ Foi adicionada uma coluna "Gender_Stats_Data_ID" e "Gender_Stats_Series_ID" para a tabela respetiva com valores sequenciais formatados.

Uniformização de Faixas Etárias:

- ▶ As faixas etárias foram uniformizadas na tabela **Gender_Stats_Data_Table** para haver coerência e uniformização com os restantes Datasets (criação de uma coluna com a faixa etária uniformizada).

Uniformização do Género:

- ▶ Foi criada uma coluna na tabela **Gender_Stats_Data_Table** referente ao género ao qual se referia cada **Indicator_Name** para haver coerência e uniformização com os restantes Datasets.

Preenchimento de Linhas Vazias:

- ▶ Na tabela **Gender_Stats_Series_table**, foram substituídas as linhas em branco pela frase "Criteria Not Available".

Tabelas Silver resultantes

Figura 8 - Tabela “Gender_Stats_Data_table”

Figura 9 - Tabela “Gender Stats Series table”

Transformações Silver-Gold

Transformação num esquema em estrela:

- ▶ Criação de um esquema em estrela.
- ▶ Para isso foi necessária a criação e a integração de chaves artificiais das tabelas de dimensão na tabela de factos.

O esquema resultante consiste em:

- ▶ Uma tabela de factos acerca dos vários valores pertencentes a cada indicador diferente, sendo esses valores somas ou médias.
- ▶ Cinco dimensões (Country, Temporal Measurement, Sources, Measurement Methods e Gender Series). A dimensão **Country** apresenta os vários países e regiões do mundo ao qual pertencem os indicadores, a **Temporal Measurement** refere-se à base temporal em que os dados foram medidos (ex: Annual), a **Sources** refere-se à fonte de dados original ao qual cada indicador foi retirado, **Measurement Methods** indica qual foi o método de agregação usado (ex: Sum) e a **Gender Series** refere-se aos detalhes de cada indicador como o seu nome, uma pequena definição, etc.

Transformações para cálculos futuros:

- ▶ Foi necessária a realização de uma pivotagem das colunas referentes a cada ano. Estas colunas referiam- se às medições realizadas num ano específico, no entanto para poderem serem avaliadas e ser possível efetuar medidas e comparações através de gráfico, estas foram transformadas em duas colunas: uma referente ao ano e outra ao valor a ele associado.
- ▶ Foram realizadas também filtragens em relação aos indicadores a serem utilizados e ao intervalo de anos necessário. Esta filtragem permite que os dados sejam guardados de forma eficiente nas respetivas tabelas a que pertencem pois devido à existência de uma grande quantidade de linhas resultante da pivotagem, o desempenho era afetado.

Tabelas Gold resultantes

```

spark.sql(
    """
    SELECT *
    FROM Projeto.dim_country
    """
).show(10)
+-----+-----+-----+
|Country_Dimension_ID|Country_Region_Situation_Name|Country_Region_Situation_Code|
+-----+-----+-----+
|      CD0|           Afghanistan|                  AFG|
|      CD1|        Africa Eastern an...|                 AFE|
|      CD2|        Africa Western an...|                 AFW|
|      CD3|                  Albania|                 ALB|
|      CD4|                  Algeria|                 DZA|
|      CD5|        American Samoa|                 ASN|
|      CD6|                  Andorra|                 AND|
|      CD7|                  Angola|                 AGO|
|      CD8|Antigua and Barbuda|                 ATG|
|      CD9|          Arab World|                 ARB|
|     CD10|             Argentina|                 ARG|
|     CD11|             Armenia|                 ARM|
|     CD12|              Aruba|                 ABW|
|     CD13|            Australia|                 AUS|
|     CD14|             Austria|                 AUT|
|     CD15|           Azerbaijan|                 AZE|
|     CD16|        Bahamas, The|                 BHS|
|     CD17|             Bahrain|                 BHR|
|     CD18|           Bangladesh|                 BGD|
|     CD19|            Barbados|                 BRB|
+-----+-----+-----+
only showing top 20 rows

```

```
spark.sql(  
    ....  
    SELECT *  
    FROM Projeto.dim_temporal_measurement  
    ....  
)  
.show()  
  
+-----+-----+-----+  
|Temporal_Measurement_ID|Periodicity|Base_Period|  
+-----+-----+-----+  
|                      TMD0|      Annual|Criteria Not Avai...|  
+-----+-----+-----+
```

```

spark.sql(
    """
    SELECT *
    FROM Projeto.dim_sources
    """
).show()

```

Sources_Dimension_ID	Notes_from_original_source	Source	Related_source_links	Other_web_links
SD0	Criteria Not Avai... Criteria Not Avai... Criteria Not Avai... Criteria Not Avai...			
SD1	Criteria Not Avai... Estimates develop... Criteria Not Avai... Criteria Not Avai...			
SD2	Criteria Not Avai... (1) United Nation... and adult mortal... Mortality rates f...			
SD3	Criteria Not Avai... World Bank staff ... Criteria Not Avai... Criteria Not Avai...			
SD4	Criteria Not Avai... United Nations Po... Criteria Not Avai... Criteria Not Avai...			

```
spark.sql(  
    """  
    SELECT *  
    FROM Projeto.dim_measurement_methods  
    """  
)  
.show()  
  
+-----+-----+-----+  
|Measurement_Methods_ID|Unit_of_measure|Aggregation_method|  
+-----+-----+-----+  
|MMD0|Criteria Not Avai...|Weighted average|  
|MMD1|Criteria Not Avai...|Criteria Not Avai...|  
|MMD2|Criteria Not Avai...|Sum|  
+-----+-----+-----+
```

```

spark.sql(
    """
    SELECT *
    FROM Projeto.gender_data
    """
).show()
+-----+-----+-----+-----+-----+-----+
|Gender_Series_ID|Country_Dimension_ID|Temporal_1_Measurement_ID|Sources_Dimension_ID|Measurement_Methods_ID|Indicator_Code|Indicator_Number|
+-----+-----+-----+-----+-----+-----+
|G542949672960|CD025|TND01|SD01|NH001|VC_IHR_PSRC_PS|3.29419604|
|G542949672961|CD025|TND01|SD01|NH001|VC_IHR_PSRC_PS|3.225055|
|G542949672962|CD025|TND01|SD01|NH001|VC_IHR_PSRC_PS|null|
|G542949672963|CD025|TND01|SD01|NH001|VC_IHR_PSRC_PS|null|
|G542949672964|CD025|TND01|SD01|NH001|VC_IHR_PSRC_PS|null|
|G542949672965|CD025|TND01|SD01|NH001|VC_IHR_PSRC_PS|null|
|G542949672966|CD025|TND01|SD01|NH001|VC_IHR_PSRC_PS|null|
|G542949672967|CD025|TND01|SD01|NH001|VC_IHR_PSRC_PS|null|
|G542949672968|CD025|TND01|SD01|NH001|VC_IHR_PSRC_PS|null|
|G542949672969|CD025|TND01|SD01|NH001|VC_IHR_PSRC_PS|null|
|G542949672970|CD025|TND01|SD01|NH001|VC_IHR_PSRC_PS|null|
|G542949672971|CD025|TND01|SD01|NH001|VC_IHR_PSRC_PS|null|
|G542949672972|CD025|TND01|SD01|NH001|VC_IHR_PSRC_PS|null|
|G542949672973|CD025|TND01|SD01|NH001|[VC_IHR_PSRC_FE_PS]|0.0|
|G542949672974|CD025|TND01|SD01|NH001|[VC_IHR_PSRC_FE_PS]|0.0|
|G542949672975|CD025|TND01|SD01|NH001|[VC_IHR_PSRC_FE_PS]|null|
|G542949672976|CD025|TND01|SD01|NH001|[VC_IHR_PSRC_FE_PS]|null|
|G542949672977|CD025|TND01|SD01|NH001|[VC_IHR_PSRC_FE_PS]|null|
|G542949672978|CD025|TND01|SD01|NH001|[VC_IHR_PSRC_FE_PS]|null|
|G542949672979|CD025|TND01|SD01|NH001|[VC_IHR_PSRC_FE_PS]|null|
+-----+-----+-----+-----+-----+-----+
only showing top 20 rows

```

Crime_Dataset

Transformações Bronze-Silver

Renomeação de nomes de colunas e tratamento de dados:

- ▶ Foi feita uma renomeação de colunas para nomes mais claros e percetíveis.
- ▶ Substituição de espaços por “_” nos nomes das colunas em todos os três Dataframes para consistência e facilidade de manipulação.
- ▶ Foi feita também uma correção de um tipo de dados.

Adição de Chaves Artificiais:

- ▶ Foi adicionada uma coluna a cada tabela que contém Id's de chave artificial

Criação de uma coluna de locais aproximados

- ▶ Tendo em conta os dados presentes numa outra coluna que continha os locais próximos de onde tinha sido efetuado um crime (ex: near of supermarket), foi criada uma nova coluna que contém os locais de uma forma mais clara para um estudo futuro (ex: supermarket).

Extração da localização:

- ▶ Adição de uma nova coluna 'Location', baseada no nome do arquivo, para identificar a localização específica dos crimes. Uma vez que o dataset inclui múltiplos ficheiros CSV (cada um representando uma localização diferente), esta coluna torna-se essencial para integrar na tabela os dados de localização, que anteriormente estavam disponíveis apenas nos nomes dos ficheiros de origem.

Uniformização de Faixas Etárias:

- ▶ As faixas etárias foram uniformizadas em ambas as tabelas para haver coerência e uniformização com os restantes Datasets (criação de uma coluna com a faixa etária uniformizada).

Tabelas Silver resultantes

Figura 10 - Figura 12 – Tabela “Stop_and_search_table”

	CrimeID	Reported_by	Falls_within	Longitude	Latitude	Near_of[LSOA_code]	LSOA_name	Outcome_type	Year	Month	Location	Crime_ID
	25769883777	[Metropolitan Poli...]	[Metropolitan Poli...]	-0.128236	51.481544	0m or near Petrol...	[E01032207]	Lambeth 004F	Suspect charged	2020	12/metropolitan[ba5f3addeff1766...]	
	25769883778	[Metropolitan Poli...]	[Metropolitan Poli...]	-0.166708	51.407854	0m or near SADLER...	[E01033399]	Merton 018D	Suspect charged	2020	12/metropolitan[81fc3fb0ade373b65...]	
	25769883779	[Metropolitan Poli...]	[Metropolitan Poli...]	-0.366375	51.588039	0m or near STATION...	[E01034499]	Ealing 031C	Investigation com...	2020	12/metropolitan[13333333333333333...]	
	25769883780	[Metropolitan Poli...]	[Metropolitan Poli...]	0.116698	51.552601	null	[E01035599]	No location null	null	null	null	12/metropolitan[0075-37a4b0f786cc...
	25769883781	[Metropolitan Poli...]	[Metropolitan Poli...]	-0.349777	51.552600	0m or near WHITTO...	[E01031399]	Ealing 001C	Investigation com...	2020	12/metropolitan[8096c343ba3b94e33...]	
	25769883782	[Metropolitan Poli...]	[Metropolitan Poli...]	-0.105566	51.564585	0m or near STATIO...	[E0102734]	Islington 005B	Suspect charged	2020	12/metropolitan[61a35963ca57e0c1...]	
	25769883783	[Metropolitan Poli...]	[Metropolitan Poli...]	-0.082698	51.471887	0m or near VESTRY...	[E0103914]	Southwark 024B	Investigation com...	2020	12/metropolitan[60ba54f7ff4fc9b946...]	
	25769883785	[Metropolitan Poli...]	[Metropolitan Poli...]	-0.097768	51.577885	0m or near PORTLA...	[E01026247]	Haringey 027C	Investigation com...	2020	12/metropolitan[6d3b6a3ea7b9cb6c...]	
	25769883787	[Metropolitan Poli...]	[Metropolitan Poli...]	0.116698	51.540538	0m or near BARKIN...	[E01008072]	Barking and Dagen...	Investigation com...	2020	12/metropolitan[c7b21985671723a...]	
	25769883788	[Metropolitan Poli...]	[Metropolitan Poli...]	0.116698	51.540538	0m or near BARKIN...	[E01008072]	Barking and Dagen...	Investigation com...	2020	12/metropolitan[f679d8ba575878970...]	
	25769883789	[Metropolitan Poli...]	[Metropolitan Poli...]	0.116698	51.540538	0m or near BARKIN...	[E01008072]	Barking and Dagen...	Investigation com...	2020	12/metropolitan[cd43cfa54981bb0...	
	25769883790	[Metropolitan Poli...]	[Metropolitan Poli...]	0.116698	51.540538	0m or near BARKIN...	[E01008072]	Barking and Dagen...	Investigation com...	2020	12/metropolitan[008c5e303499f482z...]	
	25769883791	[Metropolitan Poli...]	[Metropolitan Poli...]	0.116698	51.540538	0m or near BARKIN...	[E01008072]	Barking and Dagen...	Investigation com...	2020	12/metropolitan[008c5e303499f482z...]	
	25769883792	[Metropolitan Poli...]	[Metropolitan Poli...]	0.116698	51.540538	0m or near BARKIN...	[E01008072]	Barking and Dagen...	Investigation com...	2020	12/metropolitan[008c5e303499f482z...]	
	25769883793	[Metropolitan Poli...]	[Metropolitan Poli...]	0.116698	51.540538	0m or near BARKIN...	[E01008072]	Barking and Dagen...	Investigation com...	2020	12/metropolitan[ad43ea407767ccbd...	
	25769883794	[Metropolitan Poli...]	[Metropolitan Poli...]	0.034049	51.538831	0m or near PARKIN...	[E01035566]	Newham 014D	Investigation com...	2020	12/metropolitan[f30f1a4d1f87e184...	
	25769883795	[Metropolitan Poli...]	[Metropolitan Poli...]	0.116698	51.540538	0m or near BARKIN...	[E01008072]	Barking and Dagen...	Investigation com...	2020	12/metropolitan[18a28aa1b9c5e076...]	
	25769883796	[Metropolitan Poli...]	[Metropolitan Poli...]	0.116698	51.540538	0m or near BARKIN...	[E01008072]	Barking and Dagen...	Investigation com...	2020	12/metropolitan[5c9743340f87a7f04...	

Figura 11 - Tabela “Outcomes_table”

CrimeID	Time	Reported_by	Falls_within	LSOA_code	LSOA_name	Crime_type	Last_outcome_category	Context	Year	Month	Location	Crime_ID
19756849517	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.147739	[51.347410]Bath and North Ea...	Violence and sexua...	Status update una...	null	2021	8	Avon and somerset 1646538808d08e8993...	
19756849518	2021-08-08	Avon and Somerset...	Avon and Somerset...	-2.509930	[51.310873]Bath and North Ea...	[Anti-social behav...	null	null	2021	8	Avon and somerset	null
19756849519	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.249870	[51.22276]Bath and North Ea...	[Anti-social behav...	null	null	2021	8	Avon and somerset	
19756849520	2021-08-08	Avon and Somerset...	Avon and Somerset...	-2.516919	[51.23683]Bath and North Ea...	or near A475	Further action is...	null	2021	8	Avon and somerset [2672363e84c1f813...	
19756849521	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.249870	[51.22276]Bath and North Ea...	or near Confer...	Other theft Further action is...	null	2021	8	Avon and somerset [933a78658e854...	
19756849522	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.249870	[51.22276]Bath and North Ea...	or near Confer...	Public order Status update una...	null	2021	8	Avon and somerset [32206a0040f101733...	
19756849523	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.249870	[51.22276]Bath and North Ea...	or near Confer...	Public order Formal action is...	null	2021	8	Avon and somerset [32206a0040f101733...	
19756849524	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.249870	[51.22276]Bath and North Ea...	[Violence and sexu...	Status update una...	null	2021	8	Avon and somerset aa27259f8ae27f58...	
19756849525	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.249870	[51.22276]Bath and North Ea...	[Violence and sexu...	Unable to prosecu...	null	2021	8	Avon and somerset aa17c9e3d4a2b26...	
19756849526	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.249870	[51.22276]Bath and North Ea...	[Violence and sexu...	Further action is...	null	2021	8	Avon and somerset [ee61562b0120e@0f...	
19756849527	2021-08-08	Avon and Somerset...	Avon and Somerset...	-2.515305	[51.404171]Bath and North Ea...	[Anti-social behav...	Unable to prosecu...	null	2021	8	Avon and somerset [f963703cb093cc5c...	
19756849528	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.249870	[51.22276]Bath and North Ea...	[Violence and sexu...	Unable to prosecu...	null	2021	8	Avon and somerset [b3d999f8e659b0...	
19756849529	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.249870	[51.22276]Bath and North Ea...	[Violence and sexu...	Formal action is...	null	2021	8	Avon and somerset [06544b103d12697...	
19756849530	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.2499551	[51.318662]Bath and North Ea...	[Anti-social behav...	Unable to prosecu...	null	2021	8	Avon and somerset	null
19756849531	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.2498613	[51.160802]Bath and North Ea...	[Anti-social behav...	null	null	2021	8	Avon and somerset	
19756849532	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.2498613	[51.160802]Bath and North Ea...	[Anti-social behav...	null	null	2021	8	Avon and somerset	
19756849533	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.2498613	[51.160802]Bath and North Ea...	[Anti-social behav...	null	null	2021	8	Avon and somerset	
19756849534	2021-08-08	Avon and Somerset...	Avon and Somerset...	-1.2498613	[51.160802]Bath and North Ea...	[Anti-social behav...	Burglary Status update una...	null	2021	8	Avon and somerset	null
19756849535	2021-08-08	Avon and Somerset...	Avon and Somerset...	-2.499922	[51.417373]Bath and North Ea...	[Anti-social behav...	Burglary Formal action is...	null	2021	8	Avon and somerset [3d0e188e1641f163...	
19756849536	2021-08-08	Avon and Somerset...	Avon and Somerset...	-2.508319	[51.417573]Bath and North Ea...	[Anti-social behav...	Burglary Status update una...	null	2021	8	Avon and somerset 44514f8455cd7df5...	

Figura 12 - Tabela “Street_table”

Transformações Silver-Gold

Transformação num esquema em estrela:

- ▶ Criação de um esquema em estrela, onde grande parte das dimensões são conformes.
- ▶ Para isso foi necessária uma integração das chaves artificiais das tabelas de dimensão nas tabelas de factos.
- ▶ O código utiliza a função window do Spark para gerar identificadores sequenciais para várias dimensões degeneradas. Especificamente, cria IDs únicos para dimensões como data, local, pessoa e operação, com base em atributos ordenados como ano, latitude/longitude, gênero/categoria de idade e tipo, respectivamente. Posteriormente, essas dimensões são unidas para formar uma tabela de fatos, que centraliza dados de várias fontes mantendo relações através dos IDs gerados.

O esquema resultante consiste em:

- ▶ Uma única tabela de factos (Sobre as Operações Stop and search)
- ▶ Quatro dimensões degeneradas (Local, Data, Operation e Pessoa)

Transformações na coluna “Location”:

- ▶ Nesta última fase de GOLD, foram eliminadas do dataset linhas com as seguintes locations: 'west midlands', 'thames valley', 'south wales', 'cheshire', 'dyfed powys', 'west mercia', 'gwent', 'humberside', 'north wales', 'devon and cornwall', com o objetivo de juntamente com os restantes elementos do grupo estudar as mesmas localidades.

Preparação de dados para consulta tendo em conta os nomes de localidades no Shapefile transferido:

- ▶ Alteração dos nomes das colunas de forma a corresponderem exatamente com os nomes dos Boroughs do Shapefile.

Tabelas Gold resultantes

Dim_data

Data	Year	Month	id_data
null	null	1	
2017-05-26	2017	5	2
2017-10-17	2017	10	3
2017-05-16	2017	5	4
2017-05-31	2017	5	5
2017-04-30	2017	4	6
2017-05-29	2017	5	7
2017-05-05	2017	5	8
2017-05-10	2017	5	9
2017-05-02	2017	5	10
2017-05-06	2017	5	11
2017-05-23	2017	5	12
2017-05-17	2017	5	13
2017-05-27	2017	5	14
2017-05-24	2017	5	15
2017-05-30	2017	5	16
2017-05-11	2017	5	17
2017-05-03	2017	5	18
2017-05-21	2017	5	19
2017-10-03	2017	10	20

dim_local

Latitude	Longitude	Location	id_local
null	null	lancashire	1
null	null	avon and somerset	2
null	null	west yorkshire	3
null	null	south yorkshire	4
null	null	merseyside	5
null	null	essex	6
null	null	hampshire	7
null	null	surrey	8
null	null	cleveland	9
null	null	northumbria	10
null	null	cumbria	11
null	null	bedfordshire	12
null	null	norfolk	13
null	null	leicestershire	14
null	null	warwickshire	15
null	null	cambridgeshire	16
null	null	dorset	17
null	null	wiltshire	18
null	null	metropolitan	19
null	null	kent	20

dim_pessoa

Gender	Age_category	Self-defined_ethnicity	Officer-defined_ethnicity	id_pessoa
0-24	Other ethnic group	White	Black	1
0-24	White - Any other...	White	Black	2
0-24	Other ethnic group	Other	Black	3
0-24	Asian/Asian Briti...	Asian	Black	4
0-24	Black/African/Cari...	Black	Black	5
0-24	Black/African/Cari...	Mixed/Multiple et...	Black	6
0-24	White - Any other...	White	Black	7
0-24	White - Gypsy or ...	White	White	8
0-24	Mixed/Multiple et...	White	White	9
0-24	Mixed/Multiple et...	null	White	10
0-24	Asian/Asian Briti...	Asian	Asian	11
0-24	Other ethnic group	Black	Black	12
0-24	White - Irish	White	White	13
0-24	Other ethnic group	Asian	Asian	14
0-24	Mixed/Multiple et...	Asian	Asian	15
0-24	Asian/Asian Briti...	White	White	16
0-24	Mixed/Multiple et...	null	White	17
0-24	Black/African/Cari...	Other	Other	18
0-24	null	null	White	19
0-24	Mixed/Multiple et...	null	White	20

Dim_operation

Type	Legislation	Object_of_search	Outcome	id_operation
Person and Vehic...	null	Stolen goods	Arrest	1
Person and Vehic...	Police and Crimin...	Controlled drugs	Community resolution	2
Person and Vehic...	Firearms Act 1968...	Firearms	Penalty Notice fo...	3
Person and Vehic...	Misuse of Drugs A...	Evidence of offen...	A no further acti...	4
Person and Vehic...	Criminal Justice ...	Evidence of offen...	A no further acti...	5
Person and Vehic...	Psychoactive Subs...	Controlled drugs	Community resolution	6
Person and Vehic...	null	Stolen goods	A no further acti...	7
Person and Vehic...	Aviation Security...	null	A no further acti...	8
Person and Vehic...	Misuse of Drugs A...	Evidence of offen...	null	9
Person and Vehic...	Criminal Justice ...	null	Community resolution	10
Person and Vehic...	Criminal Justice ...	Offensive weapons	Community resolution	11
Person and Vehic...	null	Evidence of offen...	Community resolution	12
Person and Vehic...	Misuse of Drugs A...	Controlled drugs	Caution (simple o...	13
Person and Vehic...	null	null	Community resolution	14
Person and Vehic...	Police and Crimin...	Article for use i...	Community resolution	15
Person and Vehic...	Police and Crimin...	Articles for use i...	Suspect arrested	16
Person and Vehic...	Misuse of Drugs A...	Article for use i...	Arrest	17
Person and Vehic...	Police and Crimin...	Offensive weapons	A no further acti...	18
Person and Vehic...	Aviation Security...	Controlled drugs	A no further acti...	19
Person and Vehic...	Police and Crimin...	Stolen goods	Caution (simple o...	20

factos

id_data	id_local	id_pessoa	id_operation
420	38	1504	1059
420	38	1504	1059
2037	40	1152	1082
156	43	946	48
156	43	1504	545
156	43	1504	545
789	49	1412	557
1753	50	1424	599
1753	51	1424	34
1748	57	1048	599
1748	57	1048	599
1748	57	1020	599
1748	57	1020	599
616	60	946	552
1309	61	1139	599
1342	61	1424	599
1110	61	946	599
1940	61	1424	552
1187	62	1139	1059
2226	63	1142	599

Homicide_Victim

Transformações Bronze-Silver

Renomeação de Colunas e Correção de tipos de dados:

- ▶ As colunas de ambas as tabelas foram renomeadas para substituir espaços por sublinhados (_).
- ▶ Nas duas tabelas foram renomeadas colunas para nomes mais claros.
- ▶ Foi também necessária uma alteração de tipo de dados para o tipo mais adequado.

Filtragem de Dados:

- ▶ Foram removidas linhas sem informação sobre o género tanto em **Homicide_Victim_table** quanto em **Homicide_Proceeded_table**.

Tratamento de Datas:

- ▶ As colunas de data foram divididas em colunas separadas para mês e ano em ambas as tabelas, criando assim novas colunas para ano e mês.
- ▶ Foram corrigidos valores que poderiam dar origem a erros (ex: "Female" deveria ser apenas "Female", visto que o espaço pode dar origem a erros de distintos no futuro).

Adição de Chaves Artificiais:

- ▶ Foi adicionada uma coluna "Homicide_Victim_ID" e "Homicide_Proceeded_ID" para a tabela respetiva com valores sequenciais formatados

Uniformização de Localização:

- ▶ Em ambas as tabelas, foi adicionada a coluna "Location" com o valor "City of London" para coerência com o **Crime_Dataset**.

Uniformização de Faixas Etárias:

- ▶ As faixas etárias foram uniformizadas em ambas as tabelas para haver coerência e uniformização com os restantes Datasets (criação de uma coluna com a faixa etária uniformizada).

Tabelas Silver resultantes

Homicide_Victim_Id	Count_of_Victims	Age_Group	Victim_Gender	Method_of_Killing	Domestic_Abuse	Recorded_Date	Homicide_Offence_Type	Solved_Status	Borough	Officer_Observed_Ethnicity	Recorded_Month	Recorded_Year	Location	Uniform_Age_Group
HV128	1	0 to 12	Male	Other Methods of ... Not Domestic Abuse	Dec-03	Murder Solved Sutton	White	12	2003 City of London Youth 0-24 years					
HV281	1	13 to 19	Male	Knife or Sharp Im...	May-04	Murder Solved Sutton	White	5	2004 City of London Adult 35+ years					
HV698	1	35 to 44	Male	Shooting Not Domestic Abuse	Dec-06	Murder Solved Sutton	White	12	2006 City of London Adult 35+ years					
HV699	1	0 to 12	Female	Blunt Implement Not Domestic Abuse	Dec-06	Murder Solved Sutton	Black	12	2006 City of London Youth 0-24 years					
HV855	1	55 to 64	Male	[Not Known/Not Rec... Not Domestic Abuse	Nov-07	Murder Solved Sutton	White	11	2007 City of London Adult 35+ years					
HV856	1	35 to 44	Male	Knife or Sharp Im...	Nov-07	Murder Solved Sutton	Asian	5	2008 City of London Youth 0-24 years					
HV103	1	13 to 19	Male	Knife or Sharp Im...	May-08	Murder Solved Sutton	White	9	2008 City of London Adult 35+ years					
HV1022	1	0 to 12	Female	Knife or Sharp Im...	May-08	Murder Solved Sutton	Asian	5	2008 City of London Youth 0-24 years					
HV1023	1	165 and over!	Male	[Physical Assault,... Not Domestic Abuse	Sep-08	Murder Solved Sutton	White	3	2009 City of London Adult 35+ years					
HV1159	1	165 and over!	Female	[Physical Assault,... Not Domestic Abuse	Mar-09	Murder Solved Sutton	White	8	2009 City of London Adult 25-34 years					
HV1160	1	25 to 34	Male	Knife or Sharp Im...	Aug-09	Murder Solved Sutton	White	11	2009 City of London Adult 25-34 years					
HV1161	1	25 to 34	Male	[Physical Assault,... Not Domestic Abuse	Nov-09	Murder Solved Sutton	White	4	2010 City of London Adult 35+ years					
HV1282	1	35 to 44	Male	Knife or Sharp Im...	Apr-10	Murder Solved Sutton	White	5	2010 City of London Youth 0-24 years					
HV1283	1	13 to 19	Male	Knife or Sharp Im...	May-10	Murder Solved Sutton	White	12	2010 City of London Adult 35+ years					
HV1284	1	165 and over!	Female	Knife or Sharp Im...	Dec-10	Murder Solved Sutton	White	3	2011 City of London Adult 35+ years					
HV1411	1	35 to 44	Female	Knife or Sharp Im...	Mar-11	Murder Solved Sutton	White	6	2012 City of London Adult 35+ years					
HV1516	1	35 to 44	Male	Knife or Sharp Im...	Jun-12	Murder Solved Sutton	White	8	2012 City of London Adult 35+ years					
HV1517	1	45 to 54	Female	Knife or Sharp Im...	Aug-12	Murder Solved Sutton	White	10	2013 City of London Youth 0-24 years					
HV1613	1	0 to 12	Female	[Not Known/Not Rec... Not Domestic Abuse	Oct-13	Murder Solved Sutton	White	12	2016 City of London Adult 35+ years					
HV1954	1	55 to 64	Male	[Not Known/Not Rec... Not Domestic Abuse	Dec-16	Murder Unsolved Sutton	White							

Figura 11– Tabela “Homicide_Victim_table”

Homicide_Proceeded_ID	People_Accused_Count	Borough	Age_Group	Agressor_Gender	Self_Classified_Ethnicity	Offence_Type	Charged_Summons	Proceedings_Date	Proceedings_Month	Proceedings_Year	Location	Uniform_Age_Group	
HP60	1	Harrow	20 to 24	Male	Black	Murder charge or summons	Apr-03	4	2003 City of London Youth 0-24 years				
HP205	1	Harrow	25 to 29	Male	White	Murder charge or summons	Sep-03	9	2003 City of London Adult 35+ years				
HP245	1	Harrow	15 to 19	Male	White	Murder charge or summons	Nov-03	11	2003 City of London Adult 35+ years				
HP671	1	Harrow	35 to 39	Male	White	Murder charge or summons	Jul-05	7	2005 City of London Adult 35+ years				
HP728	1	Harrow	20 to 24	Male	Black	Murder charge or summons	Aug-05	8	2005 City of London Youth 0-24 years				
HP729	1	Harrow	15 to 19	Male	Black	Murder charge or summons	Aug-05	8	2005 City of London Adult 35+ years				
HP730	1	Harrow	20 to 24	Male	White	Murder charge or summons	Aug-05	8	2005 City of London Youth 0-24 years				
HP981	1	Harrow	30 to 34	Male	Asian	Murder charge or summons	Sep-06	9	2006 City of London Adult 35+ years				
HP1010	1	Harrow	20 to 24	Male	White	Murder charge or summons	Dec-06	12	2006 City of London Youth 0-24 years				
HP1404	1	Harrow	35 to 39	Male	White	Murder charge or summons	Mar-08	3	2008 City of London Adult 35+ years				
HP1638	1	Harrow	20 to 24	Male	Black	Murder charge or summons	Mar-09	3	2009 City of London Youth 0-24 years				
HP1639	1	Harrow	30 to 34	Male	Black	Murder charge or summons	Mar-09	3	2009 City of London Adult 35+ years				
HP1640	1	Harrow	15 to 19	Male	Asian	Murder charge or summons	Apr-09	4	2009 City of London Adult 35+ years				
HP1641	1	Harrow	35 to 39	Male	Black	Murder charge or summons	May-09	5	2009 City of London Adult 35+ years				
HP1780	1	Harrow	50 to 54	Male	White	Murder charge or summons	Nov-09	11	2009 City of London Adult 35+ years				
HP1809	1	Harrow	35 to 39	Male	Person declines t... Manslaughter charge or summons						2011 City of London Adult 35+ years		
HP1833	1	Harrow	30 to 34	Male	Asian	Murder charge or summons	Jan-10	1	2010 City of London Adult 35+ years				
HP1850	1	Harrow	60 to 64	Female	Person declines t... Murder charge or summons						2011 City of London Adult 35+ years		
HP1851	1	Harrow	35 to 39	Female	Mixed	Murder charge or summons	Jul-11	7	2011 City of London Adult 35+ years				
HP1972	1	Harrow	25 to 29	Male	Person declines t...	Murder charge or summons	Nov-10	11	2010 City of London Adult 35+ years				

Figura 12 – Tabela “Homicide_Proceeded_table”

Transformações Silver-Gold

Transformação num esquema em estrela:

- ▶ Criação de um esquema em estrela, onde grande parte das dimensões são conformes.
- ▶ Para isso foi necessária uma integração das chaves artificiais das tabelas de dimensão nas tabelas de factos.

O esquema resultante consiste em:

- ▶ Duas tabelas de factos (Sobre vítimas de homicídio e agressores)
- ▶ 3 dimensões degeneradas (Tempo, Local e Data).
- ▶ 1 dimensão apenas pertencente à tabela de factos de homicídio (com dados com o tipo de homicídio).

Transformações para cálculos futuros:

- ▶ Criação de colunas derivadas com base nas informações existentes, tendo em conta uma futura análise numérica. Por exemplo: cálculo da percentagem de casos relativamente ao total (homicídio/número total de homicídios) * 100, cujo somatório reflete a percentagem de casos com os filtros aplicados em relação ao total.

Preparação de dados para consulta tendo em conta os nomes de localidades no Shapefile transferido:

- ▶ Alteração dos nomes das colunas de forma a corresponderem exatamente com os nomes dos Boroughs do Shapefile.
- ▶

Tabelas Gold resultantes

Date_df.show()		
Year	Month	Id_Date
2022	10	0
2005	5	1
2007	6	2
2012	10	3
2010	12	4
2010	7	5
2015	2	6
2019	10	7
2017	3	8
2008	8	9
2017	8	10
2004	6	11
2009	11	12
2014	4	13
2005	10	14
2020	6	15
2019	5	16
2017	10	17
2018	10	18
2021	8	19

Location_df.show()		
Location	Borough	Id_Location
London	Greenwich	0
London	Tower Hamlets	1
London	Kensington and Ch...	2
London	Hackney	3
London	Hillingdon	4
London	Merton	5
London	Hammersmith and F...	6
London	Croydon	7
London	Harrow	8
London	Richmond upon Thames	9
London	Southwark	10
London	Hounslow	11
London	Lewisham	12
London	Redbridge	13
London	Newham	14
London	Westminster	15
London	Sutton	16
London	Bromley	17
London	Wandsworth	18
London	Havering	19

Person_df.show()			
Gender	Age_Group	Ethnicity	Id_Person
Male	Youth 0-24 years	Asian	0
Female	Adult 35+ years	Black	1
Male	Adult 35+ years	Asian	2
Male	Adult 35+ years	Not Reported/Not ...	3
Female	Adult 25-34 years	Not Reported/Not ...	4
Male	Adult 35+ years	Black	5
Female	Adult 35+ years	Not Reported/Not ...	6
Female	Youth 0-24 years	Black	7
Female	Youth 0-24 years	Asian	8
Male	Adult 25-34 years	Black	9
Female	Adult 35+ years	White	10
Male	Adult 35+ years	Other	11
Male	Youth 0-24 years	Not Reported/Not ...	12
Male	Youth 0-24 years	Other	13
Female	Adult 25-34 years	Other	14
Female	Adult 25-34 years	Asian	15
Male	Adult 35+ years	White	16
Female	Adult 25-34 years	White	17
Male	Adult 25-34 years	Asian	18
Female	Youth 0-24 years	White	19

Homicide_Facts_df.show()										
ID_Homicide	ID_Person	ID_Date	ID_Location	Solved_Status	Count_of_Victims	Method_of_Killing	Domestic_Abuse	Homicide_Offence_Type	Percentage_of_Total_Victims	
HV0	16	63	18	Solved	1	Knife or Sharp Im... Not Domestic Abuse	Murder	0.03544842254519674		
HV1	18	215	18	Solved	1	Not Known/Not Rec... Not Domestic Abuse	Murder	0.03544842254519674		
HV2	2	56	0	Solved	1	Physical Assault,... Not Domestic Abuse	Murder	0.03544842254519674		
HV3	16	56	20	Solved	1	Blunt Implement Not Domestic Abuse	Murder	0.03544842254519674		
HV4	19	215	11	Solved	1	Physical Assault,... Not Domestic Abuse	Murder	0.03544842254519674		
HV5	2	36	11	Solved	1	Knife or Sharp Im... Not Domestic Abuse	Murder	0.03544842254519674		
HV6	19	63	9	Solved	1	Physical Assault,... Not Domestic Abuse	Murder	0.03544842254519674		
HV7	18	22	9	Solved	1	Not Known/Not Rec... Not Domestic Abuse	Murder	0.03544842254519674		
HV8	29	194	9	Solved	1	Knife or Sharp Im... Not Domestic Abuse	Murder	0.03544842254519674		
HV9	0	63	4	Solved	1	Knife or Sharp Im... Not Domestic Abuse	Murder	0.03544842254519674		
HV10	17	123	4	Solved	1	Physical Assault,... Domestic Abuse	Murder	0.03544842254519674		
HV11	15	215	4	Solved	1	Physical Assault,... Domestic Abuse	Murder	0.03544842254519674		
HV12	10	148	4	Solved	1	Other Methods of ... Domestic Abuse	Murder	0.03544842254519674		
HV13	16	57	4	Solved	1	Knife or Sharp Im... Domestic Abuse	Murder	0.03544842254519674		
HV14	5	63	22	Solved	1	Not Known/Not Rec... Not Domestic Abuse	Murder	0.03544842254519674		
HV15	5	22	22	Unsolved	1	Shooting Not Domestic Abuse	Murder	0.03544842254519674		
HV16	1	22	22	Unsolved	1	Other Methods of ... Not Domestic Abuse	Murder	0.03544842254519674		
HV17	1	42	22	Unsolved	1	Other Methods of ... Not Domestic Abuse	Murder	0.03544842254519674		
HV18	5	42	22	Unsolved	1	Shooting Not Domestic Abuse	Murder	0.03544842254519674		
HV19	16	150	22	Unsolved	1	Other Methods of ... Not Domestic Abuse	Murder	0.03544842254519674		

Proceeded_Facts_df.show()							
ID_Proceeded	ID_Person	ID_Date	ID_Location	Offence_Type	People_Accused_Count	Charged_Summons	Percentage_of_Total_Proceeds
HP0	38	63	15	Murder	1 charge or summons	0.022988505747126436	
HP1	41	63	21	Murder	1 charge or summons	0.022988505747126436	
HP2	27	150	21	Murder	1 charge or summons	0.022988505747126436	
HP3	42	63	15	Murder	1 charge or summons	0.022988505747126436	
HP4	41	63	15	Murder	1 charge or summons	0.022988505747126436	
HP5	38	42	28	Murder	1 charge or summons	0.022988505747126436	
HP6	38	42	28	Murder	1 charge or summons	0.022988505747126436	
HP7	38	42	28	Murder	1 charge or summons	0.022988505747126436	
HP8	17	123	9	Murder	1 charge or summons	0.022988505747126436	
HP9	41	63	26	Murder	1 charge or summons	0.022988505747126436	
HP10	41	63	26	Murder	1 charge or summons	0.022988505747126436	
HP11	38	63	1	Murder	1 charge or summons	0.022988505747126436	
HP12	41	63	1	Murder	1 charge or summons	0.022988505747126436	
HP13	41	63	4	Murder	1 charge or summons	0.022988505747126436	
HP14	38	63	6	Murder	1 charge or summons	0.022988505747126436	
HP15	41	22	22	Murder	1 charge or summons	0.022988505747126436	
HP16	41	22	23	Murder	1 charge or summons	0.022988505747126436	
HP17	19	22	23	Murder	1 charge or summons	0.022988505747126436	
HP18	23	57	23	Murder	1 charge or summons	0.022988505747126436	
HP19	41	63	18	Murder	1 charge or summons	0.022988505747126436	

Child_Protection_Dataset

Transformações Bronze-Silver

Adição e Ordenação de Chave Artificial:

- Uma coluna ID artificial (Children_Careplan_Id) é adicionada e reordenada para ser a primeira coluna do DataFrame. Esta coluna é também convertida para o tipo inteiro.

Renomeação e Tratamento de Colunas:

- Alteração dos nomes de algumas colunas para nomes mais claros e intuitivos.
 - Alteração do nome de uma coluna do caracter & para a palavra "and".
 - Conversão dos valores 'N' encontrados na coluna "End Month & Year" para null.
 - Conversão dos valores em branco encontrados nas colunas "Category of Abuse", "Careplan End Reason" e para "Category Not Available" e "Criteria Not Available" respetivamente.
 - Todas as colunas da tabela são renomeadas para substituir espaços por sublinhados (_).

Filtragem de Dados:

- ▶ A coluna "Ethnic Code" é eliminada por ser possível encontrar os verdadeiros valores de A, B, C, D e E.

Uniformização de dados:

- ▶ Criação da coluna "Uniform_Age_Group", padronizando todas as idades tendo em conta os restantes datasets.
 - ▶ Criação da coluna "Borough" contendo a localização "Barnet" onde os dados foram registados e uniformizando com os datasets "Homicide_Victim" e "Homicide_Proceeded".

Criação e Renomeação de Colunas de Data:

- ▶ Criação de novas colunas para mês e ano a partir das colunas de data existentes e posteriormente renomeadas para seguir um padrão consistente.

bioRxiv preprint doi: <https://doi.org/10.1101/2023.08.15.550000>; this version posted August 15, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [aCC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

Figura 13 - Tabela "Child_protection_table"

Transformações Silver-Gold

Transformação num esquema em estrela:

- ▶ Criação de um esquema em estrela, onde grande parte das dimensões são conformes.
- ▶ Para isso foi necessária uma integração das chaves artificiais das tabelas de dimensão nas tabelas de factos.

O esquema resultante consiste em:

- ▶ Uma tabela de facto
- ▶ Três dimensões degeneradas (Date, Características da Criança, Características do Plano).
- ▶ Uma dimensão apenas pertencente à tabela de factos de homicídio (com dados com o tipo de homicídio).

Transformações para cálculos futuros:

- ▶ Criação de uma coluna com o valor de 1 para poder fazer as métricas no Tableau para posteriormente realizar as futuras análises necessárias.

cp_per_child_df.show()			
Date_Dimension_ID	Child_Dimension_ID	Care_Plan_Dimension_ID	Child_with_CP
DD1439	CD0	CPD4976	1
DD1871	CD1	CPD2177	1
DD1526	CD2	CPD4327	1
DD3508	CD3	CPD188	1
DD4337	CD4	CPD1075	1
DD2826	CD5	CPD714	1
DD3800	CD6	CPD1263	1
DD4112	CD7	CPD621	1
DD86	CD8	CPD11	1
DD3487	CD9	CPD4802	1
DD634	CD10	CPD676	1
DD926	CD11	CPD1145	1
DD3058	CD12	CPD2465	1
DD3765	CD13	CPD1461	1
DD757	CD14	CPD4493	1
DD4328	CD15	CPD3500	1
DD2185	CD16	CPD1089	1
DD3062	CD17	CPD2954	1
DD4989	CD18	CPD3010	1
DD1248	CD19	CPD250	1

only showing top 20 rows

dim_care_plan_df.show()					
Care_Plan_Dimension_ID	Children_Careplan_Id	Category_of_Abuse	Careplan_End_Reason	Current_Careplan_Protection	Borough
CPD0	49	Neglect	Criteria No Longe...	N	Barnet
CPD1	51	Neglect	Criteria No Longe...	N	Barnet
CPD2	443	Neglect	Criteria No Longe...	N	Barnet
CPD3	455	Neglect	Criteria No Longe...	N	Barnet
CPD4	1058	Category Not Avai...	Criteria No Longe...	N	Barnet
CPD5	1223	Emotional Abuse	Criteria No Longe...	N	Barnet
CPD6	1687	Neglect	Child & Family Pe...	N	Barnet
CPD7	1618	Neglect	Criteria No Longe...	N	Barnet
CPD8	1718	Neglect	Placed in Care	N	Barnet
CPD9	1914	Physical Abuse	Criteria No Longe...	N	Barnet
CPD10	2186	Neglect	Criteria Not Avai...	Y	Barnet
CPD11	2271	Category Not Avai...	Care Offering Pr...	N	Barnet
CPD12	2316	Neglect	Criteria No Longe...	N	Barnet
CPD13	2559	Emotional Abuse	Criteria No Longe...	N	Barnet
CPD14	2946	Neglect	Criteria No Longe...	N	Barnet
CPD15	3208	Neglect	Criteria No Longe...	N	Barnet
CPD16	3315	Sexual Abuse	Criteria No Longe...	N	Barnet
CPD17	3377	Physical Abuse	Criteria No Longe...	N	Barnet
CPD18	3408	Neglect	Criteria Not Avai...	Y	Barnet
CPD19	3527	Multiple: Neglect...	Criteria No Longe...	N	Barnet

only showing top 20 rows

dim_child_df.show()												
Child_Dimension_ID	Children_Careplan_Id	Gender	Disabled	Registration_Disabled	Unique	Current_Age_Group	Current_Age_Bracket	Age_Group_on_Registration	Age_Group_on_deRegistration	Uniform_Age_Group		
C00	145	Female	N	N	Y	16+	16+	1 to 5	6 to 10	6 to 10	Youth 0-24 years	
C01	379	Male	N	N	N	16+	16+	1 to 5	1 to 5	1 to 5	Youth 0-24 years	
C02	675	Male	N	N	N	16+	16+	1 to 5	6 to 10	6 to 10	Youth 0-24 years	
C03	763	Male	N	N	Y	16+	16+	1 to 5	6 to 10	6 to 10	Youth 0-24 years	
C04	782	Male	N	N	Y	16+	16+	6 to 10	6 to 10	6 to 10	Youth 0-24 years	
C05	916	Male	Y	N	N	16+	16+	6 to 10	6 to 10	6 to 10	Youth 0-24 years	
C06	4362	Female	N	N	N	16+	16+	6 to 10	6 to 10	6 to 10	Youth 0-24 years	
C07	1989	Male	M	N	Y	16+	16+	11 to 15	11 to 15	11 to 15	Youth 0-24 years	
C08	2271	Male	M	N	Y	16+	16+	11 to 15	11 to 15	11 to 15	Youth 0-24 years	
C09	2315	Male	M	N	Y	16+	16+	1 to 5	6 to 10	6 to 10	Youth 0-24 years	
C10	3087	Female	N	N	Y	11 to 15	11 to 15	1 to 5	1 to 5	1 to 5	Youth 0-24 years	
C11	3384	Female	N	N	Y	11 to 15	11 to 15	6 to 10	6 to 10	6 to 10	Youth 0-24 years	
C12	3607	Male	M	N	Y	16+	16+	11 to 15	11 to 15	11 to 15	Youth 0-24 years	
C13	3774	Male	M	N	Y	11 to 15	11 to 15	1 to 5	6 to 10	6 to 10	Youth 0-24 years	
C14	37712	Female	M	N	Y	16+	16+	6 to 10	6 to 10	6 to 10	Youth 0-24 years	
C15	4144	Male	M	N	Y	6 to 10	6 to 10	U1	U1	U1	Youth 0-24 years	
C16	4146	Male	M	N	Y	6 to 10	6 to 10	U1	U1	U1	Youth 0-24 years	
C17	4228	Female	M	N	Y	6 to 10	6 to 10	U1	U1	U1	Youth 0-24 years	
C18	4299	Female	Y	N	Y	6 to 10	6 to 10	6 to 10	6 to 10	6 to 10	Youth 0-24 years	
C19	4362	Male	M	N	Y	6 to 10	6 to 10	1 to 5	6 to 10	6 to 10	Youth 0-24 years	

dim_date_df.show()												
Date_Dimension_ID	Children_Careplan_Id	Data_Careplan_Start_month	Data_Careplan_Started_year	Data_Careplan_Ended_month	Data_Careplan_Ended_year	Start_Month_and_Year_month	Start_Month_and_Year_year	End_Month_and_Year_month	End_Month_and_Year_year			
D00	684	9	2003	2	2009	9	2003	2	2005	2	2005	
D01	755	1	2004	5	2004	1	2004	5	2004	5	2004	
D02	1028	9	1996	3	1997	9	1996	3	1997	3	1997	
D03	1323	1	2016	9	2016	1	2016	9	2016	9	2016	
D04	1449	5	2016	10	2017	5	2016	10	2017	10	2017	
D05	1523	3	2017	6	2018	3	2017	6	2018	6	2018	
D06	2241	6	2017	12	2017	6	2017	12	2017	12	2017	
D07	2386	8	1995	6	1996	8	1995	6	1996	6	1996	
D08	2524	5	2020	7	2020	5	2020	7	2020	7	2020	
D09	2658	12	2006	2	2007	12	2006	2	2007	2	2007	
D10	2888	2	2015	2	2015	2	2015	2	2015	2	2015	
D11	3150	7	2008	10	2008	7	2008	10	2008	7	2008	null
D12	3788	9	2017	6	2018	9	2017	6	2018	6	2018	
D13	3974	4	2012	6	2012	4	2012	6	2012	4	2012	
D14	4337	5	2015	9	2016	5	2015	9	2015	9	2016	
D15	4688	7	2016	1	2017	7	2016	1	2017	1	2017	
D16	4777	3	2008	8	2013	3	2008	8	2013	3	2013	
D17	4889	5	2017	12	2017	5	2017	12	2017	12	2017	
D18	4838	2	2015	3	2016	2	2015	3	2016	3	2016	
D19	4891	10	2001	9	2003	10	2001	9	2003	9	2003	

Dashboards

Gender_Stats_Dataset

Gender Data per Country

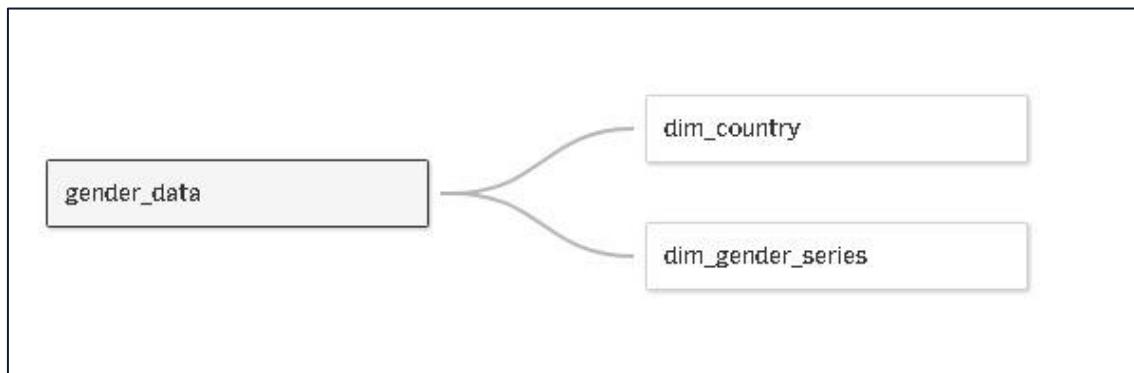


Figura 14 - Esquema de dados da Dashboard “Gender Data per Country”

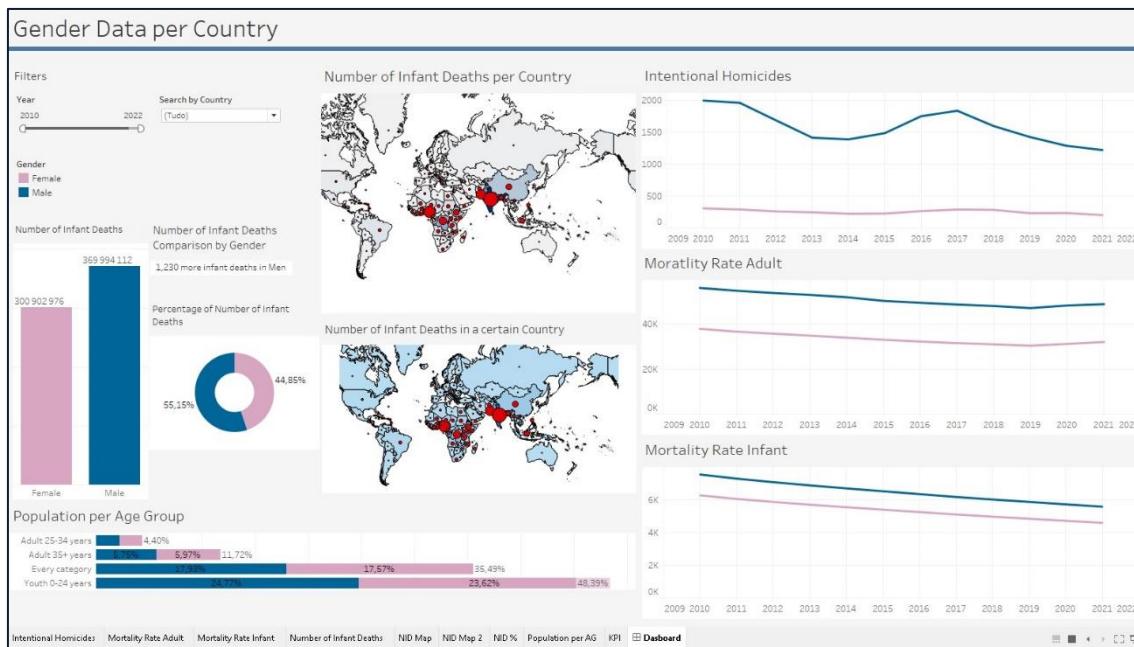


Figura 15 - Dashboard “Gender Data per Country”

Explicação da Dashboard “Gender Data per Country”

A Dashboard contém dados referentes a vários indicadores relacionados com género de vários países e regiões do mundo entre os anos 2010 e 2022.

Apesar da fonte de dados possuir originalmente cinco dimensões só foram necessárias duas, onde é possível obter os países, os indicadores, o género e os anos necessários à construção da dashboard de acordo com as questões analíticas e para a realização de filtragens.

Analizando a figura acima, podemos desde já retirar informações acerca de:

- ▶ Homicídios Intencionais
- ▶ Taxas de mortalidade quer para adulto como para jovens
- ▶ População distribuída por faixa etária
- ▶ O número e a respetiva percentagem de mortes de jovens por género
- ▶ Mapas que indicam o número de mortes de jovens totais em cada país (sendo o de cima usado como filtro)
- ▶ Uma medida que nos indica a razão entre o número de mortes de jovens em rapazes e raparigas

Tendo em conta que se trata de uma Dashboard interativa com ações editadas, esta é capaz de fornecer uma vasta gama de filtros para o utilizador final ser capaz de visualizar os dados de uma forma fácil. Assim, os filtros possíveis são os seguintes:

- ▶ Filtrar por género
- ▶ Filtrar por intervalo de anos
- ▶ Filtrar por país

Explicando brevemente cada gráfico e o que é possível ser observado lá, o primeiro mapa “Number of Infant Deaths per Country” indica uma distribuição do número de mortes de jovens em cada país (quanto maior for o círculo vermelho maior será esse número) e o mapa “Number of Infant Deaths in a certain Country” mostra o número para o país selecionado.

O gráfico de barras “Number of Infant Deaths” mostra o número de mortes de jovens por cada género sendo a sua percentagem exibida no gráfico ao lado “Percentage of Number of Infant Deaths”.

No gráfico de barras horizontal “Population per Age Group” é possível observar a percentagem de população existente num certo país por faixa etária e por género.

Os gráficos de linhas “Intentional Homicide”, “Mortality Rate Adult” e “Mortality Rate Infant” apresentam taxas para cada ano e para cada género.

Stop_and_Search

Stop and search in UK

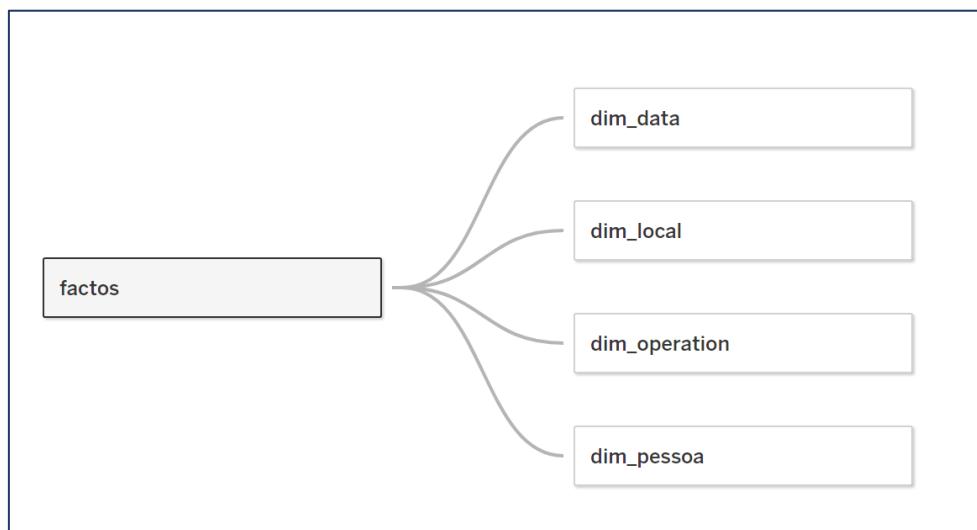


Figura 16 - Esquema de dados da Dashboard “Stop and search in UK”



Figura 17 - Dashboard “Stop and Search in UK”

Explicação da Dashboard "Stop and Search in UK"

A dashboard apresentada contém dados sobre operações de (stop and search) realizadas pela polícia no Reino Unido, cobrindo um período desde o inicio 2017 até outubro de 2023. Com base nas planilhas interativas mostradas, podemos extrair insights e aplicar filtros diretamente através da interface da dashboard.

Analizando a figura acima, podemos desde já retirar informações e filtrar acerca de:

- ▶ O número de operações por trimestre ao longo dos anos, permitindo observar tendências e padrões sazonais ou anuais.
- ▶ Os tipos de busca realizadas, como busca pessoal, busca em veículos ou ambas.
- ▶ Os objetos de busca, que incluem drogas controladas, armas ofensivas, artigos para uso em furto, bens roubados, entre outros.
- ▶ Faixa etária do cidadão abordado
- ▶ O género do cidadão abordado
- ▶ O resultado das operações, que podem resultar em prisões, resoluções comunitárias, ou não encontrar nada, o que levaria a nenhuma ação adicional.
- ▶ As legislações aplicadas durante as operações de busca,
- ▶ A etnia dos Agentes envolvidos nas operações.
- ▶ A faixa etária das pessoas sujeitas às operações de busca.
- ▶ As operações segmentadas por gênero.

Tendo em conta que se trata de uma Dashboard interativa com ações editadas, esta é capaz de fornecer uma vasta gama de filtros para o utilizador final ser capaz de visualizar os dados de uma forma fácil. Assim, os filtros possíveis são os seguintes:

- ▶ Filtrar por dia, mês e ano
- ▶ Filtrar por género
- ▶ Filtrar por tipos de busca
- ▶ Filtrar por tipo objeto procurado
- ▶ Filtrar por Etnia do Agente
- ▶ Filtrar por faixa etária
- ▶ Filtrar por Outcomes e Legislação das Operações
- ▶ Filtrar por County

O gráfico "Operations by Month" ilustra as operações ao longo do tempo, de 2017 até 2023, mostrando variações mensais que podem indicar tendências, efetividade das políticas implementadas ou mudanças nas prioridades operacionais. A consistência dos picos e vales pode sugerir sazonalidade nas operações ou resposta a eventos específicos.

No "Type of Search", observamos três categorias: busca pessoal, busca em veículos e busca combinada (pessoa e veículo). Isso pode refletir as prioridades operacionais ou a resposta a padrões de criminalidade específicos. Por exemplo, um aumento nas buscas veiculares pode estar associado a uma estratégia focada em contrabando ou roubo de veículos.

Os "Object of Search" destacam os itens mais frequentemente buscados, como drogas controladas e armas ofensivas. Isso oferece insights sobre os tipos de crime que a polícia está ativamente tentando combater.

Os resultados das operações, mostrados no gráfico "Outcome", variam de prisões a situações em que nenhuma ação adicional foi necessária. A proporção entre esses resultados pode oferecer uma perspectiva sobre a eficácia das operações de busca e abordagem.

A seção "Officer Ethnicity" e "Operations for Age Range" fornece dados demográficos sobre os oficiais e as faixas etárias das pessoas abordadas, enquanto "Operations by Gender" revela a divisão por gênero das pessoas submetidas a essas operações, o que pode levantar questões importantes sobre equidade e proporcionalidade.

O mapa interativo oferece uma visão geográfica das operações, permitindo identificar onde as abordagens estão concentradas, o que pode estar relacionado com as taxas de criminalidade ou políticas de policiamento locais.

Em suma, a dashboard permite uma análise multidimensional das práticas de busca e abordagem da polícia no Reino Unido, abrangendo aspectos operacionais, demográficos e geográficos, que são essenciais para avaliar a atuação policial e informar políticas públicas.

Homicide_Dataset

Homicide Victims in London

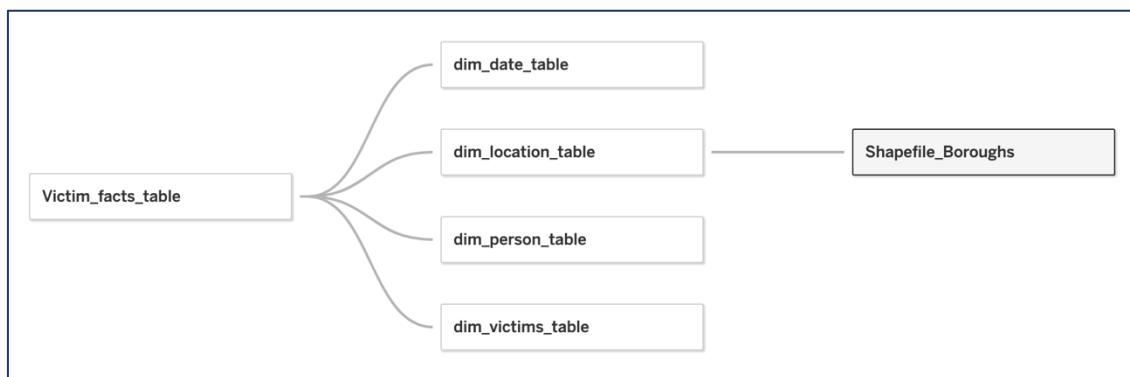


Figura 18 - Esquema de dados da Dashboard “Homicide Victims in London”

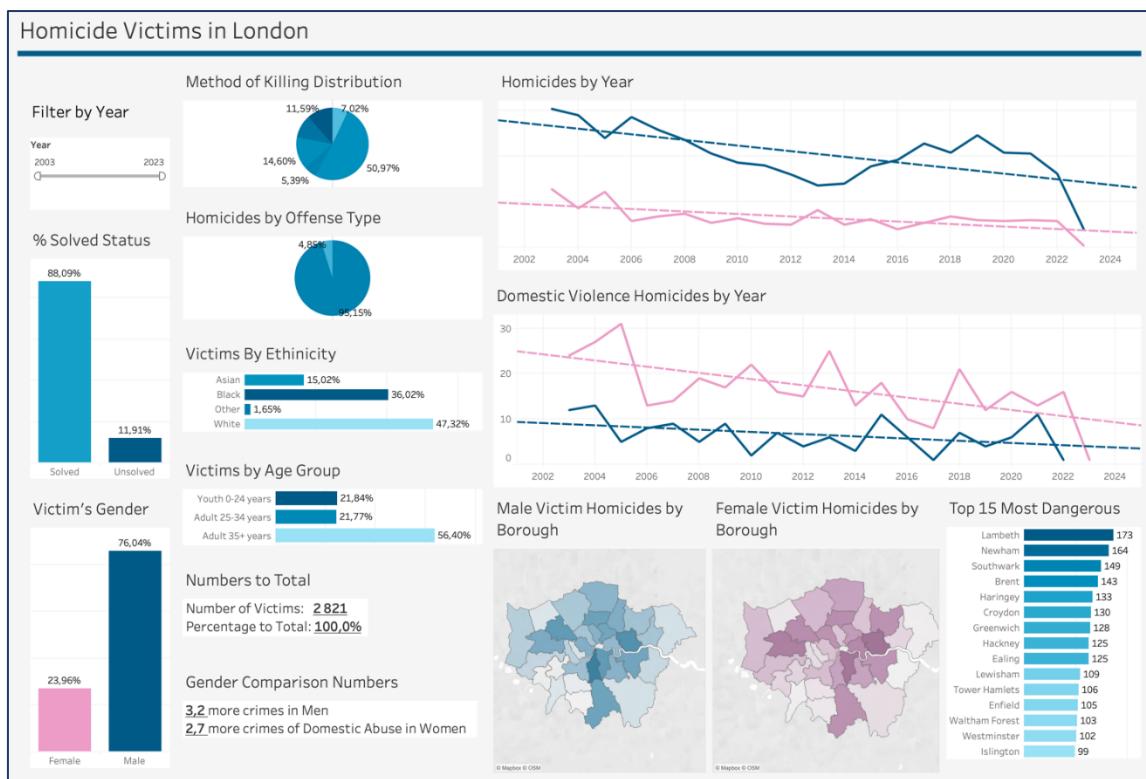


Figura 19 - Dashboard “Homicide Victims in London”

Explicação da Dashboard “Homicide Victims in London”

A Dashboard contém dados sobre diversos casos de vítimas de homicídio em Londres, cobrindo o período de 2003 a 2023.

Visto que se trata de uma dashboard com planilhas interativas, podemos retirar insights e filtrar tendo em conta as próprias planilhas.

Analisando a figura acima, podemos desde já retirar informações e filtrar acerca de:

- ▶ Estado de resolução dos crimes
- ▶ Género da vítima
- ▶ Método de homicídio
- ▶ Tipo de homicídio
- ▶ Etnia
- ▶ Faixa etária das vítimas
- ▶ Homicídios por ano
- ▶ Homicídios de violência doméstica por ano
- ▶ Valor do número da relação entre homicídios de homens com mulheres
- ▶ Homicídios por Borough, bem como visualizar o Top 15 de Boroughs mais perigosos.

Tendo em conta que se trata de uma Dashboard interativa com ações editadas, esta é capaz de fornecer uma vasta gama de filtros para o utilizador final ser capaz de visualizar os dados de uma forma fácil. Assim, os filtros possíveis são os seguintes:

- ▶ Filtrar por estado de resolução dos crimes
- ▶ Filtrar por género
- ▶ Filtrar por método de homicídio
- ▶ Filtrar por tipo de homicídio
- ▶ Filtrar por Etnia
- ▶ Filtrar por faixa etária
- ▶ Filtrar por intervalo de anos
- ▶ Filtrar por Borough

Fazendo uma análise simples sem qualquer tipo de interação (podendo fazê-la pela imagem acima), podemos observar que, entre 2003 e 2023, os homicídios foram baixando ao longo do tempo para ambos os géneros, a violência doméstica diminuiu, o top 15 de crimes mais perigosos conta com Lambeth em primeiro lugar, 89,09 % dos crimes ficaram com o estado “resolvido”, o género que sofreu mais crimes foi o masculino (com 76,04% dos homicídios).

Podemos também, tendo em conta os números apresentados, perceber que, neste intervalo de tempo, o sexo masculino sofreu de 3,2 x mais crimes do que o sexo feminino. O sexo feminino sofreu 2,7 x mais crimes de violência doméstica.

Homicide Aggressors in London



Figura 20 - Esquema de dados da Dashboard “Homicide Aggressors in London”

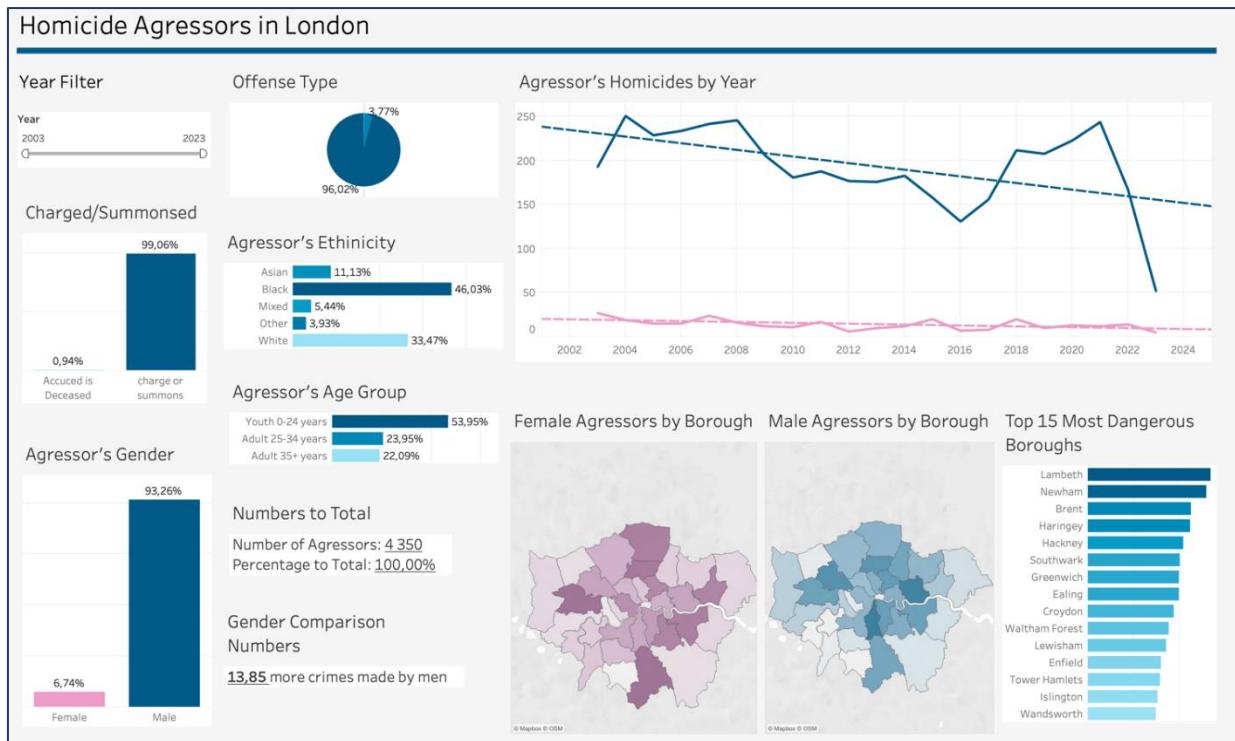


Figura 21 - Dashboard “Homicide Aggressors in London”

Explicação da Dashboard “Homicide Aggressors in London”

A Dashboard contém dados sobre diversos casos de agressores de homicídio em Londres, cobrindo também o período de 2003 a 2023.

Visto que se trata de uma Dashboard com planilhas interativas, podemos retirar insights e filtrar tendo em conta as próprias planilhas.

Analisando a figura acima, podemos desde já retirar informações e filtrar acerca de:

- ▶ Intervalo de anos
- ▶ Género do agressor
- ▶ Método de homicídio.
- ▶ Tipo de homicídio
- ▶ Etnia do agressor.
- ▶ Faixa etária das vítimas
- ▶ Homicídios por ano
- ▶ Homicídios de violência doméstica por ano
- ▶ Homicídios por Borough, bem como visualizar o Top 15 de Boroughs mais perigosos.

Tendo em conta que se trata de uma Dashboard interativa com ações editadas, esta é capaz de fornecer uma vasta gama de filtros para o utilizador final ser capaz de visualizar os dados de uma forma fácil. Assim, os filtros possíveis são os seguintes:

- ▶ Filtrar por estado de acusação dos agressores
- ▶ Filtrar por género
- ▶ Filtrar por tipo de homicídio
- ▶ Filtrar por Etnia
- ▶ Filtrar por faixa etária
- ▶ Filtrar por intervalo de anos
- ▶ Filtrar por Borough

Uma análise direta da Dashboard, sem interação adicional e no intervalo de 2003 a 2023, revela que a maioria dos agressores são do sexo masculino, representando 93,26% do total. A análise por ano mostra uma tendência decrescente nos homicídios cometidos por ambos os sexos.

Adicionalmente, 99,06% dos casos resultaram em acusação ou intimidação. O Borough de Lambeth lidera a lista dos locais mais perigosos, seguido por Newham e Brent. Através destes dados, é evidente que os homens cometem significativamente mais crimes comparados com as mulheres, com um total de 13.585 crimes a mais. Podemos observar também que o intervalo de idades mais perigoso foi “jovens de 0-24 anos” e que a etnia mais perigosa foi a etnia de pessoas negras. Tal como na dashboard de vítimas podemos visualizar que o tipo de homicídio foi, de longe, assassinato.

Child_Protection_Dataset

Children Careplan

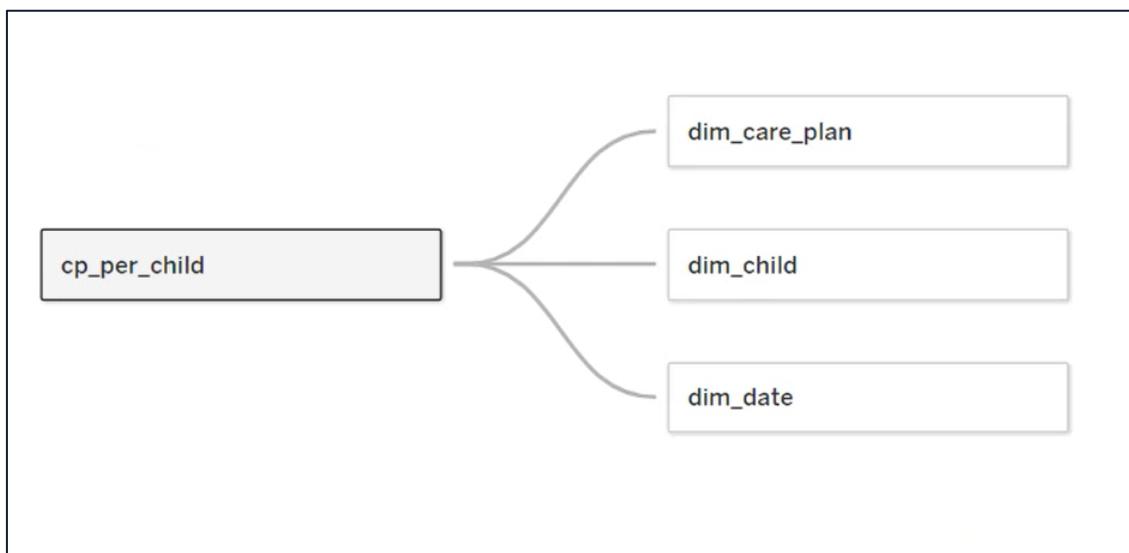


Figura 22 - Esquema de dados da Dashboard “Children Careplan”

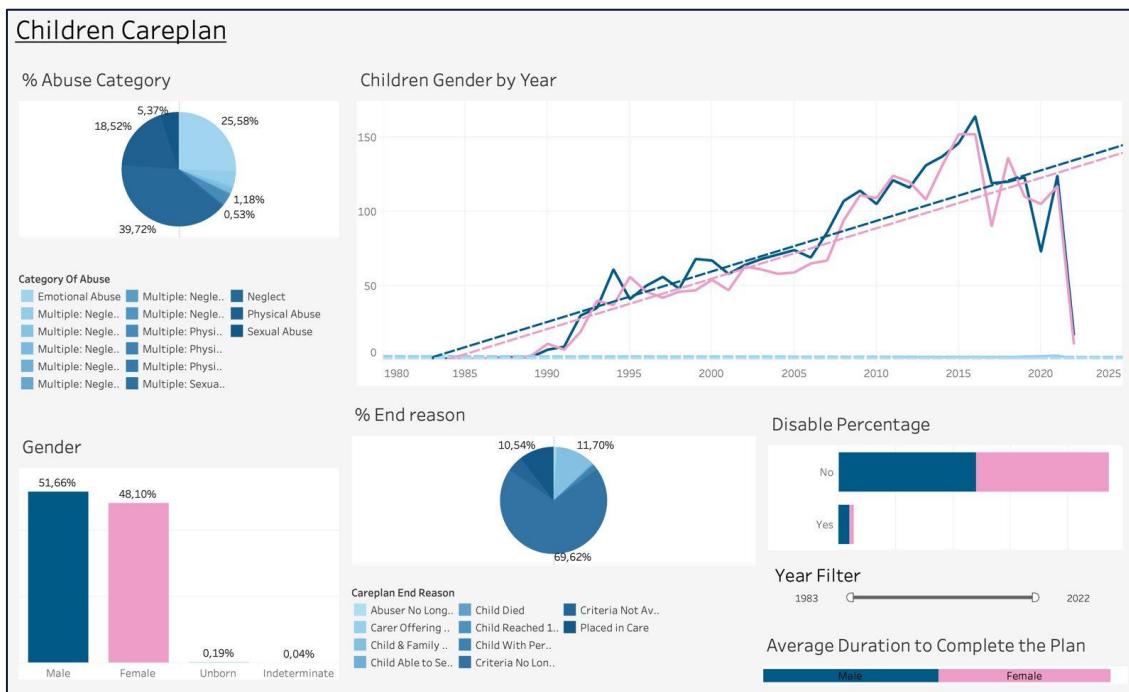


Figura 23 - Dashboard “Children Careplan”

Explicação da Dashboard “Children Careplan”

A DashBoard contem dados sobre as crianças que foram sujeitas a um plano de proteção ocorridos em Barnet, entre 1887 e 2022.

Esta dashborad foi realizada com o intuito de possuir planilhas interativas, quando uma planilha sofrer uma filtragem as outras também sofreram alterações.

Analisando a figura acima, podemos desde já retirar informações e filtrar acerca de:

- ▶ Categoria de Abuso (% Abuse Category)
- ▶ O género da criança (Gender)
- ▶ A razão para o fim do plano de proteção (% End Reason)
- ▶ A Media da duração da duração de um plano de proteção (Averege Duration to Complete the Plan)
- ▶ Número de crianças inscritas num plano por ano (Gender by Year)
- ▶ Número de crianças desabilitadas (Disable Percentage)

Considerando que esta Dashboard é interativa, é capaz de fornecer uma ampla gama de filtros pra o utilizador final, sendo este capaz de visualizar os dados de uma forma mais fácil.

Assim, os filtros possíveis são os seguintes:

- ▶ Filtrar por género
- ▶ Filtrar por categoria de abuso
- ▶ Filtrar por desabilidade
- ▶ Filtrar por razão do fim do plano
- ▶ Filtrar por intervalo de anos

Fazendo uma análise simples sem qualquer tipo de interação (podendo fazê-la pela imagem acima), podemos observar que, entre 1988 e 2022, o número de crianças que foram sujeitas a um plano de proteção tem vindo a aumentar para ambos os géneros. É possível também ver na dashboard em cima a percentagem de crianças do sexo feminino, 49.1%, e a percentagem do sexo masculino 51.9%, é possível desde já observar a existência de crianças do género masculino com maior número de inscrição num plano de proteção.

Data Profiling and Data Quality

Para a avaliação da qualidade dos dados, a equipa optou pela ferramenta recomendada pelos professores, o Talend Open Studio For Data Quality, e aplicou o seguinte conjunto de métricas, posteriormente detalhado:

- ▶ Estatísticas Básicas:
 - Com o propósito de calcular o total de registos (row count);
 - Número de registos nulos (null count);
 - Número de registos distintos (distinct count);
 - Número de registos duplicados (duplicate count);
 - Número de registos únicos (unique count);
- ▶ Estatísticas Resumidas:
 - Amplitude (range), de modo a conhecer o intervalo entre o máximo e mínimo de cada atributo numérico;
- ▶ Estatísticas Avançadas:
 - Frequência dos valores (value frequency), com o intuito de conhecer os dados mais frequentes.

Data Profiling and Dictionary

Dataset: Homicide_Victim

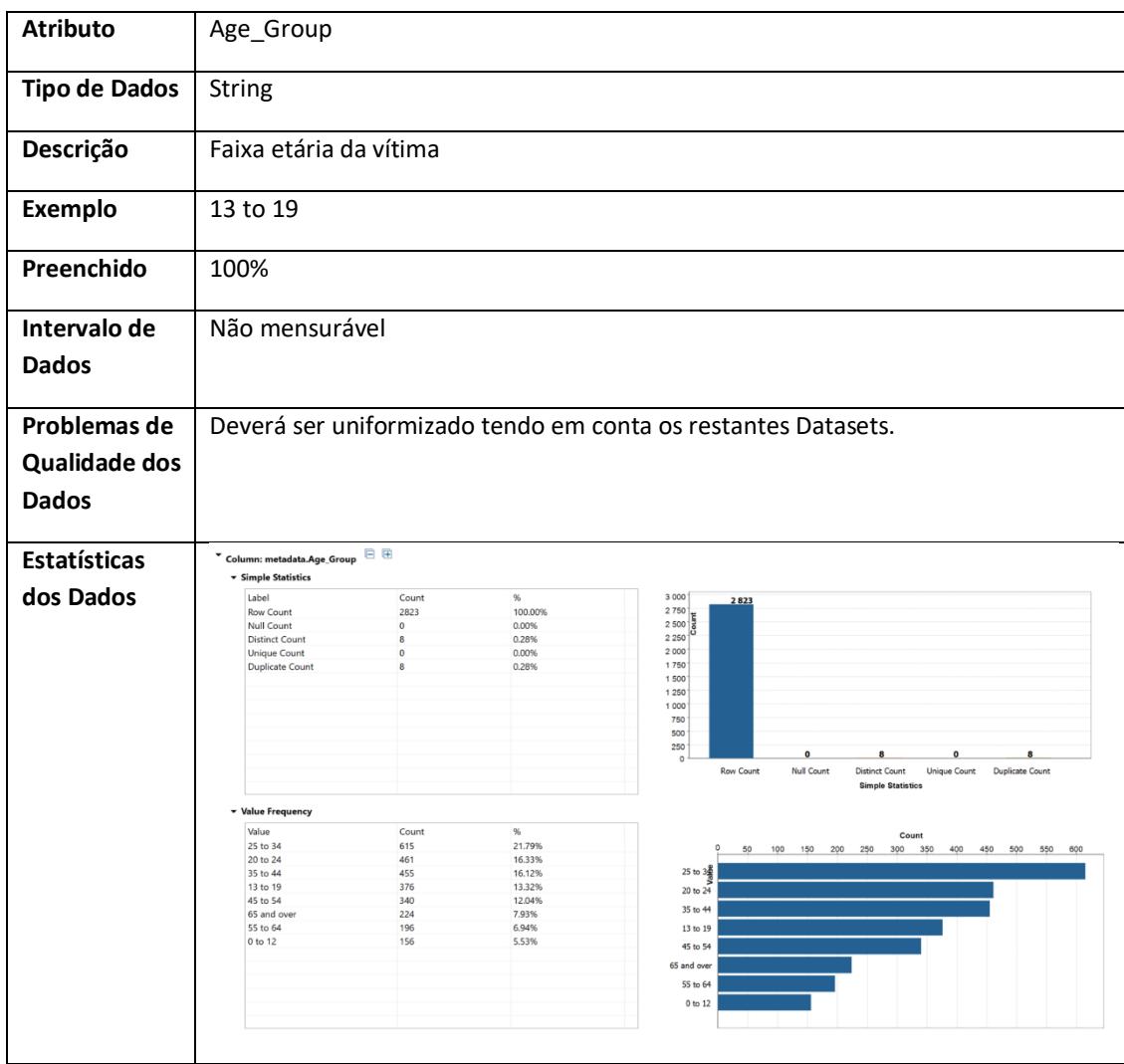
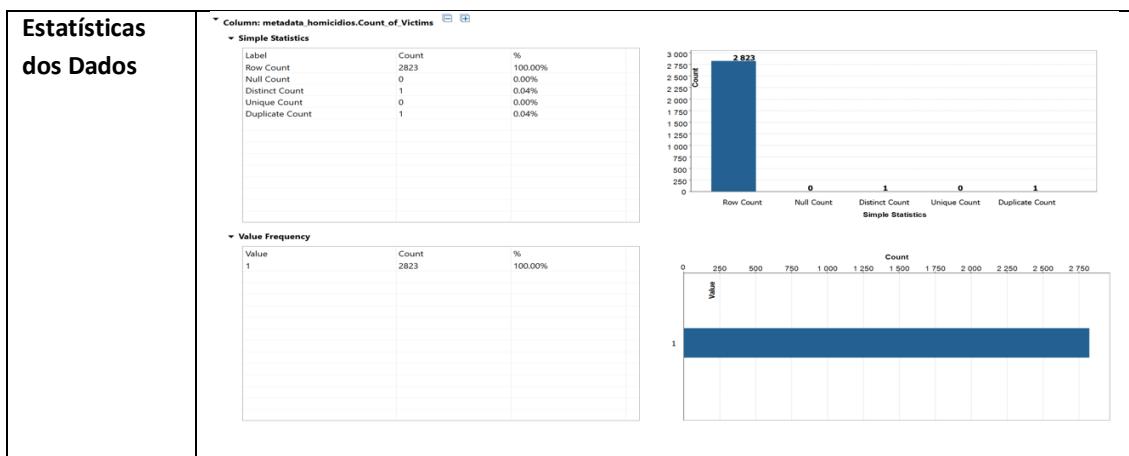
Aluno responsável: Pedro Pires (a95549)

Fonte: <https://data.london.gov.uk/dataset/mps-homicide-dashboard-data>

Descrição: O conjunto de dados, distribuído por dois ficheiros CSV, proporciona uma visão abrangente dos incidentes de homicídio ocorridos em Londres. O primeiro ficheiro ("Homicide Victim 2003 - March 2023.csv") inclui pormenores específicos de cada crime, contendo informações sobre as vítimas, os métodos empregados nos homicídios, a ligação a casos de abuso doméstico, a localização dos incidentes na cidade e o estado das investigações. O segundo ficheiro ("Homicide_Proceeded_Against_2003_March_2023.csv") centra-se nos indivíduos que foram levados a tribunal por esses crimes, registando informações sobre os arguidos, a natureza das acusações legais, a situação das acusações ou intimações, e o calendário dos procedimentos judiciais. Em conjunto, estes ficheiros constituem um recurso valioso para a análise dos padrões, características demográficas e consequências jurídicas associadas aos casos de homicídio na capital britânica durante o período especificado.

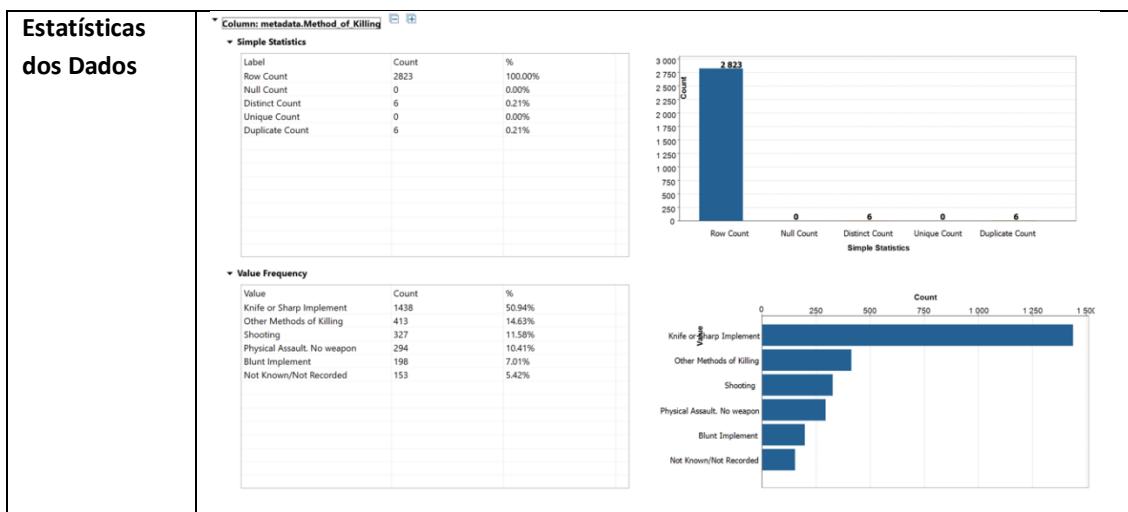
Homicide_Victim_2003.csv

Atributo	Count_of_Victims
Tipo de Dado	Integer
Descrição	Número de vitimas no caso de homicídio
Exemplo	1
Preenchido	100%
Intervalo de Dados	(1,1)
Problemas de Qualidade dos Dados	Não detetados



Atributo Sex Tipo de Dados String Descrição Género da vítima

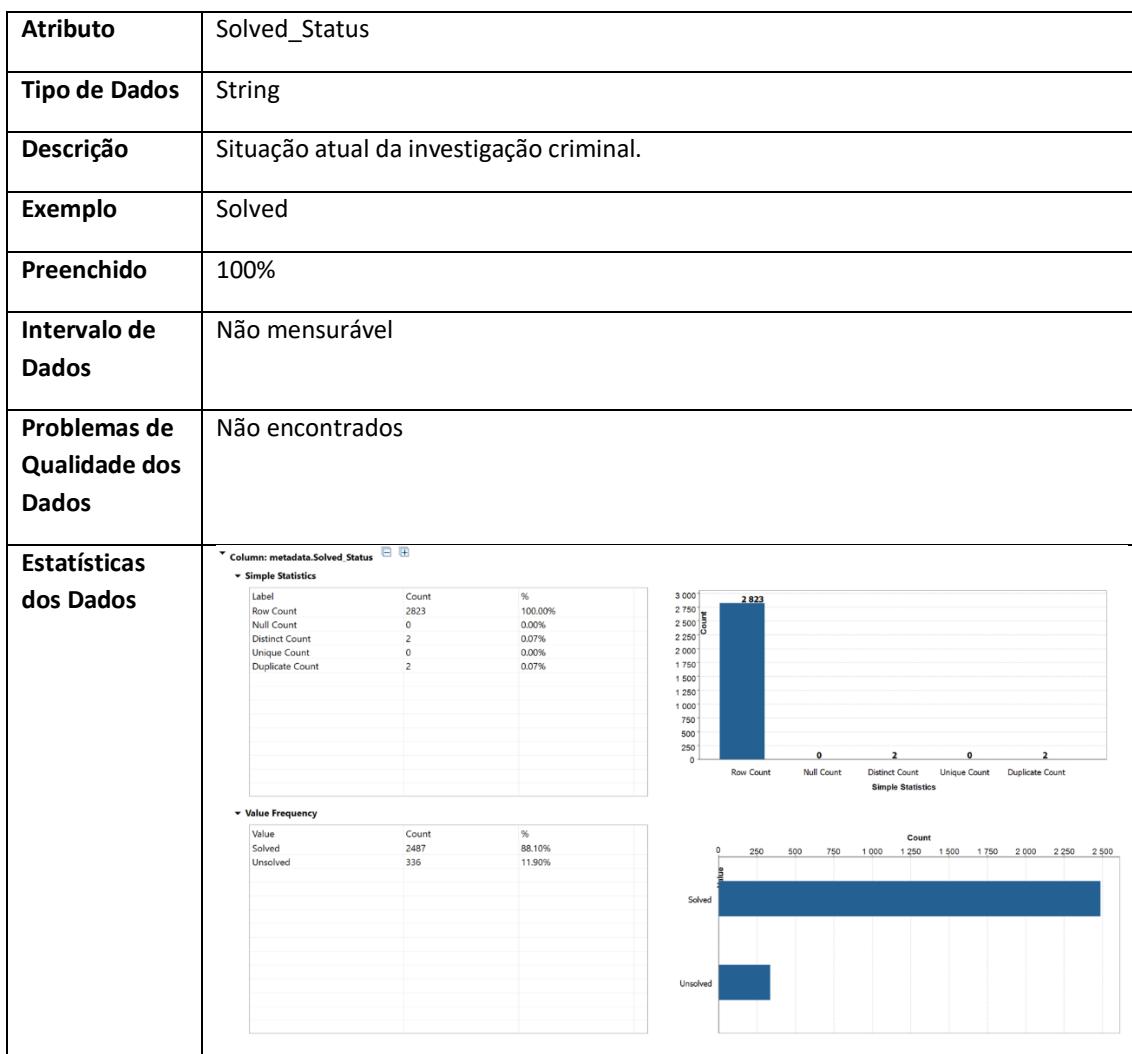
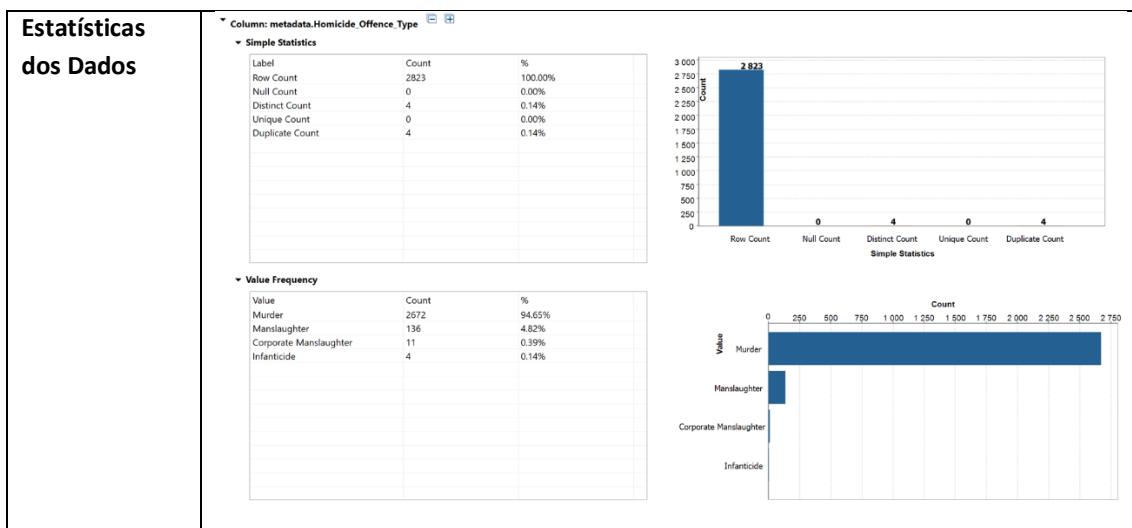
Exemplo	Female																														
Preenchido	100%																														
Intervalo de Dados	Não mensurável																														
Problemas de Qualidade dos Dados	"Unrecorded" é basicamente um null. Felizmente, são poucos os registos que contém este valor. De modo a todos os datasets terem coerência no que diz respeito às colunas relativas ao género, iremos alterar o nome do atributo para "Gender"																														
Estatísticas dos Dados	<p>Column: metadata.Sex</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>2823</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>3</td> <td>0.11%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>3</td> <td>0.11%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Male</td> <td>2145</td> <td>75.98%</td> </tr> <tr> <td>Female</td> <td>676</td> <td>23.95%</td> </tr> <tr> <td>Unrecorded</td> <td>2</td> <td>0.07%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	2823	100.00%	Null Count	0	0.00%	Distinct Count	3	0.11%	Unique Count	0	0.00%	Duplicate Count	3	0.11%	Value	Count	%	Male	2145	75.98%	Female	676	23.95%	Unrecorded	2	0.07%
Label	Count	%																													
Row Count	2823	100.00%																													
Null Count	0	0.00%																													
Distinct Count	3	0.11%																													
Unique Count	0	0.00%																													
Duplicate Count	3	0.11%																													
Value	Count	%																													
Male	2145	75.98%																													
Female	676	23.95%																													
Unrecorded	2	0.07%																													
Atributo	Method_of_Killing																														
Tipo de Dados	String																														
Descrição	Método usado no homicídio																														
Exemplo	Blunt Implement																														
Preenchido	100%																														
Intervalo de Dados	Não mensurável																														
Problemas de Qualidade dos Dados	Não encontrados.																														



Atributo	Domestic_Abuse																											
Tipo de Dados	String																											
Descrição	Indica se o homicídio foi de abuso doméstico.																											
Exemplo	Not Domestic Abuse																											
Preenchido	100%																											
Intervalo de Dados	Não mensurável																											
Problemas de Qualidade dos Dados	Não encontrados																											
Estatísticas dos Dados	<p>Column: metadata.Domestic_Abuse</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>2823</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>2</td> <td>0.07%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>2</td> <td>0.07%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Not Domestic Abuse</td> <td>2342</td> <td>82.96%</td> </tr> <tr> <td>Domestic Abuse</td> <td>481</td> <td>17.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	2823	100.00%	Null Count	0	0.00%	Distinct Count	2	0.07%	Unique Count	0	0.00%	Duplicate Count	2	0.07%	Value	Count	%	Not Domestic Abuse	2342	82.96%	Domestic Abuse	481	17.04%
Label	Count	%																										
Row Count	2823	100.00%																										
Null Count	0	0.00%																										
Distinct Count	2	0.07%																										
Unique Count	0	0.00%																										
Duplicate Count	2	0.07%																										
Value	Count	%																										
Not Domestic Abuse	2342	82.96%																										
Domestic Abuse	481	17.04%																										

Atributo	Recorded_Date																																																			
Tipo de Dados	String																																																			
Descrição	Data registada do caso																																																			
Exemplo	Jul-05																																																			
Preenchido	100%																																																			
Intervalo de Dados	Não mensurável																																																			
Problemas de Qualidade dos Dados	As datas estão guardadas num formato string e não como Date ou Datetime o que pode dificultar algumas comparações/funções temporais.																																																			
Estatísticas dos Dados	<p>Column: metadata.Recorded_Date</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>2823</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>District Count</td> <td>243</td> <td>8.61%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>243</td> <td>8.61%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Jul-05</td> <td>30</td> <td>1.06%</td> </tr> <tr> <td>Jul-03</td> <td>28</td> <td>0.99%</td> </tr> <tr> <td>Jun-07</td> <td>26</td> <td>0.92%</td> </tr> <tr> <td>Oct-03</td> <td>23</td> <td>0.81%</td> </tr> <tr> <td>Apr-04</td> <td>22</td> <td>0.78%</td> </tr> <tr> <td>Jun-03</td> <td>22</td> <td>0.78%</td> </tr> <tr> <td>Jan-04</td> <td>21</td> <td>0.74%</td> </tr> <tr> <td>Jun-17</td> <td>21</td> <td>0.74%</td> </tr> <tr> <td>Mar-03</td> <td>20</td> <td>0.71%</td> </tr> <tr> <td>Apr-06</td> <td>19</td> <td>0.67%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	2823	100.00%	Null Count	0	0.00%	District Count	243	8.61%	Unique Count	0	0.00%	Duplicate Count	243	8.61%	Value	Count	%	Jul-05	30	1.06%	Jul-03	28	0.99%	Jun-07	26	0.92%	Oct-03	23	0.81%	Apr-04	22	0.78%	Jun-03	22	0.78%	Jan-04	21	0.74%	Jun-17	21	0.74%	Mar-03	20	0.71%	Apr-06	19	0.67%
Label	Count	%																																																		
Row Count	2823	100.00%																																																		
Null Count	0	0.00%																																																		
District Count	243	8.61%																																																		
Unique Count	0	0.00%																																																		
Duplicate Count	243	8.61%																																																		
Value	Count	%																																																		
Jul-05	30	1.06%																																																		
Jul-03	28	0.99%																																																		
Jun-07	26	0.92%																																																		
Oct-03	23	0.81%																																																		
Apr-04	22	0.78%																																																		
Jun-03	22	0.78%																																																		
Jan-04	21	0.74%																																																		
Jun-17	21	0.74%																																																		
Mar-03	20	0.71%																																																		
Apr-06	19	0.67%																																																		

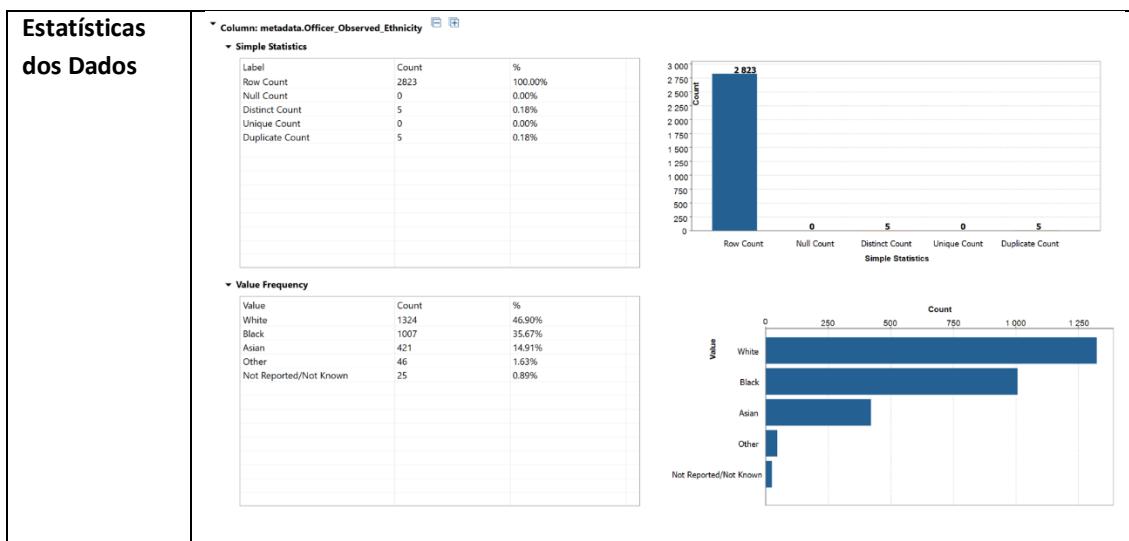
Atributo	Homicide_Offence_Type
Tipo de Dados	String
Descrição	Tipo de homicídio
Exemplo	Murder
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não encontrados



Atributo Borough	
Tipo de Dados	String

Descrição	Bairro onde ocorreu o crime																																																			
Exemplo	Lambeth																																																			
Preenchido	100%																																																			
Intervalo de Dados	Não mensurável																																																			
Problemas de Qualidade dos Dados	Não encontrados																																																			
Estatísticas dos Dados	<p>* Column: metadata.Borough </p> <p>▼ Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>2823</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>32</td> <td>1.13%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>32</td> <td>1.13%</td> </tr> </tbody> </table> <p>▼ Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Lambeth</td> <td>173</td> <td>6.13%</td> </tr> <tr> <td>Newham</td> <td>164</td> <td>5.81%</td> </tr> <tr> <td>Southwark</td> <td>149</td> <td>5.20%</td> </tr> <tr> <td>Brent</td> <td>143</td> <td>5.07%</td> </tr> <tr> <td>Haringey</td> <td>133</td> <td>4.71%</td> </tr> <tr> <td>Croydon</td> <td>130</td> <td>4.61%</td> </tr> <tr> <td>Greenwich</td> <td>128</td> <td>4.53%</td> </tr> <tr> <td>Ealing</td> <td>125</td> <td>4.43%</td> </tr> <tr> <td>Hackney</td> <td>125</td> <td>4.43%</td> </tr> <tr> <td>Lewisham</td> <td>109</td> <td>3.86%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	2823	100.00%	Null Count	0	0.00%	Distinct Count	32	1.13%	Unique Count	0	0.00%	Duplicate Count	32	1.13%	Value	Count	%	Lambeth	173	6.13%	Newham	164	5.81%	Southwark	149	5.20%	Brent	143	5.07%	Haringey	133	4.71%	Croydon	130	4.61%	Greenwich	128	4.53%	Ealing	125	4.43%	Hackney	125	4.43%	Lewisham	109	3.86%
Label	Count	%																																																		
Row Count	2823	100.00%																																																		
Null Count	0	0.00%																																																		
Distinct Count	32	1.13%																																																		
Unique Count	0	0.00%																																																		
Duplicate Count	32	1.13%																																																		
Value	Count	%																																																		
Lambeth	173	6.13%																																																		
Newham	164	5.81%																																																		
Southwark	149	5.20%																																																		
Brent	143	5.07%																																																		
Haringey	133	4.71%																																																		
Croydon	130	4.61%																																																		
Greenwich	128	4.53%																																																		
Ealing	125	4.43%																																																		
Hackney	125	4.43%																																																		
Lewisham	109	3.86%																																																		

Atributo	Officer_Observed_Ethnicity
Tipo de Dados	String
Descrição	Etnia do oficial
Exemplo	Asian
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não encontrados



Aluno responsável: Pedro Pires (a95549)

Fonte: <https://data.london.gov.uk/dataset/mps-homicide-dashboard-data>

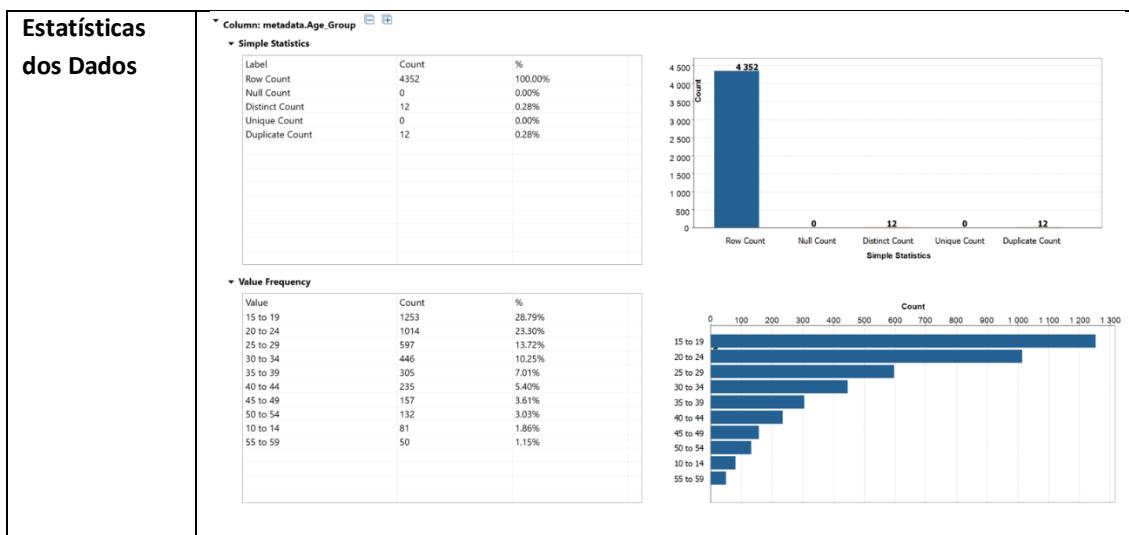
Homicide_Proceeded_Victim.csv

Atributo	Count_of_PPA
Tipo de Dados	Integer
Descrição	Número de infratores condenados e prisioneiros que estão sob o sistema de liberdade condicional
Exemplo	1
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não encontrados

Atributo	Gender																														
Tipo de Dados	String																														
Descrição	Género do oppresor.																														
Exemplo	Male																														
Preenchido	100%																														
Intervalo de Dados	Não mensurável																														
Problemas de Qualidade dos Dados	"Unrecorded" é basicamente um null, pelo que deverá ser ignorado na análise.																														
Estatísticas dos Dados	<p>Column: metadata.Gender</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>4352</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>District Count</td> <td>3</td> <td>0.07%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>3</td> <td>0.07%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Male</td> <td>4057</td> <td>93.22%</td> </tr> <tr> <td>Female</td> <td>293</td> <td>6.73%</td> </tr> <tr> <td>Unrecorded</td> <td>2</td> <td>0.05%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	4352	100.00%	Null Count	0	0.00%	District Count	3	0.07%	Unique Count	0	0.00%	Duplicate Count	3	0.07%	Value	Count	%	Male	4057	93.22%	Female	293	6.73%	Unrecorded	2	0.05%
Label	Count	%																													
Row Count	4352	100.00%																													
Null Count	0	0.00%																													
District Count	3	0.07%																													
Unique Count	0	0.00%																													
Duplicate Count	3	0.07%																													
Value	Count	%																													
Male	4057	93.22%																													
Female	293	6.73%																													
Unrecorded	2	0.05%																													
Estatísticas dos Dados	<p>Column: metadata.Count_of_PPA</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>4352</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>District Count</td> <td>1</td> <td>0.02%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>1</td> <td>0.02%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>4352</td> <td>100.00%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	4352	100.00%	Null Count	0	0.00%	District Count	1	0.02%	Unique Count	0	0.00%	Duplicate Count	1	0.02%	Value	Count	%	1	4352	100.00%						
Label	Count	%																													
Row Count	4352	100.00%																													
Null Count	0	0.00%																													
District Count	1	0.02%																													
Unique Count	0	0.00%																													
Duplicate Count	1	0.02%																													
Value	Count	%																													
1	4352	100.00%																													

Atributo	Borough																																																			
Tipo de Dados	String																																																			
Descrição	Distrito onde ocorreu o crime.																																																			
Exemplo	Brent																																																			
Preenchido	100%																																																			
Intervalo de Dados	Não mensurável																																																			
Problemas de Qualidade dos Dados	Não encontrados																																																			
Estatísticas dos Dados	<p>Column: metadata.Borough</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>4352</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>33</td> <td>0.76%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>33</td> <td>0.76%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Lambeth</td> <td>268</td> <td>6.16%</td> </tr> <tr> <td>Newham</td> <td>260</td> <td>5.97%</td> </tr> <tr> <td>Brent</td> <td>226</td> <td>5.19%</td> </tr> <tr> <td>Haringey</td> <td>224</td> <td>5.15%</td> </tr> <tr> <td>Hackney</td> <td>209</td> <td>4.80%</td> </tr> <tr> <td>Greenwich</td> <td>201</td> <td>4.62%</td> </tr> <tr> <td>Southwark</td> <td>201</td> <td>4.62%</td> </tr> <tr> <td>Ealing</td> <td>201</td> <td>4.62%</td> </tr> <tr> <td>Croydon</td> <td>188</td> <td>4.32%</td> </tr> <tr> <td>Waltham Forest</td> <td>177</td> <td>4.07%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	4352	100.00%	Null Count	0	0.00%	Distinct Count	33	0.76%	Unique Count	0	0.00%	Duplicate Count	33	0.76%	Value	Count	%	Lambeth	268	6.16%	Newham	260	5.97%	Brent	226	5.19%	Haringey	224	5.15%	Hackney	209	4.80%	Greenwich	201	4.62%	Southwark	201	4.62%	Ealing	201	4.62%	Croydon	188	4.32%	Waltham Forest	177	4.07%
Label	Count	%																																																		
Row Count	4352	100.00%																																																		
Null Count	0	0.00%																																																		
Distinct Count	33	0.76%																																																		
Unique Count	0	0.00%																																																		
Duplicate Count	33	0.76%																																																		
Value	Count	%																																																		
Lambeth	268	6.16%																																																		
Newham	260	5.97%																																																		
Brent	226	5.19%																																																		
Haringey	224	5.15%																																																		
Hackney	209	4.80%																																																		
Greenwich	201	4.62%																																																		
Southwark	201	4.62%																																																		
Ealing	201	4.62%																																																		
Croydon	188	4.32%																																																		
Waltham Forest	177	4.07%																																																		

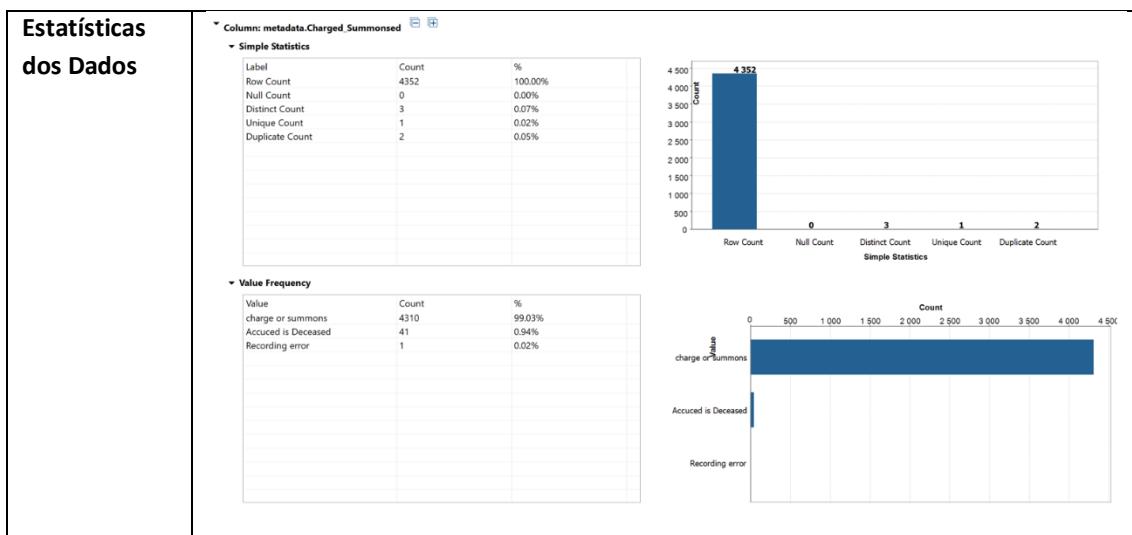
Atributo	Age_Group
Tipo de Dados	String
Descrição	Faixa etária do opressor
Exemplo	40 to 44.
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Deveria ser um intervalo numérico/número. Deverá ser uniformizado para um intervalo numérico capaz de ser comparável com os restantes Datasets.



Atributo	Offence_Type																																	
Tipo de Dados	String																																	
Descrição	Tipo de crime cometido contra a vítima																																	
Exemplo	Murder																																	
Preenchido	100%																																	
Intervalo de Dados	Não mensurável																																	
Problemas de Qualidade dos Dados	Não encontrados																																	
Estatísticas dos Dados	<p>Column: metadata.Offence_Type</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>4352</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>4</td><td>0.09%</td></tr> <tr><td>Unique Count</td><td>1</td><td>0.02%</td></tr> <tr><td>Duplicate Count</td><td>3</td><td>0.07%</td></tr> </tbody> </table> <p>Simple Statistics</p> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Murder</td><td>4179</td><td>96.02%</td></tr> <tr><td>Manslaughter</td><td>164</td><td>3.77%</td></tr> <tr><td>Corporate Manslaughter</td><td>8</td><td>0.18%</td></tr> <tr><td>Infanticide</td><td>1</td><td>0.02%</td></tr> </tbody> </table> <p>Count</p>	Label	Count	%	Row Count	4352	100.00%	Null Count	0	0.00%	Distinct Count	4	0.09%	Unique Count	1	0.02%	Duplicate Count	3	0.07%	Value	Count	%	Murder	4179	96.02%	Manslaughter	164	3.77%	Corporate Manslaughter	8	0.18%	Infanticide	1	0.02%
Label	Count	%																																
Row Count	4352	100.00%																																
Null Count	0	0.00%																																
Distinct Count	4	0.09%																																
Unique Count	1	0.02%																																
Duplicate Count	3	0.07%																																
Value	Count	%																																
Murder	4179	96.02%																																
Manslaughter	164	3.77%																																
Corporate Manslaughter	8	0.18%																																
Infanticide	1	0.02%																																

Atributo	Self_Classified_Ethnicity.																																										
Tipo de Dados	String.																																										
Descrição	Etnia do opressor																																										
Exemplo	Asian																																										
Preenchido	100%																																										
Intervalo de Dados	Não mensurável																																										
Problemas de Qualidade dos Dados	Uma quantidade considerável de regtos não permitem saber a etnia do opressor.																																										
Estatísticas dos Dados	<p>Column: metadata.Self_Classified_Ethnicity</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>4352</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>7</td> <td>0.16%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>7</td> <td>0.16%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Black</td> <td>1828</td> <td>42.00%</td> </tr> <tr> <td>White</td> <td>1329</td> <td>30.54%</td> </tr> <tr> <td>Asian</td> <td>443</td> <td>10.18%</td> </tr> <tr> <td>Person declines to define ethnicity</td> <td>294</td> <td>6.76%</td> </tr> <tr> <td>Mixed</td> <td>217</td> <td>4.99%</td> </tr> <tr> <td>Other</td> <td>156</td> <td>3.58%</td> </tr> <tr> <td>SCE Not Recorded</td> <td>85</td> <td>1.95%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	4352	100.00%	Null Count	0	0.00%	Distinct Count	7	0.16%	Unique Count	0	0.00%	Duplicate Count	7	0.16%	Value	Count	%	Black	1828	42.00%	White	1329	30.54%	Asian	443	10.18%	Person declines to define ethnicity	294	6.76%	Mixed	217	4.99%	Other	156	3.58%	SCE Not Recorded	85	1.95%
Label	Count	%																																									
Row Count	4352	100.00%																																									
Null Count	0	0.00%																																									
Distinct Count	7	0.16%																																									
Unique Count	0	0.00%																																									
Duplicate Count	7	0.16%																																									
Value	Count	%																																									
Black	1828	42.00%																																									
White	1329	30.54%																																									
Asian	443	10.18%																																									
Person declines to define ethnicity	294	6.76%																																									
Mixed	217	4.99%																																									
Other	156	3.58%																																									
SCE Not Recorded	85	1.95%																																									

Atributo	Charged_Summonsed
Tipo de Dados	String
Descrição	Acusações convocadas contra o opressor
Exemplo	Accuded is Deceased.
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	“SCE not recorded” deverá ser ignorado numa análise futura.



Atributo	Proceedings_date																																																			
Tipo de Dados	String																																																			
Descrição	Data do processo do opressor																																																			
Exemplo	Mar-19.																																																			
Preenchido	100%																																																			
Intervalo de Dados	Não mensurável																																																			
Problemas de Qualidade dos Dados	As datas estão guardadas num formato string e não como Date ou Datetime o que pode dificultar algumas comparações/funções temporais.																																																			
Estatísticas dos Dados	<p>Column: metadata.Proceedings.Date</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>4352</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>246</td> <td>5.65%</td> </tr> <tr> <td>Unique Count</td> <td>1</td> <td>0.02%</td> </tr> <tr> <td>Duplicate Count</td> <td>245</td> <td>5.63%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Jul-08</td> <td>55</td> <td>1.26%</td> </tr> <tr> <td>Nov-18</td> <td>44</td> <td>1.01%</td> </tr> <tr> <td>May-06</td> <td>39</td> <td>0.90%</td> </tr> <tr> <td>Mar-21</td> <td>36</td> <td>0.83%</td> </tr> <tr> <td>Mar-19</td> <td>35</td> <td>0.80%</td> </tr> <tr> <td>Mar-04</td> <td>34</td> <td>0.78%</td> </tr> <tr> <td>May-14</td> <td>34</td> <td>0.78%</td> </tr> <tr> <td>Apr-18</td> <td>33</td> <td>0.76%</td> </tr> <tr> <td>Jun-07</td> <td>33</td> <td>0.76%</td> </tr> <tr> <td>Jun-21</td> <td>32</td> <td>0.74%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	4352	100.00%	Null Count	0	0.00%	Distinct Count	246	5.65%	Unique Count	1	0.02%	Duplicate Count	245	5.63%	Value	Count	%	Jul-08	55	1.26%	Nov-18	44	1.01%	May-06	39	0.90%	Mar-21	36	0.83%	Mar-19	35	0.80%	Mar-04	34	0.78%	May-14	34	0.78%	Apr-18	33	0.76%	Jun-07	33	0.76%	Jun-21	32	0.74%
Label	Count	%																																																		
Row Count	4352	100.00%																																																		
Null Count	0	0.00%																																																		
Distinct Count	246	5.65%																																																		
Unique Count	1	0.02%																																																		
Duplicate Count	245	5.63%																																																		
Value	Count	%																																																		
Jul-08	55	1.26%																																																		
Nov-18	44	1.01%																																																		
May-06	39	0.90%																																																		
Mar-21	36	0.83%																																																		
Mar-19	35	0.80%																																																		
Mar-04	34	0.78%																																																		
May-14	34	0.78%																																																		
Apr-18	33	0.76%																																																		
Jun-07	33	0.76%																																																		
Jun-21	32	0.74%																																																		

Datasets: Crime_Dataset

Aluno responsável: Liandro Cruz (a100436)

Fonte: <https://data.police.uk/data/>

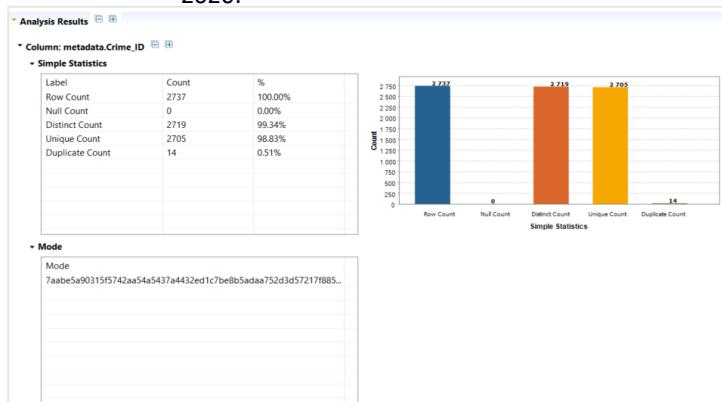
Descrição: Os três conjuntos de dados oferecem uma análise abrangente sobre a criminalidade e as medidas de aplicação da lei no UK. O primeiro, "Ano-Mês-Concelho-outcomes.csv", foca-se nos desfechos das ocorrências criminais, detalhando os resultados de cada caso, que vão desde investigações concluídas sem suspeitos identificados até casos encaminhados para outras entidades. O segundo conjunto de dados, "Ano-Mês-Concelho-stop-and-search.csv", documenta as atividades de paragem e revista, revelando as razões, procedimentos, e eficácia destas intervenções, além de dados demográficos dos indivíduos visados. Por fim, "Ano-Mês-Concelho-street.csv" regista os crimes reportados nas vias públicas, fornecendo pormenores sobre a natureza e o local dos incidentes, assim como informações limitadas sobre os desfechos destes casos.

Em conjunto, estes ficheiros compõem um retrato complexo da situação criminal e das operações policiais no UK, possibilitando uma análise pormenorizada sobre os padrões de criminalidade, a eficácia das respostas, e áreas potenciais de preocupação ou melhoria nas estratégias de aplicação da lei. A integração destas informações é crucial para uma compreensão holística da segurança pública na região, destacando a importância da transparência e da acessibilidade dos dados para a avaliação contínua das práticas e políticas de combate ao crime.

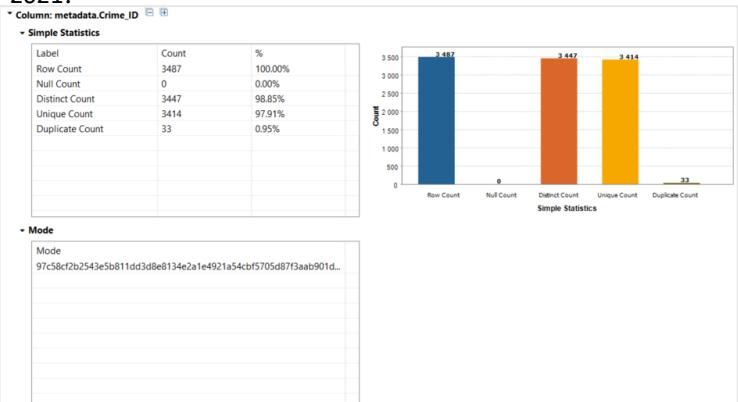
Atributo	Crime ID
Tipo de Dados	String
Descrição	Um identificador único para cada crime.
Exemplo	1
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Casing inconsistente (por exemplo, maiúsculas e minúsculas misturadas).

Dados Estatísticos:

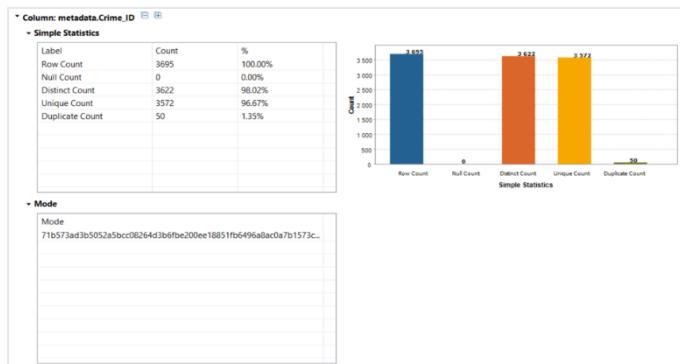
2020:



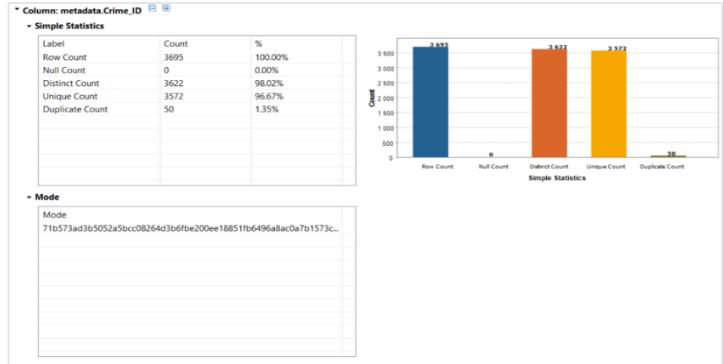
2021:



2022:



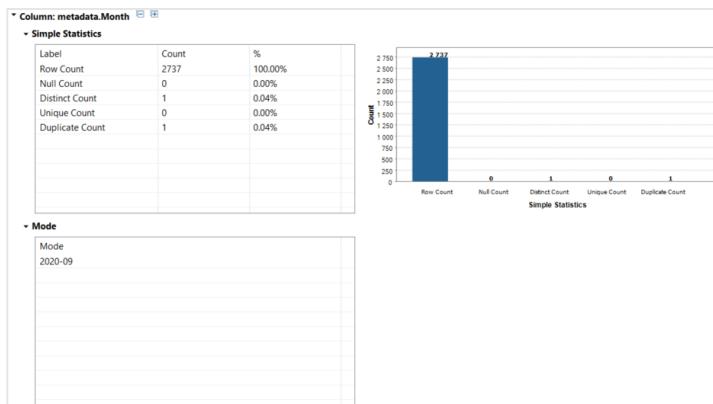
2023:



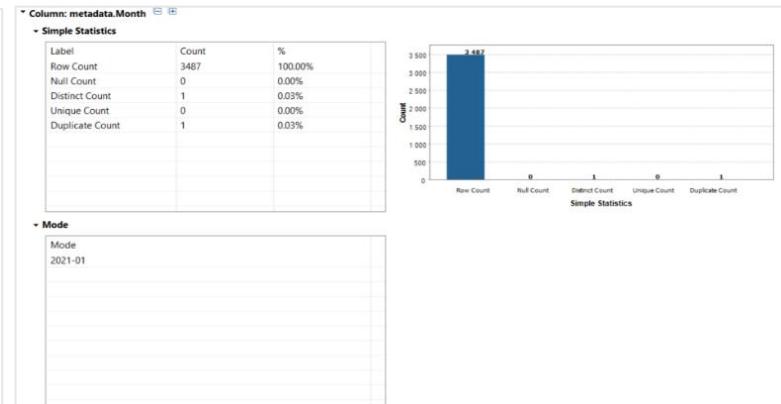
Atributo	Month
Tipo de Dados	String
Descrição	O mês em que o crime foi relatado.
Exemplo	2020-09
Preenchido	100%
Intervalo de Dados	Não mensurável.
Problemas de Qualidade dos Dados	O formato para este atributo não é o mais adequado visto que inclui o ano e é uma String. Será necessário separar em duas colunas e converter em Integer.

Estatísticas dos Dados:

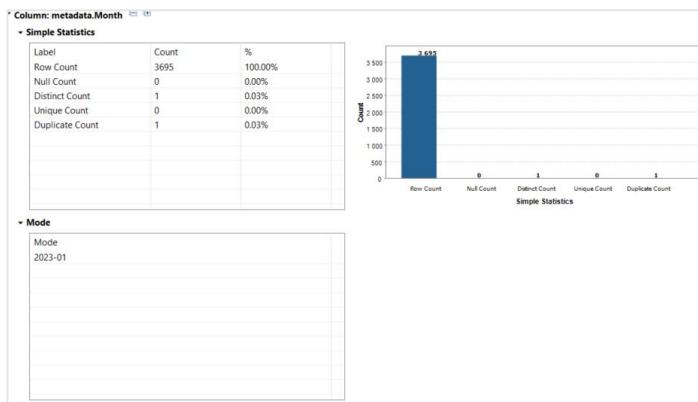
2020:



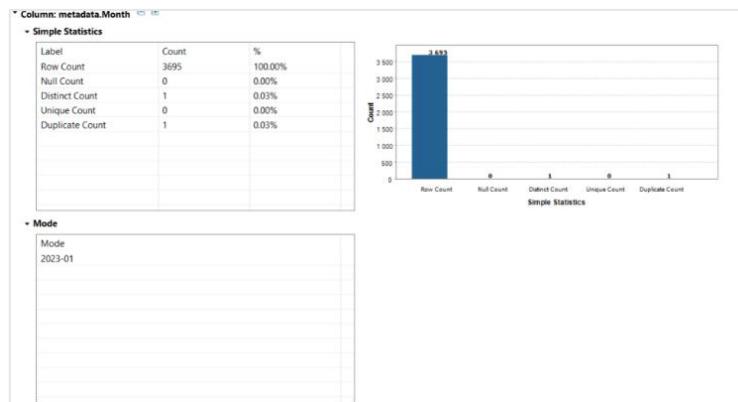
2021:



2022:



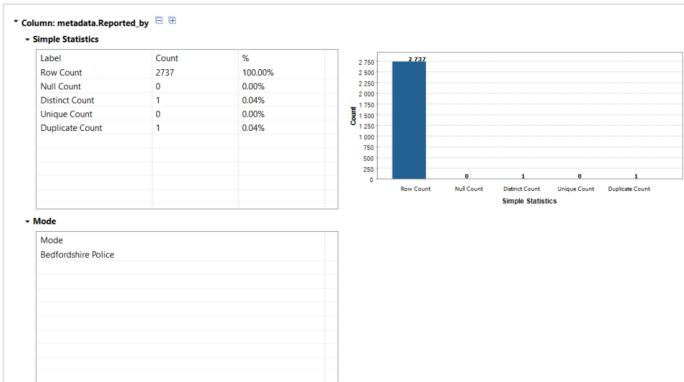
2023:



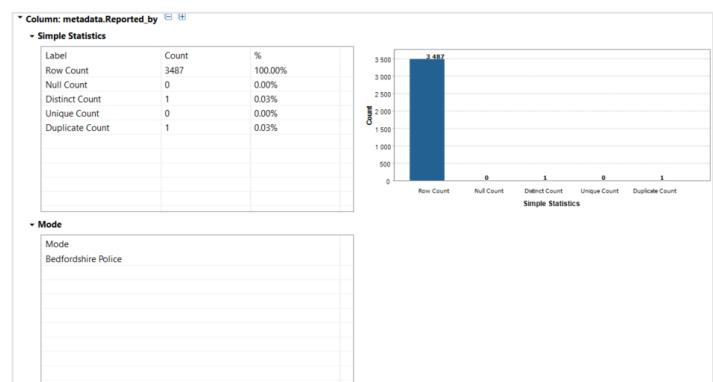
Atributo	Reported_by
Tipo de Dados	String
Descrição	Por quem o crime foi reportado
Exemplo	Bedfordshire Police
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados.

Estatísticas dos Dados:

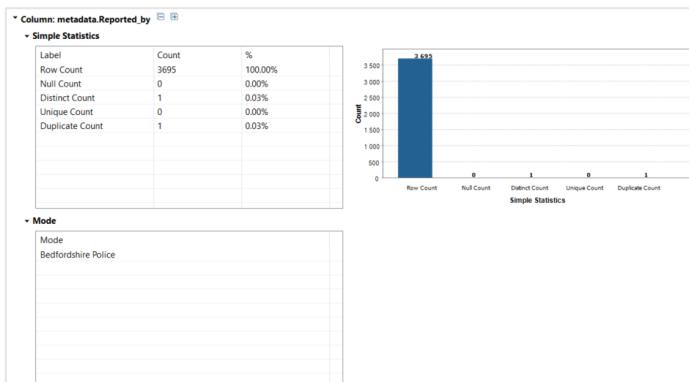
2020:



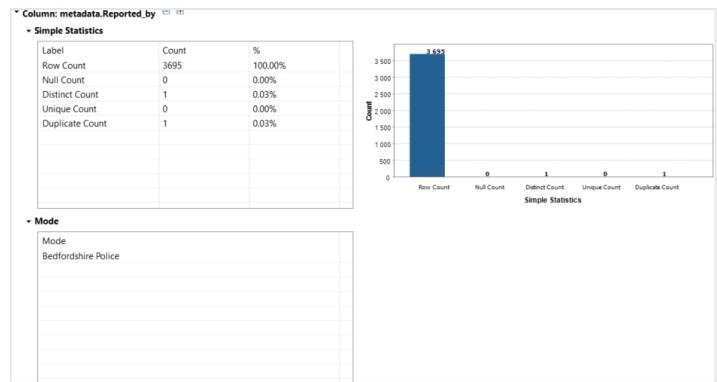
2021:



2022:



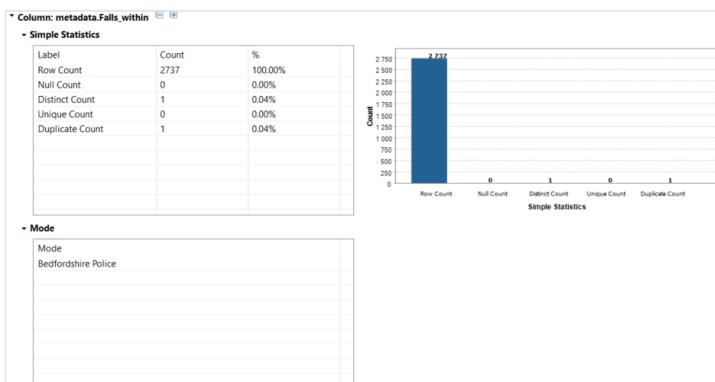
2023:



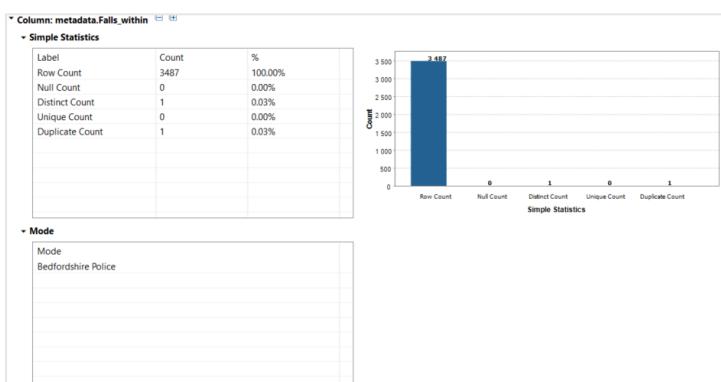
Atributo	Falls_within
Tipo de Dados	String
Descrição	A jurisdição dentro da qual o crime ocorreu
Exemplo	Bedfordshire Police
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade de Dados	Não detetados

Estatísticas dos Dados:

2020:



2021:



2022:

* Column: metadata.Falls_within

- Simple Statistics

Label	Count	%
Row Count	3695	100.00%
Null Count	0	0.00%
Distinct Count	1	0.03%
Unique Count	0	0.00%
Duplicate Count	1	0.03%

- Mode

Mode
Bedfordshire Police

2023:

* Column: metadata.Falls_within

- Simple Statistics

Label	Count	%
Row Count	3695	100.00%
Null Count	0	0.00%
Distinct Count	1	0.03%
Unique Count	0	0.00%
Duplicate Count	1	0.03%

- Mode

Mode
Bedfordshire Police

Atributo	Longitude
Tipo de Dados	Float
Descrição	A longitude do local onde o crime ocorreu.
Exemplo	-0.454961
Preenchido	97,14%
Intervalo de Dados	Não mensurável
Problema de Qualidade de Dados	Valores em falta

Estatísticas dos Dados:

2020:

* Column: metadata.Longitude

- Simple Statistics

Label	Count	%
Row Count	2737	100.00%
Null Count	0	0.00%
Distinct Count	1558	56.92%
Unique Count	1009	36.87%
Duplicate Count	549	20.06%

- Mode

Mode

2021:

* Column: metadata.Longitude

- Simple Statistics

Label	Count	%
Row Count	3487	100.00%
Null Count	84	2.41%
Distinct Count	1838	52.71%
Unique Count	1111	31.86%
Duplicate Count	727	20.85%

- Summary Statistics

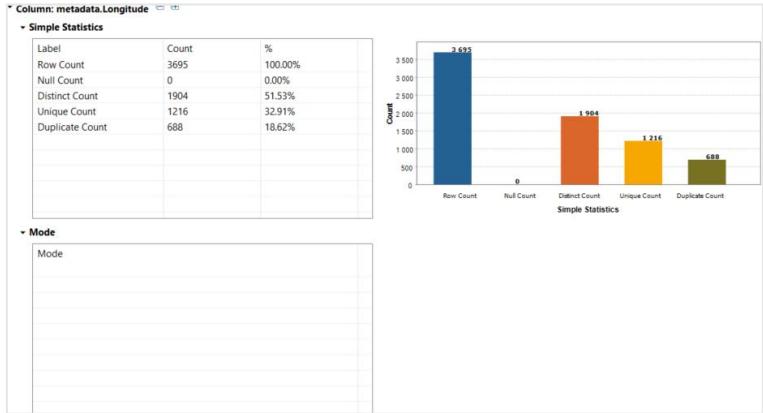
Label	Value
Mean	-0.45698714340287977
Median	-0.454961

- Mode

2022:



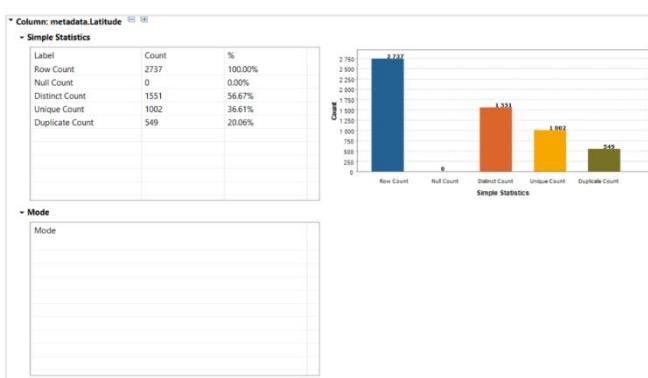
2023:



Atributo	Latitude
Tipo de Dados	Float
Descrição	A latitude do local onde o crime ocorreu
Exemplo	51.916176
Preenchido	97,14%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Valores em falta

Estatísticas dos Dados:

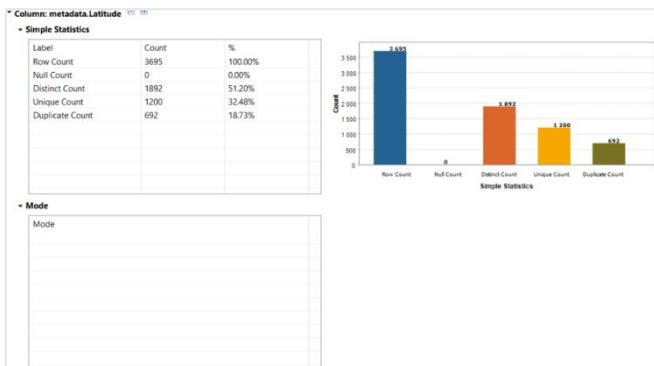
2020:



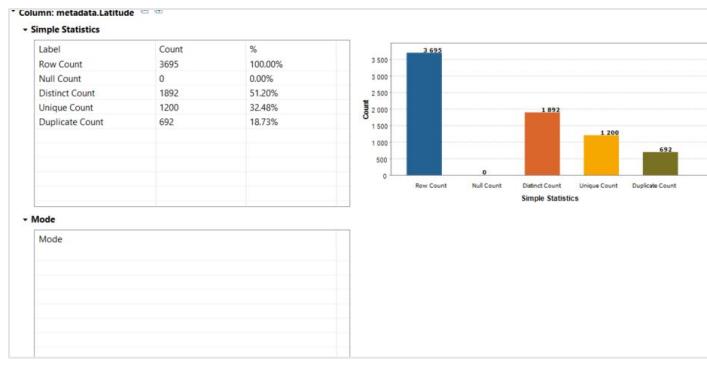
2021:



2022:



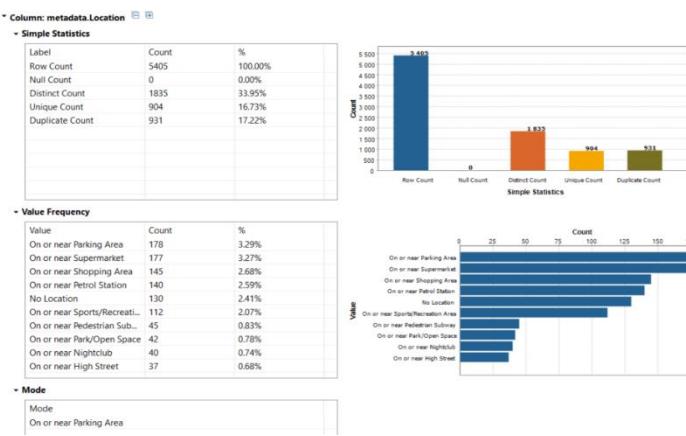
2023:



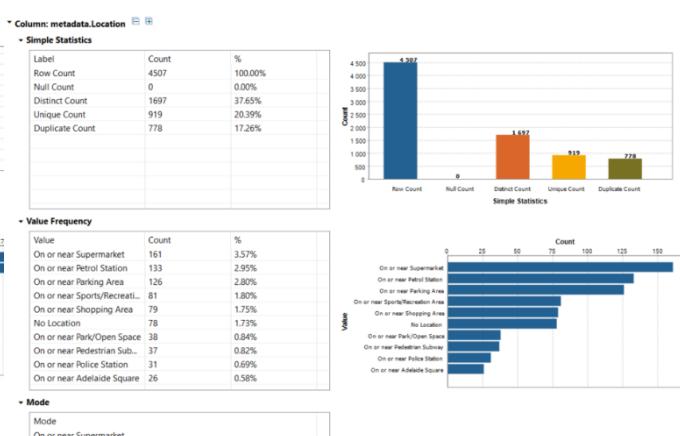
Atributo	Location
Tipo de Dados	String
Descrição	A descrição textual onde o crime ocorreu
Exemplo	On near Supermarket
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados

Estatísticas dos Dados:

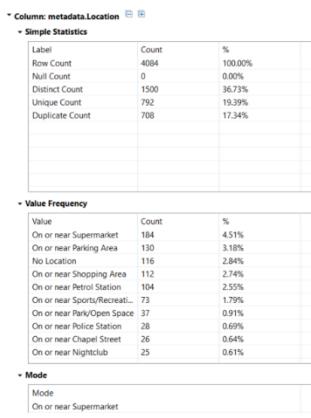
2020:



2021:



2022:



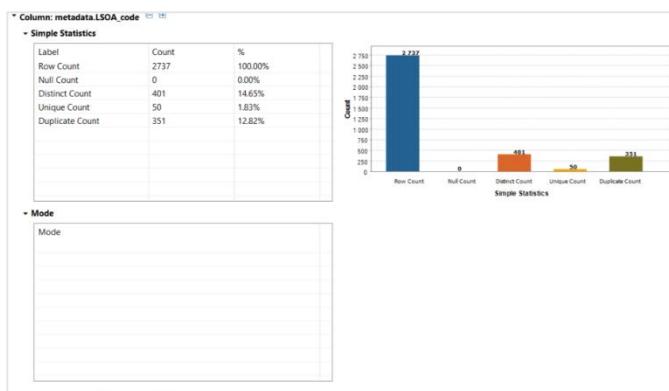
2023:



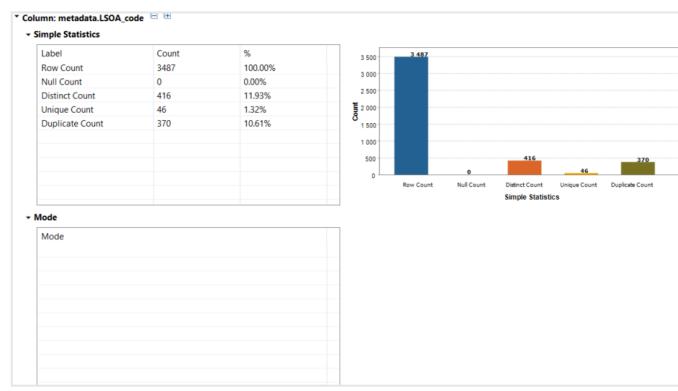
Atributo	LSOA code
Tipo de Dados	String
Descrição	O código da Área de Estatísticas de Saída de Nível Inferior (Lower Layer Super Output Area - LSOA) onde o crime ocorreu.
Exemplo	E01033801
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados

Estatísticas de Dados:

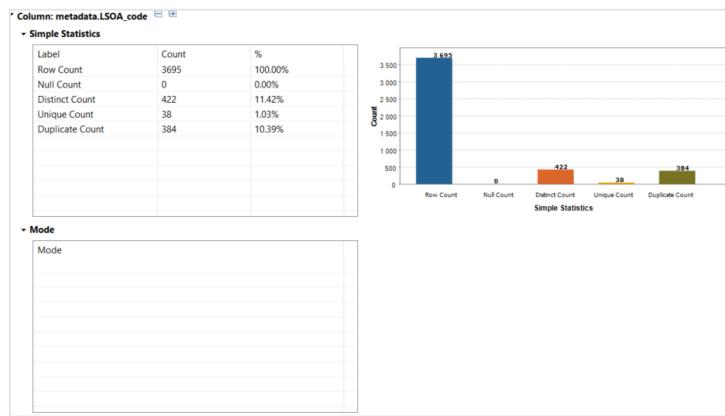
2020:



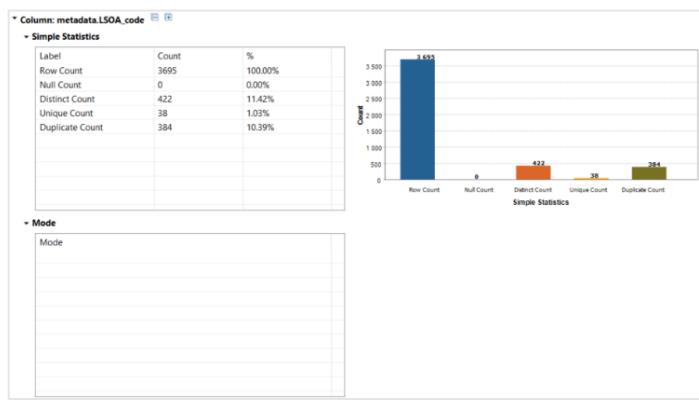
2021:



2022:



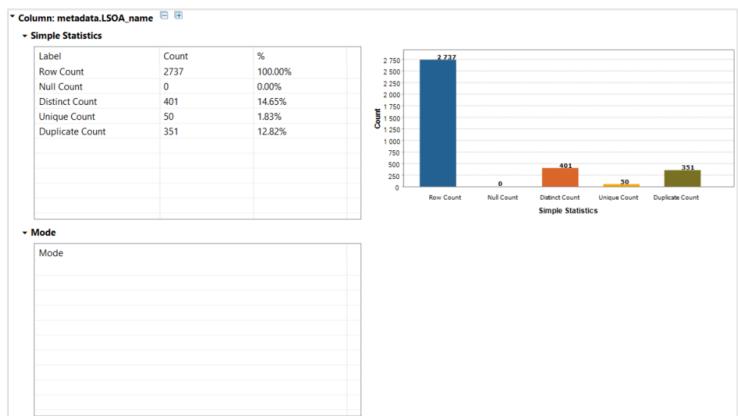
2023:



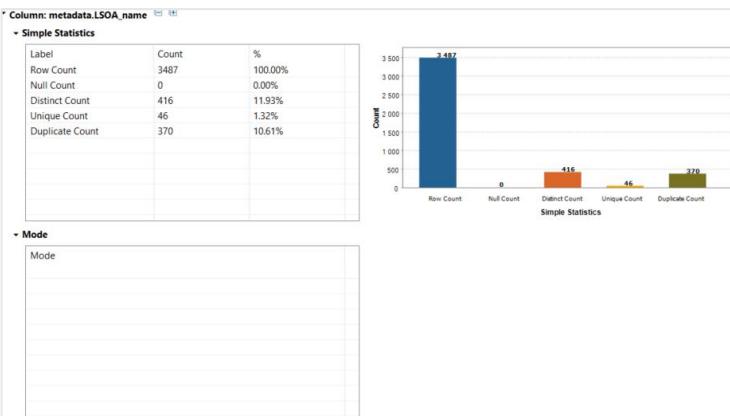
Atributo	LSOA Name
Tipo de Dados	String
Descrição	O nome da LSOA onde o crime ocorreu
Exemplo	Luton 022C
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados

Estatísticas dos Dados:

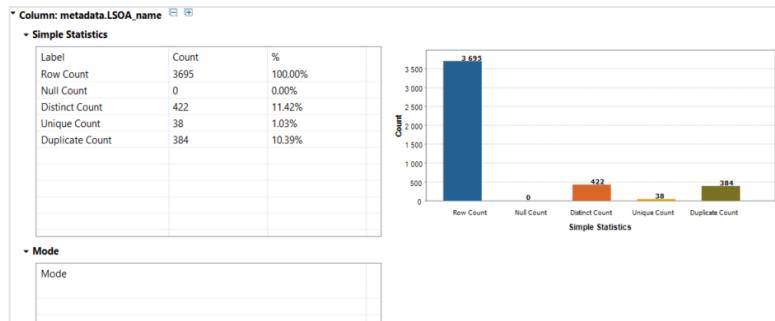
2020:



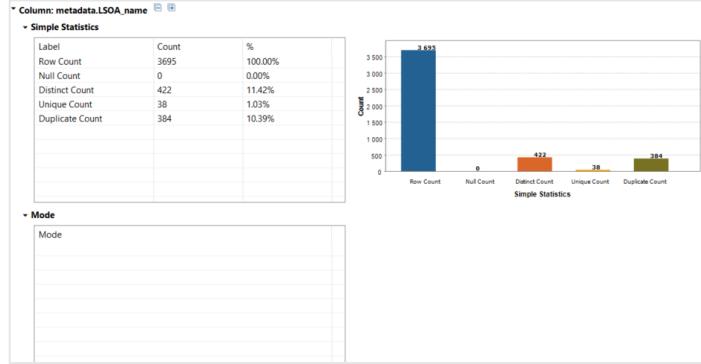
2021:



2022:



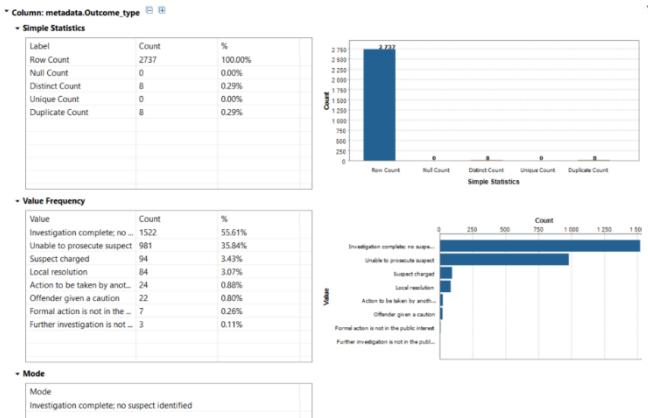
2023:



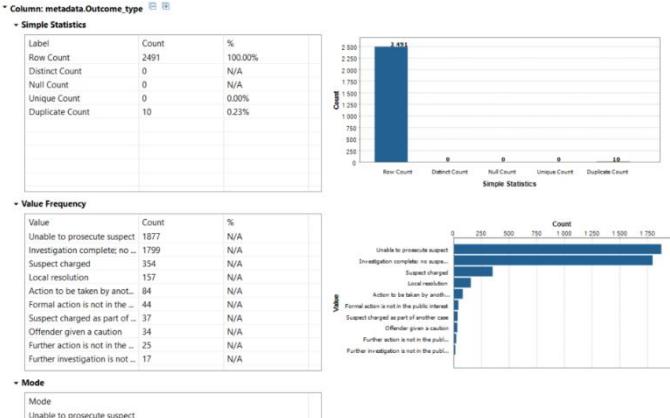
Atributo	Outcome_type
Tipo de Dados	String
Descrição	O tipo de resultado ou status da investigação do crime (por exemplo, "Ação a ser tomada por outra organização", "Incapaz de processar suspeito", "Investigação completa; nenhum suspeito identificado", etc.).
Exemplo	Unable to prosecute suspect
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade de Dados	Possíveis dados categóricos representados como strings

Estatísticas de Dados:

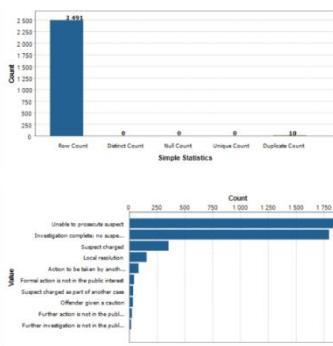
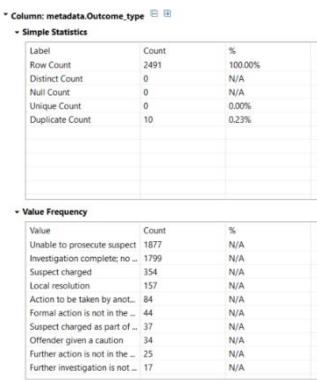
2020:



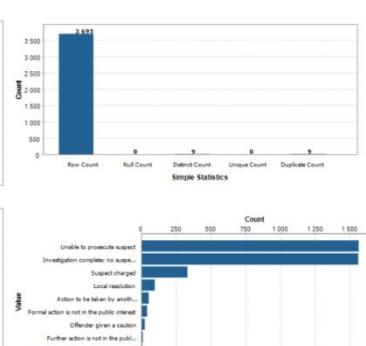
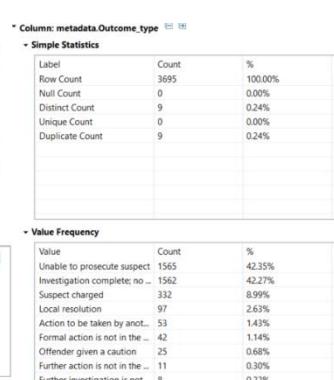
2021:



2022:



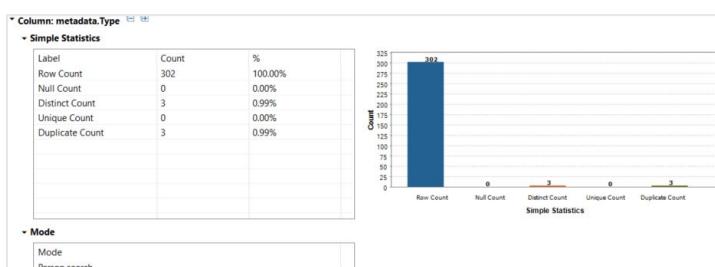
2023:



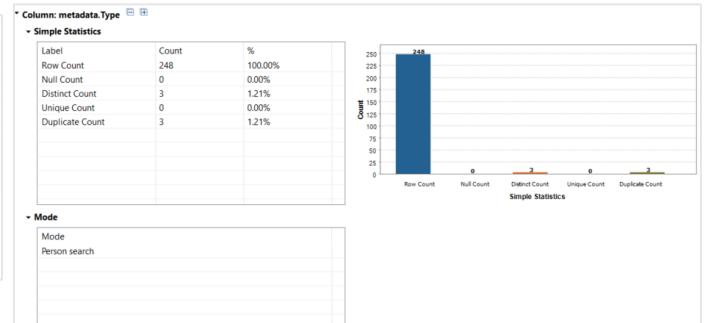
Atributo	Type
Tipo de Dados	String
Descrição	O tipo de busca realizada, que pode ser uma busca de pessoa, veículo, etc.
Exemplo	Person search
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados

Estatísticas dos Dados:

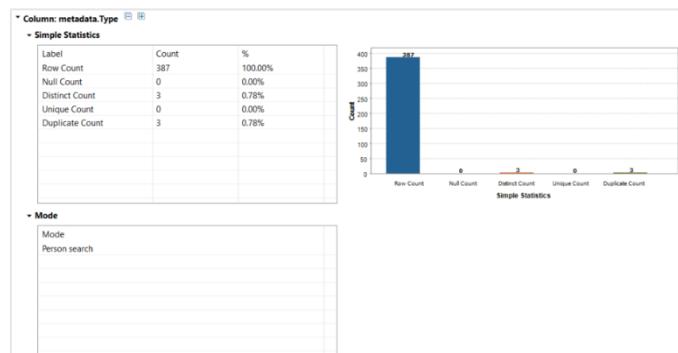
2020:



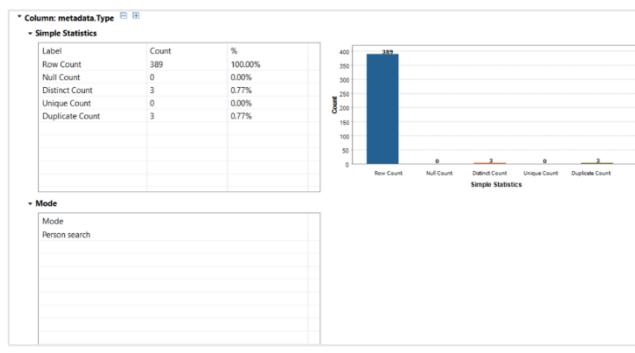
2021:



2022:



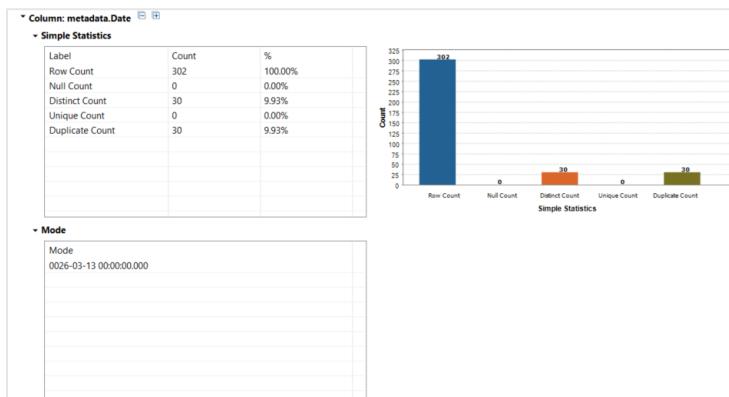
2023:



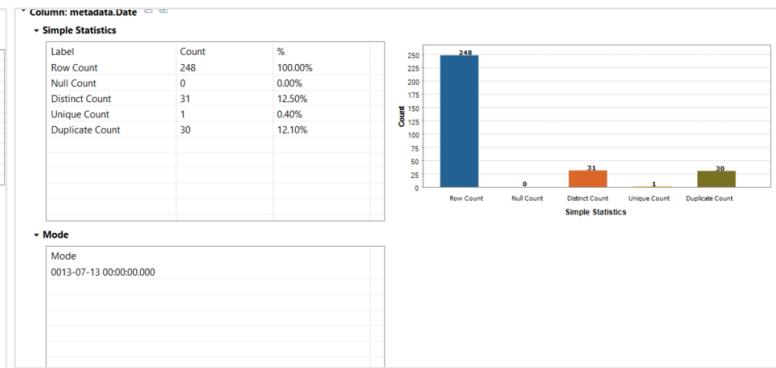
Atributo	Date
Tipo de Dados	String
Descrição	A data e hora em que a busca foi realizada.
Exemplo	1
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Apesar do formato ser em String, seria mais adequado converter em Datetime

Estatísticas dos Dados:

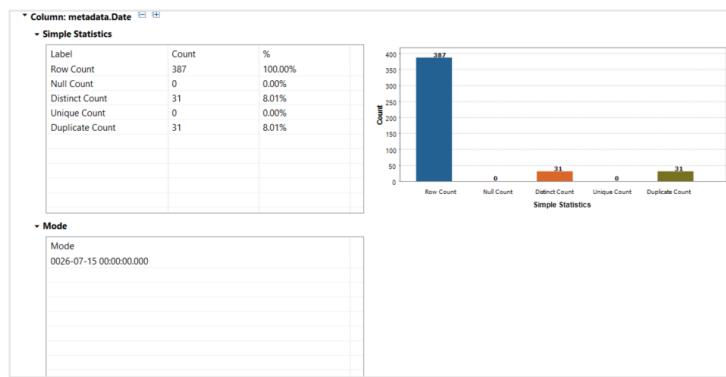
2020:



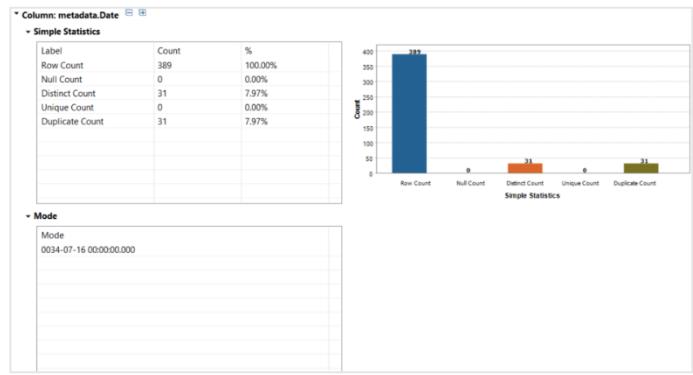
2021:



2022:



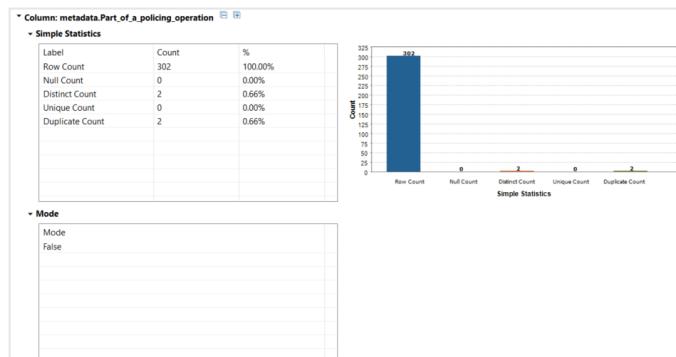
2023:



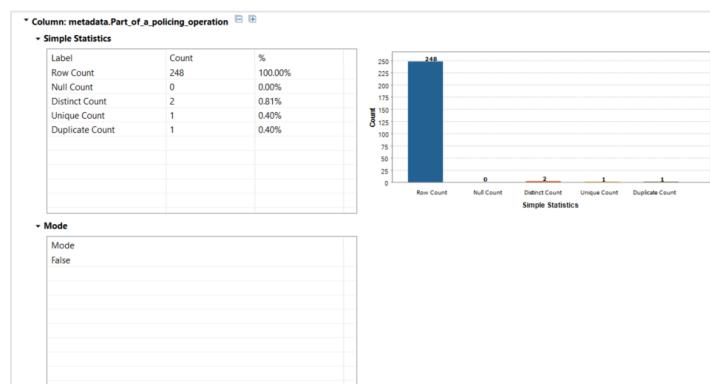
Atributo	Part of a policing operation
Tipo de Dados	Boolean
Descrição	Se a busca fazia parte de uma operação policial.
Exemplo	False
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Valores em falta

Estatísticas dos Dados:

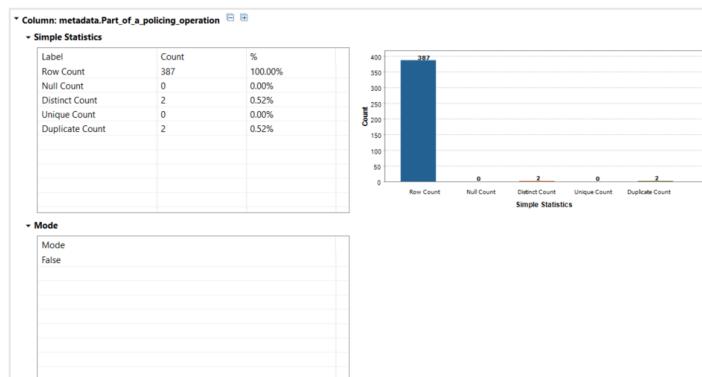
2020:



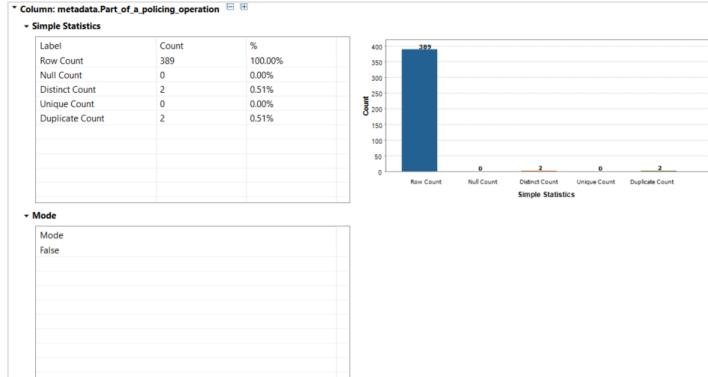
2021:



2022:



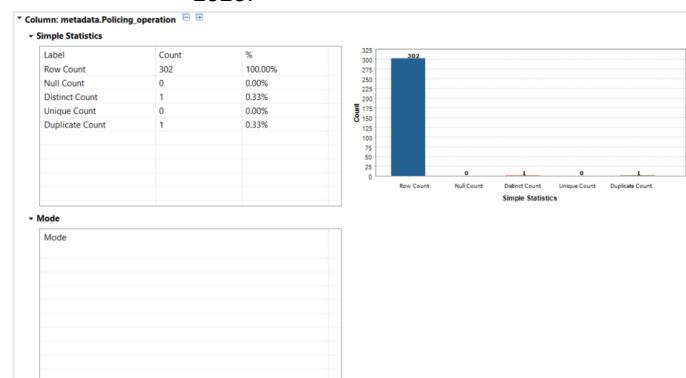
2023:



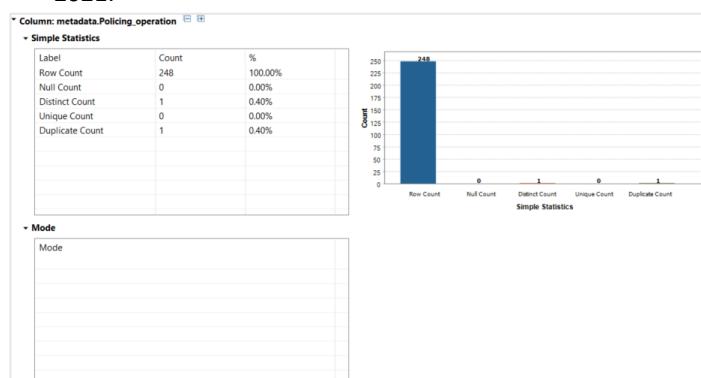
Atributo	Policing Operation
Tipo de Dados	Boolean
Descrição	Detalhes da operação policial, se aplicável
Exemplo	
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Grande parte das entradas são null, indicando que os detalhes da operação não estão disponíveis.

Estatísticas dos Dados:

2020:



2021:



2022:

Column: metadata.Policing_operation

Simple Statistics

Label	Count	%
Row Count	387	100.00%
Null Count	0	0.00%
Distinct Count	1	0.26%
Unique Count	0	0.00%
Duplicate Count	1	0.26%

Mode

Mode

Row Count Null Count Distinct Count Unique Count Duplicate Count

2023:

Column: metadata.Policing_operation

Simple Statistics

Label	Count	%
Row Count	389	100.00%
Null Count	0	0.00%
Distinct Count	1	0.26%
Unique Count	0	0.00%
Duplicate Count	1	0.26%

Mode

Mode

Row Count Null Count Distinct Count Unique Count Duplicate Count

Atributo	Gender
Tipo de Dados	Object (String)
Descrição	O género da pessoa revistada
Exemplo	Male
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detatdos

Estatísticas dos Dados:

2020:

Column: metadata.Gender

Simple Statistics

Label	Count	%
Row Count	302	100.00%
Null Count	0	0.00%
Distinct Count	3	0.99%
Unique Count	0	0.00%
Duplicate Count	3	0.99%

Mode

Mode
Male

Row Count Null Count Distinct Count Unique Count Duplicate Count

2021:

Column: metadata.Gender

Simple Statistics

Label	Count	%
Row Count	248	100.00%
Null Count	0	0.00%
Distinct Count	3	1.21%
Unique Count	0	0.00%
Duplicate Count	3	1.21%

Mode

Mode
Male

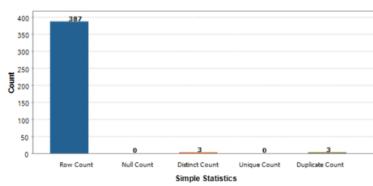
Row Count Null Count Distinct Count Unique Count Duplicate Count

2022:

* Column: metadata.Gender

Simple Statistics

Label	Count	%
Row Count	387	100.00%
Null Count	0	0.00%
Distinct Count	3	0.78%
Unique Count	0	0.00%
Duplicate Count	3	0.78%



- Mode

Mode

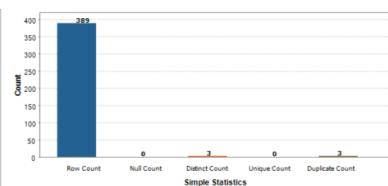
Male

2023:

* Column: metadata.Gender

Simple Statistics

Label	Count	%
Row Count	389	100.00%
Null Count	0	0.00%
Distinct Count	3	0.77%
Unique Count	0	0.00%
Duplicate Count	3	0.77%



- Mode

Mode

Male

Atributo	Age Range
Tipo de Dados	String
Descrição	A faixa etária da pessoa revistada
Exemplo	18-24
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Campo

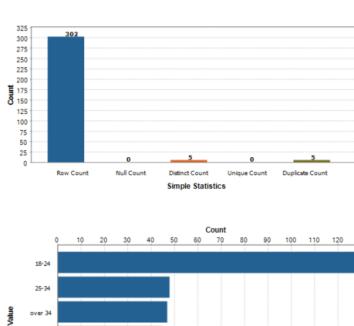
Estatísticas dos Dados:

2020:

* Column: metadata.Age_range

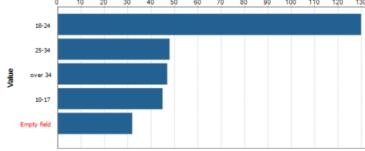
Simple Statistics

Label	Count	%
Row Count	302	100.00%
Null Count	0	0.00%
Distinct Count	5	1.66%
Unique Count	0	0.00%
Duplicate Count	5	1.66%



- Value Frequency

Value	Count	%
18-24	130	43.05%
25-34	48	15.89%
over 34	47	15.56%
10-17	45	14.90%
Empty field	32	10.60%



- Mode

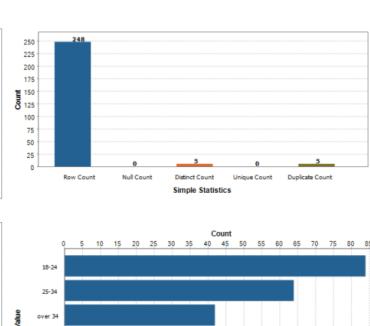
Mode
18-24

2021:

* Column: metadata.Age_range

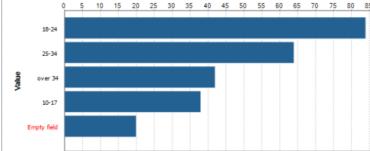
Simple Statistics

Label	Count	%
Row Count	248	100.00%
Null Count	0	0.00%
Distinct Count	5	2.02%
Unique Count	0	0.00%
Duplicate Count	5	2.02%



- Value Frequency

Value	Count	%
18-24	84	33.87%
25-34	64	25.81%
over 34	42	16.94%
10-17	38	15.32%
Empty field	20	8.06%



- Mode

Mode
18-24

2022:

* Column: metadata.Age_range

- Simple Statistics

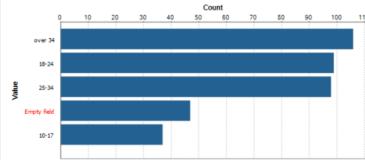
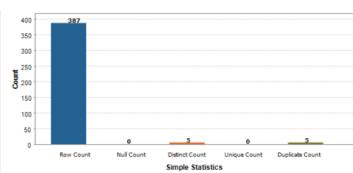
Label	Count	%
Row Count	387	100.00%
Null Count	0	0.00%
Distinct Count	5	1.29%
Unique Count	0	0.00%
Duplicate Count	5	1.29%

- Value Frequency

Value	Count	%
over 34	106	27.39%
18-24	99	25.58%
25-34	98	25.32%
Empty field	47	12.14%
10-17	37	9.56%

- Mode

Mode	over 34
------	---------



2023:

* Column: metadata.Age_range

- Simple Statistics

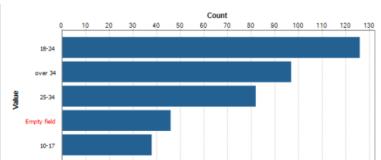
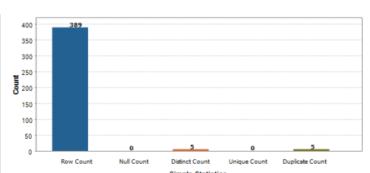
Label	Count	%
Row Count	389	100.00%
Null Count	0	0.00%
Distinct Count	5	1.29%
Unique Count	0	0.00%
Duplicate Count	5	1.29%

- Value Frequency

Value	Count	%
18-24	126	32.39%
over 34	97	24.94%
25-34	82	21.08%
Empty field	46	11.83%
10-17	38	9.77%

- Mode

Mode	18-24
------	-------



Atributo	Sel-defined ethnicity
Tipo de Dados	String
Descrição	A etnia definida pela pessoa revistada.
Exemplo	Asian/Asian British - Pakistani
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade de Dados	Não detetados

Estatísticas dos Dados:

2020:

* Column: metadata.Self_defined_ethnicity

- Simple Statistics

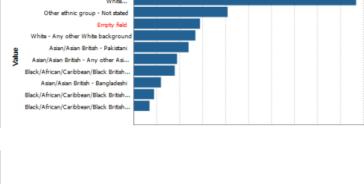
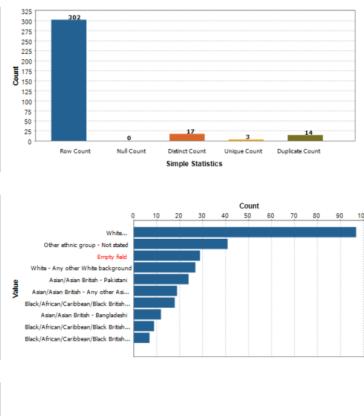
Label	Count	%
Row Count	302	100.00%
Null Count	0	0.00%
Distinct Count	17	5.63%
Unique Count	3	0.99%
Duplicate Count	14	4.64%

- Value Frequency

Value	Count	%
White - English/Welsh/Scot...	97	32.12%
Other ethnic group - Not st...	41	13.58%
Empty field	29	9.60%
White - Any other White ba...	27	8.94%
Asian/Asian British - Pakist...	24	7.95%
Asian/Asian British - Any ot...	19	6.29%
Black/African/Caribbean/Bl...	18	5.96%
Asian/Asian British - Bangl...	12	3.97%
Black/African/Caribbean/Bl...	9	2.98%
Black/African/Caribbean/Bl...	7	2.32%

- Mode

Mode	White - English/Welsh/Scottish/Northern Irish/British
------	---



2021:

* Column: metadata.Self_defined_ethnicity

- Simple Statistics

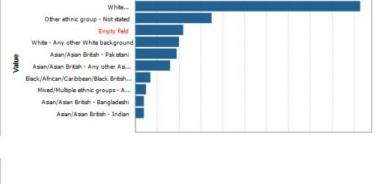
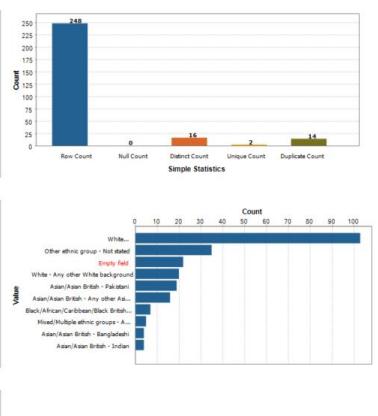
Label	Count	%
Row Count	248	100.00%
Null Count	0	0.00%
Distinct Count	16	6.45%
Unique Count	2	0.81%
Duplicate Count	14	5.65%

- Value Frequency

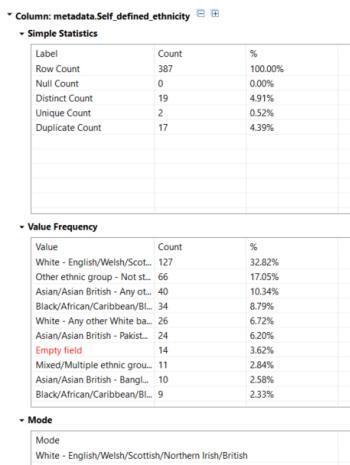
Value	Count	%
White - English/Welsh/Scot...	103	41.53%
Other ethnic group - Not st...	35	14.11%
Empty field	22	8.87%
White - Any other White ba...	20	8.06%
Asian/Asian British - Pakist...	19	7.66%
Asian/Asian British - Any ot...	16	6.45%
Black/African/Caribbean/Bl...	7	2.82%
Mixed/Multiple ethnic grou...	5	2.02%
Asian/Asian British - Bangl...	4	1.61%
Asian/Asian British - Indian	4	1.61%

- Mode

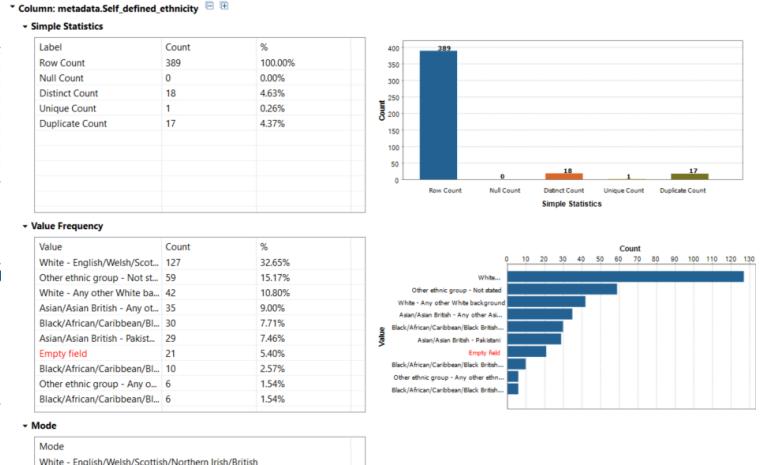
Mode	White - English/Welsh/Scottish/Northern Irish/British
------	---



2022:



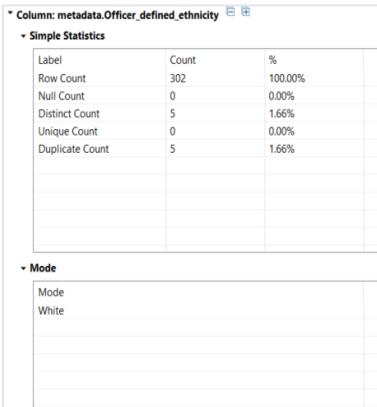
2023:



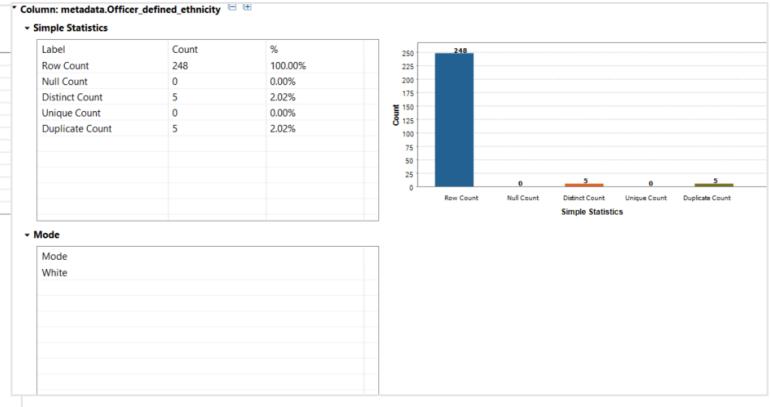
Atributo	Officer-defined ethnicity
Tipo de Dados	String
Descrição	A etnia definida pelo oficial que realizou a busca
Exemplo	White
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados

Estatísticas dos Dados:

2020:



2021:



2022:

* Column: metadata.Officer_defined_ethnicity

Simple Statistics

Label	Count	%
Row Count	387	100.00%
Null Count	0	0.00%
Distinct Count	4	1.03%
Unique Count	0	0.00%
Duplicate Count	4	1.03%

Mode

Mode
White

2023:

* Column: metadata.Officer_defined_ethnicity

Simple Statistics

Label	Count	%
Row Count	389	100.00%
Null Count	0	0.00%
Distinct Count	5	1.29%
Unique Count	0	0.00%
Duplicate Count	5	1.29%

Mode

Mode
White

Atributo	Legislation
Tipo de Dados	String
Descrição	A legislação sob a qual a busca foi realizada.
Exemplo	Misuse of Drugs Act 1971 (section 23)
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados

Estatísticas dos Dados:

2020:

* Column: metadata.Legislation

Simple Statistics

Label	Count	%
Row Count	302	100.00%
Null Count	0	0.00%
Distinct Count	2	0.66%
Unique Count	0	0.00%
Duplicate Count	2	0.66%

Value Frequency

Value	Count	%
Misuse of Drugs Act 1971 (section 23)	233	77.15%
Police and Criminal Evidence Act 1984	69	22.85%

Mode

Mode
Misuse of Drugs Act 1971 (section 23)

2021:

* Column: metadata.Legislation

Simple Statistics

Label	Count	%
Row Count	248	100.00%
Null Count	0	0.00%
Distinct Count	3	1.21%
Unique Count	0	0.00%
Duplicate Count	3	1.21%

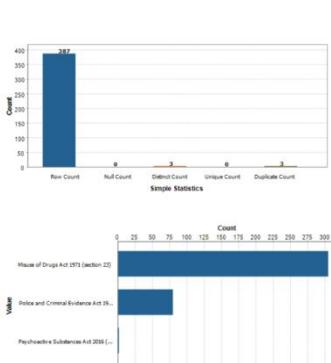
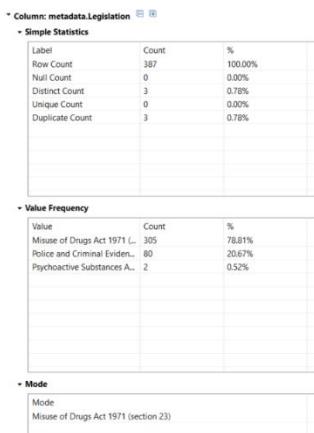
Value Frequency

Value	Count	%
Misuse of Drugs Act 1971 (section 23)	166	66.94%
Police and Criminal Evidence Act 1984	78	31.45%
Hunting Act 2004 (section 8)	4	1.61%

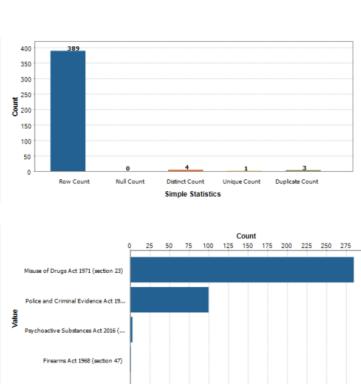
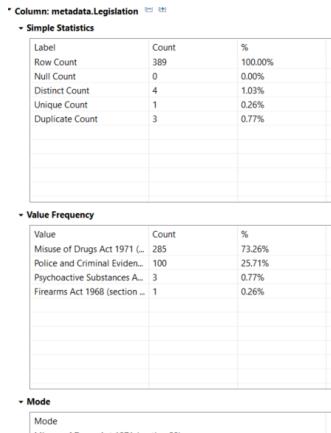
Mode

Mode
Misuse of Drugs Act 1971 (section 23)

2022:



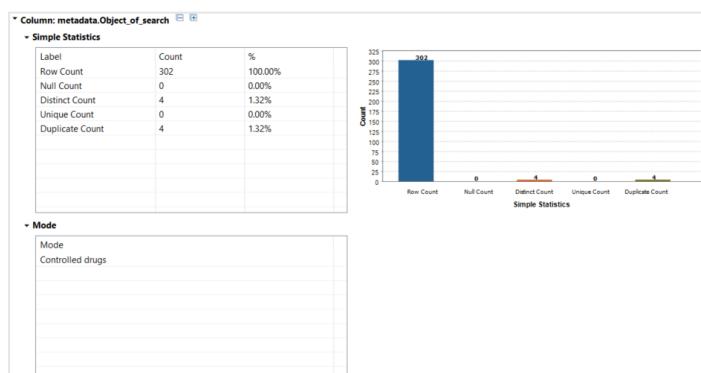
2023:



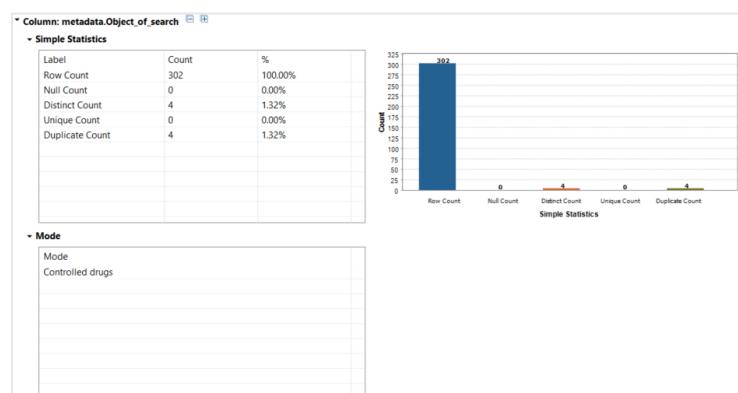
Atributo	object of search
Tipo de Dados	String
Descrição	Qual era o objetivo da busca.
Exemplo	Controlled drugs
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados

Estatísticas dos Dados:

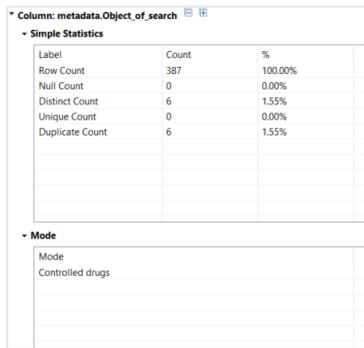
2020:



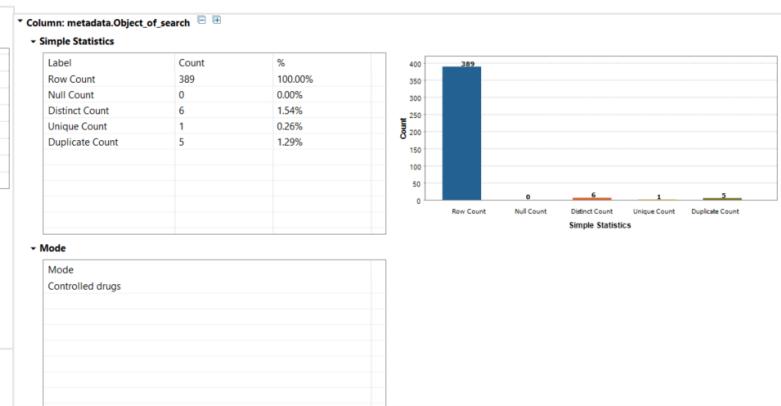
2021:



2022:



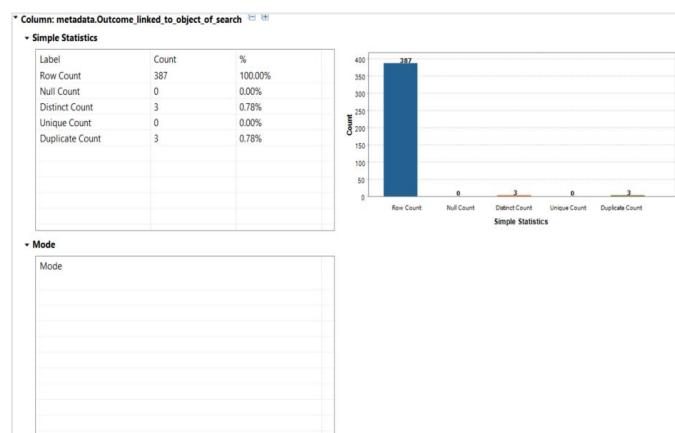
2023:



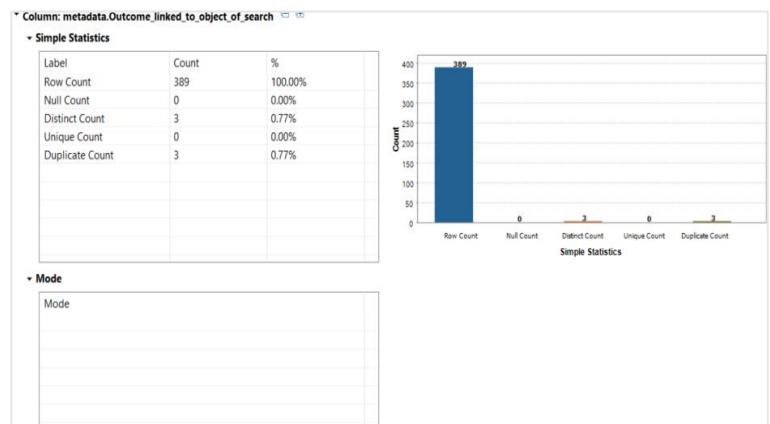
Atributo	Outcome linked to object of search
Tipo de Dados	Boolean
Descrição	Se o resultado está ligado ao objeto da busca.
Exemplo	True
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados

Estatísticas dos Dados:

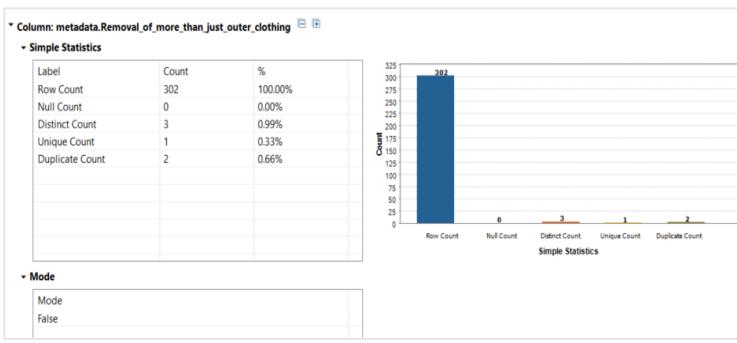
2020:



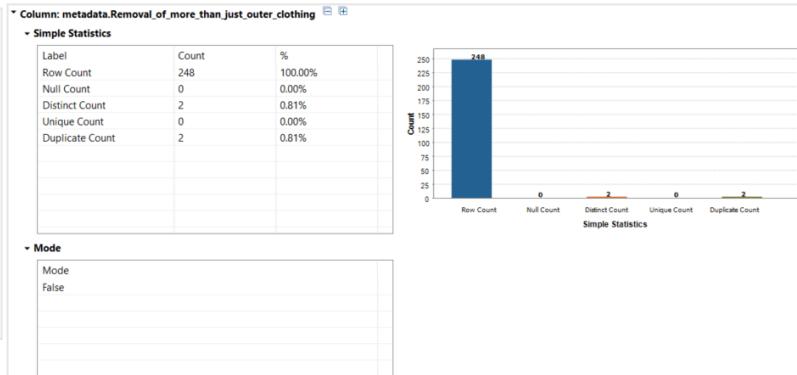
2021:



2022:



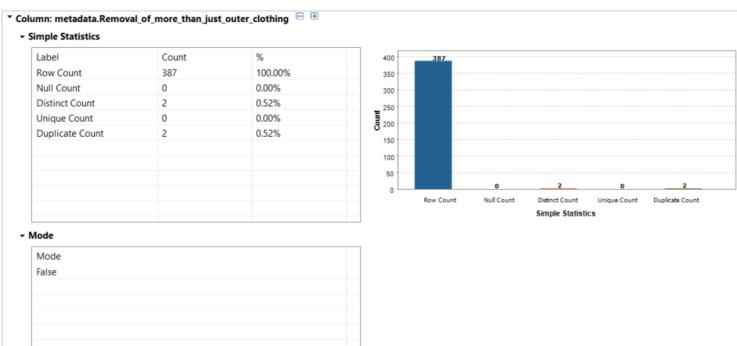
2023:



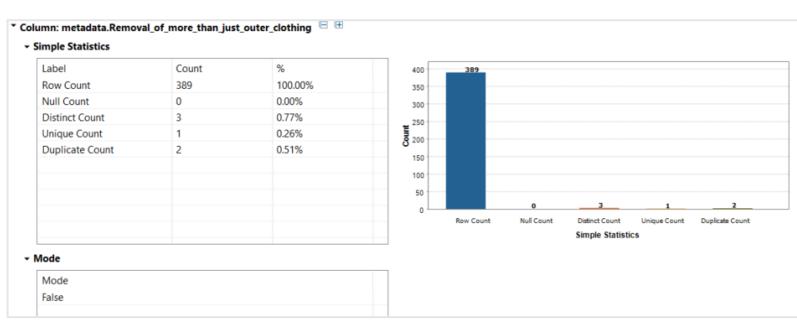
Atributo	Removal_of_more_than_just_outer_clothing
Tipo de Dados	Boolean
Descrição	Se a busca envolveu a remoção de mais do que apenas a roupa exterior.
Exemplo	False
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados

Estatísticas dos Dados:

2020:



2021:



2022:

* Column: metadata.Removal_of_more_than_just_outer_clothing

Simple Statistics

Label	Count	%
Row Count	387	100.00%
Null Count	0	0.00%
Distinct Count	2	0.52%
Unique Count	0	0.00%
Duplicate Count	2	0.52%

Mode

Mode
False

Count

2023:

* Column: metadata.Removal_of_more_than_just_outer_clothing

Simple Statistics

Label	Count	%
Row Count	389	100.00%
Null Count	0	0.00%
Distinct Count	3	0.77%
Unique Count	1	0.26%
Duplicate Count	2	0.51%

Mode

Mode
False

Count

Atributo	Crime_type
Tipo de Dados	String
Descrição	O tipo de crime cometido (por exemplo, "Violência e crimes sexuais", "Comportamento anti-social", etc.).
Exemplo	Public order
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados

Estatísticas dos Dados:

2020:

Column: metadata.Crime_type

Simple Statistics

Label	Count	%
Row Count	5405	100.00%
Null Count	0	0.00%
Distinct Count	14	0.26%
Unique Count	0	0.00%
Duplicate Count	14	0.26%

Value Frequency

Value	Count	%
Violence and sexual offence	1503	27.81%
Anti-social behaviour	1288	23.83%
Public order	469	8.68%
Criminal damage and arson	424	7.84%
Vehicle crime	399	7.38%
Other theft	366	6.77%
Burglary	270	5.00%
Shoplifting	254	4.70%
Drugs	151	2.79%
Bicycle theft	81	1.50%

Mode

Mode
Violence and sexual offences

Count

2021:

Column: metadata.Crime_type

Simple Statistics

Label	Count	%
Row Count	4084	100.00%
Null Count	0	0.00%
Distinct Count	14	0.34%
Unique Count	0	0.00%
Duplicate Count	14	0.34%

Value Frequency

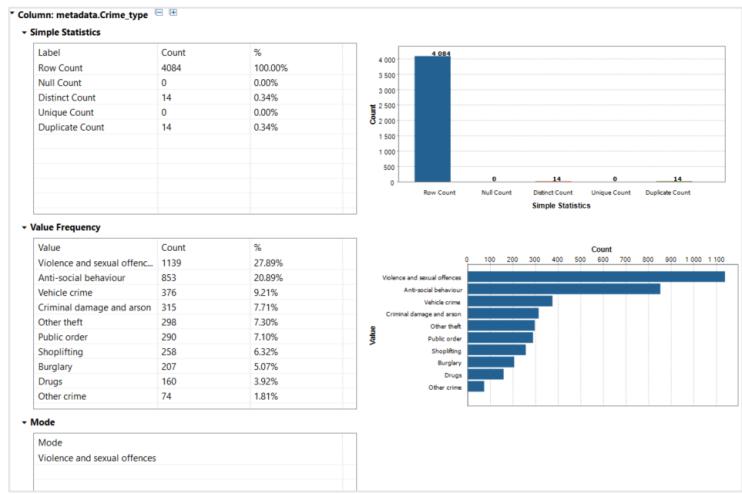
Value	Count
Violence and sexual offence	1094
Anti-social behaviour	921
Public order	758
Criminal damage and arson	685
Vehicle crime	621
Other theft	548
Burglary	485
Shoplifting	422
Drugs	359
Bicycle theft	296

Mode

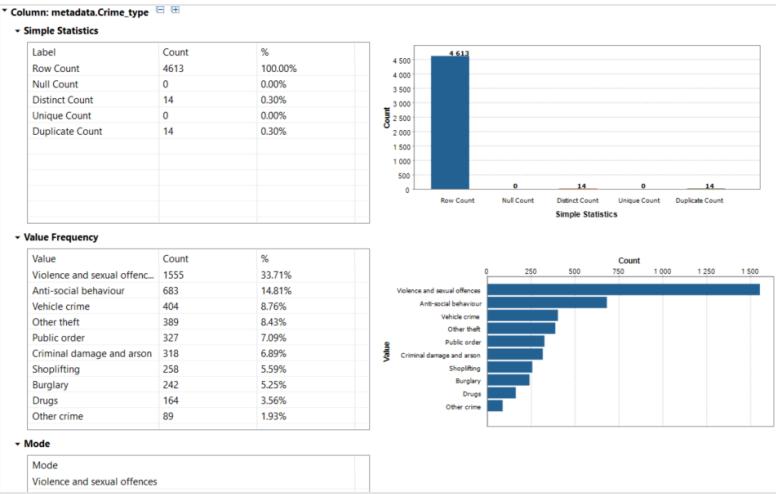
Mode
Violence and sexual offences

Count

2022:



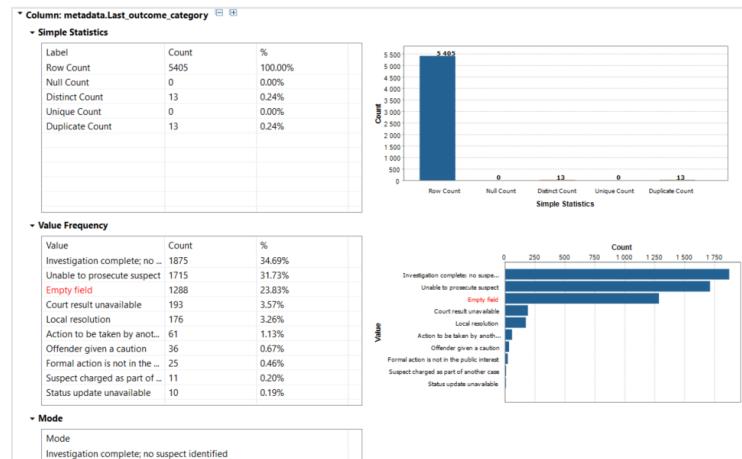
2023:



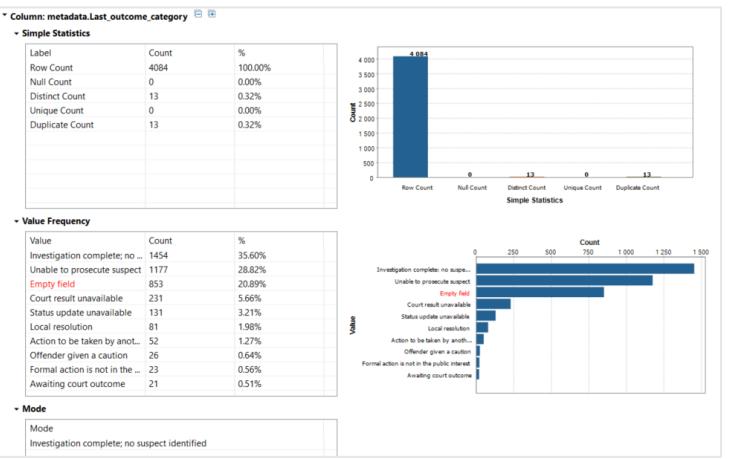
Atributo	Last outcome category
Tipo de Dados	String
Descrição	A categoria do último resultado relacionado ao crime
Exemplo	
Preenchido	76.17%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Várias entradas null

Estatísticas dos Dados:

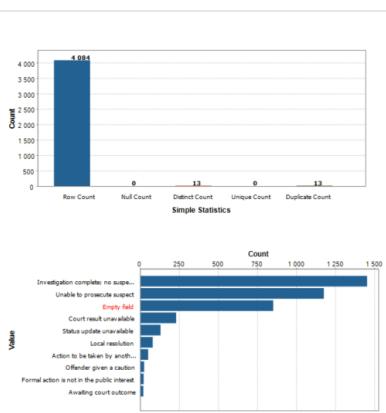
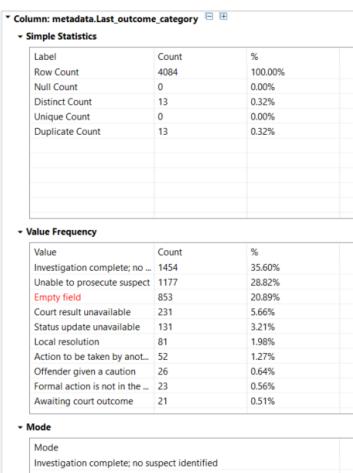
2020:



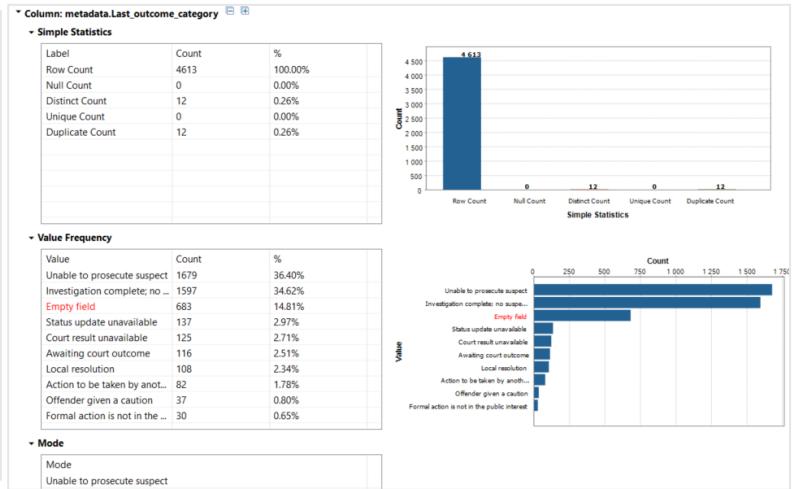
2021:



2022:



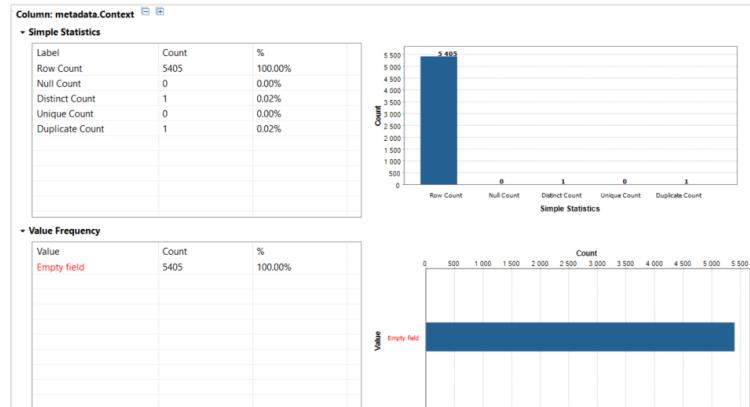
2023:



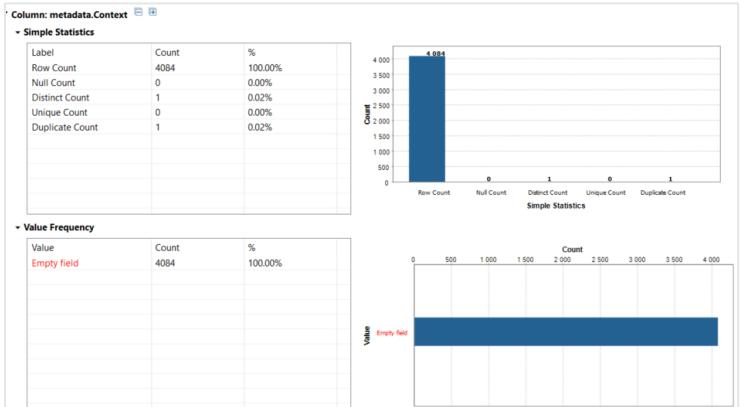
Atributo	Context
Tipo de Dados	String
Descrição	Contexto adicional sobre o crime
Exemplo	
Preenchido	0%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Grande parte das entradas são null, indicando que nenhum contexto adicional está disponível.

Estatísticas dos Dados:

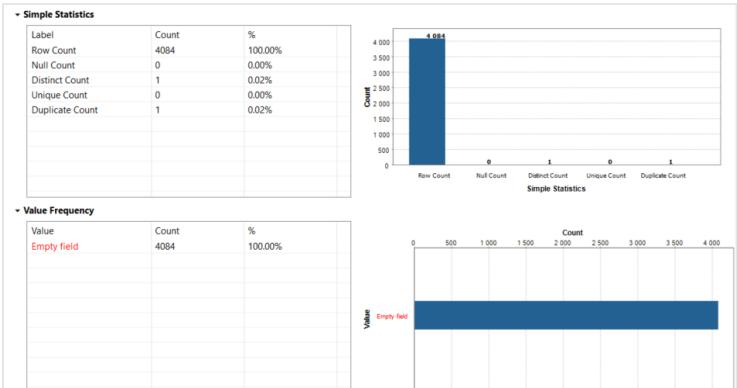
2020:



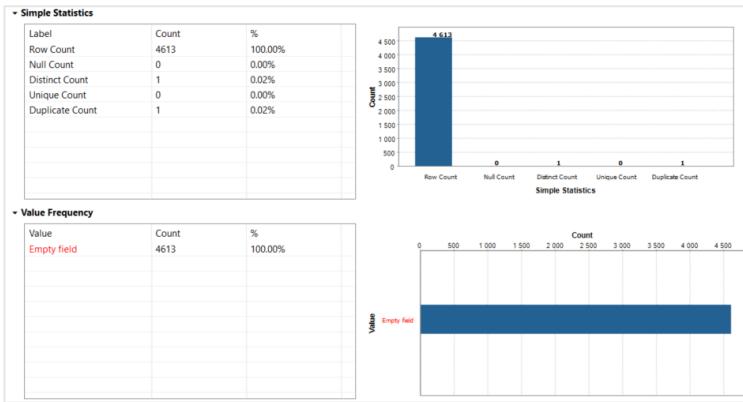
2021:



2022:



2023:



Dataset: Child_Protection_Dataset

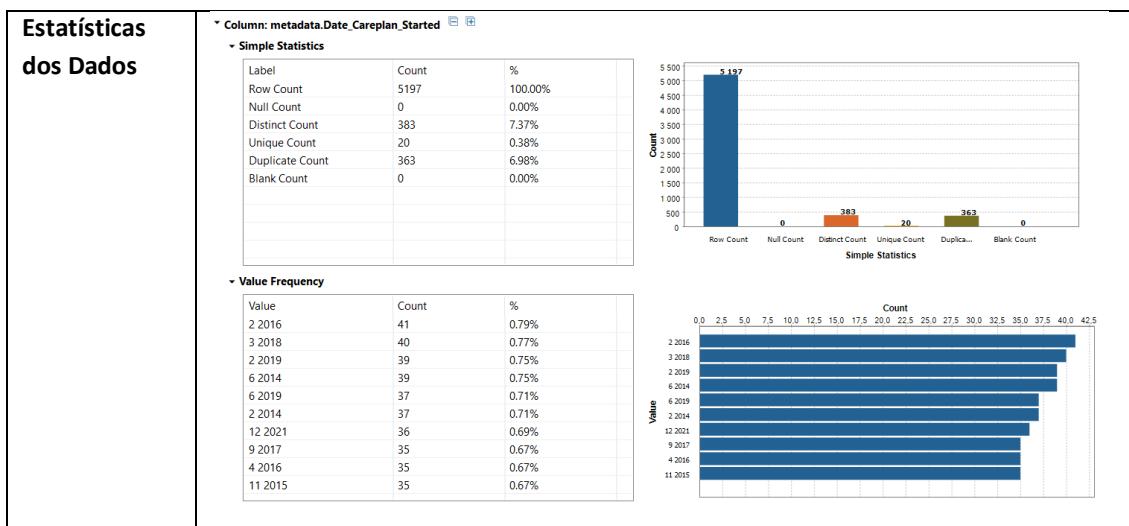
Aluno responsável: Catarina Fernandes (a101441)

Fonte:<https://www.data.gov.uk/dataset/7f2ff87b-6526-4f51-85fa-41ac4530534e/children-and-young-persons-subject-to-child-protection-plans>

Descrição: O seguinte dataset apresenta um conjunto de dados acerca de crianças e jovens sujeitos a planos de proteção infantil entre 1994 e 2021, para a cidade londrina de Barnet.

Child_Protection_Dataset.csv

Atributo	Data_Careplan_Started
Tipo de Dados	String
Descrição	Data Care Plan é um termo utilizado no Sistema de saúde do Reino Unido que se refere à data específica de um plano de cuidados
Exemplo	2 2001
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	A forma como esta representados os valores desta coluna não é a mais indicada para este tipo de dados. De forma a tornar os valores desta coluna mais coerente com todos os datasets iremos separar os valores por "Mounth_Careplan_Started" e "Year_Careplan_Started", assim dividindo o mês e o ano.



Atributo	Data_Careplan_Ended
Tipo de Dados	String
Descrição	Data Care Plan Started refere-se á data específica do fim de um plano de cuidados
Exemplo	12 2002
Preenchido	95,68%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	<p>Foram detetados dois problemas:</p> <ul style="list-style-type: none"> - O formato dos valores não é a mais adequada para o tipo de valores que contem. Neste caso, de forma a tornar os valores desta coluna mais coerente com todos os datasets, iremos separa os valores por "Mounth_Careplan_Ended" e "Year_Careplan_Ended", assim dividindo o mês e o ano. - Existência de valores em branco, no futuro essas linhas não serão eliminadas, mas sim substituídas pelo valor 0, de forma a representar os valores desconhecidos. <p>Isto acontecerá quanto fizermos a divisão do ano e o mês e passarmos o valor de string para date.</p>

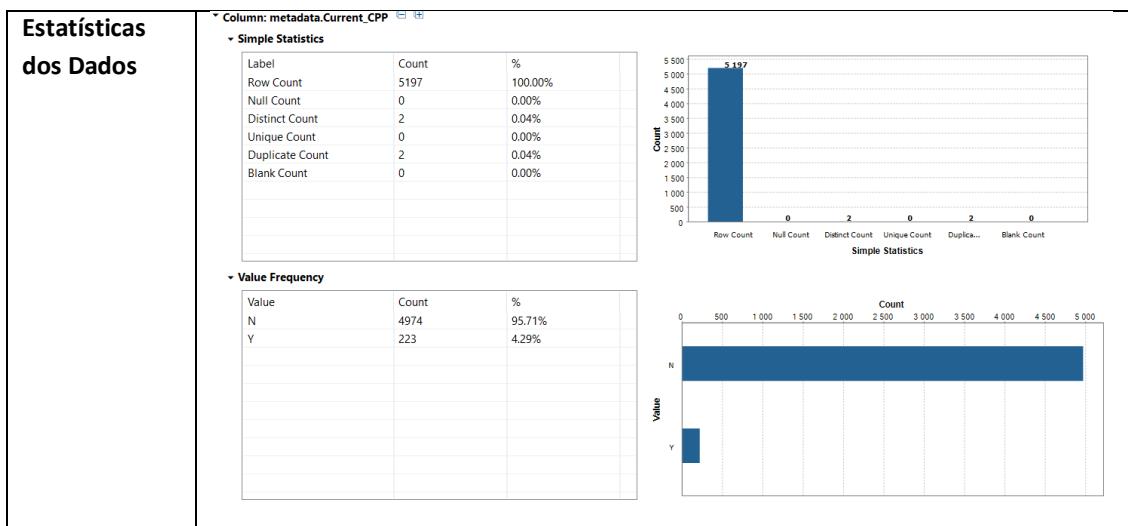
Estatísticas dos Dados	<p>* Column: metadata.Date_Careplan_Ended</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th><th>Count</th><th>%</th></tr> </thead> <tbody> <tr><td>Row Count</td><td>5197</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>354</td><td>6.81%</td></tr> <tr><td>Unique Count</td><td>14</td><td>0.27%</td></tr> <tr><td>Duplicate Count</td><td>340</td><td>6.54%</td></tr> <tr><td>Blank Count</td><td>223</td><td>4.29%</td></tr> </tbody> </table>	Label	Count	%	Row Count	5197	100.00%	Null Count	0	0.00%	Distinct Count	354	6.81%	Unique Count	14	0.27%	Duplicate Count	340	6.54%	Blank Count	223	4.29%	<table border="1"> <caption>Simple Statistics</caption> <thead> <tr> <th>Category</th> <th>Count</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>5197</td></tr> <tr><td>Null Count</td><td>0</td></tr> <tr><td>Distinct Count</td><td>354</td></tr> <tr><td>Unique Count</td><td>14</td></tr> <tr><td>Duplicate Count</td><td>340</td></tr> <tr><td>Blank Count</td><td>223</td></tr> </tbody> </table>	Category	Count	Row Count	5197	Null Count	0	Distinct Count	354	Unique Count	14	Duplicate Count	340	Blank Count	223																			
Label	Count	%																																																						
Row Count	5197	100.00%																																																						
Null Count	0	0.00%																																																						
Distinct Count	354	6.81%																																																						
Unique Count	14	0.27%																																																						
Duplicate Count	340	6.54%																																																						
Blank Count	223	4.29%																																																						
Category	Count																																																							
Row Count	5197																																																							
Null Count	0																																																							
Distinct Count	354																																																							
Unique Count	14																																																							
Duplicate Count	340																																																							
Blank Count	223																																																							
<p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>223</td><td>4.29%</td></tr> <tr><td>5 2017</td><td>47</td><td>0.90%</td></tr> <tr><td>3 2016</td><td>46</td><td>0.89%</td></tr> <tr><td>1 2017</td><td>42</td><td>0.81%</td></tr> <tr><td>3 2017</td><td>41</td><td>0.79%</td></tr> <tr><td>9 2016</td><td>40</td><td>0.77%</td></tr> <tr><td>3 2012</td><td>40</td><td>0.77%</td></tr> <tr><td>12 2017</td><td>39</td><td>0.75%</td></tr> <tr><td>10 2019</td><td>36</td><td>0.69%</td></tr> <tr><td>2 2017</td><td>36</td><td>0.69%</td></tr> </tbody> </table>	Value	Count	%	Empty field	223	4.29%	5 2017	47	0.90%	3 2016	46	0.89%	1 2017	42	0.81%	3 2017	41	0.79%	9 2016	40	0.77%	3 2012	40	0.77%	12 2017	39	0.75%	10 2019	36	0.69%	2 2017	36	0.69%	<table border="1"> <caption>Value Frequency</caption> <thead> <tr> <th>Value</th> <th>Count</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>223</td></tr> <tr><td>5 2017</td><td>47</td></tr> <tr><td>3 2016</td><td>46</td></tr> <tr><td>1 2017</td><td>42</td></tr> <tr><td>3 2017</td><td>41</td></tr> <tr><td>9 2016</td><td>40</td></tr> <tr><td>3 2012</td><td>40</td></tr> <tr><td>12 2017</td><td>39</td></tr> <tr><td>10 2019</td><td>36</td></tr> <tr><td>2 2017</td><td>36</td></tr> </tbody> </table>	Value	Count	Empty field	223	5 2017	47	3 2016	46	1 2017	42	3 2017	41	9 2016	40	3 2012	40	12 2017	39	10 2019	36	2 2017	36
Value	Count	%																																																						
Empty field	223	4.29%																																																						
5 2017	47	0.90%																																																						
3 2016	46	0.89%																																																						
1 2017	42	0.81%																																																						
3 2017	41	0.79%																																																						
9 2016	40	0.77%																																																						
3 2012	40	0.77%																																																						
12 2017	39	0.75%																																																						
10 2019	36	0.69%																																																						
2 2017	36	0.69%																																																						
Value	Count																																																							
Empty field	223																																																							
5 2017	47																																																							
3 2016	46																																																							
1 2017	42																																																							
3 2017	41																																																							
9 2016	40																																																							
3 2012	40																																																							
12 2017	39																																																							
10 2019	36																																																							
2 2017	36																																																							

Atributo	Category_of_Abuse																																																						
Tipo de Dados	String																																																						
Descrição	Termo utilizado para identificar o tipo ou tipos de diferentes abusos.																																																						
Exemplo	Emotional Abuse																																																						
Preenchido	94,38%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Foi detetado a existência de valores em branco, o que no futuro não serão eliminadas, mas sim substituídas pelo valor “Category not available”.																																																						
Estatísticas dos Dados	<p>* Column: metadata.Category_of_Abuse</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>5197</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>16</td><td>0.31%</td></tr> <tr><td>Unique Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Duplicate Count</td><td>16</td><td>0.31%</td></tr> <tr><td>Blank Count</td><td>295</td><td>5.68%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Neglect</td><td>1947</td><td>37.46%</td></tr> <tr><td>Emotional Abuse</td><td>1254</td><td>24.13%</td></tr> <tr><td>Physical Abuse</td><td>908</td><td>17.47%</td></tr> <tr><td>Empty field</td><td>295</td><td>5.68%</td></tr> <tr><td>Sexual Abuse</td><td>263</td><td>5.06%</td></tr> <tr><td>Multiple: Neglect, Emotion...</td><td>206</td><td>3.96%</td></tr> <tr><td>Multiple: Physical Abuse, E...</td><td>156</td><td>3.00%</td></tr> <tr><td>Multiple: Neglect, Physical ...</td><td>58</td><td>1.12%</td></tr> <tr><td>Multiple: Neglect, Physical ...</td><td>44</td><td>0.85%</td></tr> <tr><td>Multiple: Sexual Abuse, Em...</td><td>26</td><td>0.50%</td></tr> </tbody> </table>	Label	Count	%	Row Count	5197	100.00%	Null Count	0	0.00%	Distinct Count	16	0.31%	Unique Count	0	0.00%	Duplicate Count	16	0.31%	Blank Count	295	5.68%	Value	Count	%	Neglect	1947	37.46%	Emotional Abuse	1254	24.13%	Physical Abuse	908	17.47%	Empty field	295	5.68%	Sexual Abuse	263	5.06%	Multiple: Neglect, Emotion...	206	3.96%	Multiple: Physical Abuse, E...	156	3.00%	Multiple: Neglect, Physical ...	58	1.12%	Multiple: Neglect, Physical ...	44	0.85%	Multiple: Sexual Abuse, Em...	26	0.50%
Label	Count	%																																																					
Row Count	5197	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	16	0.31%																																																					
Unique Count	0	0.00%																																																					
Duplicate Count	16	0.31%																																																					
Blank Count	295	5.68%																																																					
Value	Count	%																																																					
Neglect	1947	37.46%																																																					
Emotional Abuse	1254	24.13%																																																					
Physical Abuse	908	17.47%																																																					
Empty field	295	5.68%																																																					
Sexual Abuse	263	5.06%																																																					
Multiple: Neglect, Emotion...	206	3.96%																																																					
Multiple: Physical Abuse, E...	156	3.00%																																																					
Multiple: Neglect, Physical ...	58	1.12%																																																					
Multiple: Neglect, Physical ...	44	0.85%																																																					
Multiple: Sexual Abuse, Em...	26	0.50%																																																					

Atributo	CP_End_Reason
Tipo de Dados	String

Descrição	Termo utilizado para identificar a causa do encerramento da proteção da criança.																																																						
Exemplo	Criteria No Longer Met																																																						
Preenchido	95,18%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Foi detetado a existência de valores em branco, o que no futuro não serão eliminadas, mas sim substituídas pelo valor “Criteria not available”.																																																						
Estatísticas dos Dados	<p>* Column: metadata.CP_End_Reason </p> <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>5197</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>10</td> <td>0.19%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>10</td> <td>0.19%</td> </tr> <tr> <td>Blank Count</td> <td>248</td> <td>4.77%</td> </tr> </tbody> </table> <ul style="list-style-type: none"> - Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Criteria No Longer Met</td> <td>3618</td> <td>69.62%</td> </tr> <tr> <td>Child & Family Permanently... <i>(partially visible)</i></td> <td>608</td> <td>11.70%</td> </tr> <tr> <td>Placed in Care</td> <td>548</td> <td>10.54%</td> </tr> <tr> <td>Empty field</td> <td>248</td> <td>4.77%</td> </tr> <tr> <td>Child With Permanent Subs...</td> <td>82</td> <td>1.58%</td> </tr> <tr> <td>Carer Offering Protection</td> <td>34</td> <td>0.65%</td> </tr> <tr> <td>Abuser No Longer in Hous...</td> <td>22</td> <td>0.42%</td> </tr> <tr> <td>Child Able to Self Protect</td> <td>16</td> <td>0.31%</td> </tr> <tr> <td>Child Reached 18 or Married</td> <td>13</td> <td>0.25%</td> </tr> <tr> <td>Child Died</td> <td>8</td> <td>0.15%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	5197	100.00%	Null Count	0	0.00%	Distinct Count	10	0.19%	Unique Count	0	0.00%	Duplicate Count	10	0.19%	Blank Count	248	4.77%	Value	Count	%	Criteria No Longer Met	3618	69.62%	Child & Family Permanently... <i>(partially visible)</i>	608	11.70%	Placed in Care	548	10.54%	Empty field	248	4.77%	Child With Permanent Subs...	82	1.58%	Carer Offering Protection	34	0.65%	Abuser No Longer in Hous...	22	0.42%	Child Able to Self Protect	16	0.31%	Child Reached 18 or Married	13	0.25%	Child Died	8	0.15%
Label	Count	%																																																					
Row Count	5197	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	10	0.19%																																																					
Unique Count	0	0.00%																																																					
Duplicate Count	10	0.19%																																																					
Blank Count	248	4.77%																																																					
Value	Count	%																																																					
Criteria No Longer Met	3618	69.62%																																																					
Child & Family Permanently... <i>(partially visible)</i>	608	11.70%																																																					
Placed in Care	548	10.54%																																																					
Empty field	248	4.77%																																																					
Child With Permanent Subs...	82	1.58%																																																					
Carer Offering Protection	34	0.65%																																																					
Abuser No Longer in Hous...	22	0.42%																																																					
Child Able to Self Protect	16	0.31%																																																					
Child Reached 18 or Married	13	0.25%																																																					
Child Died	8	0.15%																																																					

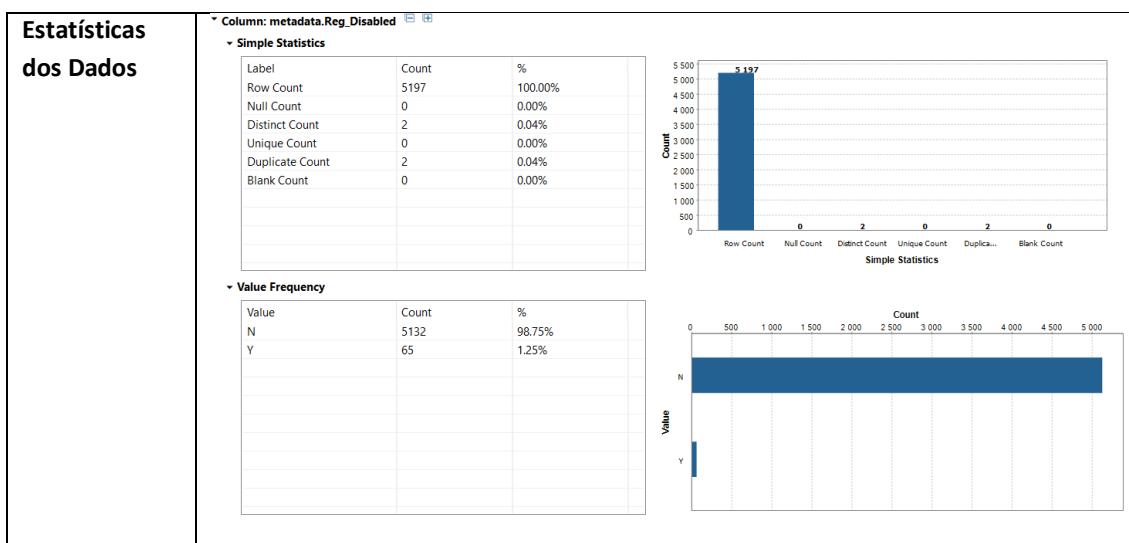
Atributo	Current_CPP
Tipo de Dados	Character
Descrição	Termo utilizado para identificar se um plano ainda está em vigor.
Exemplo	N
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados



Atributo	Gender																																				
Tipo de Dados	String																																				
Descrição	Termo utilizado para identificar o género de um individuo																																				
Exemplo	Male																																				
Preenchido	100%																																				
Intervalo de Dados	Não mensurável																																				
Problemas de Qualidade dos Dados	Não detetados																																				
Estatísticas dos Dados	<p>Column: metadata.Gender</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>5197</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>4</td> <td>0.08%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>4</td> <td>0.08%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Count</p> <p>Row Count Null Count Distinct Count Unique Count Duplicate... Blank Count</p> <p>Simple Statistics</p> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Male</td> <td>2685</td> <td>51.66%</td> </tr> <tr> <td>Female</td> <td>2500</td> <td>48.10%</td> </tr> <tr> <td>Unborn</td> <td>10</td> <td>0.19%</td> </tr> <tr> <td>Indeterminate</td> <td>2</td> <td>0.04%</td> </tr> </tbody> </table> <p>Value</p> <p>Count</p> <p>Male Female Unborn Indeterminate</p> <p>Simple Statistics</p>	Label	Count	%	Row Count	5197	100.00%	Null Count	0	0.00%	Distinct Count	4	0.08%	Unique Count	0	0.00%	Duplicate Count	4	0.08%	Blank Count	0	0.00%	Value	Count	%	Male	2685	51.66%	Female	2500	48.10%	Unborn	10	0.19%	Indeterminate	2	0.04%
Label	Count	%																																			
Row Count	5197	100.00%																																			
Null Count	0	0.00%																																			
Distinct Count	4	0.08%																																			
Unique Count	0	0.00%																																			
Duplicate Count	4	0.08%																																			
Blank Count	0	0.00%																																			
Value	Count	%																																			
Male	2685	51.66%																																			
Female	2500	48.10%																																			
Unborn	10	0.19%																																			
Indeterminate	2	0.04%																																			

Atributo	Disabled																														
Tipo de Dados	Character																														
Descrição	Termo para identificar se o individuo possui necessidades especiais.																														
Exemplo	N																														
Preenchido	100%																														
Intervalo de Dados	Não mensurável																														
Problemas de Qualidade dos Dados	Não detetados																														
Estatísticas dos Dados	<p>* Column: metadata.Disabled</p> <ul style="list-style-type: none"> -> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>5197</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>2</td> <td>0.04%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>2</td> <td>0.04%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <ul style="list-style-type: none"> -> Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>N</td> <td>4906</td> <td>94.40%</td> </tr> <tr> <td>Y</td> <td>291</td> <td>5.60%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	5197	100.00%	Null Count	0	0.00%	Distinct Count	2	0.04%	Unique Count	0	0.00%	Duplicate Count	2	0.04%	Blank Count	0	0.00%	Value	Count	%	N	4906	94.40%	Y	291	5.60%
Label	Count	%																													
Row Count	5197	100.00%																													
Null Count	0	0.00%																													
Distinct Count	2	0.04%																													
Unique Count	0	0.00%																													
Duplicate Count	2	0.04%																													
Blank Count	0	0.00%																													
Value	Count	%																													
N	4906	94.40%																													
Y	291	5.60%																													

Atributo	Reg_Disabled
Tipo de Dados	Character
Descrição	Termo utilizado para identificar se o individuo possui um registo oficial relacionado com necessidades especiais
Exemplo	N
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados



Atributo	Unique																														
Tipo de Dados	Character																														
Descrição	Termo utilizado para identificar se o registo oficial relacionado com necessidades especiais é único																														
Exemplo	N																														
Preenchido	100%																														
Intervalo dos Dados	Não mensurável																														
Problemas de Qualidade dos Dados	Não detetados																														
Estatísticas dos Dados	<p>Column: metadata.Unique</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>5197</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>2</td> <td>0.04%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>2</td> <td>0.04%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Simple Statistics</p> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Y</td> <td>4575</td> <td>88.03%</td> </tr> <tr> <td>N</td> <td>622</td> <td>11.97%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	5197	100.00%	Null Count	0	0.00%	Distinct Count	2	0.04%	Unique Count	0	0.00%	Duplicate Count	2	0.04%	Blank Count	0	0.00%	Value	Count	%	Y	4575	88.03%	N	622	11.97%
Label	Count	%																													
Row Count	5197	100.00%																													
Null Count	0	0.00%																													
Distinct Count	2	0.04%																													
Unique Count	0	0.00%																													
Duplicate Count	2	0.04%																													
Blank Count	0	0.00%																													
Value	Count	%																													
Y	4575	88.03%																													
N	622	11.97%																													

Atributo	Current_Age_Group																																							
Tipo de Dados	String																																							
Descrição	Termo utilizado para identificar o grupo de idade a que o individuo pertence.																																							
Exemplo	16+																																							
Preenchido	100%																																							
Intervalo de Dados	Não mensurável																																							
Problemas de Qualidade dos Dados	Não detetados																																							
Estatísticas dos Dados	<p>Column: metadata.Current_Age_Group</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>5197</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>5</td> <td>0.10%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>5</td> <td>0.10%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>16+</td> <td>2847</td> <td>54.78%</td> </tr> <tr> <td>11 to 15</td> <td>1129</td> <td>21.72%</td> </tr> <tr> <td>6 to 10</td> <td>796</td> <td>15.32%</td> </tr> <tr> <td>1 to 5</td> <td>384</td> <td>7.39%</td> </tr> <tr> <td>U1</td> <td>41</td> <td>0.79%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	5197	100.00%	Null Count	0	0.00%	Distinct Count	5	0.10%	Unique Count	0	0.00%	Duplicate Count	5	0.10%	Blank Count	0	0.00%	Value	Count	%	16+	2847	54.78%	11 to 15	1129	21.72%	6 to 10	796	15.32%	1 to 5	384	7.39%	U1	41	0.79%
Label	Count	%																																						
Row Count	5197	100.00%																																						
Null Count	0	0.00%																																						
Distinct Count	5	0.10%																																						
Unique Count	0	0.00%																																						
Duplicate Count	5	0.10%																																						
Blank Count	0	0.00%																																						
Value	Count	%																																						
16+	2847	54.78%																																						
11 to 15	1129	21.72%																																						
6 to 10	796	15.32%																																						
1 to 5	384	7.39%																																						
U1	41	0.79%																																						

Atributo	Ethnic_Code
Tipo de Dados	Character
Descrição	Termo utilizado para identificar o código respetivo da etnia do individuo.
Exemplo	A
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados

Estatísticas dos Dados	<p>Column: metadata.Ethnic_Code</p> <ul style="list-style-type: none"> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th><th>Count</th><th>%</th></tr> </thead> <tbody> <tr><td>Row Count</td><td>5197</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>5</td><td>0.10%</td></tr> <tr><td>Unique Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Duplicate Count</td><td>5</td><td>0.10%</td></tr> <tr><td>Blank Count</td><td>0</td><td>0.00%</td></tr> </tbody> </table>	Label	Count	%	Row Count	5197	100.00%	Null Count	0	0.00%	Distinct Count	5	0.10%	Unique Count	0	0.00%	Duplicate Count	5	0.10%	Blank Count	0	0.00%	
Label	Count	%																					
Row Count	5197	100.00%																					
Null Count	0	0.00%																					
Distinct Count	5	0.10%																					
Unique Count	0	0.00%																					
Duplicate Count	5	0.10%																					
Blank Count	0	0.00%																					
<p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>A</td><td>2696</td><td>51.88%</td></tr> <tr><td>B</td><td>925</td><td>17.80%</td></tr> <tr><td>D</td><td>886</td><td>17.05%</td></tr> <tr><td>E</td><td>426</td><td>8.20%</td></tr> <tr><td>C</td><td>264</td><td>5.08%</td></tr> </tbody> </table>	Value	Count	%	A	2696	51.88%	B	925	17.80%	D	886	17.05%	E	426	8.20%	C	264	5.08%					
Value	Count	%																					
A	2696	51.88%																					
B	925	17.80%																					
D	886	17.05%																					
E	426	8.20%																					
C	264	5.08%																					

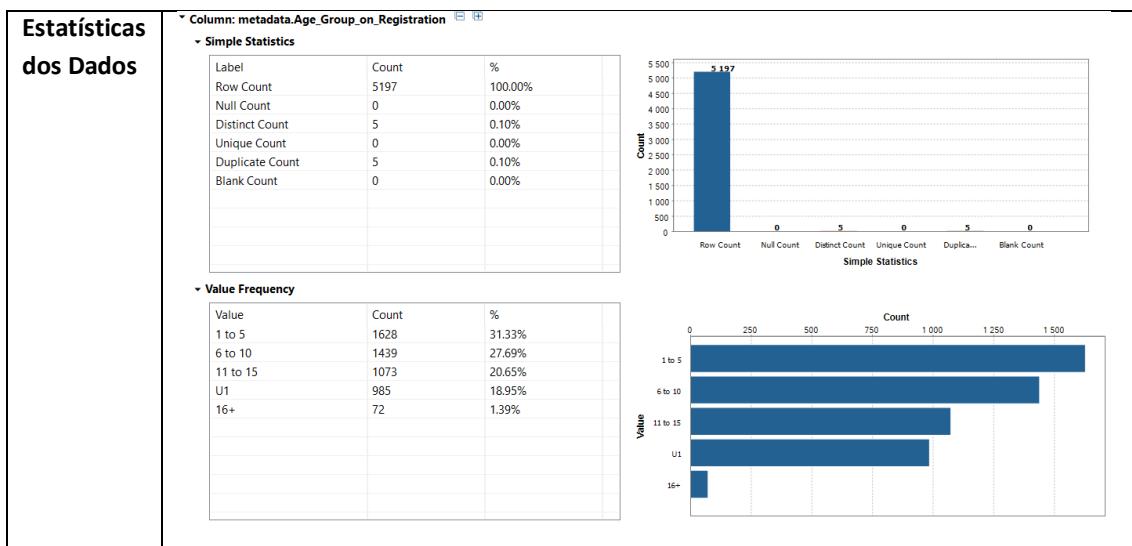
Atributo	Start_Month_Year																																		
Tipo de Dados	String																																		
Descrição	Coluna utilizada para identificar o início do mês e o ano que o indivíduo começou o plano de proteção.																																		
Exemplo	2 2016																																		
Preenchido	100%																																		
Intervalo de Dados	Não mensurável																																		
Problemas de Qualidade dos Dados	O formato que esta coluna apresenta não é a mais indicada para o tipo de dados que esta contém.																																		
Estatísticas dos Dados	<p>Column: metadata.Start_Month_Year</p> <ul style="list-style-type: none"> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>5197</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>383</td><td>7.37%</td></tr> <tr><td>Unique Count</td><td>20</td><td>0.38%</td></tr> <tr><td>Duplicate Count</td><td>363</td><td>6.98%</td></tr> <tr><td>Blank Count</td><td>0</td><td>0.00%</td></tr> </tbody> </table>	Label	Count	%	Row Count	5197	100.00%	Null Count	0	0.00%	Distinct Count	383	7.37%	Unique Count	20	0.38%	Duplicate Count	363	6.98%	Blank Count	0	0.00%													
Label	Count	%																																	
Row Count	5197	100.00%																																	
Null Count	0	0.00%																																	
Distinct Count	383	7.37%																																	
Unique Count	20	0.38%																																	
Duplicate Count	363	6.98%																																	
Blank Count	0	0.00%																																	
	<p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>2 2016</td><td>41</td><td>0.79%</td></tr> <tr><td>3 2018</td><td>40</td><td>0.77%</td></tr> <tr><td>2 2019</td><td>39</td><td>0.75%</td></tr> <tr><td>6 2014</td><td>39</td><td>0.75%</td></tr> <tr><td>6 2019</td><td>37</td><td>0.71%</td></tr> <tr><td>2 2014</td><td>37</td><td>0.71%</td></tr> <tr><td>12 2021</td><td>36</td><td>0.69%</td></tr> <tr><td>9 2017</td><td>35</td><td>0.67%</td></tr> <tr><td>4 2016</td><td>35</td><td>0.67%</td></tr> <tr><td>11 2015</td><td>35</td><td>0.67%</td></tr> </tbody> </table>	Value	Count	%	2 2016	41	0.79%	3 2018	40	0.77%	2 2019	39	0.75%	6 2014	39	0.75%	6 2019	37	0.71%	2 2014	37	0.71%	12 2021	36	0.69%	9 2017	35	0.67%	4 2016	35	0.67%	11 2015	35	0.67%	
Value	Count	%																																	
2 2016	41	0.79%																																	
3 2018	40	0.77%																																	
2 2019	39	0.75%																																	
6 2014	39	0.75%																																	
6 2019	37	0.71%																																	
2 2014	37	0.71%																																	
12 2021	36	0.69%																																	
9 2017	35	0.67%																																	
4 2016	35	0.67%																																	
11 2015	35	0.67%																																	

Atributo	End_Month_Year																																																						
Tipo de Dados	String																																																						
Descrição	Coluna utilizada para identificar o fim do mês e o ano que o indivíduo começou o plano de proteção.																																																						
Exemplo	N																																																						
Preenchido	100%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Se mudarmos a maneira como vamos colocar as datas no valor N, este terá que mudar para algo com o formato de date. Será feito a mesma coisa que em cima no End Careplan																																																						
Estatísticas dos Dados	<p>Column: metadata.End_Month_Year</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>5197</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>345</td> <td>6.64%</td> </tr> <tr> <td>Unique Count</td> <td>14</td> <td>0.27%</td> </tr> <tr> <td>Duplicate Count</td> <td>331</td> <td>6.37%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>N</td> <td>844</td> <td>16.24%</td> </tr> <tr> <td>5 2017</td> <td>47</td> <td>0.90%</td> </tr> <tr> <td>3 2016</td> <td>40</td> <td>0.77%</td> </tr> <tr> <td>3 2017</td> <td>38</td> <td>0.73%</td> </tr> <tr> <td>1 2017</td> <td>38</td> <td>0.73%</td> </tr> <tr> <td>12 2017</td> <td>38</td> <td>0.73%</td> </tr> <tr> <td>3 2012</td> <td>37</td> <td>0.71%</td> </tr> <tr> <td>9 2016</td> <td>36</td> <td>0.69%</td> </tr> <tr> <td>10 2019</td> <td>35</td> <td>0.67%</td> </tr> <tr> <td>2 2017</td> <td>35</td> <td>0.67%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	5197	100.00%	Null Count	0	0.00%	Distinct Count	345	6.64%	Unique Count	14	0.27%	Duplicate Count	331	6.37%	Blank Count	0	0.00%	Value	Count	%	N	844	16.24%	5 2017	47	0.90%	3 2016	40	0.77%	3 2017	38	0.73%	1 2017	38	0.73%	12 2017	38	0.73%	3 2012	37	0.71%	9 2016	36	0.69%	10 2019	35	0.67%	2 2017	35	0.67%
Label	Count	%																																																					
Row Count	5197	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	345	6.64%																																																					
Unique Count	14	0.27%																																																					
Duplicate Count	331	6.37%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
N	844	16.24%																																																					
5 2017	47	0.90%																																																					
3 2016	40	0.77%																																																					
3 2017	38	0.73%																																																					
1 2017	38	0.73%																																																					
12 2017	38	0.73%																																																					
3 2012	37	0.71%																																																					
9 2016	36	0.69%																																																					
10 2019	35	0.67%																																																					
2 2017	35	0.67%																																																					

Atributo	Current_Age_Bracket
Tipo de Dados	String
Descrição	Termo utilizado para identificar a faixa etária em que o indivíduo pertence atualmente. (Nota: este csv foi atualizado em 7 de Abril de 2022).
Exemplo	16+
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de	Não detetados

Qualidade dos Dados																																								
Estatísticas dos Dados	<p>Column: metadata.Current_Age_Bracket</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>5197</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>5</td> <td>0.10%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>5</td> <td>0.10%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>16+</td> <td>2848</td> <td>54.80%</td> </tr> <tr> <td>11 to 15</td> <td>1129</td> <td>21.72%</td> </tr> <tr> <td>6 to 10</td> <td>796</td> <td>15.32%</td> </tr> <tr> <td>1 to 5</td> <td>384</td> <td>7.39%</td> </tr> <tr> <td>U1</td> <td>40</td> <td>0.77%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	5197	100.00%	Null Count	0	0.00%	Distinct Count	5	0.10%	Unique Count	0	0.00%	Duplicate Count	5	0.10%	Blank Count	0	0.00%	Value	Count	%	16+	2848	54.80%	11 to 15	1129	21.72%	6 to 10	796	15.32%	1 to 5	384	7.39%	U1	40	0.77%
Label	Count	%																																						
Row Count	5197	100.00%																																						
Null Count	0	0.00%																																						
Distinct Count	5	0.10%																																						
Unique Count	0	0.00%																																						
Duplicate Count	5	0.10%																																						
Blank Count	0	0.00%																																						
Value	Count	%																																						
16+	2848	54.80%																																						
11 to 15	1129	21.72%																																						
6 to 10	796	15.32%																																						
1 to 5	384	7.39%																																						
U1	40	0.77%																																						

Atributo	Age_group_on_Registration
Tipo de Dados	String
Descrição	Colunas utilizada para identificar o grupo etário a que o individuo pertencia na data do registo inicial.
Exemplo	1 to 5
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados



Atributo	Age_group_on_deRegistration																																							
Tipo de Dados	String																																							
Descrição	Colunas utilizada para identificar o grupo etário a que o individuo pertencia na data do registo inicial.																																							
Exemplo	1 to 5																																							
Preenchido	100%																																							
Intervalo de Dados	Não mensurável																																							
Problemas de Qualidade dos Dados	Não detetados																																							
Estatísticas dos Dados	<p>Column: metadata.Age_Group_on_deRegistration</p> <ul style="list-style-type: none"> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>5197</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>5</td> <td>0.10%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>5</td> <td>0.10%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Simple Statistics</p> <ul style="list-style-type: none"> Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>1 to 5</td> <td>1606</td> <td>30.90%</td> </tr> <tr> <td>6 to 10</td> <td>1430</td> <td>27.52%</td> </tr> <tr> <td>11 to 15</td> <td>1145</td> <td>22.03%</td> </tr> <tr> <td>U1</td> <td>583</td> <td>11.22%</td> </tr> <tr> <td>16+</td> <td>433</td> <td>8.33%</td> </tr> </tbody> </table> <p>Value Frequency</p>	Label	Count	%	Row Count	5197	100.00%	Null Count	0	0.00%	Distinct Count	5	0.10%	Unique Count	0	0.00%	Duplicate Count	5	0.10%	Blank Count	0	0.00%	Value	Count	%	1 to 5	1606	30.90%	6 to 10	1430	27.52%	11 to 15	1145	22.03%	U1	583	11.22%	16+	433	8.33%
Label	Count	%																																						
Row Count	5197	100.00%																																						
Null Count	0	0.00%																																						
Distinct Count	5	0.10%																																						
Unique Count	0	0.00%																																						
Duplicate Count	5	0.10%																																						
Blank Count	0	0.00%																																						
Value	Count	%																																						
1 to 5	1606	30.90%																																						
6 to 10	1430	27.52%																																						
11 to 15	1145	22.03%																																						
U1	583	11.22%																																						
16+	433	8.33%																																						

Dataset: Gender_Stats_Dataset

Aluno responsável: Catarina Fernandes (a101441)

Fonte: <https://genderdata.worldbank.org/>

Gender_Stats_FootNote.csv

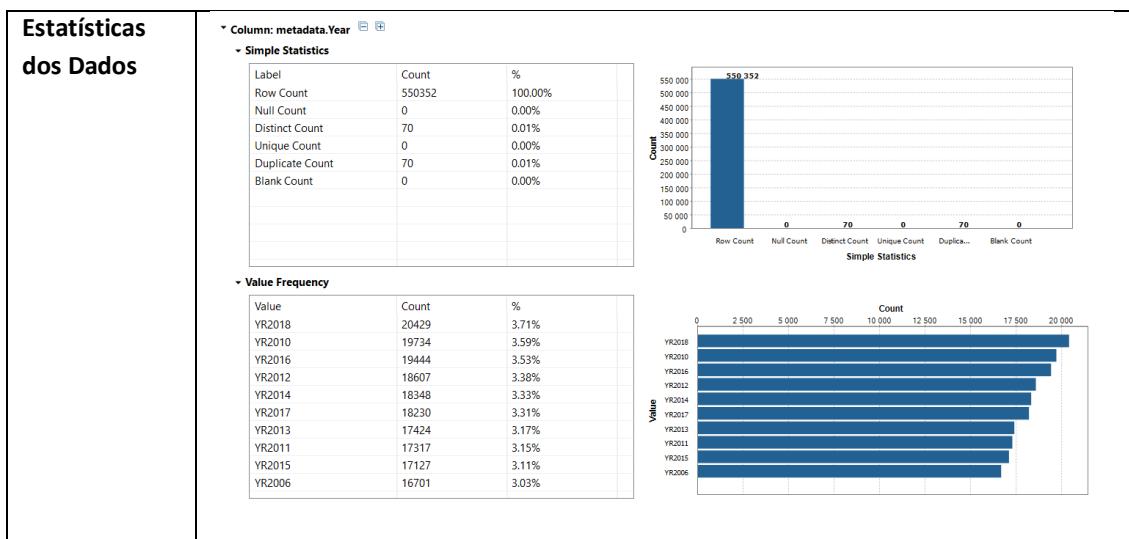
Descrição: Contém notas associadas a determinados dados, fornecendo contexto ou explicação de anomalias ou métodos de recolha de dados.

Atributo	CountryCode																																																						
Tipo de Dados	String																																																						
Descrição	Coluna utilizada para identificar o código de um país.																																																						
Exemplo	PER																																																						
Preenchido	100%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Não detetados																																																						
Estatísticas dos Dados	<p>Column: metadata_CountryCode</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>550352</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>291</td> <td>0.05%</td> </tr> <tr> <td>Unique Count</td> <td>1</td> <td>1.817E-4%</td> </tr> <tr> <td>Duplicate Count</td> <td>290</td> <td>0.05%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>PER</td> <td>4415</td> <td>0.80%</td> </tr> <tr> <td>IDN</td> <td>4342</td> <td>0.79%</td> </tr> <tr> <td>SEN</td> <td>4173</td> <td>0.76%</td> </tr> <tr> <td>NOR</td> <td>4097</td> <td>0.74%</td> </tr> <tr> <td>CYP</td> <td>4054</td> <td>0.74%</td> </tr> <tr> <td>DOM</td> <td>3980</td> <td>0.72%</td> </tr> <tr> <td>TZA</td> <td>3880</td> <td>0.71%</td> </tr> <tr> <td>USA</td> <td>3800</td> <td>0.69%</td> </tr> <tr> <td>ESP</td> <td>3783</td> <td>0.69%</td> </tr> <tr> <td>FRA</td> <td>3685</td> <td>0.67%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	550352	100.00%	Null Count	0	0.00%	Distinct Count	291	0.05%	Unique Count	1	1.817E-4%	Duplicate Count	290	0.05%	Blank Count	0	0.00%	Value	Count	%	PER	4415	0.80%	IDN	4342	0.79%	SEN	4173	0.76%	NOR	4097	0.74%	CYP	4054	0.74%	DOM	3980	0.72%	TZA	3880	0.71%	USA	3800	0.69%	ESP	3783	0.69%	FRA	3685	0.67%
Label	Count	%																																																					
Row Count	550352	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	291	0.05%																																																					
Unique Count	1	1.817E-4%																																																					
Duplicate Count	290	0.05%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
PER	4415	0.80%																																																					
IDN	4342	0.79%																																																					
SEN	4173	0.76%																																																					
NOR	4097	0.74%																																																					
CYP	4054	0.74%																																																					
DOM	3980	0.72%																																																					
TZA	3880	0.71%																																																					
USA	3800	0.69%																																																					
ESP	3783	0.69%																																																					
FRA	3685	0.67%																																																					

Atributo	SeriesCode
Tipo de Dados	String
Descrição	Código que identifica uma série de dados específicos
Exemplo	SH.DYN.MORT
Preenchido	100%

Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Não detetados																																																						
Estatísticas dos Dados	<p>Column: metadata.SeriesCode</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>550352</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>568</td> <td>0.10%</td> </tr> <tr> <td>Unique Count</td> <td>7</td> <td>1.272E-3%</td> </tr> <tr> <td>Duplicate Count</td> <td>561</td> <td>0.10%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>SH.DYN.MORT</td> <td>11737</td> <td>2.13%</td> </tr> <tr> <td>SH.DYN.MORT.MA</td> <td>11668</td> <td>2.12%</td> </tr> <tr> <td>SP.DYN.IMRT.MAIN</td> <td>11668</td> <td>2.12%</td> </tr> <tr> <td>SH.DYN.MORT.FE</td> <td>11668</td> <td>2.12%</td> </tr> <tr> <td>SP.DYN.IMRT.IN</td> <td>11668</td> <td>2.12%</td> </tr> <tr> <td>SP.DYN.IMRT.FEIN</td> <td>11668</td> <td>2.12%</td> </tr> <tr> <td>SH.DTH.IMRT.MA</td> <td>11586</td> <td>2.11%</td> </tr> <tr> <td>SH.DTH.IMRT.FE</td> <td>11586</td> <td>2.11%</td> </tr> <tr> <td>SH.DTH.IMRT</td> <td>11586</td> <td>2.11%</td> </tr> <tr> <td>SH.DTH.MORT.FE</td> <td>11217</td> <td>2.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	550352	100.00%	Null Count	0	0.00%	Distinct Count	568	0.10%	Unique Count	7	1.272E-3%	Duplicate Count	561	0.10%	Blank Count	0	0.00%	Value	Count	%	SH.DYN.MORT	11737	2.13%	SH.DYN.MORT.MA	11668	2.12%	SP.DYN.IMRT.MAIN	11668	2.12%	SH.DYN.MORT.FE	11668	2.12%	SP.DYN.IMRT.IN	11668	2.12%	SP.DYN.IMRT.FEIN	11668	2.12%	SH.DTH.IMRT.MA	11586	2.11%	SH.DTH.IMRT.FE	11586	2.11%	SH.DTH.IMRT	11586	2.11%	SH.DTH.MORT.FE	11217	2.04%
Label	Count	%																																																					
Row Count	550352	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	568	0.10%																																																					
Unique Count	7	1.272E-3%																																																					
Duplicate Count	561	0.10%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
SH.DYN.MORT	11737	2.13%																																																					
SH.DYN.MORT.MA	11668	2.12%																																																					
SP.DYN.IMRT.MAIN	11668	2.12%																																																					
SH.DYN.MORT.FE	11668	2.12%																																																					
SP.DYN.IMRT.IN	11668	2.12%																																																					
SP.DYN.IMRT.FEIN	11668	2.12%																																																					
SH.DTH.IMRT.MA	11586	2.11%																																																					
SH.DTH.IMRT.FE	11586	2.11%																																																					
SH.DTH.IMRT	11586	2.11%																																																					
SH.DTH.MORT.FE	11217	2.04%																																																					

Atributo	Year
Tipo de Dados	String
Descrição	Coluna utilizada para identificar o ano.
Exemplo	YR2018
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	O formato não é o mais adequado pelo que seria melhor converter a Integer



Atributo	DESCRIPTION																																																						
Tipo de Dados	String																																																						
Descrição	Coluna com a descrição da informação																																																						
Exemplo	UIS estimate																																																						
Preenchido	100%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Não detetados																																																						
Estatísticas dos Dados	<p>Column: metadata.DESCRIPTION</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>550352</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>178300</td> <td>32.40%</td> </tr> <tr> <td>Unique Count</td> <td>143210</td> <td>26.02%</td> </tr> <tr> <td>Duplicate Count</td> <td>35090</td> <td>6.38%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Count</p> <p>Row Count Null Count Distinct Count Unique Count Duplicate... Blank Count</p> <p>Simple Statistics</p> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>UIS Estimation</td> <td>40380</td> <td>7.34%</td> </tr> <tr> <td>UIS estimate</td> <td>23506</td> <td>4.27%</td> </tr> <tr> <td>LFS - Labour Force Survey, ...</td> <td>21881</td> <td>3.98%</td> </tr> <tr> <td>Adjusted</td> <td>20947</td> <td>3.81%</td> </tr> <tr> <td>LFS - EU Labour Force Sur... 16055</td> <td>16055</td> <td>2.92%</td> </tr> <tr> <td><0.1. Plausible bound is <0... 4104</td> <td>4104</td> <td>0.75%</td> </tr> <tr> <td>LFS - EU Labour Force Sur... 3837</td> <td>3837</td> <td>0.70%</td> </tr> <tr> <td>LFS - Labour Force Survey, 2914</td> <td>2914</td> <td>0.53%</td> </tr> <tr> <td>LFS - Labour Force Survey, 2892</td> <td>2892</td> <td>0.53%</td> </tr> <tr> <td>National estimate</td> <td>2609</td> <td>0.47%</td> </tr> </tbody> </table> <p>Count</p> <p>Value UIS Estimation UIS estimate LFS - Labour Force Survey, Repository... Adjusted LFS - EU Labour Force Sur... <0.1. Plausible bound is <0... 16055 LFS - EU Labour Force Sur... 3837 LFS - Labour Force Survey, 2914 LFS - Labour Force Survey, 2892 National estimate</p>	Label	Count	%	Row Count	550352	100.00%	Null Count	0	0.00%	Distinct Count	178300	32.40%	Unique Count	143210	26.02%	Duplicate Count	35090	6.38%	Blank Count	0	0.00%	Value	Count	%	UIS Estimation	40380	7.34%	UIS estimate	23506	4.27%	LFS - Labour Force Survey, ...	21881	3.98%	Adjusted	20947	3.81%	LFS - EU Labour Force Sur... 16055	16055	2.92%	<0.1. Plausible bound is <0... 4104	4104	0.75%	LFS - EU Labour Force Sur... 3837	3837	0.70%	LFS - Labour Force Survey, 2914	2914	0.53%	LFS - Labour Force Survey, 2892	2892	0.53%	National estimate	2609	0.47%
Label	Count	%																																																					
Row Count	550352	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	178300	32.40%																																																					
Unique Count	143210	26.02%																																																					
Duplicate Count	35090	6.38%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
UIS Estimation	40380	7.34%																																																					
UIS estimate	23506	4.27%																																																					
LFS - Labour Force Survey, ...	21881	3.98%																																																					
Adjusted	20947	3.81%																																																					
LFS - EU Labour Force Sur... 16055	16055	2.92%																																																					
<0.1. Plausible bound is <0... 4104	4104	0.75%																																																					
LFS - EU Labour Force Sur... 3837	3837	0.70%																																																					
LFS - Labour Force Survey, 2914	2914	0.53%																																																					
LFS - Labour Force Survey, 2892	2892	0.53%																																																					
National estimate	2609	0.47%																																																					

Gender_Stats Series-Time.csv

Descrição: Contém uma série de referências associadas a determinados dados, fornecendo o período de tempo em que estes foram recolhidos.

Atributo	SeriesCode																																																						
Tipo de Dados	String																																																						
Descrição	Código que identifica uma série de dados específicos																																																						
Exemplo	SP.DYN.LE60.MA.IN																																																						
Preenchido	100%																																																						
Intervalo dos Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Não detetados																																																						
Estatísticas dos Dados	<p>Column: metadata.SeriesCode</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>70</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>14</td> <td>20.00%</td> </tr> <tr> <td>Unique Count</td> <td>1</td> <td>1.43%</td> </tr> <tr> <td>Duplicate Count</td> <td>13</td> <td>18.57%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>SP.DYN.LE60.MA.IN</td> <td>13</td> <td>18.57%</td> </tr> <tr> <td>SP.DYN.LE60.FE.IN</td> <td>13</td> <td>18.57%</td> </tr> <tr> <td>IC.FRM.FEMO.ZS</td> <td>6</td> <td>8.57%</td> </tr> <tr> <td>SH.STA.ANV4.ZS</td> <td>4</td> <td>5.71%</td> </tr> <tr> <td>SH.STA.ANVC.ZS</td> <td>4</td> <td>5.71%</td> </tr> <tr> <td>SH.STA.BRTC.ZS</td> <td>4</td> <td>5.71%</td> </tr> <tr> <td>SP.DYN.CONM.ZS</td> <td>4</td> <td>5.71%</td> </tr> <tr> <td>SP.DYN.WFRT</td> <td>4</td> <td>5.71%</td> </tr> <tr> <td>SP.DYN.CONU.ZS</td> <td>4</td> <td>5.71%</td> </tr> <tr> <td>SH.FPLSATM.ZS</td> <td>4</td> <td>5.71%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	70	100.00%	Null Count	0	0.00%	Distinct Count	14	20.00%	Unique Count	1	1.43%	Duplicate Count	13	18.57%	Blank Count	0	0.00%	Value	Count	%	SP.DYN.LE60.MA.IN	13	18.57%	SP.DYN.LE60.FE.IN	13	18.57%	IC.FRM.FEMO.ZS	6	8.57%	SH.STA.ANV4.ZS	4	5.71%	SH.STA.ANVC.ZS	4	5.71%	SH.STA.BRTC.ZS	4	5.71%	SP.DYN.CONM.ZS	4	5.71%	SP.DYN.WFRT	4	5.71%	SP.DYN.CONU.ZS	4	5.71%	SH.FPLSATM.ZS	4	5.71%
Label	Count	%																																																					
Row Count	70	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	14	20.00%																																																					
Unique Count	1	1.43%																																																					
Duplicate Count	13	18.57%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
SP.DYN.LE60.MA.IN	13	18.57%																																																					
SP.DYN.LE60.FE.IN	13	18.57%																																																					
IC.FRM.FEMO.ZS	6	8.57%																																																					
SH.STA.ANV4.ZS	4	5.71%																																																					
SH.STA.ANVC.ZS	4	5.71%																																																					
SH.STA.BRTC.ZS	4	5.71%																																																					
SP.DYN.CONM.ZS	4	5.71%																																																					
SP.DYN.WFRT	4	5.71%																																																					
SP.DYN.CONU.ZS	4	5.71%																																																					
SH.FPLSATM.ZS	4	5.71%																																																					

Atributo	Year
Tipo de Dados	String
Descrição	Coluna utilizada para identificar o ano.
Exemplo	YR2010
Preenchido	100%
Intervalo dos Dados	Não mensurável
Problemas de Qualidade dos Dados	O formato não é o mais adequado pelo que seria melhor converter a Integer

Estatísticas dos Dados	<p>Column: metadata.Year</p> <p>Simple Statistics</p> <table border="1" style="width: 100%; border-collapse: collapse; font-size: 0.8em;"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>70</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>23</td><td>32.86%</td></tr> <tr><td>Unique Count</td><td>5</td><td>7.14%</td></tr> <tr><td>Duplicate Count</td><td>18</td><td>25.71%</td></tr> <tr><td>Blank Count</td><td>0</td><td>0.00%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1" style="width: 100%; border-collapse: collapse; font-size: 0.8em;"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>YR2010</td><td>10</td><td>14.29%</td></tr> <tr><td>YR2000</td><td>9</td><td>12.86%</td></tr> <tr><td>YR1990</td><td>8</td><td>11.43%</td></tr> <tr><td>YR2019</td><td>7</td><td>10.00%</td></tr> <tr><td>YR2002</td><td>3</td><td>4.29%</td></tr> <tr><td>YR2007</td><td>3</td><td>4.29%</td></tr> <tr><td>YR2018</td><td>3</td><td>4.29%</td></tr> <tr><td>YR1962</td><td>2</td><td>2.86%</td></tr> <tr><td>YR1972</td><td>2</td><td>2.86%</td></tr> <tr><td>YR1967</td><td>2</td><td>2.86%</td></tr> </tbody> </table>	Label	Count	%	Row Count	70	100.00%	Null Count	0	0.00%	Distinct Count	23	32.86%	Unique Count	5	7.14%	Duplicate Count	18	25.71%	Blank Count	0	0.00%	Value	Count	%	YR2010	10	14.29%	YR2000	9	12.86%	YR1990	8	11.43%	YR2019	7	10.00%	YR2002	3	4.29%	YR2007	3	4.29%	YR2018	3	4.29%	YR1962	2	2.86%	YR1972	2	2.86%	YR1967	2	2.86%
Label	Count	%																																																					
Row Count	70	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	23	32.86%																																																					
Unique Count	5	7.14%																																																					
Duplicate Count	18	25.71%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
YR2010	10	14.29%																																																					
YR2000	9	12.86%																																																					
YR1990	8	11.43%																																																					
YR2019	7	10.00%																																																					
YR2002	3	4.29%																																																					
YR2007	3	4.29%																																																					
YR2018	3	4.29%																																																					
YR1962	2	2.86%																																																					
YR1972	2	2.86%																																																					
YR1967	2	2.86%																																																					

Atributo	DESCRIPTION																																																						
Tipo de Dados	String																																																						
Descrição	Coluna com a descrição da informação.																																																						
Exemplo	The data refer to five-year periods running from 1 July, 1960 to 30 June, 1965.																																																						
Preenchido	100%																																																						
Intervalo dos Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Não detetados																																																						
Estatísticas dos Dados	<p>Column: metadata.DESCRIPTION</p> <p>Simple Statistics</p> <table border="1" style="width: 100%; border-collapse: collapse; font-size: 0.8em;"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>70</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>20</td><td>28.57%</td></tr> <tr><td>Unique Count</td><td>1</td><td>1.43%</td></tr> <tr><td>Duplicate Count</td><td>19</td><td>27.14%</td></tr> <tr><td>Blank Count</td><td>0</td><td>0.00%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1" style="width: 100%; border-collapse: collapse; font-size: 0.8em;"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Averages for regions/groups are bas...</td><td>10</td><td>14.29%</td></tr> <tr><td>Averages for regions/groups are bas...</td><td>9</td><td>12.86%</td></tr> <tr><td>Averages for regions/groups are bas...</td><td>8</td><td>11.43%</td></tr> <tr><td>Averages for regions/groups are bas...</td><td>8</td><td>11.43%</td></tr> <tr><td>The sample was drawn fro...</td><td>6</td><td>8.57%</td></tr> <tr><td>The data refer to five-year ...</td><td>2</td><td>2.86%</td></tr> <tr><td>Averages for regions/groups are bas...</td><td>2</td><td>2.86%</td></tr> </tbody> </table>	Label	Count	%	Row Count	70	100.00%	Null Count	0	0.00%	Distinct Count	20	28.57%	Unique Count	1	1.43%	Duplicate Count	19	27.14%	Blank Count	0	0.00%	Value	Count	%	Averages for regions/groups are bas...	10	14.29%	Averages for regions/groups are bas...	9	12.86%	Averages for regions/groups are bas...	8	11.43%	Averages for regions/groups are bas...	8	11.43%	The sample was drawn fro...	6	8.57%	The data refer to five-year ...	2	2.86%	The data refer to five-year ...	2	2.86%	The data refer to five-year ...	2	2.86%	The data refer to five-year ...	2	2.86%	Averages for regions/groups are bas...	2	2.86%
Label	Count	%																																																					
Row Count	70	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	20	28.57%																																																					
Unique Count	1	1.43%																																																					
Duplicate Count	19	27.14%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
Averages for regions/groups are bas...	10	14.29%																																																					
Averages for regions/groups are bas...	9	12.86%																																																					
Averages for regions/groups are bas...	8	11.43%																																																					
Averages for regions/groups are bas...	8	11.43%																																																					
The sample was drawn fro...	6	8.57%																																																					
The data refer to five-year ...	2	2.86%																																																					
The data refer to five-year ...	2	2.86%																																																					
The data refer to five-year ...	2	2.86%																																																					
The data refer to five-year ...	2	2.86%																																																					
Averages for regions/groups are bas...	2	2.86%																																																					

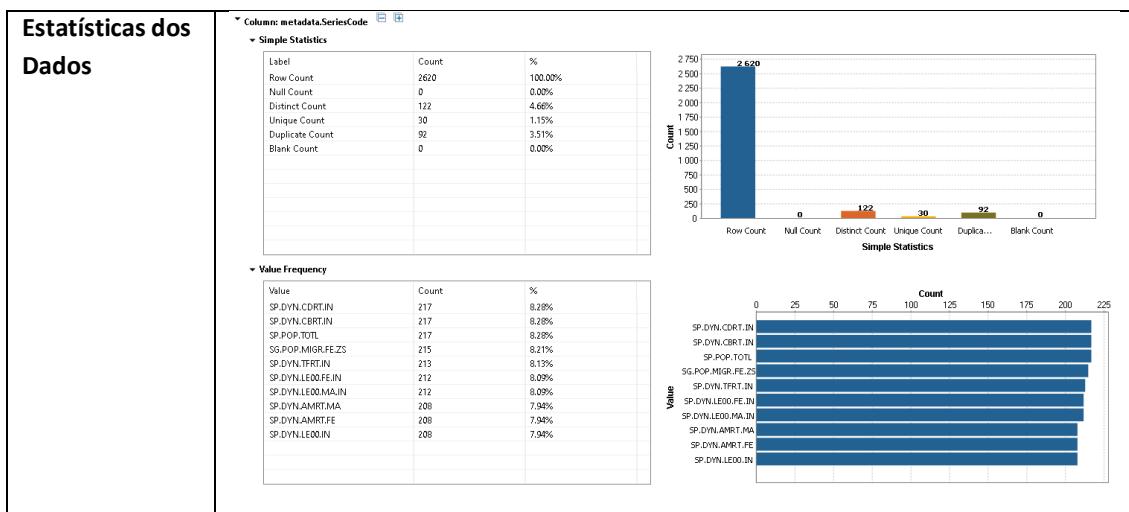
Aluno responsável: Inês Maia (a93347)

Gender_Stats Country-Series.csv

Descrição: O seguinte csv disponibiliza quais são as séries de dados disponíveis para cada país representado em código.

Atributo	_CountryCode																																																						
Tipo de Dados	String																																																						
Descrição	Código único de cada país																																																						
Exemplo	ABW																																																						
Preenchido	100%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Não detetados																																																						
Estatísticas dos Dados	<p>Column: metadata.CountryCode</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>2620</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>District Count</td> <td>218</td> <td>8.32%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>218</td> <td>8.32%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>CYP</td> <td>79</td> <td>3.02%</td> </tr> <tr> <td>FRA</td> <td>55</td> <td>2.10%</td> </tr> <tr> <td>GEO</td> <td>44</td> <td>1.68%</td> </tr> <tr> <td>MDA</td> <td>43</td> <td>1.64%</td> </tr> <tr> <td>MUS</td> <td>39</td> <td>1.49%</td> </tr> <tr> <td>TZA</td> <td>34</td> <td>1.30%</td> </tr> <tr> <td>AUS</td> <td>31</td> <td>1.18%</td> </tr> <tr> <td>CHI</td> <td>30</td> <td>1.15%</td> </tr> <tr> <td>CHN</td> <td>30</td> <td>1.15%</td> </tr> <tr> <td>IND</td> <td>30</td> <td>1.15%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	2620	100.00%	Null Count	0	0.00%	District Count	218	8.32%	Unique Count	0	0.00%	Duplicate Count	218	8.32%	Blank Count	0	0.00%	Value	Count	%	CYP	79	3.02%	FRA	55	2.10%	GEO	44	1.68%	MDA	43	1.64%	MUS	39	1.49%	TZA	34	1.30%	AUS	31	1.18%	CHI	30	1.15%	CHN	30	1.15%	IND	30	1.15%
Label	Count	%																																																					
Row Count	2620	100.00%																																																					
Null Count	0	0.00%																																																					
District Count	218	8.32%																																																					
Unique Count	0	0.00%																																																					
Duplicate Count	218	8.32%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
CYP	79	3.02%																																																					
FRA	55	2.10%																																																					
GEO	44	1.68%																																																					
MDA	43	1.64%																																																					
MUS	39	1.49%																																																					
TZA	34	1.30%																																																					
AUS	31	1.18%																																																					
CHI	30	1.15%																																																					
CHN	30	1.15%																																																					
IND	30	1.15%																																																					

Atributo	SeriesCode
Tipo de Dados	String
Descrição	Código que identifica uma série de dados específicos
Exemplo	SP.DYN.LE00.IN
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados



Atributo	DESCRIPTION																																																						
Tipo de Dados	String																																																						
Descrição	Descrição da fonte dos dados																																																						
Exemplo	Data source: United Nations World Population Prospects																																																						
Preenchido	100%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Não detetados																																																						
Estatísticas dos Dados	<p>Column: metadata.DESCRIPTION</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>2620</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>222</td><td>8.47%</td></tr> <tr><td>Unique Count</td><td>80</td><td>3.05%</td></tr> <tr><td>Duplicate Count</td><td>142</td><td>5.42%</td></tr> <tr><td>Blank Count</td><td>0</td><td>0.00%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Data source: United Nations Wor...</td><td>1349</td><td>51.49%</td></tr> <tr><td>Estimates were derived from data ...</td><td>105</td><td>4.01%</td></tr> <tr><td>Data source: Eurostat</td><td>72</td><td>2.75%</td></tr> <tr><td>Data source: HMD, Human Mortali...</td><td>72</td><td>2.75%</td></tr> <tr><td>Derived using male and female lif...</td><td>64</td><td>2.44%</td></tr> <tr><td>Estimates were derived from data ...</td><td>46</td><td>1.76%</td></tr> <tr><td>For EU Labour Force Survey, data ...</td><td>45</td><td>1.72%</td></tr> <tr><td>Data from EU Labour Force Survey...</td><td>45</td><td>1.72%</td></tr> <tr><td>Derived using ratio of age group fr...</td><td>36</td><td>1.37%</td></tr> <tr><td>Estimates were derived from data ...</td><td>23</td><td>0.88%</td></tr> </tbody> </table>	Label	Count	%	Row Count	2620	100.00%	Null Count	0	0.00%	Distinct Count	222	8.47%	Unique Count	80	3.05%	Duplicate Count	142	5.42%	Blank Count	0	0.00%	Value	Count	%	Data source: United Nations Wor...	1349	51.49%	Estimates were derived from data ...	105	4.01%	Data source: Eurostat	72	2.75%	Data source: HMD, Human Mortali...	72	2.75%	Derived using male and female lif...	64	2.44%	Estimates were derived from data ...	46	1.76%	For EU Labour Force Survey, data ...	45	1.72%	Data from EU Labour Force Survey...	45	1.72%	Derived using ratio of age group fr...	36	1.37%	Estimates were derived from data ...	23	0.88%
Label	Count	%																																																					
Row Count	2620	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	222	8.47%																																																					
Unique Count	80	3.05%																																																					
Duplicate Count	142	5.42%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
Data source: United Nations Wor...	1349	51.49%																																																					
Estimates were derived from data ...	105	4.01%																																																					
Data source: Eurostat	72	2.75%																																																					
Data source: HMD, Human Mortali...	72	2.75%																																																					
Derived using male and female lif...	64	2.44%																																																					
Estimates were derived from data ...	46	1.76%																																																					
For EU Labour Force Survey, data ...	45	1.72%																																																					
Data from EU Labour Force Survey...	45	1.72%																																																					
Derived using ratio of age group fr...	36	1.37%																																																					
Estimates were derived from data ...	23	0.88%																																																					

Gender_Stats_Country.csv

Aluno responsável: Inês Maia (a93347)

Descrição: Contém informações sobre diferentes países, incluindo moeda, região, grupo de rendimento, ano base das contas nacionais, sistema de contas nacionais, fator de conversão alternativo, ano da

pesquisa PPP (Paridade do Poder de Compra), manual de balanço de pagamentos utilizado, status de relatório de dívida externa, sistema de comércio, conceito de contabilidade governamental, padrão de disseminação de dados do FMI, últimos censos e dados de pesquisas, etc.

Atributo	Country_Code																																																						
Tipo de Dados	String																																																						
Descrição	Código único de cada país/região/situação																																																						
Exemplo	ABW																																																						
Preenchido	100%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Não detetados																																																						
Estatísticas dos Dados	<p>Column: metadata.Country_Code</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Unique Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>AFG</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>ARB</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>AFG</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>AFW</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>ARE</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>ABW</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>ARG</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>ALB</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>AND</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>AGO</td> <td>1</td> <td>0.38%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	Distinct Count	265	100.00%	Unique Count	265	100.00%	Duplicate Count	0	0.00%	Blank Count	0	0.00%	Value	Count	%	AFG	1	0.38%	ARB	1	0.38%	AFG	1	0.38%	AFW	1	0.38%	ARE	1	0.38%	ABW	1	0.38%	ARG	1	0.38%	ALB	1	0.38%	AND	1	0.38%	AGO	1	0.38%
Label	Count	%																																																					
Row Count	265	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	265	100.00%																																																					
Unique Count	265	100.00%																																																					
Duplicate Count	0	0.00%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
AFG	1	0.38%																																																					
ARB	1	0.38%																																																					
AFG	1	0.38%																																																					
AFW	1	0.38%																																																					
ARE	1	0.38%																																																					
ABW	1	0.38%																																																					
ARG	1	0.38%																																																					
ALB	1	0.38%																																																					
AND	1	0.38%																																																					
AGO	1	0.38%																																																					

Atributo	Short_Name
Tipo de Dados	String
Descrição	Nome curto de cada país/região/situação
Exemplo	Aruba
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados

Estatísticas dos Dados	<p>Column: metadata.Short_Name</p> <ul style="list-style-type: none"> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th><th>Count</th><th>%</th></tr> </thead> <tbody> <tr> <td>Row Count</td><td>265</td><td>100.00%</td></tr> <tr> <td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr> <td>Distinct Count</td><td>265</td><td>100.00%</td></tr> <tr> <td>Unique Count</td><td>265</td><td>100.00%</td></tr> <tr> <td>Duplicate Count</td><td>0</td><td>0.00%</td></tr> <tr> <td>Blank Count</td><td>0</td><td>0.00%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th><th>Count</th><th>%</th></tr> </thead> <tbody> <tr> <td>Antigua and Barbuda</td><td>1</td><td>0.38%</td></tr> <tr> <td>Afghanistan</td><td>1</td><td>0.38%</td></tr> <tr> <td>Angola</td><td>1</td><td>0.38%</td></tr> <tr> <td>Africa Eastern and Southern</td><td>1</td><td>0.38%</td></tr> <tr> <td>Andorra</td><td>1</td><td>0.38%</td></tr> <tr> <td>American Samoa</td><td>1</td><td>0.38%</td></tr> <tr> <td>Arab World</td><td>1</td><td>0.38%</td></tr> <tr> <td>Algeria</td><td>1</td><td>0.38%</td></tr> <tr> <td>Africa Western and Central</td><td>1</td><td>0.38%</td></tr> <tr> <td>Albania</td><td>1</td><td>0.38%</td></tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	Distinct Count	265	100.00%	Unique Count	265	100.00%	Duplicate Count	0	0.00%	Blank Count	0	0.00%	Value	Count	%	Antigua and Barbuda	1	0.38%	Afghanistan	1	0.38%	Angola	1	0.38%	Africa Eastern and Southern	1	0.38%	Andorra	1	0.38%	American Samoa	1	0.38%	Arab World	1	0.38%	Algeria	1	0.38%	Africa Western and Central	1	0.38%	Albania	1	0.38%	
Label	Count	%																																																						
Row Count	265	100.00%																																																						
Null Count	0	0.00%																																																						
Distinct Count	265	100.00%																																																						
Unique Count	265	100.00%																																																						
Duplicate Count	0	0.00%																																																						
Blank Count	0	0.00%																																																						
Value	Count	%																																																						
Antigua and Barbuda	1	0.38%																																																						
Afghanistan	1	0.38%																																																						
Angola	1	0.38%																																																						
Africa Eastern and Southern	1	0.38%																																																						
Andorra	1	0.38%																																																						
American Samoa	1	0.38%																																																						
Arab World	1	0.38%																																																						
Algeria	1	0.38%																																																						
Africa Western and Central	1	0.38%																																																						
Albania	1	0.38%																																																						

Atributo	Table_Name																																																							
Tipo de Dados	String																																																							
Descrição	Nome do país/região/situação conforme usado nas tabelas																																																							
Exemplo	Aruba																																																							
Preenchido	100%																																																							
Intervalo de Dados	Não mensurável																																																							
Problemas de Qualidade dos Dados	Não detetados																																																							
Estatísticas dos Dados	<p>Column: metadata.Table_Name</p> <ul style="list-style-type: none"> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Unique Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Antigua and Barbuda</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Afghanistan</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Angola</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Africa Eastern and Southern</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Andorra</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>American Samoa</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Arab World</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Algeria</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Africa Western and Central</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Albania</td> <td>1</td> <td>0.38%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	Distinct Count	265	100.00%	Unique Count	265	100.00%	Duplicate Count	0	0.00%	Blank Count	0	0.00%	Value	Count	%	Antigua and Barbuda	1	0.38%	Afghanistan	1	0.38%	Angola	1	0.38%	Africa Eastern and Southern	1	0.38%	Andorra	1	0.38%	American Samoa	1	0.38%	Arab World	1	0.38%	Algeria	1	0.38%	Africa Western and Central	1	0.38%	Albania	1	0.38%	
Label	Count	%																																																						
Row Count	265	100.00%																																																						
Null Count	0	0.00%																																																						
Distinct Count	265	100.00%																																																						
Unique Count	265	100.00%																																																						
Duplicate Count	0	0.00%																																																						
Blank Count	0	0.00%																																																						
Value	Count	%																																																						
Antigua and Barbuda	1	0.38%																																																						
Afghanistan	1	0.38%																																																						
Angola	1	0.38%																																																						
Africa Eastern and Southern	1	0.38%																																																						
Andorra	1	0.38%																																																						
American Samoa	1	0.38%																																																						
Arab World	1	0.38%																																																						
Algeria	1	0.38%																																																						
Africa Western and Central	1	0.38%																																																						
Albania	1	0.38%																																																						

Atributo	Long_Name
Tipo de Dados	String

Descrição	Nome completo do país/região/situação																																																						
Exemplo	Aruba																																																						
Preenchido	100%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Não detetados																																																						
Estatísticas dos Dados	<p>Column: metadata.Long_Name</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Unique Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Antigua and Barbuda</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Barbados</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Africa Eastern and Southern</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Arab Republic of Egypt</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>American Samoa</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Arab World</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Argentine Republic</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Aruba</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Belize</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Africa Western and Central</td> <td>1</td> <td>0.38%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	Distinct Count	265	100.00%	Unique Count	265	100.00%	Duplicate Count	0	0.00%	Blank Count	0	0.00%	Value	Count	%	Antigua and Barbuda	1	0.38%	Barbados	1	0.38%	Africa Eastern and Southern	1	0.38%	Arab Republic of Egypt	1	0.38%	American Samoa	1	0.38%	Arab World	1	0.38%	Argentine Republic	1	0.38%	Aruba	1	0.38%	Belize	1	0.38%	Africa Western and Central	1	0.38%
Label	Count	%																																																					
Row Count	265	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	265	100.00%																																																					
Unique Count	265	100.00%																																																					
Duplicate Count	0	0.00%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
Antigua and Barbuda	1	0.38%																																																					
Barbados	1	0.38%																																																					
Africa Eastern and Southern	1	0.38%																																																					
Arab Republic of Egypt	1	0.38%																																																					
American Samoa	1	0.38%																																																					
Arab World	1	0.38%																																																					
Argentine Republic	1	0.38%																																																					
Aruba	1	0.38%																																																					
Belize	1	0.38%																																																					
Africa Western and Central	1	0.38%																																																					

Atributo	_alpha_code
Tipo de Dados	String
Descrição	Código de duas letras do país/região/situação
Exemplo	AD
Preenchido	99.62%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Existência de uma linha em branco que será passada “Sem informação”, pois corresponde a um país no qual não é possível saber o seu código.

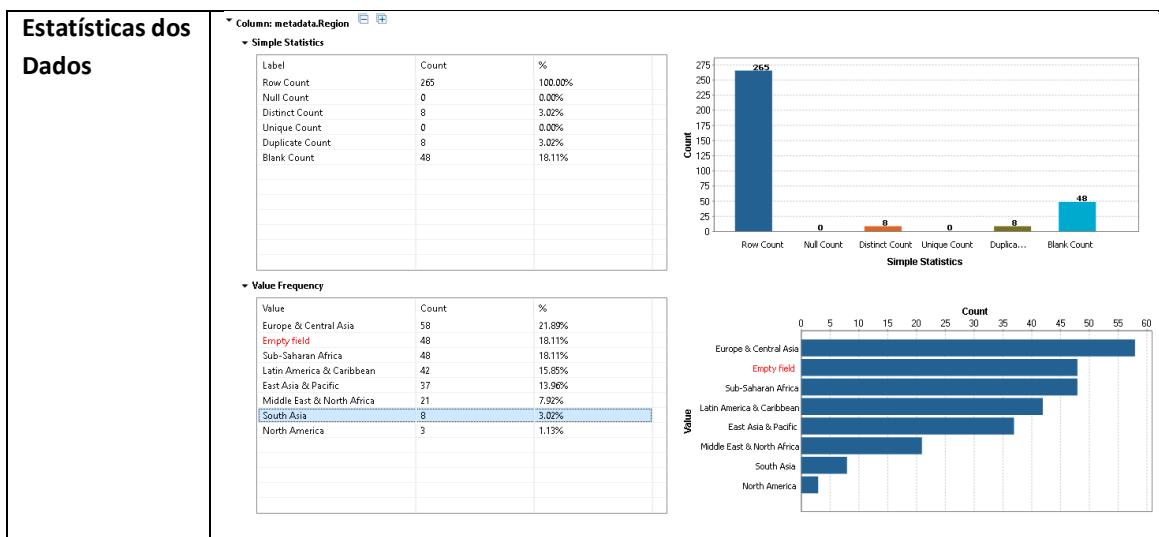


Atributo	Currency_Unit																																																						
Tipo de Dados	String																																																						
Descrição	Unidade monetária do país																																																						
Exemplo	Euro																																																						
Preenchido	81.89%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Existência de 48 linhas em branco que serão colocadas a nulo porque correspondem a regiões geográficas ou a situações específicas.																																																						
Estatísticas dos Dados	Column: metadata.Currency_Unit <ul style="list-style-type: none"> ▼ Simple Statistics <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>265</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>153</td><td>57.74%</td></tr> <tr><td>Unique Count</td><td>140</td><td>52.89%</td></tr> <tr><td>Duplicate Count</td><td>13</td><td>4.91%</td></tr> <tr><td>Blank Count</td><td>48</td><td>18.11%</td></tr> </tbody> </table> <ul style="list-style-type: none"> ▼ Value Frequency <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>48</td><td>18.11%</td></tr> <tr><td>Euro</td><td>25</td><td>9.43%</td></tr> <tr><td>U.S. dollar</td><td>14</td><td>5.28%</td></tr> <tr><td>West African CFA franc</td><td>8</td><td>3.02%</td></tr> <tr><td>Central African CFA franc</td><td>6</td><td>2.26%</td></tr> <tr><td>East Caribbean dollar</td><td>6</td><td>2.26%</td></tr> <tr><td>Australian dollar</td><td>4</td><td>1.51%</td></tr> <tr><td>Pound sterling</td><td>3</td><td>1.13%</td></tr> <tr><td>Danish krone</td><td>3</td><td>1.13%</td></tr> <tr><td>CFP franc</td><td>2</td><td>0.75%</td></tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	Distinct Count	153	57.74%	Unique Count	140	52.89%	Duplicate Count	13	4.91%	Blank Count	48	18.11%	Value	Count	%	Empty field	48	18.11%	Euro	25	9.43%	U.S. dollar	14	5.28%	West African CFA franc	8	3.02%	Central African CFA franc	6	2.26%	East Caribbean dollar	6	2.26%	Australian dollar	4	1.51%	Pound sterling	3	1.13%	Danish krone	3	1.13%	CFP franc	2	0.75%
Label	Count	%																																																					
Row Count	265	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	153	57.74%																																																					
Unique Count	140	52.89%																																																					
Duplicate Count	13	4.91%																																																					
Blank Count	48	18.11%																																																					
Value	Count	%																																																					
Empty field	48	18.11%																																																					
Euro	25	9.43%																																																					
U.S. dollar	14	5.28%																																																					
West African CFA franc	8	3.02%																																																					
Central African CFA franc	6	2.26%																																																					
East Caribbean dollar	6	2.26%																																																					
Australian dollar	4	1.51%																																																					
Pound sterling	3	1.13%																																																					
Danish krone	3	1.13%																																																					
CFP franc	2	0.75%																																																					

Atributo	Special_Notes
Tipo de Dados	String

Descrição	Notas acerca do país/visão/situação																																																						
Exemplo	22 countries, stretching from the westernmost point of Africa, across the equator, and partly along the Atlantic Ocean till the Republic of Congo in the South (https://www.worldbank.org/en/region/afr/western-and-central-africa)																																																						
Preenchido	47.55%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Existência de 139 linhas em branco que serão colocados a "Sem informação"																																																						
Estatísticas dos Dados	<p>Column: metadata.Special_Notes</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>District Count</td> <td>112</td> <td>42.26%</td> </tr> <tr> <td>Unique Count</td> <td>106</td> <td>40.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>6</td> <td>2.26%</td> </tr> <tr> <td>Blank Count</td> <td>139</td> <td>52.45%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>139</td> <td>52.45%</td> </tr> <tr> <td>Fiscal year end: March 31; reporting period: ...</td> <td>7</td> <td>2.68%</td> </tr> <tr> <td>The reporting period for national accounts: ...</td> <td>7</td> <td>2.68%</td> </tr> <tr> <td>Fiscal year end: June 30; reporting period: ...</td> <td>2</td> <td>0.75%</td> </tr> <tr> <td>Fiscal year ends on September 30; reporting period: ...</td> <td>2</td> <td>0.75%</td> </tr> <tr> <td>National accounts data are provided by: ...</td> <td>2</td> <td>0.75%</td> </tr> <tr> <td>A simple multiplier is used to convert: ...</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>26 countries, stretching from the westernmost p...</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>A simple multiplier is used to convert: ...</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>22 countries, stretching from the westernmost p...</td> <td>1</td> <td>0.38%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	District Count	112	42.26%	Unique Count	106	40.00%	Duplicate Count	6	2.26%	Blank Count	139	52.45%	Value	Count	%	Empty field	139	52.45%	Fiscal year end: March 31; reporting period: ...	7	2.68%	The reporting period for national accounts: ...	7	2.68%	Fiscal year end: June 30; reporting period: ...	2	0.75%	Fiscal year ends on September 30; reporting period: ...	2	0.75%	National accounts data are provided by: ...	2	0.75%	A simple multiplier is used to convert: ...	1	0.38%	26 countries, stretching from the westernmost p...	1	0.38%	A simple multiplier is used to convert: ...	1	0.38%	22 countries, stretching from the westernmost p...	1	0.38%
Label	Count	%																																																					
Row Count	265	100.00%																																																					
Null Count	0	0.00%																																																					
District Count	112	42.26%																																																					
Unique Count	106	40.00%																																																					
Duplicate Count	6	2.26%																																																					
Blank Count	139	52.45%																																																					
Value	Count	%																																																					
Empty field	139	52.45%																																																					
Fiscal year end: March 31; reporting period: ...	7	2.68%																																																					
The reporting period for national accounts: ...	7	2.68%																																																					
Fiscal year end: June 30; reporting period: ...	2	0.75%																																																					
Fiscal year ends on September 30; reporting period: ...	2	0.75%																																																					
National accounts data are provided by: ...	2	0.75%																																																					
A simple multiplier is used to convert: ...	1	0.38%																																																					
26 countries, stretching from the westernmost p...	1	0.38%																																																					
A simple multiplier is used to convert: ...	1	0.38%																																																					
22 countries, stretching from the westernmost p...	1	0.38%																																																					

Atributo	Region
Tipo de Dados	String
Descrição	Região geográfica do país
Exemplo	South Asia
Preenchido	81.89%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Existência de 48 linhas em branco que serão colocados a nulos, pois essas linhas já correspondem a regiões geográficas ou a situações específicas.

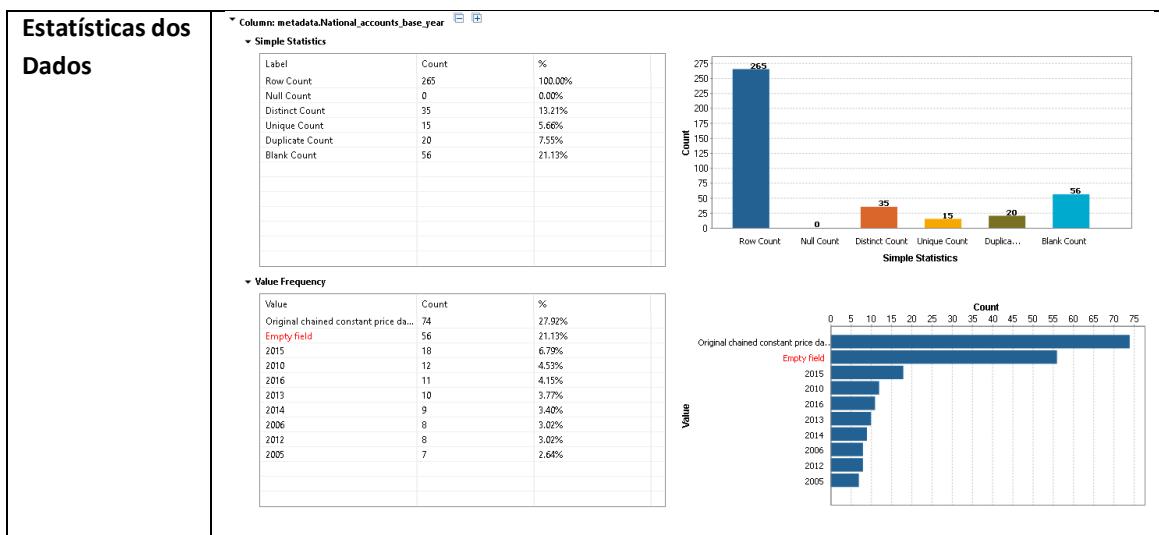


Atributo	Income_Group																																							
Tipo de Dados	String																																							
Descrição	Classificação do país com base no rendimento																																							
Exemplo	High income																																							
Preenchido	81.51%																																							
Intervalo de Dados	Não mensurável																																							
Problemas de Qualidade dos Dados	Existência de 49 linhas em branco. 48 serão colocados a nulos, pois correspondem a regiões geográficas ou a situações específicas e a restante será substituída pela informação necessária acerca desse país.																																							
Estatísticas dos Dados	<p>Column: metadata.Income_Group</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>265</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>5</td><td>1.89%</td></tr> <tr><td>Unique Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Duplicate Count</td><td>5</td><td>1.89%</td></tr> <tr><td>Blank Count</td><td>49</td><td>18.40%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>High income</td><td>82</td><td>30.94%</td></tr> <tr><td>Upper middle income</td><td>54</td><td>20.38%</td></tr> <tr><td>Lower middle income</td><td>54</td><td>20.38%</td></tr> <tr><td>Empty field</td><td>49</td><td>18.40%</td></tr> <tr><td>Low income</td><td>26</td><td>9.81%</td></tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	Distinct Count	5	1.89%	Unique Count	0	0.00%	Duplicate Count	5	1.89%	Blank Count	49	18.40%	Value	Count	%	High income	82	30.94%	Upper middle income	54	20.38%	Lower middle income	54	20.38%	Empty field	49	18.40%	Low income	26	9.81%
Label	Count	%																																						
Row Count	265	100.00%																																						
Null Count	0	0.00%																																						
Distinct Count	5	1.89%																																						
Unique Count	0	0.00%																																						
Duplicate Count	5	1.89%																																						
Blank Count	49	18.40%																																						
Value	Count	%																																						
High income	82	30.94%																																						
Upper middle income	54	20.38%																																						
Lower middle income	54	20.38%																																						
Empty field	49	18.40%																																						
Low income	26	9.81%																																						

Atributo	WB_2_code
Tipo de Dados	String

Descrição	Código de dois dígitos do World Bank																																																						
Exemplo	AD																																																						
Preenchido	100%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Não detetados																																																						
Estatísticas dos Dados	<p>Column: metadata.WB_2_code</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Unique Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>1A</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>4E</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>AD</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>AE</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>AF</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>1W</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>AG</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>BS</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>AL</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>7E</td> <td>1</td> <td>0.38%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	Distinct Count	265	100.00%	Unique Count	265	100.00%	Duplicate Count	0	0.00%	Blank Count	0	0.00%	Value	Count	%	1A	1	0.38%	4E	1	0.38%	AD	1	0.38%	AE	1	0.38%	AF	1	0.38%	1W	1	0.38%	AG	1	0.38%	BS	1	0.38%	AL	1	0.38%	7E	1	0.38%
Label	Count	%																																																					
Row Count	265	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	265	100.00%																																																					
Unique Count	265	100.00%																																																					
Duplicate Count	0	0.00%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
1A	1	0.38%																																																					
4E	1	0.38%																																																					
AD	1	0.38%																																																					
AE	1	0.38%																																																					
AF	1	0.38%																																																					
1W	1	0.38%																																																					
AG	1	0.38%																																																					
BS	1	0.38%																																																					
AL	1	0.38%																																																					
7E	1	0.38%																																																					

Atributo	National_accounts_base_year
Tipo de Dados	String
Descrição	Ano base para as contas nacionais
Exemplo	2015
Preenchido	78.7%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Apesar de existirem dados numéricos estes necessitam de continuar no formato String devido aos restantes dados representados por frases. Existem 56 linhas em branco que serão passadas a "Sem informação"

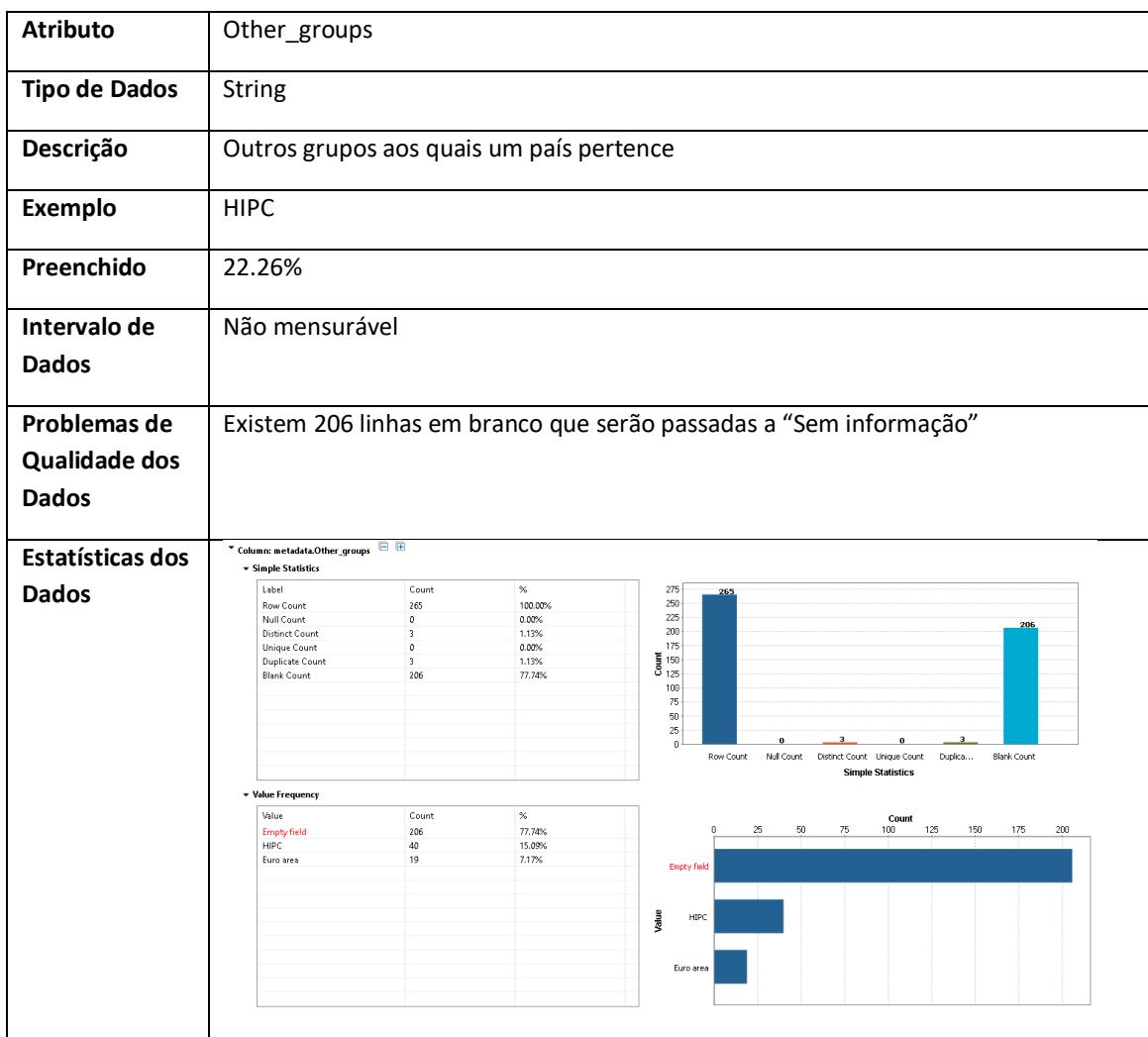
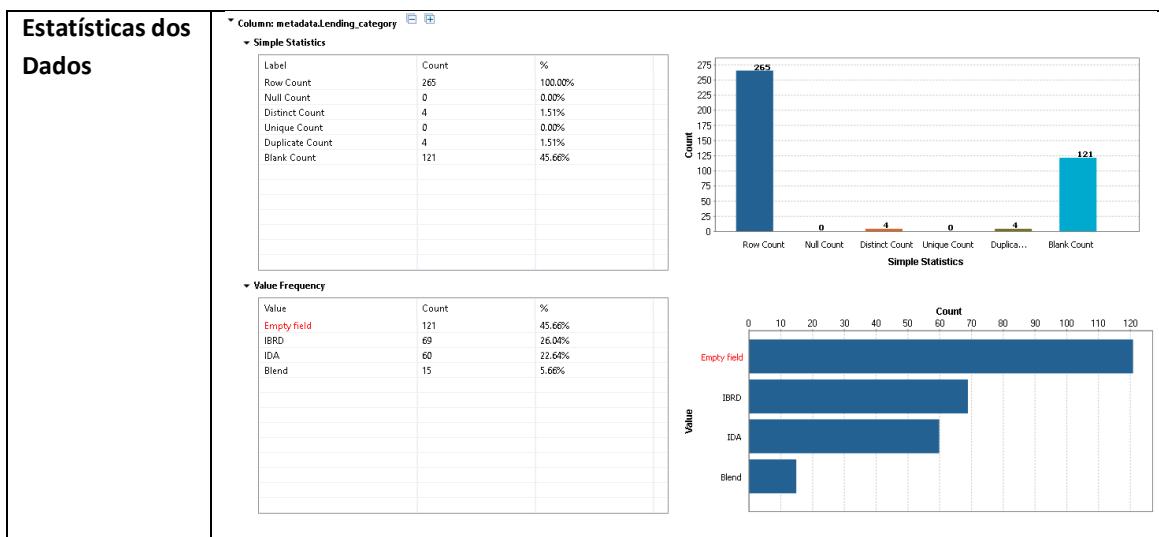


Atributo	National_accounts_reference_year																																																						
Tipo de Dados	String																																																						
Descrição	Ano de referência para as contas nacionais																																																						
Exemplo	2015																																																						
Preenchido	27.92%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Como só existem valores numéricos, estes serão convertidos em Integer. Existem 191 linhas em branco que serão passadas a null.																																																						
Estatísticas dos Dados	<p>Column: metadata.National_accounts_reference_year</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>265</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>18</td><td>6.79%</td></tr> <tr><td>Unique Count</td><td>9</td><td>3.40%</td></tr> <tr><td>Duplicate Count</td><td>9</td><td>3.40%</td></tr> <tr><td>Blank Count</td><td>191</td><td>72.08%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>191</td><td>72.08%</td></tr> <tr><td>2015</td><td>41</td><td>15.47%</td></tr> <tr><td>2010</td><td>7</td><td>2.64%</td></tr> <tr><td>2005</td><td>5</td><td>1.86%</td></tr> <tr><td>2018</td><td>3</td><td>1.13%</td></tr> <tr><td>2012</td><td>3</td><td>1.13%</td></tr> <tr><td>2016</td><td>2</td><td>0.75%</td></tr> <tr><td>2014</td><td>2</td><td>0.75%</td></tr> <tr><td>2021</td><td>2</td><td>0.75%</td></tr> <tr><td>2000</td><td>1</td><td>0.38%</td></tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	Distinct Count	18	6.79%	Unique Count	9	3.40%	Duplicate Count	9	3.40%	Blank Count	191	72.08%	Value	Count	%	Empty field	191	72.08%	2015	41	15.47%	2010	7	2.64%	2005	5	1.86%	2018	3	1.13%	2012	3	1.13%	2016	2	0.75%	2014	2	0.75%	2021	2	0.75%	2000	1	0.38%
Label	Count	%																																																					
Row Count	265	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	18	6.79%																																																					
Unique Count	9	3.40%																																																					
Duplicate Count	9	3.40%																																																					
Blank Count	191	72.08%																																																					
Value	Count	%																																																					
Empty field	191	72.08%																																																					
2015	41	15.47%																																																					
2010	7	2.64%																																																					
2005	5	1.86%																																																					
2018	3	1.13%																																																					
2012	3	1.13%																																																					
2016	2	0.75%																																																					
2014	2	0.75%																																																					
2021	2	0.75%																																																					
2000	1	0.38%																																																					

Atributo	SNA_price_valuation
Tipo de Dados	String

Descrição	Método de avaliação de preço																																	
Exemplo	Value added at basic prices (VAB)																																	
Preenchido	78.11%																																	
Intervalo de Dados	Não mensurável																																	
Problemas de Qualidade dos Dados	Existem 58 linhas em branco que serão passadas a “Sem informação”																																	
Estatísticas dos Dados	<p>Column: metadata.SNA_price_valuation</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>District Count</td> <td>3</td> <td>1.13%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>3</td> <td>1.13%</td> </tr> <tr> <td>Blank Count</td> <td>58</td> <td>21.89%</td> </tr> </tbody> </table> <p>Simple Statistics</p> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Value added at basic prices (VAB)</td> <td>193</td> <td>72.83%</td> </tr> <tr> <td>Empty field</td> <td>58</td> <td>21.89%</td> </tr> <tr> <td>Value added at producer prices (V...</td> <td>14</td> <td>5.28%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	District Count	3	1.13%	Unique Count	0	0.00%	Duplicate Count	3	1.13%	Blank Count	58	21.89%	Value	Count	%	Value added at basic prices (VAB)	193	72.83%	Empty field	58	21.89%	Value added at producer prices (V...	14	5.28%
Label	Count	%																																
Row Count	265	100.00%																																
Null Count	0	0.00%																																
District Count	3	1.13%																																
Unique Count	0	0.00%																																
Duplicate Count	3	1.13%																																
Blank Count	58	21.89%																																
Value	Count	%																																
Value added at basic prices (VAB)	193	72.83%																																
Empty field	58	21.89%																																
Value added at producer prices (V...	14	5.28%																																

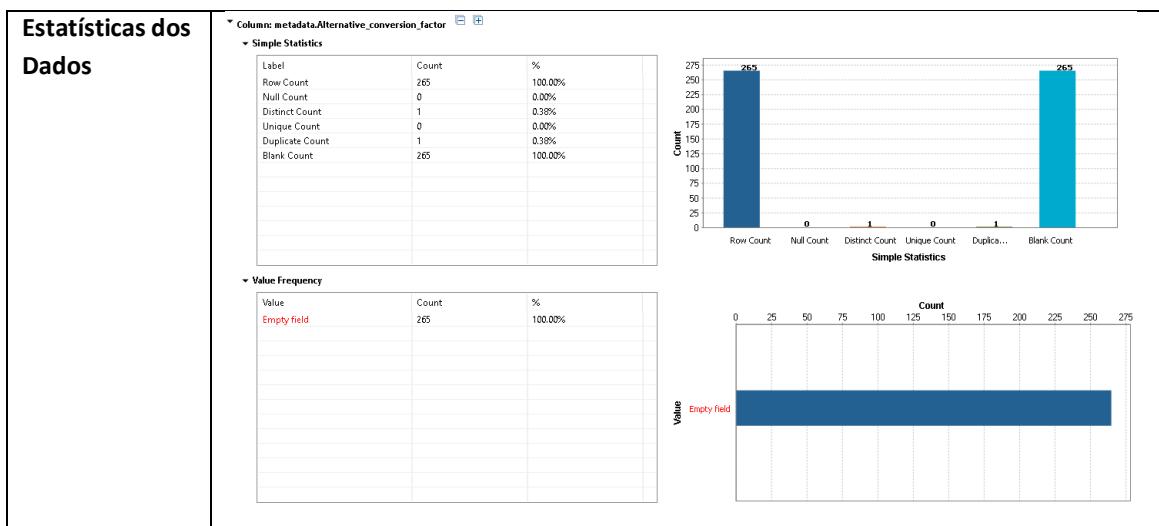
Atributo	Lending_category
Tipo de Dados	String
Descrição	Categoria de empréstimo do país
Exemplo	IDA
Preenchido	54.34%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Existem 121 linhas em branco que serão passadas a “Sem informação”



Atributo System_of_National_Accounts	
Tipo de Dados	String

Descrição	O sistema de contas nacionais usado																																							
Exemplo	Country uses the 2008 System of National Accounts methodology																																							
Preenchido	78.49%																																							
Intervalo de Dados	Não mensurável																																							
Problemas de Qualidade dos Dados	Existem 57 linhas em branco em que algumas correspondem a regiões geográficas, outras a situações específicas e as restantes a países. Estas serão passadas a null																																							
Estatísticas dos Dados	<p>Column: metadata.System_of_National_Accounts</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>5</td> <td>1.89%</td> </tr> <tr> <td>Unique Count</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Duplicate Count</td> <td>4</td> <td>1.51%</td> </tr> <tr> <td>Blank Count</td> <td>57</td> <td>21.51%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Country uses the 2008 System of ...</td> <td>142</td> <td>52.50%</td> </tr> <tr> <td>Country uses the 1993 System of ...</td> <td>59</td> <td>22.26%</td> </tr> <tr> <td>Empty field</td> <td>57</td> <td>21.51%</td> </tr> <tr> <td>Country uses the 1968 System of ...</td> <td>6</td> <td>2.26%</td> </tr> <tr> <td>Country uses the 2010 European s...</td> <td>1</td> <td>0.38%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	Distinct Count	5	1.89%	Unique Count	1	0.38%	Duplicate Count	4	1.51%	Blank Count	57	21.51%	Value	Count	%	Country uses the 2008 System of ...	142	52.50%	Country uses the 1993 System of ...	59	22.26%	Empty field	57	21.51%	Country uses the 1968 System of ...	6	2.26%	Country uses the 2010 European s...	1	0.38%
Label	Count	%																																						
Row Count	265	100.00%																																						
Null Count	0	0.00%																																						
Distinct Count	5	1.89%																																						
Unique Count	1	0.38%																																						
Duplicate Count	4	1.51%																																						
Blank Count	57	21.51%																																						
Value	Count	%																																						
Country uses the 2008 System of ...	142	52.50%																																						
Country uses the 1993 System of ...	59	22.26%																																						
Empty field	57	21.51%																																						
Country uses the 1968 System of ...	6	2.26%																																						
Country uses the 2010 European s...	1	0.38%																																						

Atributo	Alternative_conversion_factor
Tipo de Dados	String
Descrição	Fator de conversão alternativo.
Exemplo	
Preenchido	0%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Todas as linhas estão em branco, o que indica que não foi usada um fator de conversão alternativo. Assim, estas linhas poderão ser convertidas em nulos.

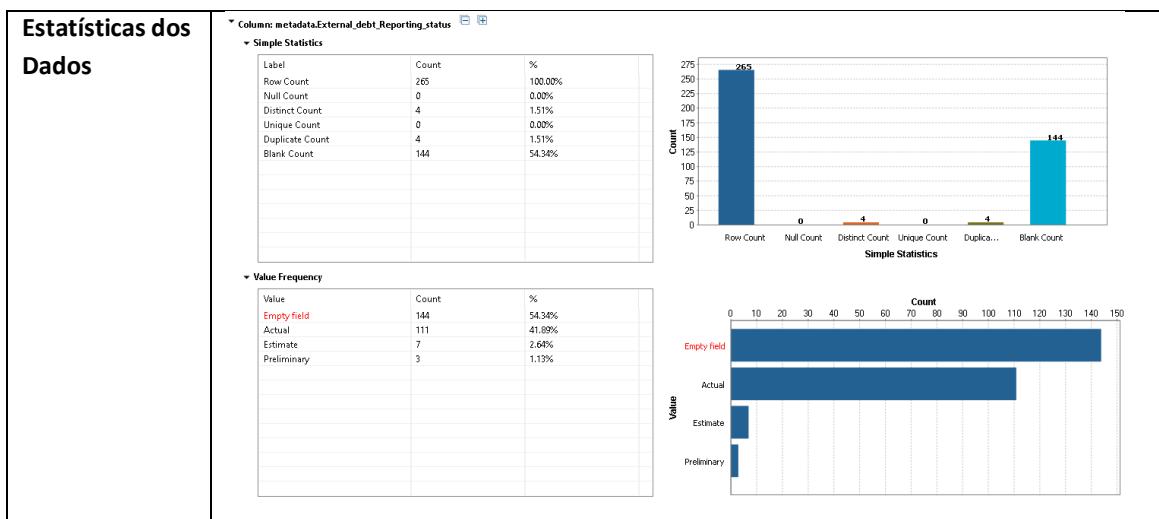


Atributo	PPP_survey_year																											
Tipo de Dados	String																											
Descrição	Ano da pesquisa de Paridade do Poder de Compra																											
Exemplo																												
Preenchido	0%																											
Intervalo de Dados	Não mensurável																											
Problemas de Qualidade dos Dados	Todas as linhas estão em branco, que serão passadas a nulo.																											
Estatísticas dos Dados	<p>Column: metadata.PPP_survey_year</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>1</td> <td>0.38%</td> </tr> <tr> <td>Blank Count</td> <td>265</td> <td>100.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>265</td> <td>100.00%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	Distinct Count	1	0.38%	Unique Count	0	0.00%	Duplicate Count	1	0.38%	Blank Count	265	100.00%	Value	Count	%	Empty field	265	100.00%
Label	Count	%																										
Row Count	265	100.00%																										
Null Count	0	0.00%																										
Distinct Count	1	0.38%																										
Unique Count	0	0.00%																										
Duplicate Count	1	0.38%																										
Blank Count	265	100.00%																										
Value	Count	%																										
Empty field	265	100.00%																										

Atributo	Balance_of_Payments_Manual_in_use
Tipo de Dados	String

Descrição	Manual do Balanço de Pagamentos em uso																																	
Exemplo	BPM6																																	
Preenchido	73.58%																																	
Intervalo de Dados	Não mensurável																																	
Problemas de Qualidade dos Dados	Existem 70 linhas em branco que serão passadas a “Sem informação”																																	
Estatísticas dos Dados	<p>* Column: metadata.Balance_of_Payments_Manual_in_use</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>District Count</td> <td>3</td> <td>1.13%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>3</td> <td>1.13%</td> </tr> <tr> <td>Blank Count</td> <td>70</td> <td>26.42%</td> </tr> </tbody> </table> <p>Simple Statistics</p> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>BPM6</td> <td>174</td> <td>65.66%</td> </tr> <tr> <td>Empty field</td> <td>70</td> <td>26.42%</td> </tr> <tr> <td>BPM5 (Converted into BPM6 by IMF)</td> <td>21</td> <td>7.92%</td> </tr> </tbody> </table> <p>Value</p>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	District Count	3	1.13%	Unique Count	0	0.00%	Duplicate Count	3	1.13%	Blank Count	70	26.42%	Value	Count	%	BPM6	174	65.66%	Empty field	70	26.42%	BPM5 (Converted into BPM6 by IMF)	21	7.92%
Label	Count	%																																
Row Count	265	100.00%																																
Null Count	0	0.00%																																
District Count	3	1.13%																																
Unique Count	0	0.00%																																
Duplicate Count	3	1.13%																																
Blank Count	70	26.42%																																
Value	Count	%																																
BPM6	174	65.66%																																
Empty field	70	26.42%																																
BPM5 (Converted into BPM6 by IMF)	21	7.92%																																

Atributo	External_debt_Reportig_status
Tipo de Dados	String
Descrição	Status do relatório de dívida externa
Exemplo	Actual
Preenchido	45.66%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Existem 144 linhas em branco que serão passadas a “Sem informação”

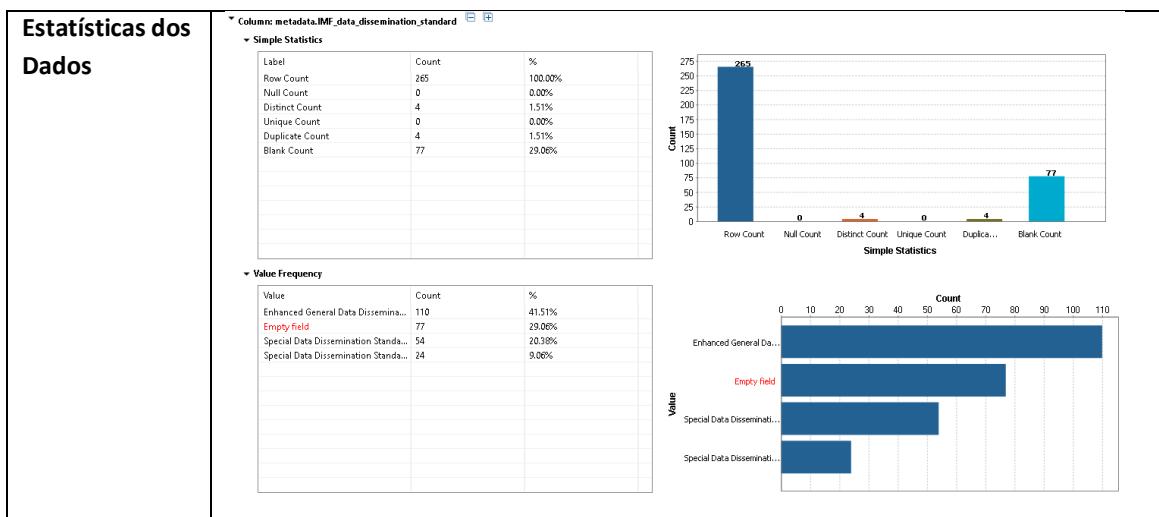


Atributo	System_of_trade																																	
Tipo de Dados	String																																	
Descrição	Sistema de comércio utilizado pelo país																																	
Exemplo	General trade system																																	
Preenchido	64.09%																																	
Intervalo de Dados	Não mensurável																																	
Problemas de Qualidade dos Dados	Existem 93 linhas em branco que serão passadas a "Sem informação"																																	
Estatísticas dos Dados	<p>Column: metadata.System_of_trade</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>265</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>3</td><td>1.13%</td></tr> <tr><td>Unique Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Duplicate Count</td><td>3</td><td>1.13%</td></tr> <tr><td>Blank Count</td><td>93</td><td>35.09%</td></tr> </tbody> </table> <p>Simple Statistics</p> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>General trade system</td><td>108</td><td>40.75%</td></tr> <tr><td>Empty field</td><td>93</td><td>35.09%</td></tr> <tr><td>Special trade system</td><td>64</td><td>24.15%</td></tr> </tbody> </table> <p>Value Frequency</p>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	Distinct Count	3	1.13%	Unique Count	0	0.00%	Duplicate Count	3	1.13%	Blank Count	93	35.09%	Value	Count	%	General trade system	108	40.75%	Empty field	93	35.09%	Special trade system	64	24.15%
Label	Count	%																																
Row Count	265	100.00%																																
Null Count	0	0.00%																																
Distinct Count	3	1.13%																																
Unique Count	0	0.00%																																
Duplicate Count	3	1.13%																																
Blank Count	93	35.09%																																
Value	Count	%																																
General trade system	108	40.75%																																
Empty field	93	35.09%																																
Special trade system	64	24.15%																																

Atributo	Government_Accounting_concept
Tipo de Dados	String

Descrição	Conceito de contabilidade governamental																																	
Exemplo	Consolidated central government																																	
Preenchido	59.62%																																	
Intervalo de Dados	Não mensurável																																	
Problemas de Qualidade dos Dados	Existem 107 linhas em branco que serão passadas a “Sem informação”																																	
Estatísticas dos Dados	<p>Column: metadata.Government_Accounting_concept</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>District Count</td> <td>3</td> <td>1.13%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>3</td> <td>1.13%</td> </tr> <tr> <td>Blank Count</td> <td>107</td> <td>40.38%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>107</td> <td>40.38%</td> </tr> <tr> <td>Consolidated central government</td> <td>91</td> <td>34.34%</td> </tr> <tr> <td>Budgetary central government</td> <td>67</td> <td>25.28%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	District Count	3	1.13%	Unique Count	0	0.00%	Duplicate Count	3	1.13%	Blank Count	107	40.38%	Value	Count	%	Empty field	107	40.38%	Consolidated central government	91	34.34%	Budgetary central government	67	25.28%
Label	Count	%																																
Row Count	265	100.00%																																
Null Count	0	0.00%																																
District Count	3	1.13%																																
Unique Count	0	0.00%																																
Duplicate Count	3	1.13%																																
Blank Count	107	40.38%																																
Value	Count	%																																
Empty field	107	40.38%																																
Consolidated central government	91	34.34%																																
Budgetary central government	67	25.28%																																

Atributo	IMF_data_dissemination_standard
Tipo de Dados	String
Descrição	Padrão de disseminação de dados do FMI
Exemplo	Special Data Dissemination Standard (SDDS)
Preenchido	70.94%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Existem 77 linhas em branco que serão passadas a “Sem informação”

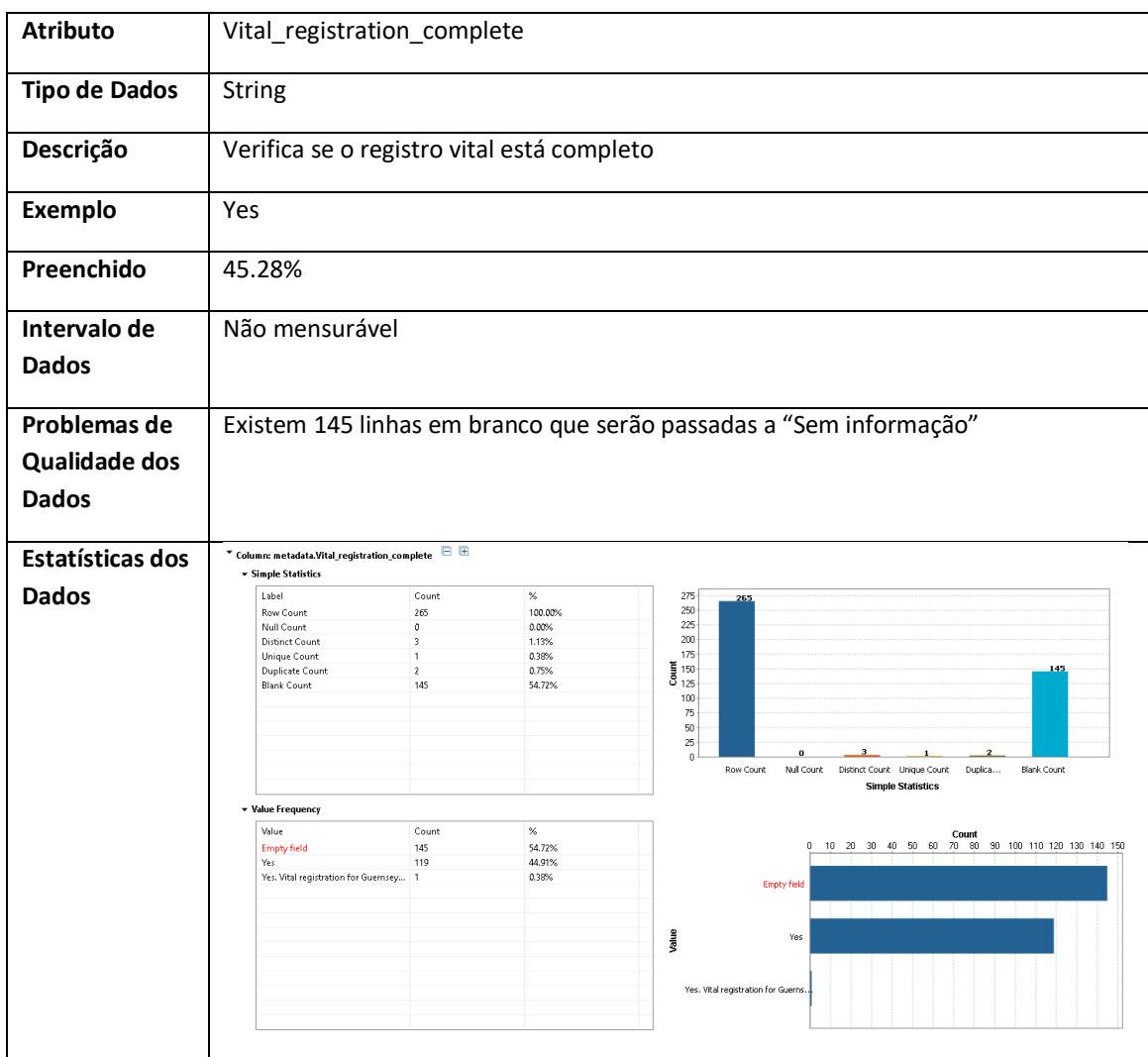
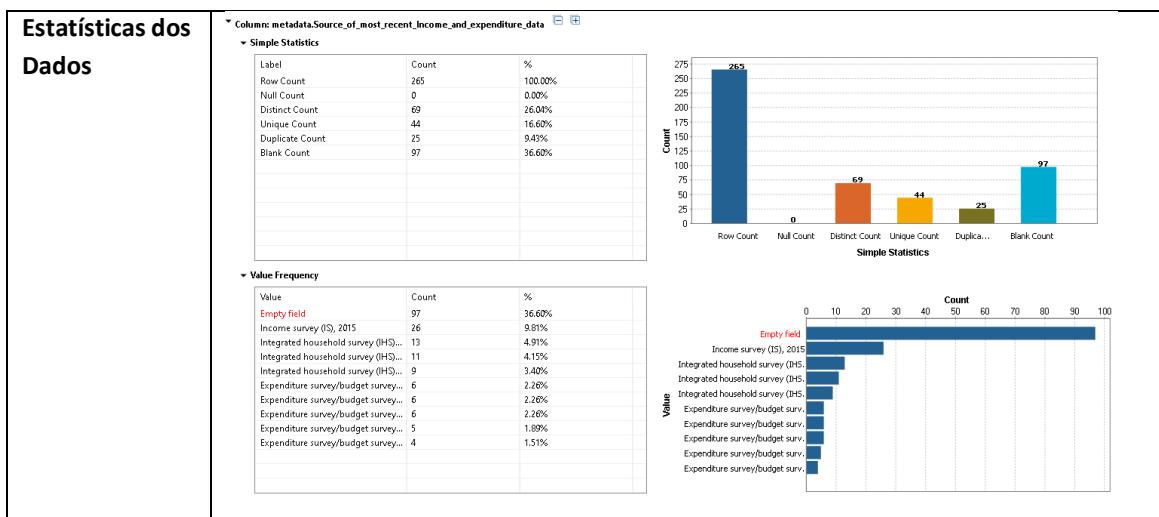


Atributo	Latest_population_census																																																						
Tipo de Dados	String																																																						
Descrição	Data do último censo populacional																																																						
Exemplo	2011																																																						
Preenchido	70.94%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os valores numéricos correspondentes aos anos encontram-se acompanhados de Strings. Existem 51 linhas em branco que serão passadas a "Sem informação"																																																						
Estatísticas dos Dados	<p>Column: metadata.latest_population_census</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>265</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>33</td><td>12.45%</td></tr> <tr><td>Unique Count</td><td>12</td><td>4.55%</td></tr> <tr><td>Duplicate Count</td><td>21</td><td>7.82%</td></tr> <tr><td>Blank Count</td><td>51</td><td>19.25%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>2020 (expected)</td><td>62</td><td>23.40%</td></tr> <tr><td>Empty field</td><td>51</td><td>19.25%</td></tr> <tr><td>2011</td><td>33</td><td>12.45%</td></tr> <tr><td>2012</td><td>16</td><td>6.04%</td></tr> <tr><td>2017</td><td>11</td><td>4.15%</td></tr> <tr><td>2016</td><td>10</td><td>3.77%</td></tr> <tr><td>2014</td><td>9</td><td>3.40%</td></tr> <tr><td>2019</td><td>8</td><td>3.02%</td></tr> <tr><td>2011. Population figures compiled by the UN</td><td>8</td><td>3.02%</td></tr> <tr><td>2013</td><td>7</td><td>2.64%</td></tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	Distinct Count	33	12.45%	Unique Count	12	4.55%	Duplicate Count	21	7.82%	Blank Count	51	19.25%	Value	Count	%	2020 (expected)	62	23.40%	Empty field	51	19.25%	2011	33	12.45%	2012	16	6.04%	2017	11	4.15%	2016	10	3.77%	2014	9	3.40%	2019	8	3.02%	2011. Population figures compiled by the UN	8	3.02%	2013	7	2.64%
Label	Count	%																																																					
Row Count	265	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	33	12.45%																																																					
Unique Count	12	4.55%																																																					
Duplicate Count	21	7.82%																																																					
Blank Count	51	19.25%																																																					
Value	Count	%																																																					
2020 (expected)	62	23.40%																																																					
Empty field	51	19.25%																																																					
2011	33	12.45%																																																					
2012	16	6.04%																																																					
2017	11	4.15%																																																					
2016	10	3.77%																																																					
2014	9	3.40%																																																					
2019	8	3.02%																																																					
2011. Population figures compiled by the UN	8	3.02%																																																					
2013	7	2.64%																																																					

Atributo	Latest_household_survey
Tipo de Dados	String

Descrição	Data da última pesquisa domiciliar																																																						
Exemplo	Multiple Indicator Cluster Survey, 2006																																																						
Preenchido	58.11%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os valores numéricos correspondentes aos anos encontram-se acompanhados de Strings. Existem 111 linhas em branco que serão passadas a “Sem informação”																																																						
Estatísticas dos Dados	<p>Column: metadata.latest_household_survey</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>81</td> <td>30.57%</td> </tr> <tr> <td>Unique Count</td> <td>56</td> <td>21.13%</td> </tr> <tr> <td>Duplicate Count</td> <td>25</td> <td>9.43%</td> </tr> <tr> <td>Blank Count</td> <td>111</td> <td>41.89%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>111</td> <td>41.89%</td> </tr> <tr> <td>Multiple Indicator Cluster Survey, ...</td> <td>13</td> <td>4.91%</td> </tr> <tr> <td>World Health Survey, 2003</td> <td>10</td> <td>3.77%</td> </tr> <tr> <td>Multiple Indicator Cluster Survey, ...</td> <td>9</td> <td>3.40%</td> </tr> <tr> <td>Multiple Indicator Cluster Survey, ...</td> <td>7</td> <td>2.64%</td> </tr> <tr> <td>Demographic and Health Survey, ...</td> <td>6</td> <td>2.26%</td> </tr> <tr> <td>Multiple Indicator Cluster Survey, ...</td> <td>5</td> <td>1.89%</td> </tr> <tr> <td>Multiple Indicator Cluster Survey, ...</td> <td>5</td> <td>1.89%</td> </tr> <tr> <td>Multiple Indicator Cluster Survey, ...</td> <td>4</td> <td>1.51%</td> </tr> <tr> <td>Multiple Indicator Cluster Survey, ...</td> <td>4</td> <td>1.51%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	Distinct Count	81	30.57%	Unique Count	56	21.13%	Duplicate Count	25	9.43%	Blank Count	111	41.89%	Value	Count	%	Empty field	111	41.89%	Multiple Indicator Cluster Survey, ...	13	4.91%	World Health Survey, 2003	10	3.77%	Multiple Indicator Cluster Survey, ...	9	3.40%	Multiple Indicator Cluster Survey, ...	7	2.64%	Demographic and Health Survey, ...	6	2.26%	Multiple Indicator Cluster Survey, ...	5	1.89%	Multiple Indicator Cluster Survey, ...	5	1.89%	Multiple Indicator Cluster Survey, ...	4	1.51%	Multiple Indicator Cluster Survey, ...	4	1.51%
Label	Count	%																																																					
Row Count	265	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	81	30.57%																																																					
Unique Count	56	21.13%																																																					
Duplicate Count	25	9.43%																																																					
Blank Count	111	41.89%																																																					
Value	Count	%																																																					
Empty field	111	41.89%																																																					
Multiple Indicator Cluster Survey, ...	13	4.91%																																																					
World Health Survey, 2003	10	3.77%																																																					
Multiple Indicator Cluster Survey, ...	9	3.40%																																																					
Multiple Indicator Cluster Survey, ...	7	2.64%																																																					
Demographic and Health Survey, ...	6	2.26%																																																					
Multiple Indicator Cluster Survey, ...	5	1.89%																																																					
Multiple Indicator Cluster Survey, ...	5	1.89%																																																					
Multiple Indicator Cluster Survey, ...	4	1.51%																																																					
Multiple Indicator Cluster Survey, ...	4	1.51%																																																					

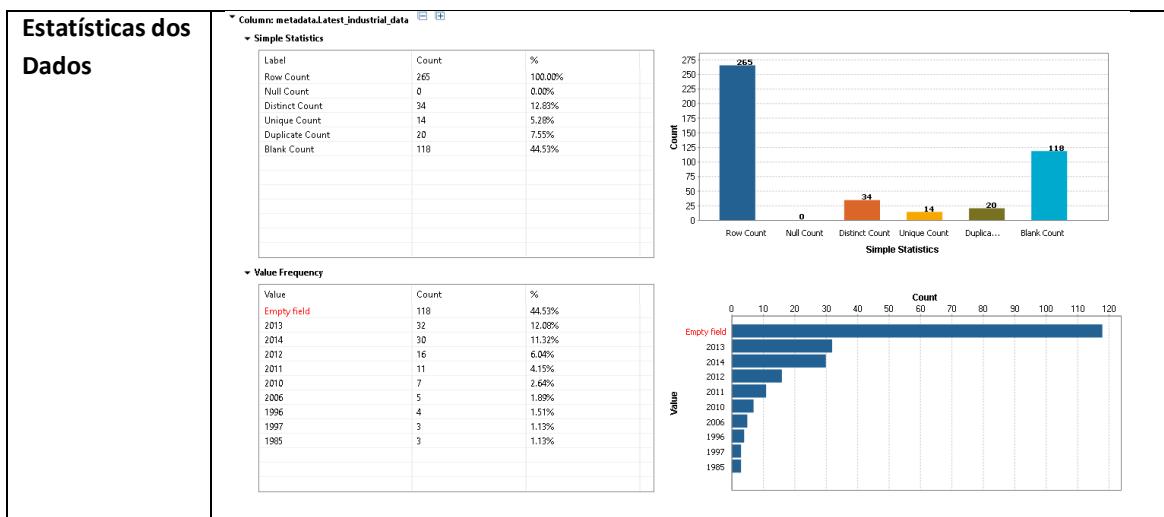
Atributo	Source_of_most_recent_Income_and_expenditure_data
Tipo de Dados	String
Descrição	Fonte dos dados mais recentes do rendimento e despesas
Exemplo	Income survey (IS), 2015
Preenchido	63.4%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os valores numéricos correspondentes aos anos encontram-se acompanhados de Strings. Existem 97 linhas em branco que serão passadas a “Sem informação”



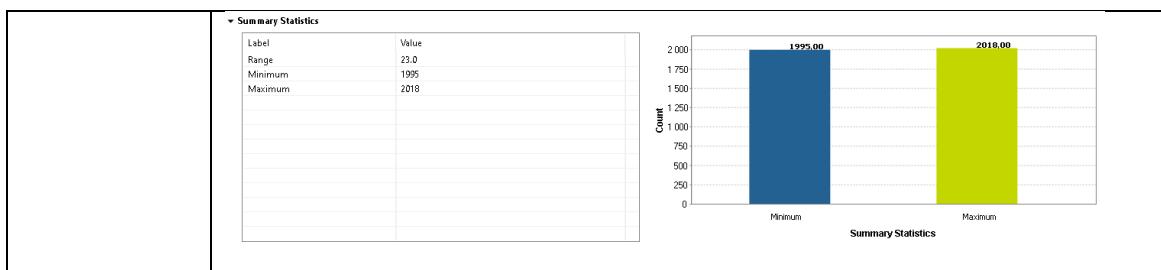
Atributo	Latest_agricultural_census
Tipo de Dados	String

Descrição	Data do último censo agrícola																																																						
Exemplo	2015																																																						
Preenchido	48.3%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os valores numéricos correspondentes aos anos encontram-se acompanhados de Strings. Existem 137 linhas em branco que serão passadas a “Sem informação”																																																						
Estatísticas dos Dados	<p>* Column: metadata.latest_agricultural_census</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>District Count</td> <td>31</td> <td>11.70%</td> </tr> <tr> <td>Unique Count</td> <td>12</td> <td>4.53%</td> </tr> <tr> <td>Duplicate Count</td> <td>19</td> <td>7.17%</td> </tr> <tr> <td>Blank Count</td> <td>137</td> <td>51.70%</td> </tr> </tbody> </table> <p>Simple Statistics</p> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>137</td> <td>51.70%</td> </tr> <tr> <td>2010</td> <td>35</td> <td>13.21%</td> </tr> <tr> <td>2007</td> <td>11</td> <td>4.15%</td> </tr> <tr> <td>2012</td> <td>10</td> <td>3.77%</td> </tr> <tr> <td>2015</td> <td>9</td> <td>3.40%</td> </tr> <tr> <td>2011</td> <td>7</td> <td>2.64%</td> </tr> <tr> <td>2014</td> <td>6</td> <td>2.26%</td> </tr> <tr> <td>2013</td> <td>6</td> <td>2.26%</td> </tr> <tr> <td>2009</td> <td>5</td> <td>1.89%</td> </tr> <tr> <td>2008</td> <td>5</td> <td>1.89%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	0	0.00%	District Count	31	11.70%	Unique Count	12	4.53%	Duplicate Count	19	7.17%	Blank Count	137	51.70%	Value	Count	%	Empty field	137	51.70%	2010	35	13.21%	2007	11	4.15%	2012	10	3.77%	2015	9	3.40%	2011	7	2.64%	2014	6	2.26%	2013	6	2.26%	2009	5	1.89%	2008	5	1.89%
Label	Count	%																																																					
Row Count	265	100.00%																																																					
Null Count	0	0.00%																																																					
District Count	31	11.70%																																																					
Unique Count	12	4.53%																																																					
Duplicate Count	19	7.17%																																																					
Blank Count	137	51.70%																																																					
Value	Count	%																																																					
Empty field	137	51.70%																																																					
2010	35	13.21%																																																					
2007	11	4.15%																																																					
2012	10	3.77%																																																					
2015	9	3.40%																																																					
2011	7	2.64%																																																					
2014	6	2.26%																																																					
2013	6	2.26%																																																					
2009	5	1.89%																																																					
2008	5	1.89%																																																					

Atributo	Latest_industrial_data
Tipo de Dados	String
Descrição	Data dos últimos dados industriais
Exemplo	1985
Preenchido	55.47%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Como só existem valores numéricos, estes serão convertidos em Integer. Existem 118 linhas em branco que serão passadas a “Sem informação”



Atributo	Latest_trade_data																																																			
Tipo de Dados	Integer																																																			
Descrição	Data dos últimos dados comerciais																																																			
Exemplo	2017																																																			
Preenchido	72.08%																																																			
Intervalo de Dados	23.0																																																			
Problemas de Qualidade dos Dados	Não detetados																																																			
Estatísticas dos Dados	<p>Column: metadata.latest_trade_data</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>265</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>74</td> <td>27.92%</td> </tr> <tr> <td>District Count</td> <td>17</td> <td>6.42%</td> </tr> <tr> <td>Unique Count</td> <td>7</td> <td>2.64%</td> </tr> <tr> <td>Duplicate Count</td> <td>10</td> <td>3.77%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>2018</td> <td>129</td> <td>48.69%</td> </tr> <tr> <td>Null field</td> <td>74</td> <td>27.92%</td> </tr> <tr> <td>2017</td> <td>24</td> <td>9.06%</td> </tr> <tr> <td>2016</td> <td>8</td> <td>3.02%</td> </tr> <tr> <td>2015</td> <td>8</td> <td>3.02%</td> </tr> <tr> <td>2009</td> <td>4</td> <td>1.51%</td> </tr> <tr> <td>2012</td> <td>4</td> <td>1.51%</td> </tr> <tr> <td>2013</td> <td>3</td> <td>1.13%</td> </tr> <tr> <td>2010</td> <td>2</td> <td>0.75%</td> </tr> <tr> <td>2014</td> <td>2</td> <td>0.75%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	265	100.00%	Null Count	74	27.92%	District Count	17	6.42%	Unique Count	7	2.64%	Duplicate Count	10	3.77%	Value	Count	%	2018	129	48.69%	Null field	74	27.92%	2017	24	9.06%	2016	8	3.02%	2015	8	3.02%	2009	4	1.51%	2012	4	1.51%	2013	3	1.13%	2010	2	0.75%	2014	2	0.75%
Label	Count	%																																																		
Row Count	265	100.00%																																																		
Null Count	74	27.92%																																																		
District Count	17	6.42%																																																		
Unique Count	7	2.64%																																																		
Duplicate Count	10	3.77%																																																		
Value	Count	%																																																		
2018	129	48.69%																																																		
Null field	74	27.92%																																																		
2017	24	9.06%																																																		
2016	8	3.02%																																																		
2015	8	3.02%																																																		
2009	4	1.51%																																																		
2012	4	1.51%																																																		
2013	3	1.13%																																																		
2010	2	0.75%																																																		
2014	2	0.75%																																																		



Gender_Stats Data.csv

Aluno responsável: Inês Maia (a93347)

Descrição: Este csv contém dados anuais desde 1960 a 2022 com indicadores referentes a várias estatísticas relacionadas com género.

Atributo	_Country_Name																																																						
Tipo de Dados	String																																																						
Descrição	Nome completo do país/região/situação																																																						
Exemplo	Albania																																																						
Preenchido	100%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	O nome dado à coluna não é intuitivo devido ao facto de para além de países serem incluídas regiões e situações específicas.																																																						
Estatísticas dos Dados	<p>Column: metadata_Country_Name</p> <p>Simple Statistics</p> <table border="1" style="width: 100px; margin-left: auto; margin-right: auto;"> <tr><td>Label</td><td>Count</td><td>%</td></tr> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>265</td><td>0.09%</td></tr> <tr><td>Unique Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Duplicate Count</td><td>265</td><td>0.09%</td></tr> <tr><td>Blank Count</td><td>0</td><td>0.00%</td></tr> </table> <div style="text-align: center; margin-top: 10px;"> <p>Simple Statistics</p> </div> <p>Value Frequency</p> <table border="1" style="width: 100px; margin-left: auto; margin-right: auto;"> <tr><td>Value</td><td>Count</td><td>%</td></tr> <tr><td>Albania</td><td>1153</td><td>0.38%</td></tr> <tr><td>Africa Western and Central</td><td>1153</td><td>0.38%</td></tr> <tr><td>Algeria</td><td>1153</td><td>0.38%</td></tr> <tr><td>Arab World</td><td>1153</td><td>0.38%</td></tr> <tr><td>American Samoa</td><td>1153</td><td>0.38%</td></tr> <tr><td>Andorra</td><td>1153</td><td>0.38%</td></tr> <tr><td>Africa Eastern and Southern</td><td>1153</td><td>0.38%</td></tr> <tr><td>Angola</td><td>1153</td><td>0.38%</td></tr> <tr><td>Afghanistan</td><td>1153</td><td>0.38%</td></tr> <tr><td>Antigua and Barbuda</td><td>1153</td><td>0.38%</td></tr> </table> <div style="text-align: center; margin-top: 10px;"> <p>Value</p> <p>Count</p> </div>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	265	0.09%	Unique Count	0	0.00%	Duplicate Count	265	0.09%	Blank Count	0	0.00%	Value	Count	%	Albania	1153	0.38%	Africa Western and Central	1153	0.38%	Algeria	1153	0.38%	Arab World	1153	0.38%	American Samoa	1153	0.38%	Andorra	1153	0.38%	Africa Eastern and Southern	1153	0.38%	Angola	1153	0.38%	Afghanistan	1153	0.38%	Antigua and Barbuda	1153	0.38%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	265	0.09%																																																					
Unique Count	0	0.00%																																																					
Duplicate Count	265	0.09%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
Albania	1153	0.38%																																																					
Africa Western and Central	1153	0.38%																																																					
Algeria	1153	0.38%																																																					
Arab World	1153	0.38%																																																					
American Samoa	1153	0.38%																																																					
Andorra	1153	0.38%																																																					
Africa Eastern and Southern	1153	0.38%																																																					
Angola	1153	0.38%																																																					
Afghanistan	1153	0.38%																																																					
Antigua and Barbuda	1153	0.38%																																																					

Atributo	Country_Code
-----------------	--------------

Tipo de Dados	String																																																						
Descrição	Código do país/região/situação																																																						
Exemplo	AGO																																																						
Preenchido	100%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	O nome dado à coluna não é intuitivo devido ao facto de para além de países serem incluídas regiões e situações específicas.																																																						
Estatísticas dos Dados	<p>* Column: metadata.Country_Code</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>265</td> <td>0.09%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>265</td> <td>0.09%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>AGO</td> <td>1153</td> <td>0.38%</td> </tr> <tr> <td>AND</td> <td>1153</td> <td>0.38%</td> </tr> <tr> <td>ALB</td> <td>1153</td> <td>0.38%</td> </tr> <tr> <td>ARG</td> <td>1153</td> <td>0.38%</td> </tr> <tr> <td>ABW</td> <td>1153</td> <td>0.38%</td> </tr> <tr> <td>ARE</td> <td>1153</td> <td>0.38%</td> </tr> <tr> <td>AFW</td> <td>1153</td> <td>0.38%</td> </tr> <tr> <td>AFG</td> <td>1153</td> <td>0.38%</td> </tr> <tr> <td>ARB</td> <td>1153</td> <td>0.38%</td> </tr> <tr> <td>APE</td> <td>1153</td> <td>0.38%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	265	0.09%	Unique Count	0	0.00%	Duplicate Count	265	0.09%	Blank Count	0	0.00%	Value	Count	%	AGO	1153	0.38%	AND	1153	0.38%	ALB	1153	0.38%	ARG	1153	0.38%	ABW	1153	0.38%	ARE	1153	0.38%	AFW	1153	0.38%	AFG	1153	0.38%	ARB	1153	0.38%	APE	1153	0.38%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	265	0.09%																																																					
Unique Count	0	0.00%																																																					
Duplicate Count	265	0.09%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
AGO	1153	0.38%																																																					
AND	1153	0.38%																																																					
ALB	1153	0.38%																																																					
ARG	1153	0.38%																																																					
ABW	1153	0.38%																																																					
ARE	1153	0.38%																																																					
AFW	1153	0.38%																																																					
AFG	1153	0.38%																																																					
ARB	1153	0.38%																																																					
APE	1153	0.38%																																																					

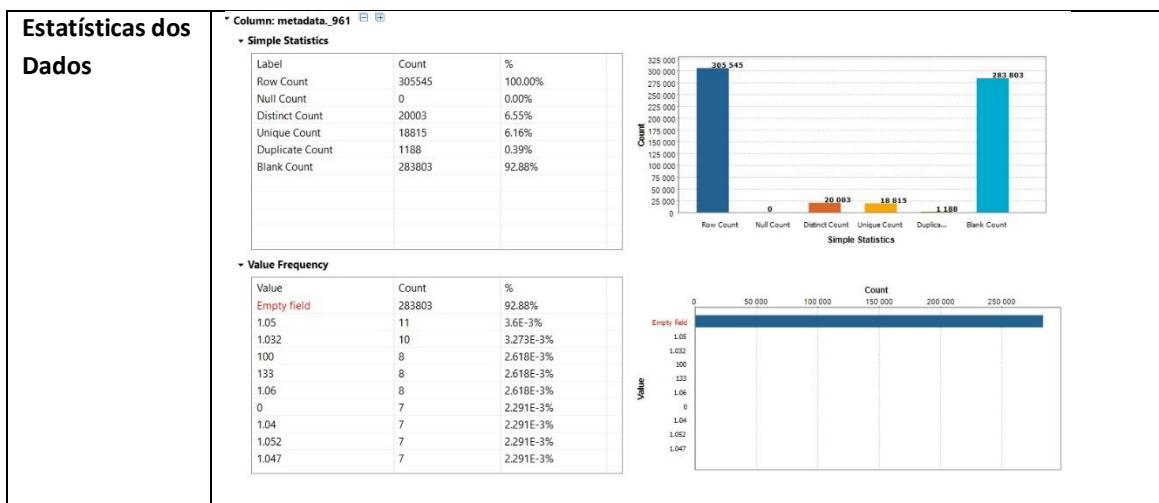
Atributo	Indicator_Name
Tipo de Dados	String
Descrição	Nome do indicador estatístico
Exemplo	A woman can apply for a passport in the same way as a man (1=yes; 0=no)
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados.

Estatísticas dos Dados	Column: metadata.Indicator_Name																																					
	Simple Statistics																																					
<table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>1153</td><td>0.38%</td></tr> <tr><td>Unique Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Duplicate Count</td><td>1153</td><td>0.38%</td></tr> <tr><td>Blank Count</td><td>0</td><td>0.00%</td></tr> </tbody> </table>			Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	1153	0.38%	Unique Count	0	0.00%	Duplicate Count	1153	0.38%	Blank Count	0	0.00%															
Label	Count	%																																				
Row Count	305545	100.00%																																				
Null Count	0	0.00%																																				
Distinct Count	1153	0.38%																																				
Unique Count	0	0.00%																																				
Duplicate Count	1153	0.38%																																				
Blank Count	0	0.00%																																				
Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>A woman can sign a contract in L...</td><td>265</td><td>0.09%</td></tr> <tr><td>A woman can open a bank accoun...</td><td>265</td><td>0.09%</td></tr> <tr><td>A woman can apply for a passpo...</td><td>265</td><td>0.09%</td></tr> <tr><td>A woman can travel outside her...</td><td>265</td><td>0.09%</td></tr> <tr><td>A woman can get a job in the sa...</td><td>265</td><td>0.09%</td></tr> <tr><td>A woman can choose where to live...</td><td>265</td><td>0.09%</td></tr> <tr><td>A woman can travel outside the c...</td><td>265</td><td>0.09%</td></tr> <tr><td>A woman can register a busines...</td><td>265</td><td>0.09%</td></tr> <tr><td>A woman can obtain a judgm...</td><td>265</td><td>0.09%</td></tr> <tr><td>A woman can be head of household ...</td><td>265</td><td>0.09%</td></tr> </tbody> </table>			Value	Count	%	A woman can sign a contract in L...	265	0.09%	A woman can open a bank accoun...	265	0.09%	A woman can apply for a passpo...	265	0.09%	A woman can travel outside her...	265	0.09%	A woman can get a job in the sa...	265	0.09%	A woman can choose where to live...	265	0.09%	A woman can travel outside the c...	265	0.09%	A woman can register a busines...	265	0.09%	A woman can obtain a judgm...	265	0.09%	A woman can be head of household ...	265	0.09%			
Value	Count	%																																				
A woman can sign a contract in L...	265	0.09%																																				
A woman can open a bank accoun...	265	0.09%																																				
A woman can apply for a passpo...	265	0.09%																																				
A woman can travel outside her...	265	0.09%																																				
A woman can get a job in the sa...	265	0.09%																																				
A woman can choose where to live...	265	0.09%																																				
A woman can travel outside the c...	265	0.09%																																				
A woman can register a busines...	265	0.09%																																				
A woman can obtain a judgm...	265	0.09%																																				
A woman can be head of household ...	265	0.09%																																				

Atributo	Indicator_Code																																			
Tipo de Dados	String																																			
Descrição	Código único para cada indicador																																			
Exemplo	SG.APL.PSPT.EQ																																			
Preenchido	100%																																			
Intervalo de Dados	Não mensurável																																			
Problemas de Qualidade dos Dados	Não detetados.																																			
Estatísticas dos Dados	Column: metadata.Indicator_Code																																			
Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>1153</td><td>0.38%</td></tr> <tr><td>Unique Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Duplicate Count</td><td>1153</td><td>0.38%</td></tr> <tr><td>Blank Count</td><td>0</td><td>0.00%</td></tr> </tbody> </table>		Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	1153	0.38%	Unique Count	0	0.00%	Duplicate Count	1153	0.38%	Blank Count	0	0.00%														
Label	Count	%																																		
Row Count	305545	100.00%																																		
Null Count	0	0.00%																																		
Distinct Count	1153	0.38%																																		
Unique Count	0	0.00%																																		
Duplicate Count	1153	0.38%																																		
Blank Count	0	0.00%																																		
Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>HD.HC1AMRT</td><td>265</td><td>0.09%</td></tr> <tr><td>HD.HC1EYRS.MA</td><td>265</td><td>0.09%</td></tr> <tr><td>HD.HC1AMRT.FE</td><td>265</td><td>0.09%</td></tr> <tr><td>HD.HC1HLOS.FE</td><td>265</td><td>0.09%</td></tr> <tr><td>HD.HC1HLOS</td><td>265</td><td>0.09%</td></tr> <tr><td>HD.HC1HLOS.MA</td><td>265</td><td>0.09%</td></tr> <tr><td>HD.HC1EYRS</td><td>265</td><td>0.09%</td></tr> <tr><td>FP.CPITOTLZG</td><td>265</td><td>0.09%</td></tr> <tr><td>HD.HC1EYRS.FE</td><td>265</td><td>0.09%</td></tr> <tr><td>HD.HC1AMRT.MA</td><td>265</td><td>0.09%</td></tr> </tbody> </table>		Value	Count	%	HD.HC1AMRT	265	0.09%	HD.HC1EYRS.MA	265	0.09%	HD.HC1AMRT.FE	265	0.09%	HD.HC1HLOS.FE	265	0.09%	HD.HC1HLOS	265	0.09%	HD.HC1HLOS.MA	265	0.09%	HD.HC1EYRS	265	0.09%	FP.CPITOTLZG	265	0.09%	HD.HC1EYRS.FE	265	0.09%	HD.HC1AMRT.MA	265	0.09%		
Value	Count	%																																		
HD.HC1AMRT	265	0.09%																																		
HD.HC1EYRS.MA	265	0.09%																																		
HD.HC1AMRT.FE	265	0.09%																																		
HD.HC1HLOS.FE	265	0.09%																																		
HD.HC1HLOS	265	0.09%																																		
HD.HC1HLOS.MA	265	0.09%																																		
HD.HC1EYRS	265	0.09%																																		
FP.CPITOTLZG	265	0.09%																																		
HD.HC1EYRS.FE	265	0.09%																																		
HD.HC1AMRT.MA	265	0.09%																																		

Atributo	1960																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1960																																																						
Exemplo	140.180525679236																																																						
Preenchido	7.13%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_960</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>20005</td> <td>6.55%</td> </tr> <tr> <td>Unique Count</td> <td>18782</td> <td>6.15%</td> </tr> <tr> <td>Duplicate Count</td> <td>1223</td> <td>0.40%</td> </tr> <tr> <td>Blank Count</td> <td>283751</td> <td>92.87%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>283751</td> <td>92.87%</td> </tr> <tr> <td>1.05</td> <td>11</td> <td>3.6E-3%</td> </tr> <tr> <td>88</td> <td>10</td> <td>3.273E-3%</td> </tr> <tr> <td>100</td> <td>10</td> <td>3.273E-3%</td> </tr> <tr> <td>1.032</td> <td>10</td> <td>3.273E-3%</td> </tr> <tr> <td>34</td> <td>9</td> <td>2.946E-3%</td> </tr> <tr> <td>1.06</td> <td>8</td> <td>2.618E-3%</td> </tr> <tr> <td>1.061</td> <td>8</td> <td>2.618E-3%</td> </tr> <tr> <td>1.047</td> <td>8</td> <td>2.618E-3%</td> </tr> <tr> <td>10</td> <td>8</td> <td>2.618E-3%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	20005	6.55%	Unique Count	18782	6.15%	Duplicate Count	1223	0.40%	Blank Count	283751	92.87%	Value	Count	%	Empty field	283751	92.87%	1.05	11	3.6E-3%	88	10	3.273E-3%	100	10	3.273E-3%	1.032	10	3.273E-3%	34	9	2.946E-3%	1.06	8	2.618E-3%	1.061	8	2.618E-3%	1.047	8	2.618E-3%	10	8	2.618E-3%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	20005	6.55%																																																					
Unique Count	18782	6.15%																																																					
Duplicate Count	1223	0.40%																																																					
Blank Count	283751	92.87%																																																					
Value	Count	%																																																					
Empty field	283751	92.87%																																																					
1.05	11	3.6E-3%																																																					
88	10	3.273E-3%																																																					
100	10	3.273E-3%																																																					
1.032	10	3.273E-3%																																																					
34	9	2.946E-3%																																																					
1.06	8	2.618E-3%																																																					
1.061	8	2.618E-3%																																																					
1.047	8	2.618E-3%																																																					
10	8	2.618E-3%																																																					

Atributo	1961
Tipo de Dados	String
Descrição	Contém dados para o ano 1961
Exemplo	140.810248007317
Preenchido	7.12%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null

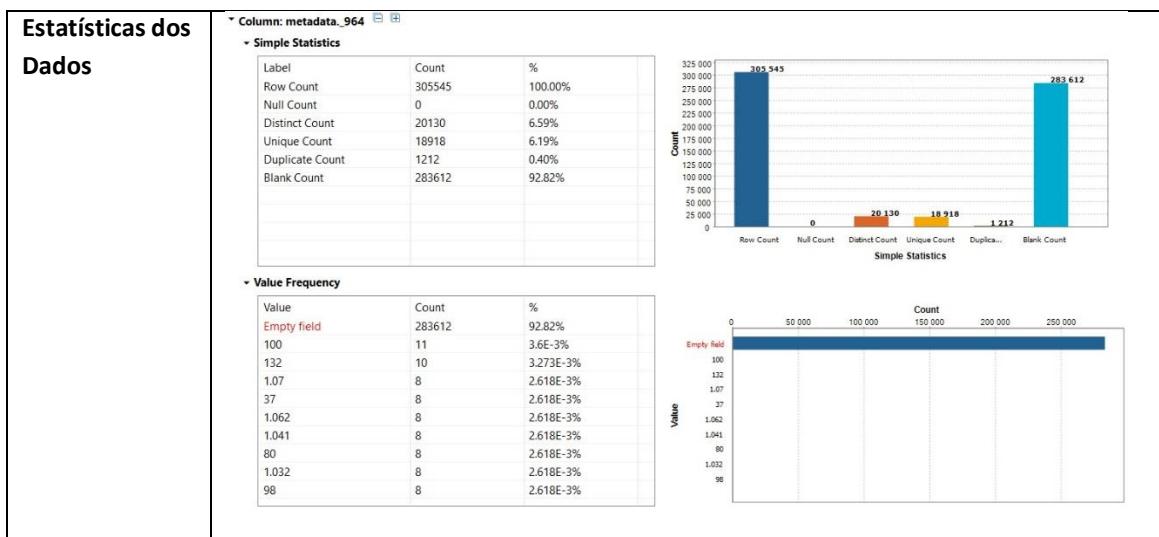


Atributo	1962																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1962																																																						
Exemplo	141.664168114272																																																						
Preenchido	7.28%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>* Column: metadata_962</p> <p>+ Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>20453</td><td>6.69%</td></tr> <tr><td>Unique Count</td><td>19283</td><td>6.31%</td></tr> <tr><td>Duplicate Count</td><td>1170</td><td>0.38%</td></tr> <tr><td>Blank Count</td><td>283312</td><td>92.72%</td></tr> </tbody> </table> <p>Simple Statistics</p> <p>+ Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>283312</td><td>92.72%</td></tr> <tr><td>100</td><td>12</td><td>3.927E-3%</td></tr> <tr><td>70</td><td>10</td><td>3.273E-3%</td></tr> <tr><td>1.06</td><td>9</td><td>2.946E-3%</td></tr> <tr><td>1.032</td><td>9</td><td>2.946E-3%</td></tr> <tr><td>44</td><td>8</td><td>2.618E-3%</td></tr> <tr><td>134</td><td>8</td><td>2.618E-3%</td></tr> <tr><td>1.05</td><td>8</td><td>2.618E-3%</td></tr> <tr><td>1.035</td><td>8</td><td>2.618E-3%</td></tr> <tr><td>120</td><td>8</td><td>2.618E-3%</td></tr> </tbody> </table> <p>Value Frequency</p>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	20453	6.69%	Unique Count	19283	6.31%	Duplicate Count	1170	0.38%	Blank Count	283312	92.72%	Value	Count	%	Empty field	283312	92.72%	100	12	3.927E-3%	70	10	3.273E-3%	1.06	9	2.946E-3%	1.032	9	2.946E-3%	44	8	2.618E-3%	134	8	2.618E-3%	1.05	8	2.618E-3%	1.035	8	2.618E-3%	120	8	2.618E-3%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	20453	6.69%																																																					
Unique Count	19283	6.31%																																																					
Duplicate Count	1170	0.38%																																																					
Blank Count	283312	92.72%																																																					
Value	Count	%																																																					
Empty field	283312	92.72%																																																					
100	12	3.927E-3%																																																					
70	10	3.273E-3%																																																					
1.06	9	2.946E-3%																																																					
1.032	9	2.946E-3%																																																					
44	8	2.618E-3%																																																					
134	8	2.618E-3%																																																					
1.05	8	2.618E-3%																																																					
1.035	8	2.618E-3%																																																					
120	8	2.618E-3%																																																					

Atributo	1963
Tipo de Dados	String
Descrição	Contém dados para o ano 1963

Exemplo	142.324951230209																																																						
Preenchido	7.17%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>* Column: metadata_963</p> <p>- Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>20108</td> <td>6.58%</td> </tr> <tr> <td>Unique Count</td> <td>18898</td> <td>6.19%</td> </tr> <tr> <td>Duplicate Count</td> <td>1210</td> <td>0.40%</td> </tr> <tr> <td>Blank Count</td> <td>283651</td> <td>92.83%</td> </tr> </tbody> </table> <p>Simple Statistics</p> <p>- Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>283651</td> <td>92.83%</td> </tr> <tr> <td>100</td> <td>15</td> <td>4.909E-3%</td> </tr> <tr> <td>1.032</td> <td>10</td> <td>3.273E-3%</td> </tr> <tr> <td>1.05</td> <td>9</td> <td>2.946E-3%</td> </tr> <tr> <td>1.061</td> <td>9</td> <td>2.946E-3%</td> </tr> <tr> <td>80</td> <td>9</td> <td>2.946E-3%</td> </tr> <tr> <td>110</td> <td>8</td> <td>2.618E-3%</td> </tr> <tr> <td>127</td> <td>8</td> <td>2.618E-3%</td> </tr> <tr> <td>180</td> <td>8</td> <td>2.618E-3%</td> </tr> <tr> <td>1.056</td> <td>8</td> <td>2.618E-3%</td> </tr> </tbody> </table> <p>Value Frequency</p>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	20108	6.58%	Unique Count	18898	6.19%	Duplicate Count	1210	0.40%	Blank Count	283651	92.83%	Value	Count	%	Empty field	283651	92.83%	100	15	4.909E-3%	1.032	10	3.273E-3%	1.05	9	2.946E-3%	1.061	9	2.946E-3%	80	9	2.946E-3%	110	8	2.618E-3%	127	8	2.618E-3%	180	8	2.618E-3%	1.056	8	2.618E-3%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	20108	6.58%																																																					
Unique Count	18898	6.19%																																																					
Duplicate Count	1210	0.40%																																																					
Blank Count	283651	92.83%																																																					
Value	Count	%																																																					
Empty field	283651	92.83%																																																					
100	15	4.909E-3%																																																					
1.032	10	3.273E-3%																																																					
1.05	9	2.946E-3%																																																					
1.061	9	2.946E-3%																																																					
80	9	2.946E-3%																																																					
110	8	2.618E-3%																																																					
127	8	2.618E-3%																																																					
180	8	2.618E-3%																																																					
1.056	8	2.618E-3%																																																					

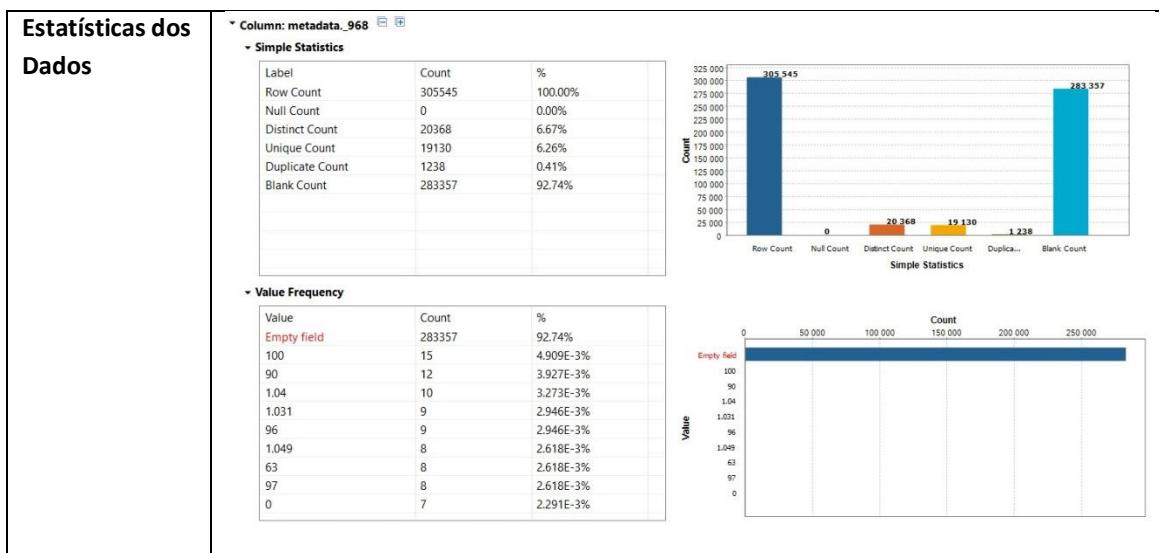
Atributo	1964
Tipo de Dados	String
Descrição	Contém dados para o ano 1964
Exemplo	143.456932955638
Preenchido	7.18%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



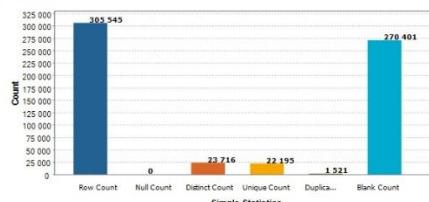
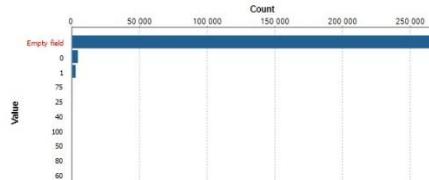
Atributo	1965																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1965																																																						
Exemplo	144.331149937632																																																						
Preenchido	7.2%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	Column: metadata_965 <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>20207</td><td>6.61%</td></tr> <tr><td>Unique Count</td><td>18983</td><td>6.21%</td></tr> <tr><td>Duplicate Count</td><td>1224</td><td>0.40%</td></tr> <tr><td>Blank Count</td><td>283531</td><td>92.80%</td></tr> </tbody> </table> <ul style="list-style-type: none"> - Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>283531</td><td>92.80%</td></tr> <tr><td>100</td><td>11</td><td>3.6E-3%</td></tr> <tr><td>140</td><td>11</td><td>3.6E-3%</td></tr> <tr><td>1.05</td><td>8</td><td>2.618E-3%</td></tr> <tr><td>96</td><td>8</td><td>2.618E-3%</td></tr> <tr><td>0</td><td>7</td><td>2.291E-3%</td></tr> <tr><td>1.042</td><td>7</td><td>2.291E-3%</td></tr> <tr><td>1.032</td><td>7</td><td>2.291E-3%</td></tr> <tr><td>1.056</td><td>7</td><td>2.291E-3%</td></tr> <tr><td>1.036</td><td>7</td><td>2.291E-3%</td></tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	20207	6.61%	Unique Count	18983	6.21%	Duplicate Count	1224	0.40%	Blank Count	283531	92.80%	Value	Count	%	Empty field	283531	92.80%	100	11	3.6E-3%	140	11	3.6E-3%	1.05	8	2.618E-3%	96	8	2.618E-3%	0	7	2.291E-3%	1.042	7	2.291E-3%	1.032	7	2.291E-3%	1.056	7	2.291E-3%	1.036	7	2.291E-3%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	20207	6.61%																																																					
Unique Count	18983	6.21%																																																					
Duplicate Count	1224	0.40%																																																					
Blank Count	283531	92.80%																																																					
Value	Count	%																																																					
Empty field	283531	92.80%																																																					
100	11	3.6E-3%																																																					
140	11	3.6E-3%																																																					
1.05	8	2.618E-3%																																																					
96	8	2.618E-3%																																																					
0	7	2.291E-3%																																																					
1.042	7	2.291E-3%																																																					
1.032	7	2.291E-3%																																																					
1.056	7	2.291E-3%																																																					
1.036	7	2.291E-3%																																																					

Atributo	1966																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1966																																																						
Exemplo	145.295946187855																																																						
Preenchido	7.26%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_966</p> <ul style="list-style-type: none"> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>20335</td> <td>6.66%</td> </tr> <tr> <td>Unique Count</td> <td>19089</td> <td>6.25%</td> </tr> <tr> <td>Duplicate Count</td> <td>1246</td> <td>0.41%</td> </tr> <tr> <td>Blank Count</td> <td>283377</td> <td>92.74%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>283377</td> <td>92.74%</td> </tr> <tr> <td>100</td> <td>11</td> <td>3.6E-3%</td> </tr> <tr> <td>1.049</td> <td>10</td> <td>3.273E-3%</td> </tr> <tr> <td>0</td> <td>9</td> <td>2.946E-3%</td> </tr> <tr> <td>88</td> <td>9</td> <td>2.946E-3%</td> </tr> <tr> <td>26</td> <td>9</td> <td>2.946E-3%</td> </tr> <tr> <td>110</td> <td>8</td> <td>2.618E-3%</td> </tr> <tr> <td>1.07</td> <td>8</td> <td>2.618E-3%</td> </tr> <tr> <td>1.04</td> <td>8</td> <td>2.618E-3%</td> </tr> <tr> <td>1.062</td> <td>8</td> <td>2.618E-3%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	20335	6.66%	Unique Count	19089	6.25%	Duplicate Count	1246	0.41%	Blank Count	283377	92.74%	Value	Count	%	Empty field	283377	92.74%	100	11	3.6E-3%	1.049	10	3.273E-3%	0	9	2.946E-3%	88	9	2.946E-3%	26	9	2.946E-3%	110	8	2.618E-3%	1.07	8	2.618E-3%	1.04	8	2.618E-3%	1.062	8	2.618E-3%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	20335	6.66%																																																					
Unique Count	19089	6.25%																																																					
Duplicate Count	1246	0.41%																																																					
Blank Count	283377	92.74%																																																					
Value	Count	%																																																					
Empty field	283377	92.74%																																																					
100	11	3.6E-3%																																																					
1.049	10	3.273E-3%																																																					
0	9	2.946E-3%																																																					
88	9	2.946E-3%																																																					
26	9	2.946E-3%																																																					
110	8	2.618E-3%																																																					
1.07	8	2.618E-3%																																																					
1.04	8	2.618E-3%																																																					
1.062	8	2.618E-3%																																																					

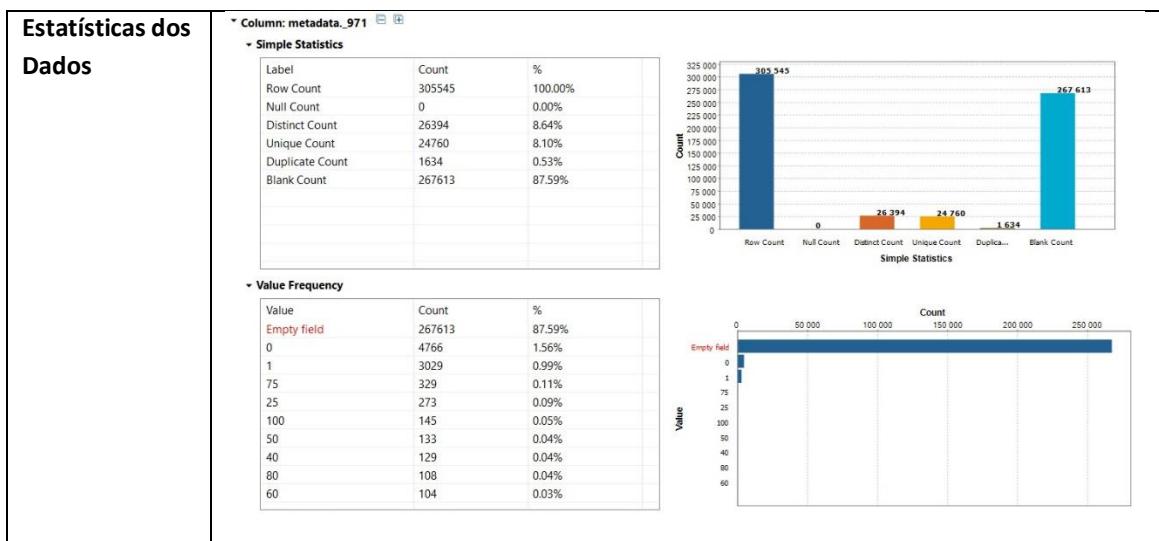
Atributo	1968
Tipo de Dados	String
Descrição	Contém dados para o ano 1968
Exemplo	145.748064357066
Preenchido	7.26%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	1969																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1969																																																						
Exemplo	148.012004326763																																																						
Preenchido	7.29%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_969</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>20449</td><td>6.69%</td></tr> <tr><td>Unique Count</td><td>19190</td><td>6.28%</td></tr> <tr><td>Duplicate Count</td><td>1259</td><td>0.41%</td></tr> <tr><td>Blank Count</td><td>283268</td><td>92.71%</td></tr> </tbody> </table> <p>Count</p> <p>Simple Statistics</p> <p>Row Count Null Count Distinct Count Unique Count Duplicate... Blank Count</p> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>283268</td><td>92.71%</td></tr> <tr><td>100</td><td>13</td><td>4.255E-3%</td></tr> <tr><td>1.049</td><td>10</td><td>3.273E-3%</td></tr> <tr><td>0</td><td>8</td><td>2.618E-3%</td></tr> <tr><td>144</td><td>8</td><td>2.618E-3%</td></tr> <tr><td>1.05</td><td>8</td><td>2.618E-3%</td></tr> <tr><td>1.04</td><td>8</td><td>2.618E-3%</td></tr> <tr><td>94</td><td>8</td><td>2.618E-3%</td></tr> <tr><td>1.036</td><td>8</td><td>2.618E-3%</td></tr> <tr><td>1.031</td><td>7</td><td>2.291E-3%</td></tr> </tbody> </table> <p>Value</p> <p>Count</p> <p>Empty field 100 1.049 0 144 1.05 1.04 94 1.036 1.031</p>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	20449	6.69%	Unique Count	19190	6.28%	Duplicate Count	1259	0.41%	Blank Count	283268	92.71%	Value	Count	%	Empty field	283268	92.71%	100	13	4.255E-3%	1.049	10	3.273E-3%	0	8	2.618E-3%	144	8	2.618E-3%	1.05	8	2.618E-3%	1.04	8	2.618E-3%	94	8	2.618E-3%	1.036	8	2.618E-3%	1.031	7	2.291E-3%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	20449	6.69%																																																					
Unique Count	19190	6.28%																																																					
Duplicate Count	1259	0.41%																																																					
Blank Count	283268	92.71%																																																					
Value	Count	%																																																					
Empty field	283268	92.71%																																																					
100	13	4.255E-3%																																																					
1.049	10	3.273E-3%																																																					
0	8	2.618E-3%																																																					
144	8	2.618E-3%																																																					
1.05	8	2.618E-3%																																																					
1.04	8	2.618E-3%																																																					
94	8	2.618E-3%																																																					
1.036	8	2.618E-3%																																																					
1.031	7	2.291E-3%																																																					

Atributo	1970																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1970																																																						
Exemplo	149.579704242624																																																						
Preenchido	11.5%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata._970</p> <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>23716</td> <td>7.76%</td> </tr> <tr> <td>Unique Count</td> <td>22195</td> <td>7.26%</td> </tr> <tr> <td>Duplicate Count</td> <td>1521</td> <td>0.50%</td> </tr> <tr> <td>Blank Count</td> <td>270401</td> <td>88.50%</td> </tr> </tbody> </table>  <ul style="list-style-type: none"> - Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>270401</td> <td>88.50%</td> </tr> <tr> <td>0</td> <td>4712</td> <td>1.54%</td> </tr> <tr> <td>1</td> <td>3019</td> <td>0.99%</td> </tr> <tr> <td>75</td> <td>325</td> <td>0.11%</td> </tr> <tr> <td>25</td> <td>287</td> <td>0.09%</td> </tr> <tr> <td>40</td> <td>136</td> <td>0.04%</td> </tr> <tr> <td>100</td> <td>135</td> <td>0.04%</td> </tr> <tr> <td>50</td> <td>131</td> <td>0.04%</td> </tr> <tr> <td>80</td> <td>111</td> <td>0.04%</td> </tr> <tr> <td>60</td> <td>96</td> <td>0.03%</td> </tr> </tbody> </table> 	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	23716	7.76%	Unique Count	22195	7.26%	Duplicate Count	1521	0.50%	Blank Count	270401	88.50%	Value	Count	%	Empty field	270401	88.50%	0	4712	1.54%	1	3019	0.99%	75	325	0.11%	25	287	0.09%	40	136	0.04%	100	135	0.04%	50	131	0.04%	80	111	0.04%	60	96	0.03%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	23716	7.76%																																																					
Unique Count	22195	7.26%																																																					
Duplicate Count	1521	0.50%																																																					
Blank Count	270401	88.50%																																																					
Value	Count	%																																																					
Empty field	270401	88.50%																																																					
0	4712	1.54%																																																					
1	3019	0.99%																																																					
75	325	0.11%																																																					
25	287	0.09%																																																					
40	136	0.04%																																																					
100	135	0.04%																																																					
50	131	0.04%																																																					
80	111	0.04%																																																					
60	96	0.03%																																																					

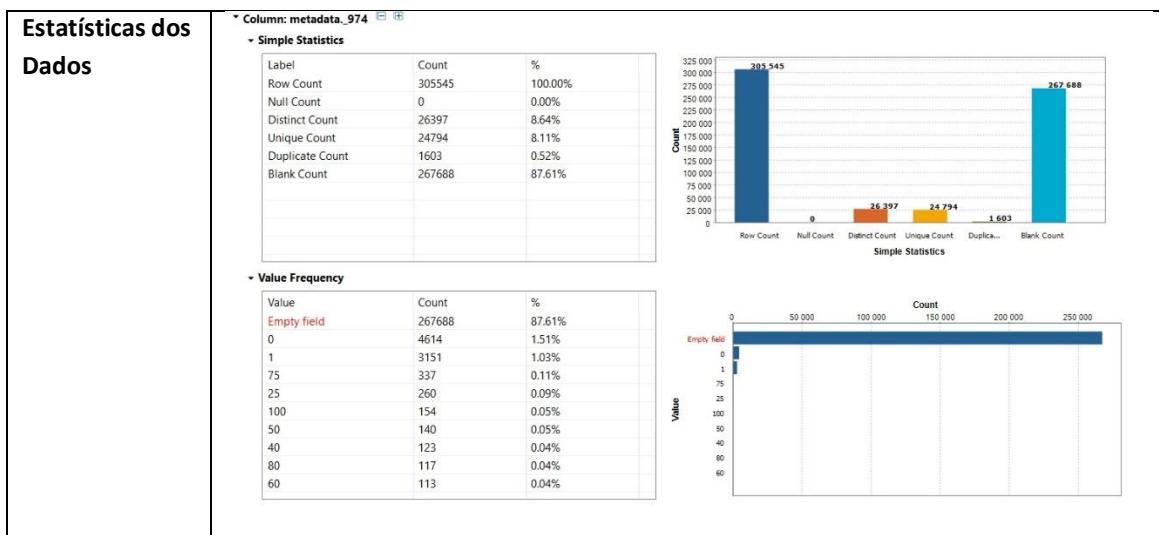
Atributo	1971
Tipo de Dados	String
Descrição	Contém dados para o ano 1971
Exemplo	152.846769500502
Preenchido	12.41%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	1972																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1972																																																						
Exemplo	153.951116872101																																																						
Preenchido	12.48%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>* Column: metadata_972</p> <ul style="list-style-type: none"> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>26699</td> <td>8.74%</td> </tr> <tr> <td>Unique Count</td> <td>25117</td> <td>8.22%</td> </tr> <tr> <td>Duplicate Count</td> <td>1582</td> <td>0.52%</td> </tr> <tr> <td>Blank Count</td> <td>267398</td> <td>87.52%</td> </tr> </tbody> </table> <p>Count</p> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>267398</td> <td>87.52%</td> </tr> <tr> <td>0</td> <td>4717</td> <td>1.54%</td> </tr> <tr> <td>1</td> <td>3058</td> <td>1.00%</td> </tr> <tr> <td>75</td> <td>328</td> <td>0.11%</td> </tr> <tr> <td>25</td> <td>261</td> <td>0.09%</td> </tr> <tr> <td>100</td> <td>146</td> <td>0.05%</td> </tr> <tr> <td>50</td> <td>142</td> <td>0.05%</td> </tr> <tr> <td>40</td> <td>126</td> <td>0.04%</td> </tr> <tr> <td>80</td> <td>110</td> <td>0.04%</td> </tr> <tr> <td>60</td> <td>110</td> <td>0.04%</td> </tr> </tbody> </table> <p>Count</p>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	26699	8.74%	Unique Count	25117	8.22%	Duplicate Count	1582	0.52%	Blank Count	267398	87.52%	Value	Count	%	Empty field	267398	87.52%	0	4717	1.54%	1	3058	1.00%	75	328	0.11%	25	261	0.09%	100	146	0.05%	50	142	0.05%	40	126	0.04%	80	110	0.04%	60	110	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	26699	8.74%																																																					
Unique Count	25117	8.22%																																																					
Duplicate Count	1582	0.52%																																																					
Blank Count	267398	87.52%																																																					
Value	Count	%																																																					
Empty field	267398	87.52%																																																					
0	4717	1.54%																																																					
1	3058	1.00%																																																					
75	328	0.11%																																																					
25	261	0.09%																																																					
100	146	0.05%																																																					
50	142	0.05%																																																					
40	126	0.04%																																																					
80	110	0.04%																																																					
60	110	0.04%																																																					

Atributo	1973																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1973																																																						
Exemplo	152.57050224758																																																						
Preenchido	12.37%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_973</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>26331</td> <td>8.62%</td> </tr> <tr> <td>Unique Count</td> <td>24745</td> <td>8.10%</td> </tr> <tr> <td>Duplicate Count</td> <td>1586</td> <td>0.52%</td> </tr> <tr> <td>Blank Count</td> <td>267749</td> <td>87.63%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>267749</td> <td>87.63%</td> </tr> <tr> <td>0</td> <td>4640</td> <td>1.52%</td> </tr> <tr> <td>1</td> <td>3120</td> <td>1.02%</td> </tr> <tr> <td>75</td> <td>329</td> <td>0.11%</td> </tr> <tr> <td>25</td> <td>264</td> <td>0.09%</td> </tr> <tr> <td>100</td> <td>156</td> <td>0.05%</td> </tr> <tr> <td>50</td> <td>143</td> <td>0.05%</td> </tr> <tr> <td>40</td> <td>118</td> <td>0.04%</td> </tr> <tr> <td>80</td> <td>117</td> <td>0.04%</td> </tr> <tr> <td>60</td> <td>104</td> <td>0.03%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	26331	8.62%	Unique Count	24745	8.10%	Duplicate Count	1586	0.52%	Blank Count	267749	87.63%	Value	Count	%	Empty field	267749	87.63%	0	4640	1.52%	1	3120	1.02%	75	329	0.11%	25	264	0.09%	100	156	0.05%	50	143	0.05%	40	118	0.04%	80	117	0.04%	60	104	0.03%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	26331	8.62%																																																					
Unique Count	24745	8.10%																																																					
Duplicate Count	1586	0.52%																																																					
Blank Count	267749	87.63%																																																					
Value	Count	%																																																					
Empty field	267749	87.63%																																																					
0	4640	1.52%																																																					
1	3120	1.02%																																																					
75	329	0.11%																																																					
25	264	0.09%																																																					
100	156	0.05%																																																					
50	143	0.05%																																																					
40	118	0.04%																																																					
80	117	0.04%																																																					
60	104	0.03%																																																					

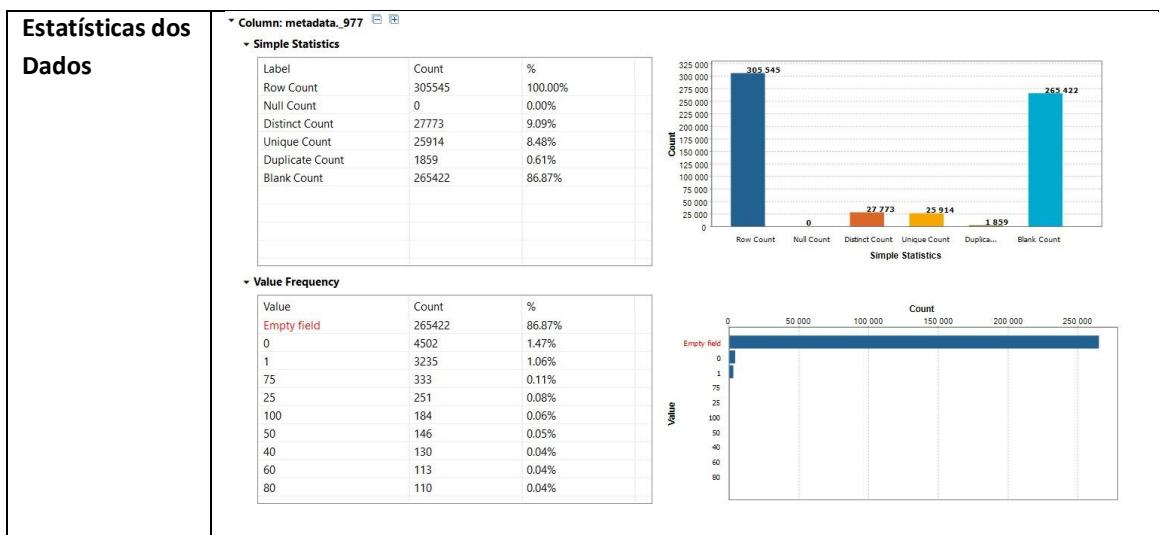
Atributo	1974
Tipo de Dados	String
Descrição	Contém dados para o ano 1974
Exemplo	152.504672544408
Preenchido	12.39%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	1975																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1975																																																						
Exemplo	151.86791415169																																																						
Preenchido	12.86%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	Column: metadata_975 <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>26929</td><td>8.81%</td></tr> <tr><td>Unique Count</td><td>25085</td><td>8.21%</td></tr> <tr><td>Duplicate Count</td><td>1844</td><td>0.60%</td></tr> <tr><td>Blank Count</td><td>266254</td><td>87.14%</td></tr> </tbody> </table> <ul style="list-style-type: none"> - Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>266254</td><td>87.14%</td></tr> <tr><td>0</td><td>4579</td><td>1.50%</td></tr> <tr><td>1</td><td>3184</td><td>1.04%</td></tr> <tr><td>75</td><td>339</td><td>0.11%</td></tr> <tr><td>25</td><td>257</td><td>0.08%</td></tr> <tr><td>100</td><td>167</td><td>0.05%</td></tr> <tr><td>50</td><td>135</td><td>0.04%</td></tr> <tr><td>40</td><td>129</td><td>0.04%</td></tr> <tr><td>80</td><td>115</td><td>0.04%</td></tr> <tr><td>60</td><td>112</td><td>0.04%</td></tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	26929	8.81%	Unique Count	25085	8.21%	Duplicate Count	1844	0.60%	Blank Count	266254	87.14%	Value	Count	%	Empty field	266254	87.14%	0	4579	1.50%	1	3184	1.04%	75	339	0.11%	25	257	0.08%	100	167	0.05%	50	135	0.04%	40	129	0.04%	80	115	0.04%	60	112	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	26929	8.81%																																																					
Unique Count	25085	8.21%																																																					
Duplicate Count	1844	0.60%																																																					
Blank Count	266254	87.14%																																																					
Value	Count	%																																																					
Empty field	266254	87.14%																																																					
0	4579	1.50%																																																					
1	3184	1.04%																																																					
75	339	0.11%																																																					
25	257	0.08%																																																					
100	167	0.05%																																																					
50	135	0.04%																																																					
40	129	0.04%																																																					
80	115	0.04%																																																					
60	112	0.04%																																																					

Atributo	1976																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1976																																																						
Exemplo	150.533608687468																																																						
Preenchido	13.03%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_976</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>27435</td> <td>8.98%</td> </tr> <tr> <td>Unique Count</td> <td>25610</td> <td>8.38%</td> </tr> <tr> <td>Duplicate Count</td> <td>1825</td> <td>0.60%</td> </tr> <tr> <td>Blank Count</td> <td>265739</td> <td>86.97%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>265739</td> <td>86.97%</td> </tr> <tr> <td>0</td> <td>4550</td> <td>1.49%</td> </tr> <tr> <td>1</td> <td>3204</td> <td>1.05%</td> </tr> <tr> <td>75</td> <td>335</td> <td>0.11%</td> </tr> <tr> <td>25</td> <td>254</td> <td>0.08%</td> </tr> <tr> <td>100</td> <td>170</td> <td>0.06%</td> </tr> <tr> <td>50</td> <td>137</td> <td>0.04%</td> </tr> <tr> <td>40</td> <td>130</td> <td>0.04%</td> </tr> <tr> <td>60</td> <td>120</td> <td>0.04%</td> </tr> <tr> <td>80</td> <td>112</td> <td>0.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	27435	8.98%	Unique Count	25610	8.38%	Duplicate Count	1825	0.60%	Blank Count	265739	86.97%	Value	Count	%	Empty field	265739	86.97%	0	4550	1.49%	1	3204	1.05%	75	335	0.11%	25	254	0.08%	100	170	0.06%	50	137	0.04%	40	130	0.04%	60	120	0.04%	80	112	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	27435	8.98%																																																					
Unique Count	25610	8.38%																																																					
Duplicate Count	1825	0.60%																																																					
Blank Count	265739	86.97%																																																					
Value	Count	%																																																					
Empty field	265739	86.97%																																																					
0	4550	1.49%																																																					
1	3204	1.05%																																																					
75	335	0.11%																																																					
25	254	0.08%																																																					
100	170	0.06%																																																					
50	137	0.04%																																																					
40	130	0.04%																																																					
60	120	0.04%																																																					
80	112	0.04%																																																					

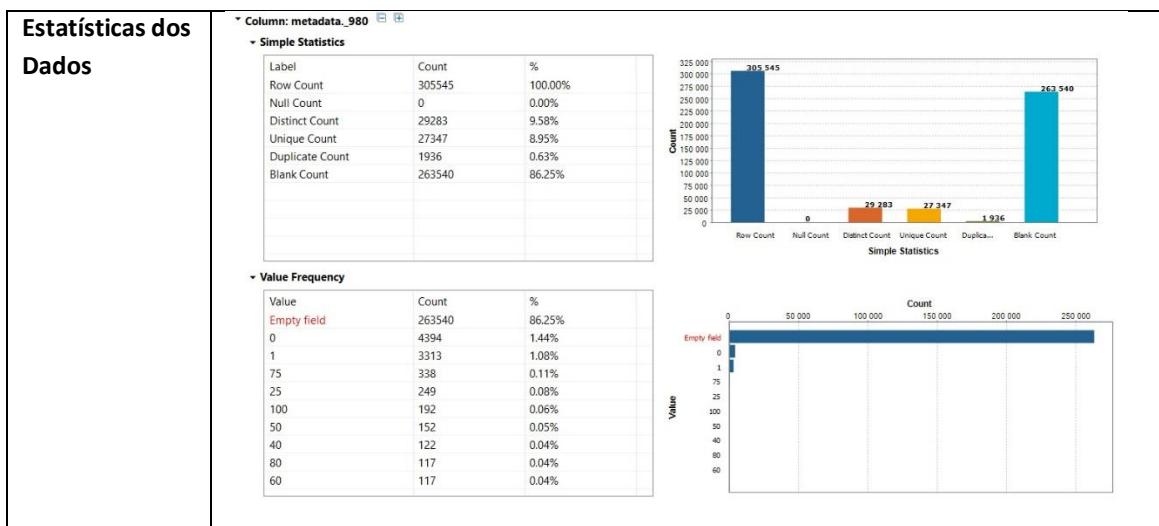
Atributo	1977
Tipo de Dados	String
Descrição	Contém dados para o ano 1977
Exemplo	148.93109556429
Preenchido	13.13%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	1978																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1978																																																						
Exemplo	146.644103471495																																																						
Preenchido	13.05%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	* Column: metadata_978 <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>27438</td><td>8.98%</td></tr> <tr><td>Unique Count</td><td>25542</td><td>8.36%</td></tr> <tr><td>Duplicate Count</td><td>1896</td><td>0.62%</td></tr> <tr><td>Blank Count</td><td>265673</td><td>86.95%</td></tr> </tbody> </table> <ul style="list-style-type: none"> - Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>265673</td><td>86.95%</td></tr> <tr><td>0</td><td>4463</td><td>1.46%</td></tr> <tr><td>1</td><td>3273</td><td>1.07%</td></tr> <tr><td>75</td><td>339</td><td>0.11%</td></tr> <tr><td>25</td><td>249</td><td>0.08%</td></tr> <tr><td>100</td><td>188</td><td>0.06%</td></tr> <tr><td>50</td><td>143</td><td>0.05%</td></tr> <tr><td>40</td><td>124</td><td>0.04%</td></tr> <tr><td>60</td><td>116</td><td>0.04%</td></tr> <tr><td>80</td><td>113</td><td>0.04%</td></tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	27438	8.98%	Unique Count	25542	8.36%	Duplicate Count	1896	0.62%	Blank Count	265673	86.95%	Value	Count	%	Empty field	265673	86.95%	0	4463	1.46%	1	3273	1.07%	75	339	0.11%	25	249	0.08%	100	188	0.06%	50	143	0.05%	40	124	0.04%	60	116	0.04%	80	113	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	27438	8.98%																																																					
Unique Count	25542	8.36%																																																					
Duplicate Count	1896	0.62%																																																					
Blank Count	265673	86.95%																																																					
Value	Count	%																																																					
Empty field	265673	86.95%																																																					
0	4463	1.46%																																																					
1	3273	1.07%																																																					
75	339	0.11%																																																					
25	249	0.08%																																																					
100	188	0.06%																																																					
50	143	0.05%																																																					
40	124	0.04%																																																					
60	116	0.04%																																																					
80	113	0.04%																																																					

Atributo	1979																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1979																																																						
Exemplo	145.484720348173																																																						
Preenchido	13.13%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_979</p> <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>27717</td> <td>9.07%</td> </tr> <tr> <td>Unique Count</td> <td>25802</td> <td>8.44%</td> </tr> <tr> <td>Duplicate Count</td> <td>1915</td> <td>0.63%</td> </tr> <tr> <td>Blank Count</td> <td>265424</td> <td>86.87%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>265424</td> <td>86.87%</td> </tr> <tr> <td>0</td> <td>4419</td> <td>1.45%</td> </tr> <tr> <td>1</td> <td>3291</td> <td>1.08%</td> </tr> <tr> <td>75</td> <td>335</td> <td>0.11%</td> </tr> <tr> <td>25</td> <td>244</td> <td>0.08%</td> </tr> <tr> <td>100</td> <td>190</td> <td>0.06%</td> </tr> <tr> <td>50</td> <td>146</td> <td>0.05%</td> </tr> <tr> <td>40</td> <td>124</td> <td>0.04%</td> </tr> <tr> <td>60</td> <td>115</td> <td>0.04%</td> </tr> <tr> <td>80</td> <td>112</td> <td>0.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	27717	9.07%	Unique Count	25802	8.44%	Duplicate Count	1915	0.63%	Blank Count	265424	86.87%	Value	Count	%	Empty field	265424	86.87%	0	4419	1.45%	1	3291	1.08%	75	335	0.11%	25	244	0.08%	100	190	0.06%	50	146	0.05%	40	124	0.04%	60	115	0.04%	80	112	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	27717	9.07%																																																					
Unique Count	25802	8.44%																																																					
Duplicate Count	1915	0.63%																																																					
Blank Count	265424	86.87%																																																					
Value	Count	%																																																					
Empty field	265424	86.87%																																																					
0	4419	1.45%																																																					
1	3291	1.08%																																																					
75	335	0.11%																																																					
25	244	0.08%																																																					
100	190	0.06%																																																					
50	146	0.05%																																																					
40	124	0.04%																																																					
60	115	0.04%																																																					
80	112	0.04%																																																					

Atributo	1980
Tipo de Dados	String
Descrição	Contém dados para o ano 1980
Exemplo	143.75343754838
Preenchido	13.75%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	1981																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1981																																																						
Exemplo	142.558131825617																																																						
Preenchido	13.53%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_981</p> <ul style="list-style-type: none"> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>28507</td><td>9.33%</td></tr> <tr><td>Unique Count</td><td>26546</td><td>8.69%</td></tr> <tr><td>Duplicate Count</td><td>1961</td><td>0.64%</td></tr> <tr><td>Blank Count</td><td>264194</td><td>86.47%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>264194</td><td>86.47%</td></tr> <tr><td>0</td><td>4360</td><td>1.43%</td></tr> <tr><td>1</td><td>3352</td><td>1.10%</td></tr> <tr><td>75</td><td>339</td><td>0.11%</td></tr> <tr><td>25</td><td>245</td><td>0.08%</td></tr> <tr><td>100</td><td>204</td><td>0.07%</td></tr> <tr><td>50</td><td>151</td><td>0.05%</td></tr> <tr><td>40</td><td>127</td><td>0.04%</td></tr> <tr><td>60</td><td>117</td><td>0.04%</td></tr> <tr><td>80</td><td>112</td><td>0.04%</td></tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	28507	9.33%	Unique Count	26546	8.69%	Duplicate Count	1961	0.64%	Blank Count	264194	86.47%	Value	Count	%	Empty field	264194	86.47%	0	4360	1.43%	1	3352	1.10%	75	339	0.11%	25	245	0.08%	100	204	0.07%	50	151	0.05%	40	127	0.04%	60	117	0.04%	80	112	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	28507	9.33%																																																					
Unique Count	26546	8.69%																																																					
Duplicate Count	1961	0.64%																																																					
Blank Count	264194	86.47%																																																					
Value	Count	%																																																					
Empty field	264194	86.47%																																																					
0	4360	1.43%																																																					
1	3352	1.10%																																																					
75	339	0.11%																																																					
25	245	0.08%																																																					
100	204	0.07%																																																					
50	151	0.05%																																																					
40	127	0.04%																																																					
60	117	0.04%																																																					
80	112	0.04%																																																					

Atributo	1982																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1982																																																						
Exemplo	141.191950820555																																																						
Preenchido	13.52%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_982</p> <ul style="list-style-type: none"> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>28515</td> <td>9.33%</td> </tr> <tr> <td>Unique Count</td> <td>26559</td> <td>8.69%</td> </tr> <tr> <td>Duplicate Count</td> <td>1956</td> <td>0.64%</td> </tr> <tr> <td>Blank Count</td> <td>264243</td> <td>86.48%</td> </tr> </tbody> </table> <ul style="list-style-type: none"> Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>264243</td> <td>86.48%</td> </tr> <tr> <td>0</td> <td>4325</td> <td>1.42%</td> </tr> <tr> <td>1</td> <td>3371</td> <td>1.10%</td> </tr> <tr> <td>75</td> <td>341</td> <td>0.11%</td> </tr> <tr> <td>25</td> <td>245</td> <td>0.08%</td> </tr> <tr> <td>100</td> <td>211</td> <td>0.07%</td> </tr> <tr> <td>50</td> <td>150</td> <td>0.05%</td> </tr> <tr> <td>40</td> <td>129</td> <td>0.04%</td> </tr> <tr> <td>60</td> <td>118</td> <td>0.04%</td> </tr> <tr> <td>80</td> <td>111</td> <td>0.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	28515	9.33%	Unique Count	26559	8.69%	Duplicate Count	1956	0.64%	Blank Count	264243	86.48%	Value	Count	%	Empty field	264243	86.48%	0	4325	1.42%	1	3371	1.10%	75	341	0.11%	25	245	0.08%	100	211	0.07%	50	150	0.05%	40	129	0.04%	60	118	0.04%	80	111	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	28515	9.33%																																																					
Unique Count	26559	8.69%																																																					
Duplicate Count	1956	0.64%																																																					
Blank Count	264243	86.48%																																																					
Value	Count	%																																																					
Empty field	264243	86.48%																																																					
0	4325	1.42%																																																					
1	3371	1.10%																																																					
75	341	0.11%																																																					
25	245	0.08%																																																					
100	211	0.07%																																																					
50	150	0.05%																																																					
40	129	0.04%																																																					
60	118	0.04%																																																					
80	111	0.04%																																																					

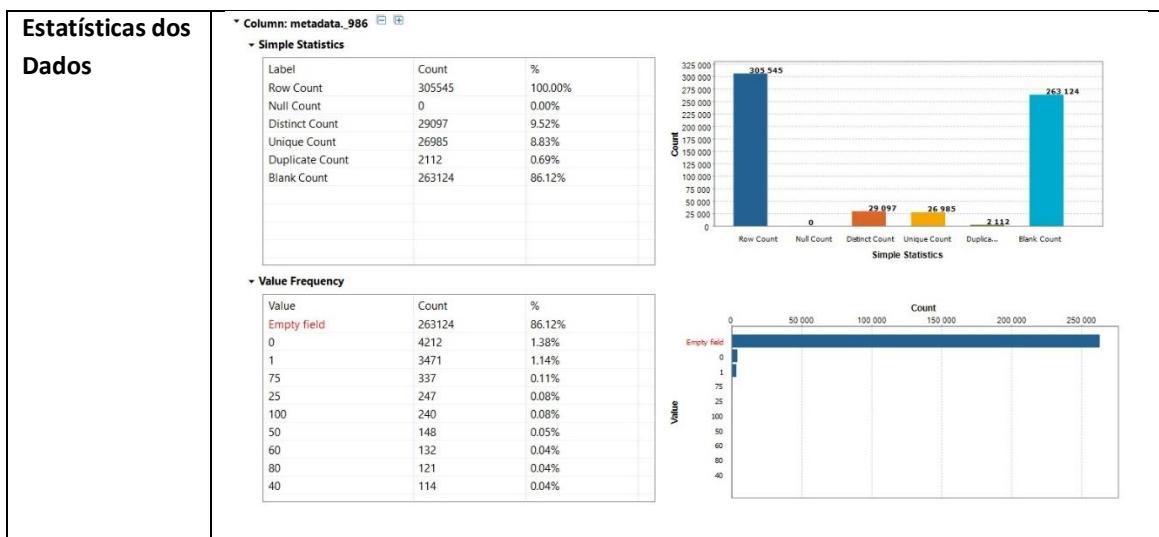
Atributo	1983
Tipo de Dados	String
Descrição	Contém dados para o ano 1983
Exemplo	139.763806631955
Preenchido	13.54%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null

Estatísticas dos Dados	Column: metadata_983		
	Simple Statistics		Value Frequency
	Label	Count	%
	Row Count	305545	100.00%
	Null Count	0	0.00%
	Distinct Count	28507	9.33%
	Unique Count	26514	8.68%
	Duplicate Count	1993	0.65%
	Blank Count	264170	86.46%

Atributo	1984																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1984																																																						
Exemplo	138.785843258905																																																						
Preenchido	13.58%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_984</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>28554</td> <td>9.35%</td> </tr> <tr> <td>Unique Count</td> <td>26544</td> <td>8.69%</td> </tr> <tr> <td>Duplicate Count</td> <td>2010</td> <td>0.66%</td> </tr> <tr> <td>Blank Count</td> <td>264043</td> <td>86.42%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>264043</td> <td>86.42%</td> </tr> <tr> <td>0</td> <td>4257</td> <td>1.39%</td> </tr> <tr> <td>1</td> <td>3423</td> <td>1.12%</td> </tr> <tr> <td>75</td> <td>347</td> <td>0.11%</td> </tr> <tr> <td>25</td> <td>238</td> <td>0.08%</td> </tr> <tr> <td>100</td> <td>217</td> <td>0.07%</td> </tr> <tr> <td>50</td> <td>153</td> <td>0.05%</td> </tr> <tr> <td>60</td> <td>126</td> <td>0.04%</td> </tr> <tr> <td>40</td> <td>118</td> <td>0.04%</td> </tr> <tr> <td>80</td> <td>114</td> <td>0.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	28554	9.35%	Unique Count	26544	8.69%	Duplicate Count	2010	0.66%	Blank Count	264043	86.42%	Value	Count	%	Empty field	264043	86.42%	0	4257	1.39%	1	3423	1.12%	75	347	0.11%	25	238	0.08%	100	217	0.07%	50	153	0.05%	60	126	0.04%	40	118	0.04%	80	114	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	28554	9.35%																																																					
Unique Count	26544	8.69%																																																					
Duplicate Count	2010	0.66%																																																					
Blank Count	264043	86.42%																																																					
Value	Count	%																																																					
Empty field	264043	86.42%																																																					
0	4257	1.39%																																																					
1	3423	1.12%																																																					
75	347	0.11%																																																					
25	238	0.08%																																																					
100	217	0.07%																																																					
50	153	0.05%																																																					
60	126	0.04%																																																					
40	118	0.04%																																																					
80	114	0.04%																																																					

Atributo	1985																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1985																																																						
Exemplo	137.953223122652																																																						
Preenchido	14.01%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_985</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>29620</td> <td>9.69%</td> </tr> <tr> <td>Unique Count</td> <td>27535</td> <td>9.01%</td> </tr> <tr> <td>Duplicate Count</td> <td>2085</td> <td>0.68%</td> </tr> <tr> <td>Blank Count</td> <td>262741</td> <td>85.99%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>262741</td> <td>85.99%</td> </tr> <tr> <td>0</td> <td>4238</td> <td>1.39%</td> </tr> <tr> <td>1</td> <td>3449</td> <td>1.13%</td> </tr> <tr> <td>75</td> <td>338</td> <td>0.11%</td> </tr> <tr> <td>25</td> <td>244</td> <td>0.08%</td> </tr> <tr> <td>100</td> <td>227</td> <td>0.07%</td> </tr> <tr> <td>50</td> <td>151</td> <td>0.05%</td> </tr> <tr> <td>60</td> <td>126</td> <td>0.04%</td> </tr> <tr> <td>80</td> <td>115</td> <td>0.04%</td> </tr> <tr> <td>40</td> <td>113</td> <td>0.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	29620	9.69%	Unique Count	27535	9.01%	Duplicate Count	2085	0.68%	Blank Count	262741	85.99%	Value	Count	%	Empty field	262741	85.99%	0	4238	1.39%	1	3449	1.13%	75	338	0.11%	25	244	0.08%	100	227	0.07%	50	151	0.05%	60	126	0.04%	80	115	0.04%	40	113	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	29620	9.69%																																																					
Unique Count	27535	9.01%																																																					
Duplicate Count	2085	0.68%																																																					
Blank Count	262741	85.99%																																																					
Value	Count	%																																																					
Empty field	262741	85.99%																																																					
0	4238	1.39%																																																					
1	3449	1.13%																																																					
75	338	0.11%																																																					
25	244	0.08%																																																					
100	227	0.07%																																																					
50	151	0.05%																																																					
60	126	0.04%																																																					
80	115	0.04%																																																					
40	113	0.04%																																																					

Atributo	1986
Tipo de Dados	String
Descrição	Contém dados para o ano 1986
Exemplo	136.12847932477
Preenchido	13.88%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	1987																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1987																																																						
Exemplo	135.346146015478																																																						
Preenchido	13.91%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	Column: metadata_987 <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>29147</td><td>9.54%</td></tr> <tr><td>Unique Count</td><td>27112</td><td>8.87%</td></tr> <tr><td>Duplicate Count</td><td>2035</td><td>0.67%</td></tr> <tr><td>Blank Count</td><td>263052</td><td>86.09%</td></tr> </tbody> </table> Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>263052</td><td>86.09%</td></tr> <tr><td>0</td><td>4184</td><td>1.37%</td></tr> <tr><td>1</td><td>3477</td><td>1.14%</td></tr> <tr><td>75</td><td>337</td><td>0.11%</td></tr> <tr><td>25</td><td>243</td><td>0.08%</td></tr> <tr><td>100</td><td>243</td><td>0.08%</td></tr> <tr><td>50</td><td>147</td><td>0.05%</td></tr> <tr><td>60</td><td>128</td><td>0.04%</td></tr> <tr><td>80</td><td>125</td><td>0.04%</td></tr> <tr><td>40</td><td>113</td><td>0.04%</td></tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	29147	9.54%	Unique Count	27112	8.87%	Duplicate Count	2035	0.67%	Blank Count	263052	86.09%	Value	Count	%	Empty field	263052	86.09%	0	4184	1.37%	1	3477	1.14%	75	337	0.11%	25	243	0.08%	100	243	0.08%	50	147	0.05%	60	128	0.04%	80	125	0.04%	40	113	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	29147	9.54%																																																					
Unique Count	27112	8.87%																																																					
Duplicate Count	2035	0.67%																																																					
Blank Count	263052	86.09%																																																					
Value	Count	%																																																					
Empty field	263052	86.09%																																																					
0	4184	1.37%																																																					
1	3477	1.14%																																																					
75	337	0.11%																																																					
25	243	0.08%																																																					
100	243	0.08%																																																					
50	147	0.05%																																																					
60	128	0.04%																																																					
80	125	0.04%																																																					
40	113	0.04%																																																					

Atributo	1988																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1988																																																						
Exemplo	133.911118066723																																																						
Preenchido	13.78%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_988</p> <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>28826</td> <td>9.43%</td> </tr> <tr> <td>Unique Count</td> <td>26762</td> <td>8.76%</td> </tr> <tr> <td>Duplicate Count</td> <td>2064</td> <td>0.68%</td> </tr> <tr> <td>Blank Count</td> <td>263442</td> <td>86.22%</td> </tr> </tbody> </table> <ul style="list-style-type: none"> - Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>263442</td> <td>86.22%</td> </tr> <tr> <td>0</td> <td>4131</td> <td>1.35%</td> </tr> <tr> <td>1</td> <td>3527</td> <td>1.15%</td> </tr> <tr> <td>75</td> <td>337</td> <td>0.11%</td> </tr> <tr> <td>100</td> <td>240</td> <td>0.08%</td> </tr> <tr> <td>25</td> <td>230</td> <td>0.08%</td> </tr> <tr> <td>50</td> <td>148</td> <td>0.05%</td> </tr> <tr> <td>60</td> <td>132</td> <td>0.04%</td> </tr> <tr> <td>80</td> <td>126</td> <td>0.04%</td> </tr> <tr> <td>40</td> <td>113</td> <td>0.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	28826	9.43%	Unique Count	26762	8.76%	Duplicate Count	2064	0.68%	Blank Count	263442	86.22%	Value	Count	%	Empty field	263442	86.22%	0	4131	1.35%	1	3527	1.15%	75	337	0.11%	100	240	0.08%	25	230	0.08%	50	148	0.05%	60	132	0.04%	80	126	0.04%	40	113	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	28826	9.43%																																																					
Unique Count	26762	8.76%																																																					
Duplicate Count	2064	0.68%																																																					
Blank Count	263442	86.22%																																																					
Value	Count	%																																																					
Empty field	263442	86.22%																																																					
0	4131	1.35%																																																					
1	3527	1.15%																																																					
75	337	0.11%																																																					
100	240	0.08%																																																					
25	230	0.08%																																																					
50	148	0.05%																																																					
60	132	0.04%																																																					
80	126	0.04%																																																					
40	113	0.04%																																																					

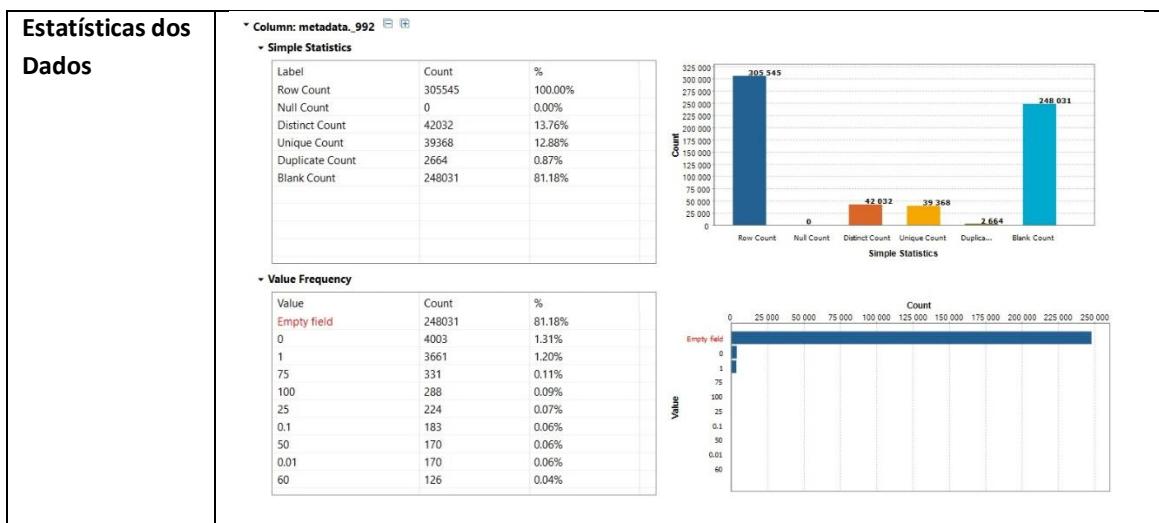
Atributo	1989
Tipo de Dados	String
Descrição	Contém dados para o ano 1989
Exemplo	132.422604235282
Preenchido	13.84%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	1990																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1990																																																						
Exemplo	132.845311378448																																																						
Preenchido	16.31%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_990</p> <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>35040</td><td>11.47%</td></tr> <tr><td>Unique Count</td><td>32682</td><td>10.70%</td></tr> <tr><td>Duplicate Count</td><td>2358</td><td>0.77%</td></tr> <tr><td>Blank Count</td><td>255696</td><td>83.69%</td></tr> </tbody> </table> <p>Simple Statistics</p> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>255696</td><td>83.69%</td></tr> <tr><td>0</td><td>4088</td><td>1.34%</td></tr> <tr><td>1</td><td>3596</td><td>1.18%</td></tr> <tr><td>75</td><td>336</td><td>0.11%</td></tr> <tr><td>100</td><td>274</td><td>0.09%</td></tr> <tr><td>25</td><td>229</td><td>0.07%</td></tr> <tr><td>0.1</td><td>204</td><td>0.07%</td></tr> <tr><td>0.01</td><td>186</td><td>0.06%</td></tr> <tr><td>50</td><td>156</td><td>0.05%</td></tr> <tr><td>60</td><td>136</td><td>0.04%</td></tr> </tbody> </table> <p>Value Frequency</p>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	35040	11.47%	Unique Count	32682	10.70%	Duplicate Count	2358	0.77%	Blank Count	255696	83.69%	Value	Count	%	Empty field	255696	83.69%	0	4088	1.34%	1	3596	1.18%	75	336	0.11%	100	274	0.09%	25	229	0.07%	0.1	204	0.07%	0.01	186	0.06%	50	156	0.05%	60	136	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	35040	11.47%																																																					
Unique Count	32682	10.70%																																																					
Duplicate Count	2358	0.77%																																																					
Blank Count	255696	83.69%																																																					
Value	Count	%																																																					
Empty field	255696	83.69%																																																					
0	4088	1.34%																																																					
1	3596	1.18%																																																					
75	336	0.11%																																																					
100	274	0.09%																																																					
25	229	0.07%																																																					
0.1	204	0.07%																																																					
0.01	186	0.06%																																																					
50	156	0.05%																																																					
60	136	0.04%																																																					

Atributo	1991																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1991																																																						
Exemplo	131.922595134796																																																						
Preenchido	18.75%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>* Column: metadata_991</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>41919</td> <td>13.72%</td> </tr> <tr> <td>Unique Count</td> <td>39298</td> <td>12.86%</td> </tr> <tr> <td>Duplicate Count</td> <td>2621</td> <td>0.86%</td> </tr> <tr> <td>Blank Count</td> <td>248261</td> <td>81.25%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>248261</td> <td>81.25%</td> </tr> <tr> <td>0</td> <td>4046</td> <td>1.32%</td> </tr> <tr> <td>1</td> <td>3622</td> <td>1.19%</td> </tr> <tr> <td>75</td> <td>336</td> <td>0.11%</td> </tr> <tr> <td>100</td> <td>274</td> <td>0.09%</td> </tr> <tr> <td>25</td> <td>219</td> <td>0.07%</td> </tr> <tr> <td>0.1</td> <td>205</td> <td>0.07%</td> </tr> <tr> <td>0.01</td> <td>176</td> <td>0.06%</td> </tr> <tr> <td>50</td> <td>160</td> <td>0.05%</td> </tr> <tr> <td>80</td> <td>127</td> <td>0.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	41919	13.72%	Unique Count	39298	12.86%	Duplicate Count	2621	0.86%	Blank Count	248261	81.25%	Value	Count	%	Empty field	248261	81.25%	0	4046	1.32%	1	3622	1.19%	75	336	0.11%	100	274	0.09%	25	219	0.07%	0.1	205	0.07%	0.01	176	0.06%	50	160	0.05%	80	127	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	41919	13.72%																																																					
Unique Count	39298	12.86%																																																					
Duplicate Count	2621	0.86%																																																					
Blank Count	248261	81.25%																																																					
Value	Count	%																																																					
Empty field	248261	81.25%																																																					
0	4046	1.32%																																																					
1	3622	1.19%																																																					
75	336	0.11%																																																					
100	274	0.09%																																																					
25	219	0.07%																																																					
0.1	205	0.07%																																																					
0.01	176	0.06%																																																					
50	160	0.05%																																																					
80	127	0.04%																																																					

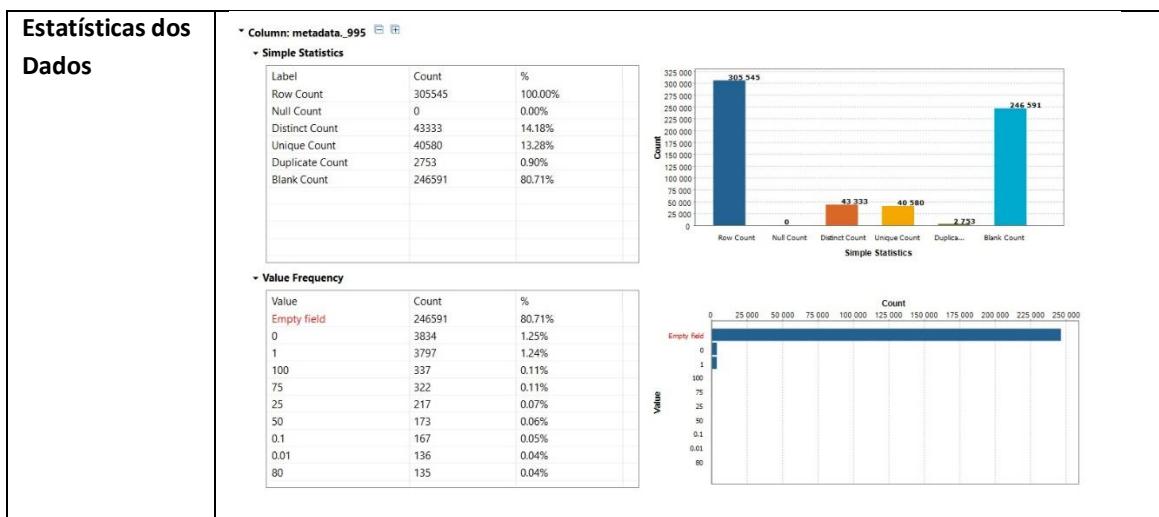
Atributo	1992
Tipo de Dados	String
Descrição	Contém dados para o ano 1992
Exemplo	128.650226745135
Preenchido	18.82%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	1993																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1993																																																						
Exemplo	129.800646275656																																																						
Preenchido	18.69%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>* Column: metadata_993</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>41701</td><td>13.65%</td></tr> <tr><td>Unique Count</td><td>39069</td><td>12.79%</td></tr> <tr><td>Duplicate Count</td><td>2632</td><td>0.86%</td></tr> <tr><td>Blank Count</td><td>248446</td><td>81.31%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>248446</td><td>81.31%</td></tr> <tr><td>0</td><td>3942</td><td>1.29%</td></tr> <tr><td>1</td><td>3705</td><td>1.21%</td></tr> <tr><td>75</td><td>337</td><td>0.11%</td></tr> <tr><td>100</td><td>305</td><td>0.10%</td></tr> <tr><td>25</td><td>224</td><td>0.07%</td></tr> <tr><td>0.1</td><td>182</td><td>0.06%</td></tr> <tr><td>50</td><td>168</td><td>0.05%</td></tr> <tr><td>0.01</td><td>156</td><td>0.05%</td></tr> <tr><td>80</td><td>130</td><td>0.04%</td></tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	41701	13.65%	Unique Count	39069	12.79%	Duplicate Count	2632	0.86%	Blank Count	248446	81.31%	Value	Count	%	Empty field	248446	81.31%	0	3942	1.29%	1	3705	1.21%	75	337	0.11%	100	305	0.10%	25	224	0.07%	0.1	182	0.06%	50	168	0.05%	0.01	156	0.05%	80	130	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	41701	13.65%																																																					
Unique Count	39069	12.79%																																																					
Duplicate Count	2632	0.86%																																																					
Blank Count	248446	81.31%																																																					
Value	Count	%																																																					
Empty field	248446	81.31%																																																					
0	3942	1.29%																																																					
1	3705	1.21%																																																					
75	337	0.11%																																																					
100	305	0.10%																																																					
25	224	0.07%																																																					
0.1	182	0.06%																																																					
50	168	0.05%																																																					
0.01	156	0.05%																																																					
80	130	0.04%																																																					

Atributo	1994																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1994																																																						
Exemplo	128.792351238038																																																						
Preenchido	18.84%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata.994</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>42065</td> <td>13.77%</td> </tr> <tr> <td>Unique Count</td> <td>39265</td> <td>12.85%</td> </tr> <tr> <td>Duplicate Count</td> <td>2800</td> <td>0.92%</td> </tr> <tr> <td>Blank Count</td> <td>247995</td> <td>81.16%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>247995</td> <td>81.16%</td> </tr> <tr> <td>0</td> <td>3881</td> <td>1.27%</td> </tr> <tr> <td>1</td> <td>3753</td> <td>1.23%</td> </tr> <tr> <td>75</td> <td>321</td> <td>0.11%</td> </tr> <tr> <td>100</td> <td>315</td> <td>0.10%</td> </tr> <tr> <td>25</td> <td>220</td> <td>0.07%</td> </tr> <tr> <td>50</td> <td>174</td> <td>0.06%</td> </tr> <tr> <td>0.1</td> <td>171</td> <td>0.06%</td> </tr> <tr> <td>0.01</td> <td>145</td> <td>0.05%</td> </tr> <tr> <td>80</td> <td>129</td> <td>0.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	42065	13.77%	Unique Count	39265	12.85%	Duplicate Count	2800	0.92%	Blank Count	247995	81.16%	Value	Count	%	Empty field	247995	81.16%	0	3881	1.27%	1	3753	1.23%	75	321	0.11%	100	315	0.10%	25	220	0.07%	50	174	0.06%	0.1	171	0.06%	0.01	145	0.05%	80	129	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	42065	13.77%																																																					
Unique Count	39265	12.85%																																																					
Duplicate Count	2800	0.92%																																																					
Blank Count	247995	81.16%																																																					
Value	Count	%																																																					
Empty field	247995	81.16%																																																					
0	3881	1.27%																																																					
1	3753	1.23%																																																					
75	321	0.11%																																																					
100	315	0.10%																																																					
25	220	0.07%																																																					
50	174	0.06%																																																					
0.1	171	0.06%																																																					
0.01	145	0.05%																																																					
80	129	0.04%																																																					

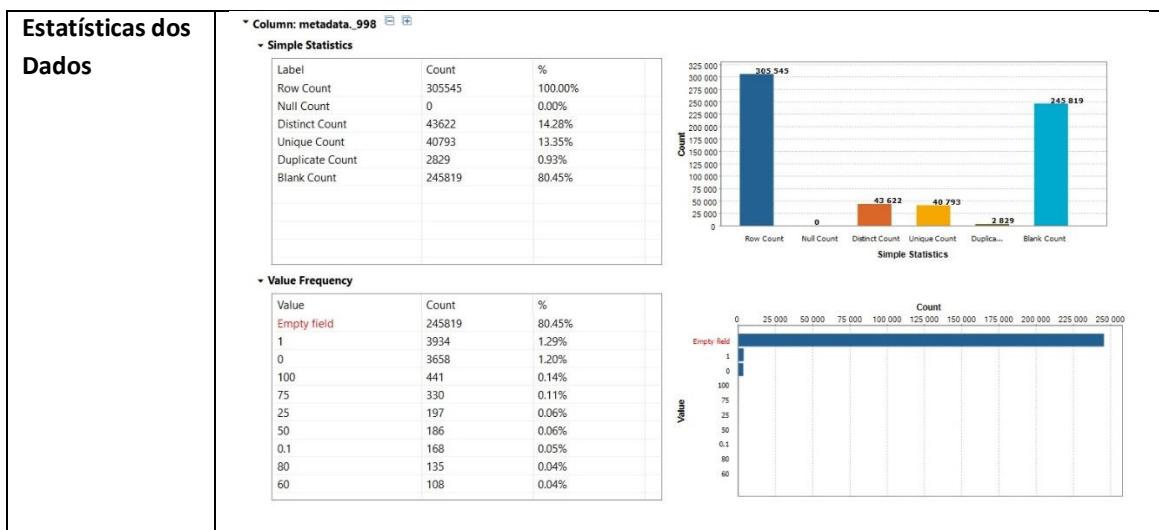
Atributo	1995
Tipo de Dados	String
Descrição	Contém dados para o ano 1995
Exemplo	127.306550787239
Preenchido	19.29%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	1996																																																																																										
Tipo de Dados	String																																																																																										
Descrição	Contém dados para o ano 1996																																																																																										
Exemplo	128.239448076508																																																																																										
Preenchido	18.99%																																																																																										
Intervalo de Dados	Não mensurável																																																																																										
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																																																										
Estatísticas dos Dados	<p>- Column: metadata_996</p> <p>- Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>42100</td> <td>13.78%</td> </tr> <tr> <td>Unique Count</td> <td>39239</td> <td>12.84%</td> </tr> <tr> <td>Duplicate Count</td> <td>2861</td> <td>0.94%</td> </tr> <tr> <td>Blank Count</td> <td>247534</td> <td>81.01%</td> </tr> </tbody> </table> <p>Count</p> <table border="1"> <thead> <tr> <th>Category</th> <th>Count</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305,545</td> </tr> <tr> <td>Null Count</td> <td>0</td> </tr> <tr> <td>Distinct Count</td> <td>42,100</td> </tr> <tr> <td>Unique Count</td> <td>39,239</td> </tr> <tr> <td>Duplicate Count</td> <td>2,861</td> </tr> <tr> <td>Blank Count</td> <td>247,534</td> </tr> </tbody> </table> <p>Simple Statistics</p> <p>- Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>247534</td> <td>81.01%</td> </tr> <tr> <td>1</td> <td>3850</td> <td>1.26%</td> </tr> <tr> <td>0</td> <td>3723</td> <td>1.22%</td> </tr> <tr> <td>100</td> <td>340</td> <td>0.11%</td> </tr> <tr> <td>75</td> <td>325</td> <td>0.11%</td> </tr> <tr> <td>25</td> <td>200</td> <td>0.07%</td> </tr> <tr> <td>50</td> <td>179</td> <td>0.06%</td> </tr> <tr> <td>0.1</td> <td>173</td> <td>0.06%</td> </tr> <tr> <td>80</td> <td>135</td> <td>0.04%</td> </tr> <tr> <td>0.01</td> <td>123</td> <td>0.04%</td> </tr> </tbody> </table> <p>Value</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>247,534</td> </tr> <tr> <td>1</td> <td>3850</td> </tr> <tr> <td>0</td> <td>3723</td> </tr> <tr> <td>100</td> <td>340</td> </tr> <tr> <td>75</td> <td>325</td> </tr> <tr> <td>25</td> <td>200</td> </tr> <tr> <td>50</td> <td>179</td> </tr> <tr> <td>0.1</td> <td>173</td> </tr> <tr> <td>80</td> <td>135</td> </tr> <tr> <td>0.01</td> <td>123</td> </tr> </tbody> </table> <p>Simple Statistics</p>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	42100	13.78%	Unique Count	39239	12.84%	Duplicate Count	2861	0.94%	Blank Count	247534	81.01%	Category	Count	Row Count	305,545	Null Count	0	Distinct Count	42,100	Unique Count	39,239	Duplicate Count	2,861	Blank Count	247,534	Value	Count	%	Empty field	247534	81.01%	1	3850	1.26%	0	3723	1.22%	100	340	0.11%	75	325	0.11%	25	200	0.07%	50	179	0.06%	0.1	173	0.06%	80	135	0.04%	0.01	123	0.04%	Value	Count	Empty field	247,534	1	3850	0	3723	100	340	75	325	25	200	50	179	0.1	173	80	135	0.01	123
Label	Count	%																																																																																									
Row Count	305545	100.00%																																																																																									
Null Count	0	0.00%																																																																																									
Distinct Count	42100	13.78%																																																																																									
Unique Count	39239	12.84%																																																																																									
Duplicate Count	2861	0.94%																																																																																									
Blank Count	247534	81.01%																																																																																									
Category	Count																																																																																										
Row Count	305,545																																																																																										
Null Count	0																																																																																										
Distinct Count	42,100																																																																																										
Unique Count	39,239																																																																																										
Duplicate Count	2,861																																																																																										
Blank Count	247,534																																																																																										
Value	Count	%																																																																																									
Empty field	247534	81.01%																																																																																									
1	3850	1.26%																																																																																									
0	3723	1.22%																																																																																									
100	340	0.11%																																																																																									
75	325	0.11%																																																																																									
25	200	0.07%																																																																																									
50	179	0.06%																																																																																									
0.1	173	0.06%																																																																																									
80	135	0.04%																																																																																									
0.01	123	0.04%																																																																																									
Value	Count																																																																																										
Empty field	247,534																																																																																										
1	3850																																																																																										
0	3723																																																																																										
100	340																																																																																										
75	325																																																																																										
25	200																																																																																										
50	179																																																																																										
0.1	173																																																																																										
80	135																																																																																										
0.01	123																																																																																										

Atributo	1997																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1997																																																						
Exemplo	125.26199777544																																																						
Preenchido	18.97%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_1997</p> <ul style="list-style-type: none"> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>42198</td> <td>13.81%</td> </tr> <tr> <td>Unique Count</td> <td>39375</td> <td>12.89%</td> </tr> <tr> <td>Duplicate Count</td> <td>2823</td> <td>0.92%</td> </tr> <tr> <td>Blank Count</td> <td>247570</td> <td>81.03%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>247570</td> <td>81.03%</td> </tr> <tr> <td>1</td> <td>3898</td> <td>1.28%</td> </tr> <tr> <td>0</td> <td>3684</td> <td>1.21%</td> </tr> <tr> <td>100</td> <td>352</td> <td>0.12%</td> </tr> <tr> <td>75</td> <td>321</td> <td>0.11%</td> </tr> <tr> <td>25</td> <td>201</td> <td>0.07%</td> </tr> <tr> <td>50</td> <td>186</td> <td>0.06%</td> </tr> <tr> <td>0.1</td> <td>178</td> <td>0.06%</td> </tr> <tr> <td>80</td> <td>137</td> <td>0.04%</td> </tr> <tr> <td>60</td> <td>113</td> <td>0.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	42198	13.81%	Unique Count	39375	12.89%	Duplicate Count	2823	0.92%	Blank Count	247570	81.03%	Value	Count	%	Empty field	247570	81.03%	1	3898	1.28%	0	3684	1.21%	100	352	0.12%	75	321	0.11%	25	201	0.07%	50	186	0.06%	0.1	178	0.06%	80	137	0.04%	60	113	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	42198	13.81%																																																					
Unique Count	39375	12.89%																																																					
Duplicate Count	2823	0.92%																																																					
Blank Count	247570	81.03%																																																					
Value	Count	%																																																					
Empty field	247570	81.03%																																																					
1	3898	1.28%																																																					
0	3684	1.21%																																																					
100	352	0.12%																																																					
75	321	0.11%																																																					
25	201	0.07%																																																					
50	186	0.06%																																																					
0.1	178	0.06%																																																					
80	137	0.04%																																																					
60	113	0.04%																																																					

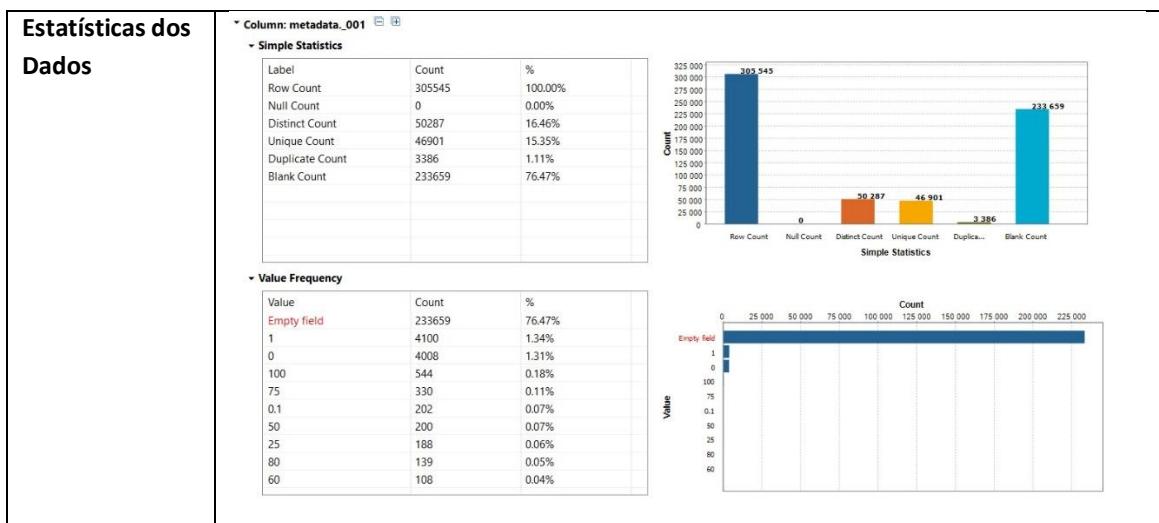
Atributo	1998
Tipo de Dados	String
Descrição	Contém dados para o ano 1998
Exemplo	123.860302659944
Preenchido	19.55%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	1999																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 1999																																																						
Exemplo	124.958023564541																																																						
Preenchido	20.88%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	Column: metadata_999 <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>47279</td><td>15.47%</td></tr> <tr><td>Unique Count</td><td>44351</td><td>14.52%</td></tr> <tr><td>Duplicate Count</td><td>2928</td><td>0.96%</td></tr> <tr><td>Blank Count</td><td>241745</td><td>79.12%</td></tr> </tbody> </table> <p>Simple Statistics</p> <ul style="list-style-type: none"> - Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>241745</td><td>79.12%</td></tr> <tr><td>1</td><td>3980</td><td>1.30%</td></tr> <tr><td>0</td><td>3625</td><td>1.19%</td></tr> <tr><td>100</td><td>511</td><td>0.17%</td></tr> <tr><td>75</td><td>334</td><td>0.11%</td></tr> <tr><td>25</td><td>187</td><td>0.06%</td></tr> <tr><td>50</td><td>183</td><td>0.06%</td></tr> <tr><td>0.1</td><td>167</td><td>0.05%</td></tr> <tr><td>80</td><td>137</td><td>0.04%</td></tr> <tr><td>60</td><td>109</td><td>0.04%</td></tr> </tbody> </table> <p>Value</p>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	47279	15.47%	Unique Count	44351	14.52%	Duplicate Count	2928	0.96%	Blank Count	241745	79.12%	Value	Count	%	Empty field	241745	79.12%	1	3980	1.30%	0	3625	1.19%	100	511	0.17%	75	334	0.11%	25	187	0.06%	50	183	0.06%	0.1	167	0.05%	80	137	0.04%	60	109	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	47279	15.47%																																																					
Unique Count	44351	14.52%																																																					
Duplicate Count	2928	0.96%																																																					
Blank Count	241745	79.12%																																																					
Value	Count	%																																																					
Empty field	241745	79.12%																																																					
1	3980	1.30%																																																					
0	3625	1.19%																																																					
100	511	0.17%																																																					
75	334	0.11%																																																					
25	187	0.06%																																																					
50	183	0.06%																																																					
0.1	167	0.05%																																																					
80	137	0.04%																																																					
60	109	0.04%																																																					

Atributo	2000																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 2000																																																						
Exemplo	123.134769131908																																																						
Preenchido	28.14%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>* Column: metadata_000</p> <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>62173</td> <td>20.35%</td> </tr> <tr> <td>Unique Count</td> <td>58464</td> <td>19.13%</td> </tr> <tr> <td>Duplicate Count</td> <td>3709</td> <td>1.21%</td> </tr> <tr> <td>Blank Count</td> <td>219569</td> <td>71.86%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>219569</td> <td>71.86%</td> </tr> <tr> <td>0</td> <td>4102</td> <td>1.34%</td> </tr> <tr> <td>1</td> <td>4071</td> <td>1.33%</td> </tr> <tr> <td>100</td> <td>534</td> <td>0.17%</td> </tr> <tr> <td>75</td> <td>332</td> <td>0.11%</td> </tr> <tr> <td>0.1</td> <td>215</td> <td>0.07%</td> </tr> <tr> <td>25</td> <td>196</td> <td>0.06%</td> </tr> <tr> <td>50</td> <td>195</td> <td>0.06%</td> </tr> <tr> <td>80</td> <td>151</td> <td>0.05%</td> </tr> <tr> <td>40</td> <td>114</td> <td>0.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	62173	20.35%	Unique Count	58464	19.13%	Duplicate Count	3709	1.21%	Blank Count	219569	71.86%	Value	Count	%	Empty field	219569	71.86%	0	4102	1.34%	1	4071	1.33%	100	534	0.17%	75	332	0.11%	0.1	215	0.07%	25	196	0.06%	50	195	0.06%	80	151	0.05%	40	114	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	62173	20.35%																																																					
Unique Count	58464	19.13%																																																					
Duplicate Count	3709	1.21%																																																					
Blank Count	219569	71.86%																																																					
Value	Count	%																																																					
Empty field	219569	71.86%																																																					
0	4102	1.34%																																																					
1	4071	1.33%																																																					
100	534	0.17%																																																					
75	332	0.11%																																																					
0.1	215	0.07%																																																					
25	196	0.06%																																																					
50	195	0.06%																																																					
80	151	0.05%																																																					
40	114	0.04%																																																					

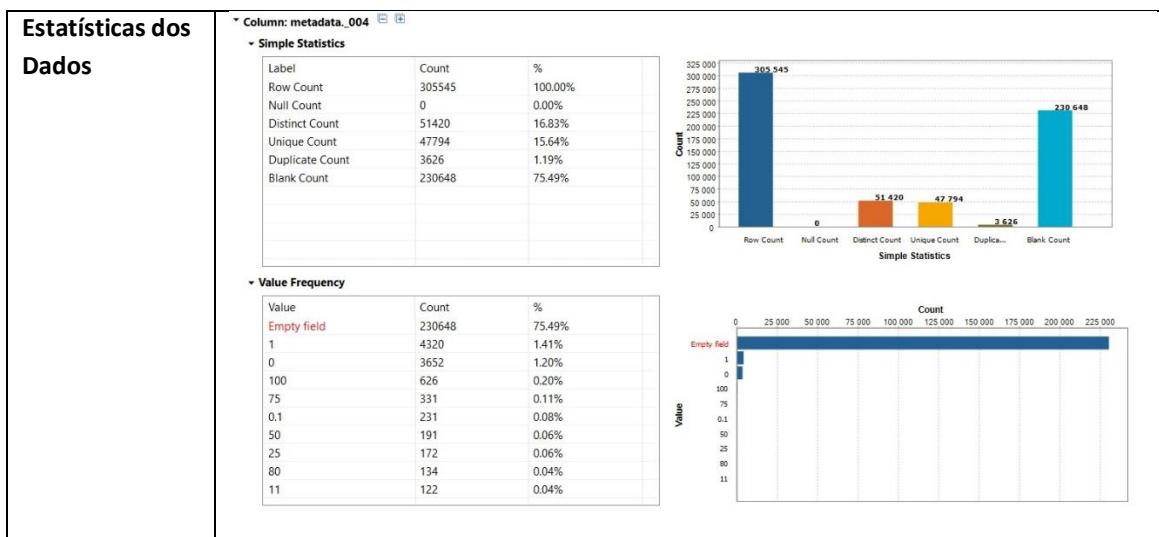
Atributo	2001
Tipo de Dados	String
Descrição	Contém dados para o ano 2001
Exemplo	123.629170417788
Preenchido	23.53%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	2002																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 2002																																																						
Exemplo	121.936911479602																																																						
Preenchido	23.55%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_002</p> <ul style="list-style-type: none"> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>50651</td><td>16.58%</td></tr> <tr><td>Unique Count</td><td>47287</td><td>15.48%</td></tr> <tr><td>Duplicate Count</td><td>3364</td><td>1.10%</td></tr> <tr><td>Blank Count</td><td>233600</td><td>76.45%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>233600</td><td>76.45%</td></tr> <tr><td>1</td><td>4169</td><td>1.36%</td></tr> <tr><td>0</td><td>3846</td><td>1.26%</td></tr> <tr><td>100</td><td>558</td><td>0.18%</td></tr> <tr><td>75</td><td>333</td><td>0.11%</td></tr> <tr><td>0.1</td><td>209</td><td>0.07%</td></tr> <tr><td>50</td><td>189</td><td>0.06%</td></tr> <tr><td>25</td><td>179</td><td>0.06%</td></tr> <tr><td>80</td><td>141</td><td>0.05%</td></tr> <tr><td>40</td><td>110</td><td>0.04%</td></tr> </tbody> </table> <p>Count</p> <p>Simple Statistics</p> <p>Count</p> <p>Value</p>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	50651	16.58%	Unique Count	47287	15.48%	Duplicate Count	3364	1.10%	Blank Count	233600	76.45%	Value	Count	%	Empty field	233600	76.45%	1	4169	1.36%	0	3846	1.26%	100	558	0.18%	75	333	0.11%	0.1	209	0.07%	50	189	0.06%	25	179	0.06%	80	141	0.05%	40	110	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	50651	16.58%																																																					
Unique Count	47287	15.48%																																																					
Duplicate Count	3364	1.10%																																																					
Blank Count	233600	76.45%																																																					
Value	Count	%																																																					
Empty field	233600	76.45%																																																					
1	4169	1.36%																																																					
0	3846	1.26%																																																					
100	558	0.18%																																																					
75	333	0.11%																																																					
0.1	209	0.07%																																																					
50	189	0.06%																																																					
25	179	0.06%																																																					
80	141	0.05%																																																					
40	110	0.04%																																																					

Atributo	2003																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 2003																																																						
Exemplo	119.557516351678																																																						
Preenchido	24.19%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_003</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>50365</td><td>16.48%</td></tr> <tr><td>Unique Count</td><td>46742</td><td>15.30%</td></tr> <tr><td>Duplicate Count</td><td>3623</td><td>1.19%</td></tr> <tr><td>Blank Count</td><td>231622</td><td>75.81%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>231622</td><td>75.81%</td></tr> <tr><td>1</td><td>4245</td><td>1.39%</td></tr> <tr><td>0</td><td>3750</td><td>1.23%</td></tr> <tr><td>100</td><td>579</td><td>0.19%</td></tr> <tr><td>75</td><td>336</td><td>0.11%</td></tr> <tr><td>0.1</td><td>225</td><td>0.07%</td></tr> <tr><td>50</td><td>191</td><td>0.06%</td></tr> <tr><td>25</td><td>190</td><td>0.06%</td></tr> <tr><td>80</td><td>145</td><td>0.05%</td></tr> <tr><td>40</td><td>117</td><td>0.04%</td></tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	50365	16.48%	Unique Count	46742	15.30%	Duplicate Count	3623	1.19%	Blank Count	231622	75.81%	Value	Count	%	Empty field	231622	75.81%	1	4245	1.39%	0	3750	1.23%	100	579	0.19%	75	336	0.11%	0.1	225	0.07%	50	191	0.06%	25	190	0.06%	80	145	0.05%	40	117	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	50365	16.48%																																																					
Unique Count	46742	15.30%																																																					
Duplicate Count	3623	1.19%																																																					
Blank Count	231622	75.81%																																																					
Value	Count	%																																																					
Empty field	231622	75.81%																																																					
1	4245	1.39%																																																					
0	3750	1.23%																																																					
100	579	0.19%																																																					
75	336	0.11%																																																					
0.1	225	0.07%																																																					
50	191	0.06%																																																					
25	190	0.06%																																																					
80	145	0.05%																																																					
40	117	0.04%																																																					

Atributo	2004
Tipo de Dados	String
Descrição	Contém dados para o ano 2004
Exemplo	118.823211926383
Preenchido	24.51%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null

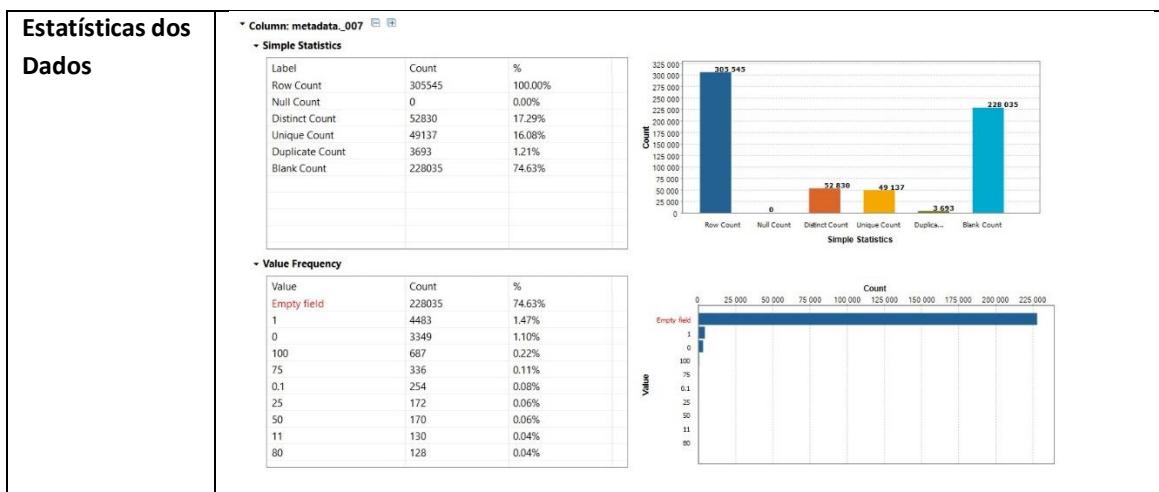


Atributo	2005																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 2005																																																						
Exemplo	117.584240435768																																																						
Preenchido	25.99%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_005</p> <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>53467</td><td>17.50%</td></tr> <tr><td>Unique Count</td><td>49499</td><td>16.20%</td></tr> <tr><td>Duplicate Count</td><td>3968</td><td>1.30%</td></tr> <tr><td>Blank Count</td><td>226136</td><td>74.01%</td></tr> </tbody> </table> <p>Simple Statistics</p> <ul style="list-style-type: none"> - Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>226136</td><td>74.01%</td></tr> <tr><td>1</td><td>4378</td><td>1.43%</td></tr> <tr><td>0</td><td>3542</td><td>1.16%</td></tr> <tr><td>100</td><td>636</td><td>0.21%</td></tr> <tr><td>75</td><td>331</td><td>0.11%</td></tr> <tr><td>0.1</td><td>241</td><td>0.08%</td></tr> <tr><td>50</td><td>189</td><td>0.06%</td></tr> <tr><td>25</td><td>175</td><td>0.06%</td></tr> <tr><td>80</td><td>134</td><td>0.04%</td></tr> <tr><td>11</td><td>127</td><td>0.04%</td></tr> </tbody> </table> <p>Value</p> <p>Count</p>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	53467	17.50%	Unique Count	49499	16.20%	Duplicate Count	3968	1.30%	Blank Count	226136	74.01%	Value	Count	%	Empty field	226136	74.01%	1	4378	1.43%	0	3542	1.16%	100	636	0.21%	75	331	0.11%	0.1	241	0.08%	50	189	0.06%	25	175	0.06%	80	134	0.04%	11	127	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	53467	17.50%																																																					
Unique Count	49499	16.20%																																																					
Duplicate Count	3968	1.30%																																																					
Blank Count	226136	74.01%																																																					
Value	Count	%																																																					
Empty field	226136	74.01%																																																					
1	4378	1.43%																																																					
0	3542	1.16%																																																					
100	636	0.21%																																																					
75	331	0.11%																																																					
0.1	241	0.08%																																																					
50	189	0.06%																																																					
25	175	0.06%																																																					
80	134	0.04%																																																					
11	127	0.04%																																																					

Atributo	2006
Tipo de Dados	String

Descrição	Contém dados para o ano 2006																																																						
Exemplo	116.401306961845																																																						
Preenchido	25.4%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>* Column: metadata_006</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>District Count</td> <td>52194</td> <td>17.08%</td> </tr> <tr> <td>Unique Count</td> <td>48466</td> <td>15.86%</td> </tr> <tr> <td>Duplicate Count</td> <td>3728</td> <td>1.22%</td> </tr> <tr> <td>Blank Count</td> <td>227928</td> <td>74.60%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>227928</td> <td>74.60%</td> </tr> <tr> <td>1</td> <td>4415</td> <td>1.44%</td> </tr> <tr> <td>0</td> <td>3428</td> <td>1.12%</td> </tr> <tr> <td>100</td> <td>651</td> <td>0.21%</td> </tr> <tr> <td>75</td> <td>327</td> <td>0.11%</td> </tr> <tr> <td>0.1</td> <td>237</td> <td>0.08%</td> </tr> <tr> <td>50</td> <td>189</td> <td>0.06%</td> </tr> <tr> <td>25</td> <td>170</td> <td>0.06%</td> </tr> <tr> <td>11</td> <td>132</td> <td>0.04%</td> </tr> <tr> <td>80</td> <td>131</td> <td>0.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	District Count	52194	17.08%	Unique Count	48466	15.86%	Duplicate Count	3728	1.22%	Blank Count	227928	74.60%	Value	Count	%	Empty field	227928	74.60%	1	4415	1.44%	0	3428	1.12%	100	651	0.21%	75	327	0.11%	0.1	237	0.08%	50	189	0.06%	25	170	0.06%	11	132	0.04%	80	131	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
District Count	52194	17.08%																																																					
Unique Count	48466	15.86%																																																					
Duplicate Count	3728	1.22%																																																					
Blank Count	227928	74.60%																																																					
Value	Count	%																																																					
Empty field	227928	74.60%																																																					
1	4415	1.44%																																																					
0	3428	1.12%																																																					
100	651	0.21%																																																					
75	327	0.11%																																																					
0.1	237	0.08%																																																					
50	189	0.06%																																																					
25	170	0.06%																																																					
11	132	0.04%																																																					
80	131	0.04%																																																					

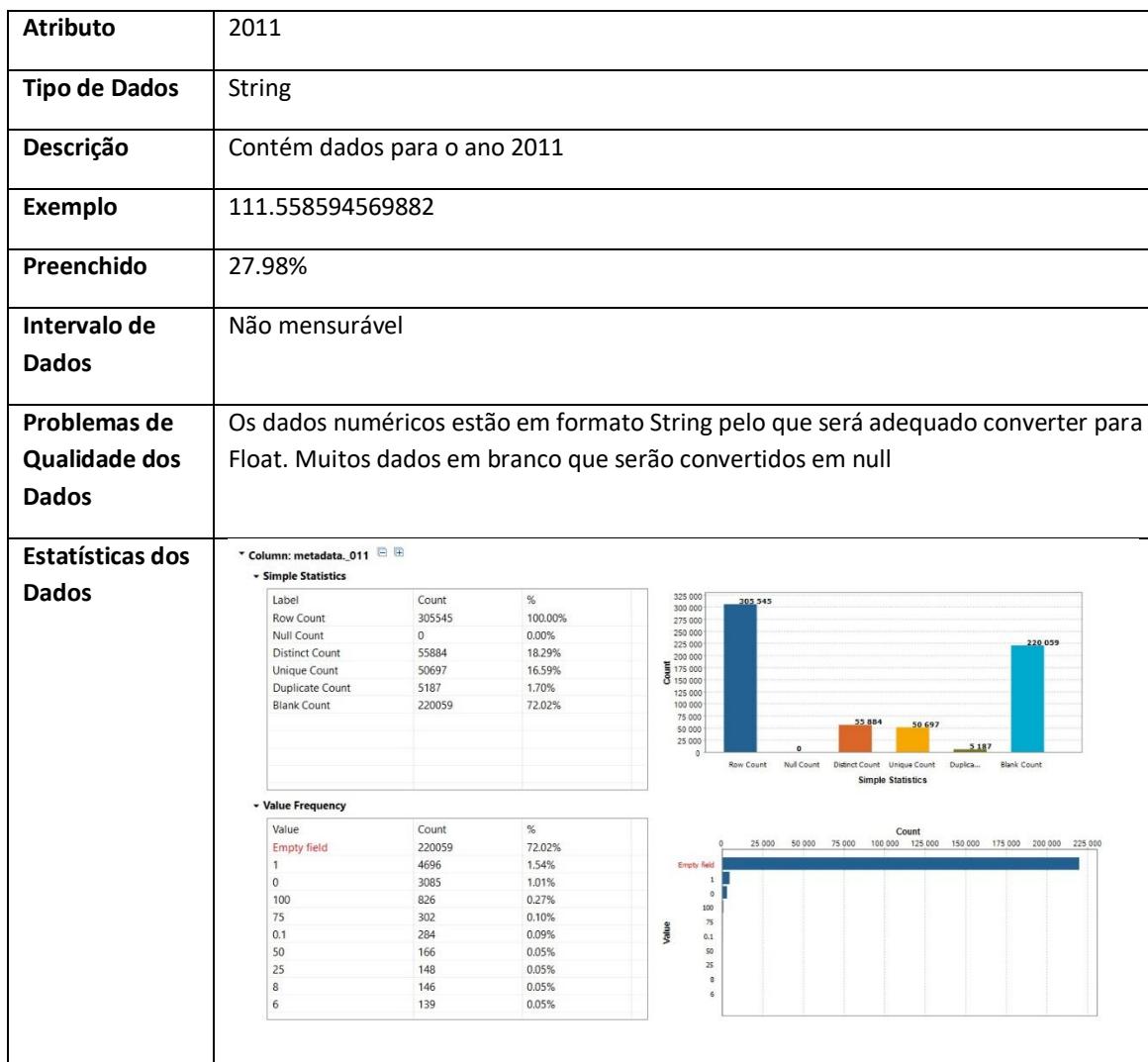
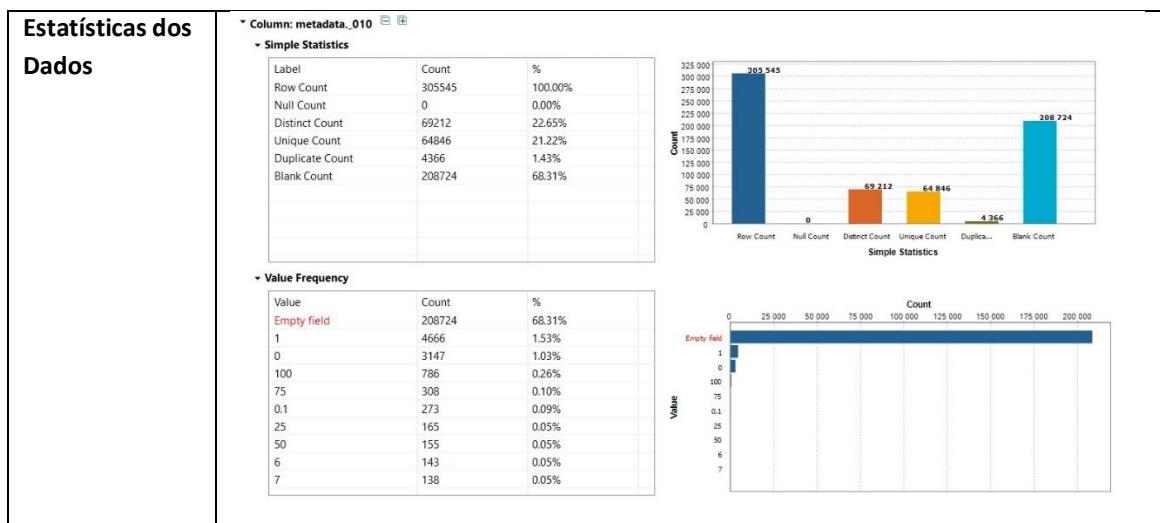
Atributo	2007
Tipo de Dados	String
Descrição	Contém dados para o ano 2007
Exemplo	115.060877421007
Preenchido	25.37%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	2008																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 2008																																																						
Exemplo	114.933408732692																																																						
Preenchido	25.27%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>* Column: metadata_008</p> <p>+ Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>52741</td><td>17.26%</td></tr> <tr><td>Unique Count</td><td>49030</td><td>16.05%</td></tr> <tr><td>Duplicate Count</td><td>3711</td><td>1.21%</td></tr> <tr><td>Blank Count</td><td>228332</td><td>74.73%</td></tr> </tbody> </table> <p>305 545</p> <p>+ Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>228332</td><td>74.73%</td></tr> <tr><td>1</td><td>4554</td><td>1.49%</td></tr> <tr><td>0</td><td>3252</td><td>1.06%</td></tr> <tr><td>100</td><td>708</td><td>0.23%</td></tr> <tr><td>75</td><td>322</td><td>0.11%</td></tr> <tr><td>0.1</td><td>245</td><td>0.08%</td></tr> <tr><td>25</td><td>183</td><td>0.06%</td></tr> <tr><td>50</td><td>170</td><td>0.06%</td></tr> <tr><td>80</td><td>129</td><td>0.04%</td></tr> <tr><td>8</td><td>123</td><td>0.04%</td></tr> </tbody> </table> <p>228 332</p>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	52741	17.26%	Unique Count	49030	16.05%	Duplicate Count	3711	1.21%	Blank Count	228332	74.73%	Value	Count	%	Empty field	228332	74.73%	1	4554	1.49%	0	3252	1.06%	100	708	0.23%	75	322	0.11%	0.1	245	0.08%	25	183	0.06%	50	170	0.06%	80	129	0.04%	8	123	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	52741	17.26%																																																					
Unique Count	49030	16.05%																																																					
Duplicate Count	3711	1.21%																																																					
Blank Count	228332	74.73%																																																					
Value	Count	%																																																					
Empty field	228332	74.73%																																																					
1	4554	1.49%																																																					
0	3252	1.06%																																																					
100	708	0.23%																																																					
75	322	0.11%																																																					
0.1	245	0.08%																																																					
25	183	0.06%																																																					
50	170	0.06%																																																					
80	129	0.04%																																																					
8	123	0.04%																																																					

Atributo	2009																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 2009																																																						
Exemplo	112.602370949035																																																						
Preenchido	25.54%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_009</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>53247</td> <td>17.43%</td> </tr> <tr> <td>Unique Count</td> <td>49493</td> <td>16.20%</td> </tr> <tr> <td>Duplicate Count</td> <td>3754</td> <td>1.23%</td> </tr> <tr> <td>Blank Count</td> <td>227505</td> <td>74.46%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>227505</td> <td>74.46%</td> </tr> <tr> <td>1</td> <td>4615</td> <td>1.51%</td> </tr> <tr> <td>0</td> <td>3172</td> <td>1.04%</td> </tr> <tr> <td>100</td> <td>745</td> <td>0.24%</td> </tr> <tr> <td>75</td> <td>311</td> <td>0.10%</td> </tr> <tr> <td>0.1</td> <td>252</td> <td>0.08%</td> </tr> <tr> <td>50</td> <td>157</td> <td>0.05%</td> </tr> <tr> <td>25</td> <td>157</td> <td>0.05%</td> </tr> <tr> <td>7</td> <td>138</td> <td>0.05%</td> </tr> <tr> <td>8</td> <td>137</td> <td>0.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	53247	17.43%	Unique Count	49493	16.20%	Duplicate Count	3754	1.23%	Blank Count	227505	74.46%	Value	Count	%	Empty field	227505	74.46%	1	4615	1.51%	0	3172	1.04%	100	745	0.24%	75	311	0.10%	0.1	252	0.08%	50	157	0.05%	25	157	0.05%	7	138	0.05%	8	137	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	53247	17.43%																																																					
Unique Count	49493	16.20%																																																					
Duplicate Count	3754	1.23%																																																					
Blank Count	227505	74.46%																																																					
Value	Count	%																																																					
Empty field	227505	74.46%																																																					
1	4615	1.51%																																																					
0	3172	1.04%																																																					
100	745	0.24%																																																					
75	311	0.10%																																																					
0.1	252	0.08%																																																					
50	157	0.05%																																																					
25	157	0.05%																																																					
7	138	0.05%																																																					
8	137	0.04%																																																					

Atributo	2010
Tipo de Dados	String
Descrição	Contém dados para o ano 2010
Exemplo	111.758713212606
Preenchido	31.69%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	2012
Tipo de Dados	String

Descrição	Contém dados para o ano 2012																																																						
Exemplo	109.532169608089																																																						
Preenchido	26.99%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>* Column: metadata_012</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>54887</td> <td>17.96%</td> </tr> <tr> <td>Unique Count</td> <td>50810</td> <td>16.63%</td> </tr> <tr> <td>Duplicate Count</td> <td>4077</td> <td>1.33%</td> </tr> <tr> <td>Blank Count</td> <td>223078</td> <td>73.01%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>223078</td> <td>73.01%</td> </tr> <tr> <td>1</td> <td>4733</td> <td>1.55%</td> </tr> <tr> <td>0</td> <td>3039</td> <td>0.99%</td> </tr> <tr> <td>100</td> <td>838</td> <td>0.27%</td> </tr> <tr> <td>75</td> <td>295</td> <td>0.10%</td> </tr> <tr> <td>0.1</td> <td>288</td> <td>0.09%</td> </tr> <tr> <td>7</td> <td>182</td> <td>0.06%</td> </tr> <tr> <td>50</td> <td>165</td> <td>0.05%</td> </tr> <tr> <td>8</td> <td>150</td> <td>0.05%</td> </tr> <tr> <td>6</td> <td>147</td> <td>0.05%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	54887	17.96%	Unique Count	50810	16.63%	Duplicate Count	4077	1.33%	Blank Count	223078	73.01%	Value	Count	%	Empty field	223078	73.01%	1	4733	1.55%	0	3039	0.99%	100	838	0.27%	75	295	0.10%	0.1	288	0.09%	7	182	0.06%	50	165	0.05%	8	150	0.05%	6	147	0.05%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	54887	17.96%																																																					
Unique Count	50810	16.63%																																																					
Duplicate Count	4077	1.33%																																																					
Blank Count	223078	73.01%																																																					
Value	Count	%																																																					
Empty field	223078	73.01%																																																					
1	4733	1.55%																																																					
0	3039	0.99%																																																					
100	838	0.27%																																																					
75	295	0.10%																																																					
0.1	288	0.09%																																																					
7	182	0.06%																																																					
50	165	0.05%																																																					
8	150	0.05%																																																					
6	147	0.05%																																																					

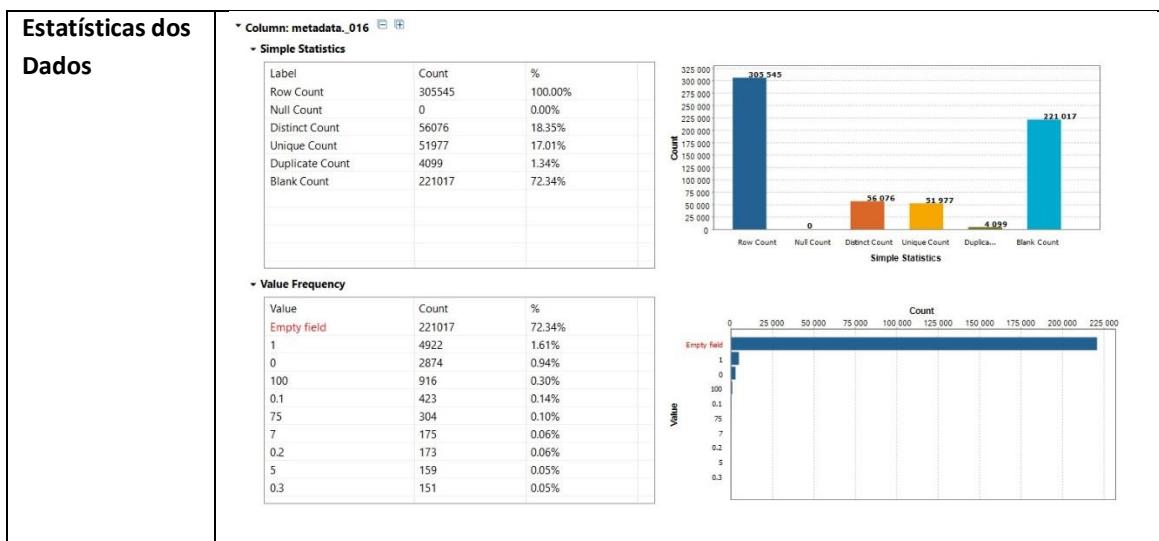
Atributo	2013
Tipo de Dados	String
Descrição	Contém dados para o ano 2013
Exemplo	108.092795855758
Preenchido	26.4%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	2014																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 2014																																																						
Exemplo	107.697714993556																																																						
Preenchido	31.78%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_014</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>58503</td><td>19.15%</td></tr> <tr><td>Unique Count</td><td>51393</td><td>16.82%</td></tr> <tr><td>Duplicate Count</td><td>7110</td><td>2.33%</td></tr> <tr><td>Blank Count</td><td>208431</td><td>68.22%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>208431</td><td>68.22%</td></tr> <tr><td>1</td><td>4827</td><td>1.58%</td></tr> <tr><td>0</td><td>3016</td><td>0.99%</td></tr> <tr><td>100</td><td>941</td><td>0.31%</td></tr> <tr><td>75</td><td>296</td><td>0.10%</td></tr> <tr><td>0.1</td><td>283</td><td>0.09%</td></tr> <tr><td>7</td><td>191</td><td>0.06%</td></tr> <tr><td>50</td><td>166</td><td>0.05%</td></tr> <tr><td>6</td><td>153</td><td>0.05%</td></tr> <tr><td>5</td><td>149</td><td>0.05%</td></tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	58503	19.15%	Unique Count	51393	16.82%	Duplicate Count	7110	2.33%	Blank Count	208431	68.22%	Value	Count	%	Empty field	208431	68.22%	1	4827	1.58%	0	3016	0.99%	100	941	0.31%	75	296	0.10%	0.1	283	0.09%	7	191	0.06%	50	166	0.05%	6	153	0.05%	5	149	0.05%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	58503	19.15%																																																					
Unique Count	51393	16.82%																																																					
Duplicate Count	7110	2.33%																																																					
Blank Count	208431	68.22%																																																					
Value	Count	%																																																					
Empty field	208431	68.22%																																																					
1	4827	1.58%																																																					
0	3016	0.99%																																																					
100	941	0.31%																																																					
75	296	0.10%																																																					
0.1	283	0.09%																																																					
7	191	0.06%																																																					
50	166	0.05%																																																					
6	153	0.05%																																																					
5	149	0.05%																																																					

Atributo	2015																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 2015																																																						
Exemplo	105.50183295561																																																						
Preenchido	31.14%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>* Column: metadata_015</p> <p>- Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>67368</td> <td>22.05%</td> </tr> <tr> <td>Unique Count</td> <td>62884</td> <td>20.58%</td> </tr> <tr> <td>Duplicate Count</td> <td>4484</td> <td>1.47%</td> </tr> <tr> <td>Blank Count</td> <td>210405</td> <td>68.86%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>210405</td> <td>68.86%</td> </tr> <tr> <td>1</td> <td>4848</td> <td>1.59%</td> </tr> <tr> <td>0</td> <td>2941</td> <td>0.96%</td> </tr> <tr> <td>100</td> <td>897</td> <td>0.29%</td> </tr> <tr> <td>75</td> <td>306</td> <td>0.10%</td> </tr> <tr> <td>0.1</td> <td>297</td> <td>0.10%</td> </tr> <tr> <td>7</td> <td>162</td> <td>0.05%</td> </tr> <tr> <td>50</td> <td>154</td> <td>0.05%</td> </tr> <tr> <td>5</td> <td>141</td> <td>0.05%</td> </tr> <tr> <td>25</td> <td>137</td> <td>0.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	67368	22.05%	Unique Count	62884	20.58%	Duplicate Count	4484	1.47%	Blank Count	210405	68.86%	Value	Count	%	Empty field	210405	68.86%	1	4848	1.59%	0	2941	0.96%	100	897	0.29%	75	306	0.10%	0.1	297	0.10%	7	162	0.05%	50	154	0.05%	5	141	0.05%	25	137	0.04%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	67368	22.05%																																																					
Unique Count	62884	20.58%																																																					
Duplicate Count	4484	1.47%																																																					
Blank Count	210405	68.86%																																																					
Value	Count	%																																																					
Empty field	210405	68.86%																																																					
1	4848	1.59%																																																					
0	2941	0.96%																																																					
100	897	0.29%																																																					
75	306	0.10%																																																					
0.1	297	0.10%																																																					
7	162	0.05%																																																					
50	154	0.05%																																																					
5	141	0.05%																																																					
25	137	0.04%																																																					

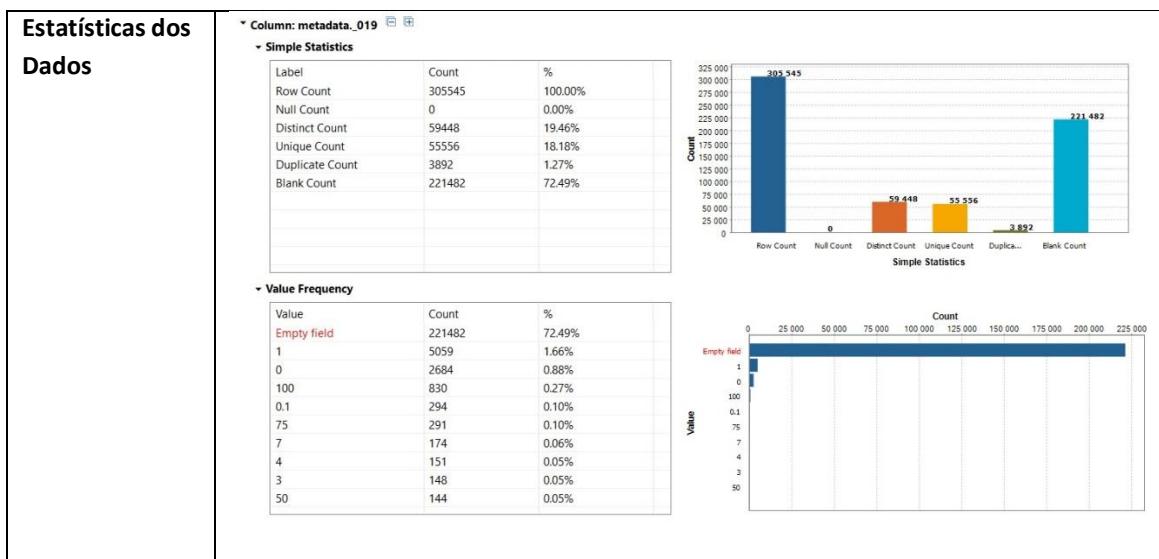
Atributo	2016
Tipo de Dados	String
Descrição	Contém dados para o ano 2016
Exemplo	103.222825498629
Preenchido	27.66%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	2017																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 2017																																																						
Exemplo	100.96352629023																																																						
Preenchido	32.51%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>* Column: metadata_017</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>59527</td><td>19.48%</td></tr> <tr><td>Unique Count</td><td>51672</td><td>16.91%</td></tr> <tr><td>Duplicate Count</td><td>7855</td><td>2.57%</td></tr> <tr><td>Blank Count</td><td>206227</td><td>67.49%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>206227</td><td>67.49%</td></tr> <tr><td>1</td><td>4934</td><td>1.61%</td></tr> <tr><td>0</td><td>2879</td><td>0.94%</td></tr> <tr><td>100</td><td>941</td><td>0.31%</td></tr> <tr><td>75</td><td>296</td><td>0.10%</td></tr> <tr><td>0.1</td><td>287</td><td>0.09%</td></tr> <tr><td>7</td><td>158</td><td>0.05%</td></tr> <tr><td>5</td><td>152</td><td>0.05%</td></tr> <tr><td>50</td><td>146</td><td>0.05%</td></tr> <tr><td>4</td><td>146</td><td>0.05%</td></tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	59527	19.48%	Unique Count	51672	16.91%	Duplicate Count	7855	2.57%	Blank Count	206227	67.49%	Value	Count	%	Empty field	206227	67.49%	1	4934	1.61%	0	2879	0.94%	100	941	0.31%	75	296	0.10%	0.1	287	0.09%	7	158	0.05%	5	152	0.05%	50	146	0.05%	4	146	0.05%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	59527	19.48%																																																					
Unique Count	51672	16.91%																																																					
Duplicate Count	7855	2.57%																																																					
Blank Count	206227	67.49%																																																					
Value	Count	%																																																					
Empty field	206227	67.49%																																																					
1	4934	1.61%																																																					
0	2879	0.94%																																																					
100	941	0.31%																																																					
75	296	0.10%																																																					
0.1	287	0.09%																																																					
7	158	0.05%																																																					
5	152	0.05%																																																					
50	146	0.05%																																																					
4	146	0.05%																																																					

Atributo	2018																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 2018																																																						
Exemplo	99.4917026482318																																																						
Preenchido	27.98%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_018</p> <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>57787</td> <td>18.91%</td> </tr> <tr> <td>Unique Count</td> <td>53627</td> <td>17.55%</td> </tr> <tr> <td>Duplicate Count</td> <td>4160</td> <td>1.36%</td> </tr> <tr> <td>Blank Count</td> <td>220060</td> <td>72.02%</td> </tr> </tbody> </table> <p>Simple Statistics</p> <ul style="list-style-type: none"> - Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>220060</td> <td>72.02%</td> </tr> <tr> <td>1</td> <td>5013</td> <td>1.64%</td> </tr> <tr> <td>0</td> <td>2797</td> <td>0.92%</td> </tr> <tr> <td>100</td> <td>860</td> <td>0.28%</td> </tr> <tr> <td>0.1</td> <td>309</td> <td>0.10%</td> </tr> <tr> <td>75</td> <td>303</td> <td>0.10%</td> </tr> <tr> <td>4</td> <td>184</td> <td>0.06%</td> </tr> <tr> <td>7</td> <td>166</td> <td>0.05%</td> </tr> <tr> <td>6</td> <td>146</td> <td>0.05%</td> </tr> <tr> <td>5</td> <td>146</td> <td>0.05%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	57787	18.91%	Unique Count	53627	17.55%	Duplicate Count	4160	1.36%	Blank Count	220060	72.02%	Value	Count	%	Empty field	220060	72.02%	1	5013	1.64%	0	2797	0.92%	100	860	0.28%	0.1	309	0.10%	75	303	0.10%	4	184	0.06%	7	166	0.05%	6	146	0.05%	5	146	0.05%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	57787	18.91%																																																					
Unique Count	53627	17.55%																																																					
Duplicate Count	4160	1.36%																																																					
Blank Count	220060	72.02%																																																					
Value	Count	%																																																					
Empty field	220060	72.02%																																																					
1	5013	1.64%																																																					
0	2797	0.92%																																																					
100	860	0.28%																																																					
0.1	309	0.10%																																																					
75	303	0.10%																																																					
4	184	0.06%																																																					
7	166	0.05%																																																					
6	146	0.05%																																																					
5	146	0.05%																																																					

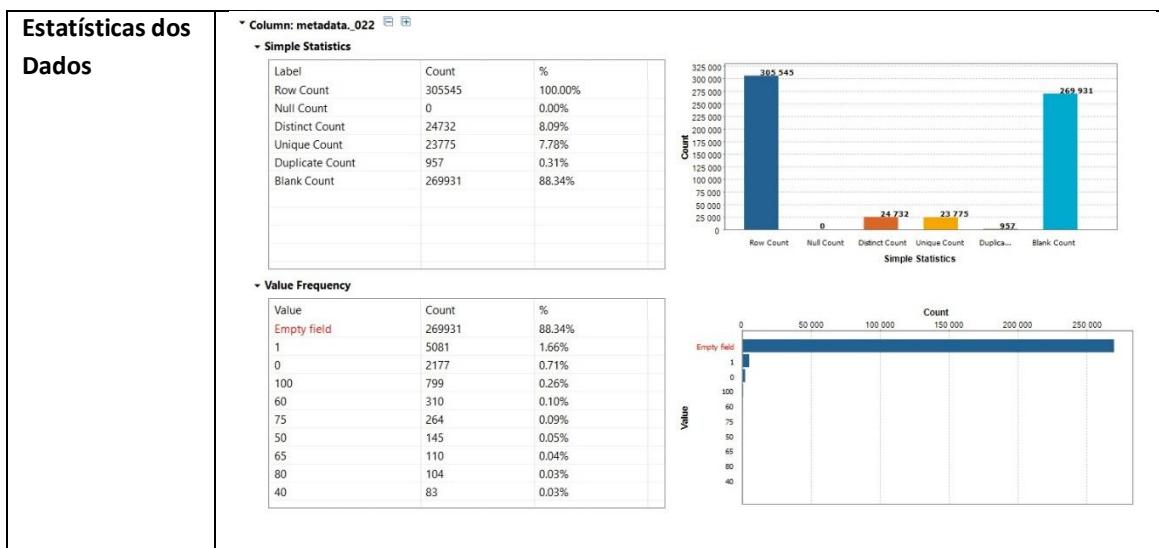
Atributo	2019
Tipo de Dados	String
Descrição	Contém dados para o ano 2019
Exemplo	99.0855722587841
Preenchido	27.51%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Atributo	2020																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 2020																																																						
Exemplo	97.4135847695511																																																						
Preenchido	22.6%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	* Column: metadata_020 <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>305545</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>50001</td><td>16.36%</td></tr> <tr><td>Unique Count</td><td>46634</td><td>15.26%</td></tr> <tr><td>Duplicate Count</td><td>3367</td><td>1.10%</td></tr> <tr><td>Blank Count</td><td>236502</td><td>77.40%</td></tr> </tbody> </table> <ul style="list-style-type: none"> - Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>236502</td><td>77.40%</td></tr> <tr><td>1</td><td>5048</td><td>1.65%</td></tr> <tr><td>0</td><td>2594</td><td>0.85%</td></tr> <tr><td>100</td><td>832</td><td>0.27%</td></tr> <tr><td>75</td><td>284</td><td>0.09%</td></tr> <tr><td>0.1</td><td>215</td><td>0.07%</td></tr> <tr><td>50</td><td>138</td><td>0.05%</td></tr> <tr><td>80</td><td>123</td><td>0.04%</td></tr> <tr><td>60</td><td>106</td><td>0.03%</td></tr> <tr><td>40</td><td>95</td><td>0.03%</td></tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	50001	16.36%	Unique Count	46634	15.26%	Duplicate Count	3367	1.10%	Blank Count	236502	77.40%	Value	Count	%	Empty field	236502	77.40%	1	5048	1.65%	0	2594	0.85%	100	832	0.27%	75	284	0.09%	0.1	215	0.07%	50	138	0.05%	80	123	0.04%	60	106	0.03%	40	95	0.03%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	50001	16.36%																																																					
Unique Count	46634	15.26%																																																					
Duplicate Count	3367	1.10%																																																					
Blank Count	236502	77.40%																																																					
Value	Count	%																																																					
Empty field	236502	77.40%																																																					
1	5048	1.65%																																																					
0	2594	0.85%																																																					
100	832	0.27%																																																					
75	284	0.09%																																																					
0.1	215	0.07%																																																					
50	138	0.05%																																																					
80	123	0.04%																																																					
60	106	0.03%																																																					
40	95	0.03%																																																					

Atributo	2021																																																						
Tipo de Dados	String																																																						
Descrição	Contém dados para o ano 2021																																																						
Exemplo	96.1814423061056																																																						
Preenchido	28.41%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null																																																						
Estatísticas dos Dados	<p>Column: metadata_021</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>305545</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>45258</td> <td>14.81%</td> </tr> <tr> <td>Unique Count</td> <td>36707</td> <td>12.01%</td> </tr> <tr> <td>Duplicate Count</td> <td>8551</td> <td>2.80%</td> </tr> <tr> <td>Blank Count</td> <td>218738</td> <td>71.59%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>218738</td> <td>71.59%</td> </tr> <tr> <td>1</td> <td>5084</td> <td>1.66%</td> </tr> <tr> <td>0</td> <td>2315</td> <td>0.76%</td> </tr> <tr> <td>100</td> <td>822</td> <td>0.27%</td> </tr> <tr> <td>75</td> <td>284</td> <td>0.09%</td> </tr> <tr> <td>0.1</td> <td>195</td> <td>0.06%</td> </tr> <tr> <td>50</td> <td>134</td> <td>0.04%</td> </tr> <tr> <td>80</td> <td>122</td> <td>0.04%</td> </tr> <tr> <td>60</td> <td>105</td> <td>0.03%</td> </tr> <tr> <td>25</td> <td>92</td> <td>0.03%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	305545	100.00%	Null Count	0	0.00%	Distinct Count	45258	14.81%	Unique Count	36707	12.01%	Duplicate Count	8551	2.80%	Blank Count	218738	71.59%	Value	Count	%	Empty field	218738	71.59%	1	5084	1.66%	0	2315	0.76%	100	822	0.27%	75	284	0.09%	0.1	195	0.06%	50	134	0.04%	80	122	0.04%	60	105	0.03%	25	92	0.03%
Label	Count	%																																																					
Row Count	305545	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	45258	14.81%																																																					
Unique Count	36707	12.01%																																																					
Duplicate Count	8551	2.80%																																																					
Blank Count	218738	71.59%																																																					
Value	Count	%																																																					
Empty field	218738	71.59%																																																					
1	5084	1.66%																																																					
0	2315	0.76%																																																					
100	822	0.27%																																																					
75	284	0.09%																																																					
0.1	195	0.06%																																																					
50	134	0.04%																																																					
80	122	0.04%																																																					
60	105	0.03%																																																					
25	92	0.03%																																																					

Atributo	2022
Tipo de Dados	String
Descrição	Contém dados para o ano 2022
Exemplo	80.3863043476913
Preenchido	11.66%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Os dados numéricos estão em formato String pelo que será adequado converter para Float. Muitos dados em branco que serão convertidos em null



Gender_Stats Series.csv

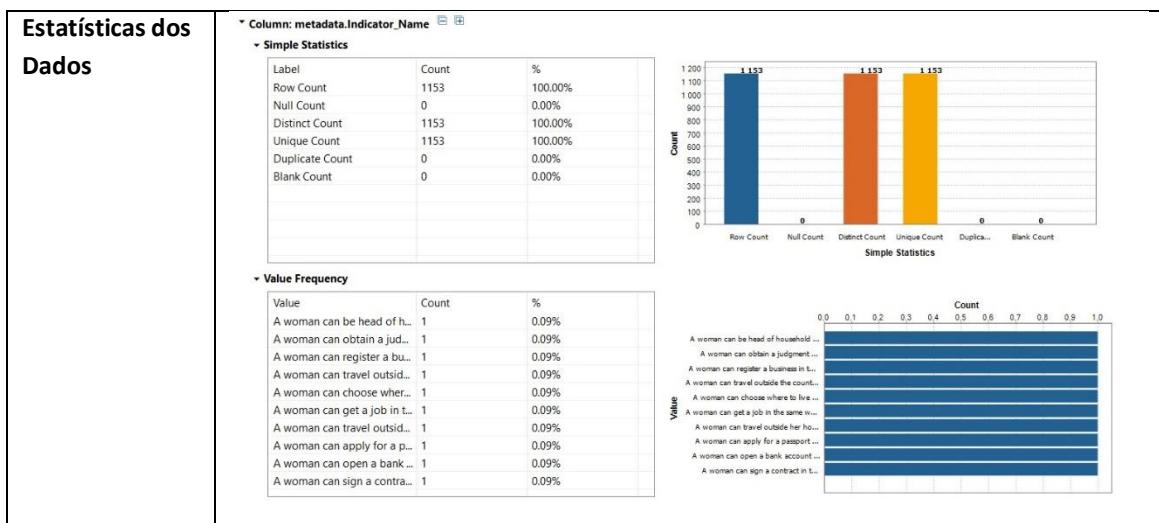
Aluno responsável: Inês Maia (a93347)

Descrição: O seguinte csv descreve uma série de dados de um tópico específico ou estatística.

Atributo	_Series_Code																																																						
Tipo de Dados	String																																																						
Descrição	Código único para cada série de dados																																																						
Exemplo	HD.HCI.AMRT																																																						
Preenchido	100%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Não detetados.																																																						
Estatísticas dos Dados	<p>* Column: metadata_Series_Code</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>1153</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>1153</td><td>100.00%</td></tr> <tr><td>Unique Count</td><td>1153</td><td>100.00%</td></tr> <tr><td>Duplicate Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Blank Count</td><td>0</td><td>0.00%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>HD.HCLAMRT.MA</td><td>1</td><td>0.09%</td></tr> <tr><td>HD.HCLEVRS.FE</td><td>1</td><td>0.09%</td></tr> <tr><td>FP.CPL.TOTLZG</td><td>1</td><td>0.09%</td></tr> <tr><td>HD.HCLEVRS</td><td>1</td><td>0.09%</td></tr> <tr><td>HD.HCLI.HLOS.MA</td><td>1</td><td>0.09%</td></tr> <tr><td>HD.HCLI.HLOS</td><td>1</td><td>0.09%</td></tr> <tr><td>HD.HCLI.HLOS.FE</td><td>1</td><td>0.09%</td></tr> <tr><td>HD.HCLAMRT.FE</td><td>1</td><td>0.09%</td></tr> <tr><td>HD.HCLEVRS.MA</td><td>1</td><td>0.09%</td></tr> <tr><td>HD.HCLAMRT</td><td>1</td><td>0.09%</td></tr> </tbody> </table> <p>Count</p> <p>Row Count Null Count Distinct Count Unique Count Duplicate... Blank Count</p> <p>Simple Statistics</p> <p>Value</p> <p>HD.HCLAMRT.MA HD.HCLEVRS.FE FP.CPL.TOTLZG HD.HCLEVRS HD.HCLI.HLOS.MA HD.HCLI.HLOS HD.HCLI.HLOS.FE HD.HCLAMRT.FE HD.HCLEVRS.MA HD.HCLAMRT</p> <p>Count</p>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	1153	100.00%	Unique Count	1153	100.00%	Duplicate Count	0	0.00%	Blank Count	0	0.00%	Value	Count	%	HD.HCLAMRT.MA	1	0.09%	HD.HCLEVRS.FE	1	0.09%	FP.CPL.TOTLZG	1	0.09%	HD.HCLEVRS	1	0.09%	HD.HCLI.HLOS.MA	1	0.09%	HD.HCLI.HLOS	1	0.09%	HD.HCLI.HLOS.FE	1	0.09%	HD.HCLAMRT.FE	1	0.09%	HD.HCLEVRS.MA	1	0.09%	HD.HCLAMRT	1	0.09%
Label	Count	%																																																					
Row Count	1153	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	1153	100.00%																																																					
Unique Count	1153	100.00%																																																					
Duplicate Count	0	0.00%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
HD.HCLAMRT.MA	1	0.09%																																																					
HD.HCLEVRS.FE	1	0.09%																																																					
FP.CPL.TOTLZG	1	0.09%																																																					
HD.HCLEVRS	1	0.09%																																																					
HD.HCLI.HLOS.MA	1	0.09%																																																					
HD.HCLI.HLOS	1	0.09%																																																					
HD.HCLI.HLOS.FE	1	0.09%																																																					
HD.HCLAMRT.FE	1	0.09%																																																					
HD.HCLEVRS.MA	1	0.09%																																																					
HD.HCLAMRT	1	0.09%																																																					

Atributo	Topic																																																						
Tipo de Dados	String																																																						
Descrição	Tópico específico ao qual a série pertence																																																						
Exemplo	Assets																																																						
Preenchido	100%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Não detetados.																																																						
Estatísticas dos Dados	<p>* Column: metadata.Topic</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>1153</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>43</td> <td>3.73%</td> </tr> <tr> <td>Unique Count</td> <td>9</td> <td>0.78%</td> </tr> <tr> <td>Duplicate Count</td> <td>34</td> <td>2.95%</td> </tr> <tr> <td>Blank Count</td> <td>0</td> <td>0.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Assets</td> <td>243</td> <td>21.08%</td> </tr> <tr> <td>Gender</td> <td>153</td> <td>13.27%</td> </tr> <tr> <td>Health: Risk factors</td> <td>103</td> <td>8.93%</td> </tr> <tr> <td>Health: Population: Structure</td> <td>66</td> <td>5.72%</td> </tr> <tr> <td>Violence</td> <td>53</td> <td>4.60%</td> </tr> <tr> <td>Education</td> <td>48</td> <td>4.16%</td> </tr> <tr> <td>Social Protection & Labor: ...</td> <td>44</td> <td>3.82%</td> </tr> <tr> <td>Education: Outcomes</td> <td>43</td> <td>3.73%</td> </tr> <tr> <td>Health: Mortality</td> <td>40</td> <td>3.47%</td> </tr> <tr> <td>Education: Participation</td> <td>35</td> <td>3.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	43	3.73%	Unique Count	9	0.78%	Duplicate Count	34	2.95%	Blank Count	0	0.00%	Value	Count	%	Assets	243	21.08%	Gender	153	13.27%	Health: Risk factors	103	8.93%	Health: Population: Structure	66	5.72%	Violence	53	4.60%	Education	48	4.16%	Social Protection & Labor: ...	44	3.82%	Education: Outcomes	43	3.73%	Health: Mortality	40	3.47%	Education: Participation	35	3.04%
Label	Count	%																																																					
Row Count	1153	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	43	3.73%																																																					
Unique Count	9	0.78%																																																					
Duplicate Count	34	2.95%																																																					
Blank Count	0	0.00%																																																					
Value	Count	%																																																					
Assets	243	21.08%																																																					
Gender	153	13.27%																																																					
Health: Risk factors	103	8.93%																																																					
Health: Population: Structure	66	5.72%																																																					
Violence	53	4.60%																																																					
Education	48	4.16%																																																					
Social Protection & Labor: ...	44	3.82%																																																					
Education: Outcomes	43	3.73%																																																					
Health: Mortality	40	3.47%																																																					
Education: Participation	35	3.04%																																																					

Atributo	Indicator_Name
Tipo de Dados	String
Descrição	Nome do indicador estatístico
Exemplo	Account (% age 15+)
Preenchido	100%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Não detetados.

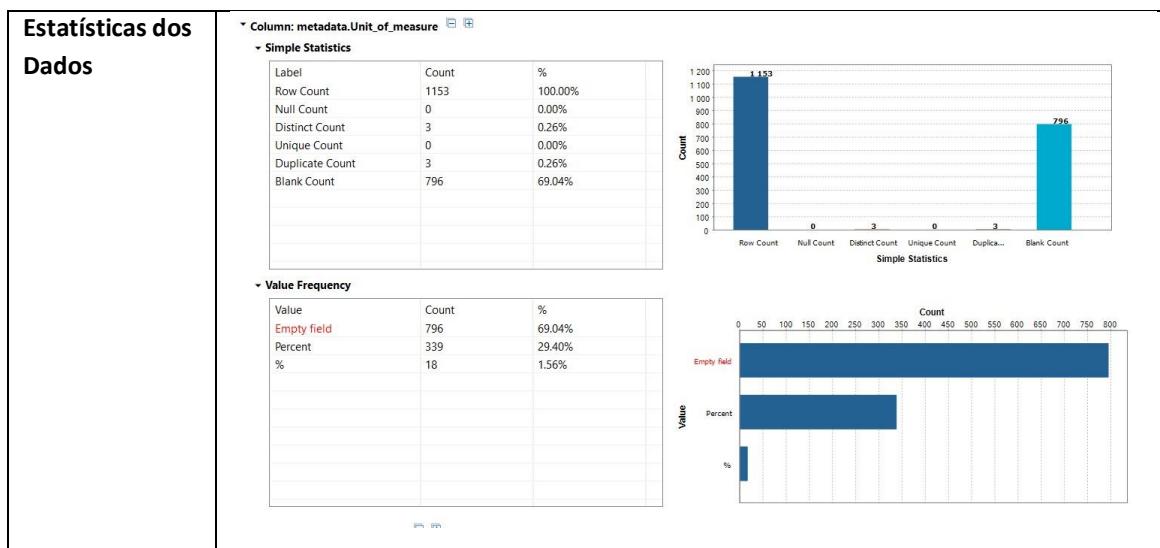


Atributo	Short_definition																																																						
Tipo de Dados	String																																																						
Descrição	Definição curta do indicador																																																						
Exemplo	Among respondents who report having an inactive account, the percentage who report that they don't use their account because the bank or financial institution is too far away.																																																						
Preenchido	25.67%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Existem 857 linhas em branco que serão passadas a "Sem informação"																																																						
Estatísticas dos Dados	<p>* Column: metadata.Short_definition</p> <p>- Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>1153</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>287</td><td>24.89%</td></tr> <tr><td>Unique Count</td><td>284</td><td>24.63%</td></tr> <tr><td>Duplicate Count</td><td>3</td><td>0.26%</td></tr> <tr><td>Blank Count</td><td>857</td><td>74.33%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Empty field</td><td>857</td><td>74.33%</td></tr> <tr><td>Age population, female ref...</td><td>6</td><td>0.52%</td></tr> <tr><td>Age population, male refer...</td><td>6</td><td>0.52%</td></tr> <tr><td>Among respondents who r...</td><td>1</td><td>0.09%</td></tr> <tr><td>Among respondents who r...</td><td>1</td><td>0.09%</td></tr> <tr><td>Among respondents who r...</td><td>1</td><td>0.09%</td></tr> <tr><td>Among respondents report...</td><td>1</td><td>0.09%</td></tr> <tr><td>Among respondents who r...</td><td>1</td><td>0.09%</td></tr> <tr><td>Among respondents who r...</td><td>1</td><td>0.09%</td></tr> <tr><td>Among respondents who r...</td><td>1</td><td>0.09%</td></tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	287	24.89%	Unique Count	284	24.63%	Duplicate Count	3	0.26%	Blank Count	857	74.33%	Value	Count	%	Empty field	857	74.33%	Age population, female ref...	6	0.52%	Age population, male refer...	6	0.52%	Among respondents who r...	1	0.09%	Among respondents who r...	1	0.09%	Among respondents who r...	1	0.09%	Among respondents report...	1	0.09%	Among respondents who r...	1	0.09%	Among respondents who r...	1	0.09%	Among respondents who r...	1	0.09%
Label	Count	%																																																					
Row Count	1153	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	287	24.89%																																																					
Unique Count	284	24.63%																																																					
Duplicate Count	3	0.26%																																																					
Blank Count	857	74.33%																																																					
Value	Count	%																																																					
Empty field	857	74.33%																																																					
Age population, female ref...	6	0.52%																																																					
Age population, male refer...	6	0.52%																																																					
Among respondents who r...	1	0.09%																																																					
Among respondents who r...	1	0.09%																																																					
Among respondents who r...	1	0.09%																																																					
Among respondents report...	1	0.09%																																																					
Among respondents who r...	1	0.09%																																																					
Among respondents who r...	1	0.09%																																																					
Among respondents who r...	1	0.09%																																																					

Atributo	Long_definition
-----------------	-----------------

Tipo de Dados	String																																																						
Descrição	Definição extensa do indicador																																																						
Exemplo	The percentage of respondents who report having an account (by themselves or together with someone else) at a bank or another type of financial institution (see the definition for "financial institution account") or report personally using a mobile money service in the past year (see the definition for "mobile money account").																																																						
Preenchido	99.83%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Existem 2 linhas em branco que serão passadas a "Sem informação"																																																						
Estatísticas dos Dados	<p>Column metadata.Long_definition</p> <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>1153</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>902</td> <td>78.23%</td> </tr> <tr> <td>Unique Count</td> <td>797</td> <td>69.12%</td> </tr> <tr> <td>Duplicate Count</td> <td>105</td> <td>9.11%</td> </tr> <tr> <td>Blank Count</td> <td>2</td> <td>0.17%</td> </tr> </tbody> </table> <ul style="list-style-type: none"> - Value Frequency <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Female share of graduates ...</td> <td>13</td> <td>1.13%</td> </tr> <tr> <td>Age population, female ref...</td> <td>6</td> <td>0.52%</td> </tr> <tr> <td>Employment to population...</td> <td>6</td> <td>0.52%</td> </tr> <tr> <td>Age population, male refer...</td> <td>6</td> <td>0.52%</td> </tr> <tr> <td>Employment to population...</td> <td>6</td> <td>0.52%</td> </tr> <tr> <td>Labor force participation ...</td> <td>6</td> <td>0.52%</td> </tr> <tr> <td>Labor force participation r...</td> <td>6</td> <td>0.52%</td> </tr> <tr> <td>Percentage of women age...</td> <td>5</td> <td>0.43%</td> </tr> <tr> <td>Percentage of women age...</td> <td>5</td> <td>0.43%</td> </tr> <tr> <td>Percentage of women age...</td> <td>5</td> <td>0.43%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	902	78.23%	Unique Count	797	69.12%	Duplicate Count	105	9.11%	Blank Count	2	0.17%	Value	Count	%	Female share of graduates ...	13	1.13%	Age population, female ref...	6	0.52%	Employment to population...	6	0.52%	Age population, male refer...	6	0.52%	Employment to population...	6	0.52%	Labor force participation ...	6	0.52%	Labor force participation r...	6	0.52%	Percentage of women age...	5	0.43%	Percentage of women age...	5	0.43%	Percentage of women age...	5	0.43%
Label	Count	%																																																					
Row Count	1153	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	902	78.23%																																																					
Unique Count	797	69.12%																																																					
Duplicate Count	105	9.11%																																																					
Blank Count	2	0.17%																																																					
Value	Count	%																																																					
Female share of graduates ...	13	1.13%																																																					
Age population, female ref...	6	0.52%																																																					
Employment to population...	6	0.52%																																																					
Age population, male refer...	6	0.52%																																																					
Employment to population...	6	0.52%																																																					
Labor force participation ...	6	0.52%																																																					
Labor force participation r...	6	0.52%																																																					
Percentage of women age...	5	0.43%																																																					
Percentage of women age...	5	0.43%																																																					
Percentage of women age...	5	0.43%																																																					

Atributo	Unit_of_measure
Tipo de Dados	String
Descrição	Unidade de medida do indicador
Exemplo	Percent
Preenchido	30.96%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Existem 796 linhas em brancos correspondentes aos indicadores que não são percentagens, que serão passadas a "Sem informação"



Atributo	Periodicity																																				
Tipo de Dados	String																																				
Descrição	A frequência com que o indicador é medido																																				
Exemplo	Triennial																																				
Preenchido	93.93%																																				
Intervalo de Dados	Não mensurável																																				
Problemas de Qualidade dos Dados	Existem 70 linhas em brancos que serão passadas a "Sem informação"																																				
Estatísticas dos Dados	<p>* Column: metadata.Periodicity</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>1153</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>4</td> <td>0.35%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>4</td> <td>0.35%</td> </tr> <tr> <td>Blank Count</td> <td>70</td> <td>6.07%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Annual</td> <td>806</td> <td>69.90%</td> </tr> <tr> <td>Triennial</td> <td>274</td> <td>23.76%</td> </tr> <tr> <td>Empty field</td> <td>70</td> <td>6.07%</td> </tr> <tr> <td>Biennial</td> <td>3</td> <td>0.26%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	4	0.35%	Unique Count	0	0.00%	Duplicate Count	4	0.35%	Blank Count	70	6.07%	Value	Count	%	Annual	806	69.90%	Triennial	274	23.76%	Empty field	70	6.07%	Biennial	3	0.26%
Label	Count	%																																			
Row Count	1153	100.00%																																			
Null Count	0	0.00%																																			
Distinct Count	4	0.35%																																			
Unique Count	0	0.00%																																			
Duplicate Count	4	0.35%																																			
Blank Count	70	6.07%																																			
Value	Count	%																																			
Annual	806	69.90%																																			
Triennial	274	23.76%																																			
Empty field	70	6.07%																																			
Biennial	3	0.26%																																			

Atributo	Base_Period
Tipo de Dados	String

Descrição	Período de tempo usado como base para o indicador																											
Exemplo																												
Preenchido	0%																											
Intervalo de Dados	Não mensurável																											
Problemas de Qualidade dos Dados	Todas as linhas encontram-se em branco que serão passadas a “Sem informação”																											
Estatísticas dos Dados	<p>Column: metadata.Base_Period</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>1153</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>1</td> <td>0.09%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>1</td> <td>0.09%</td> </tr> <tr> <td>Blank Count</td> <td>1153</td> <td>100.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>1153</td> <td>100.00%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	1	0.09%	Unique Count	0	0.00%	Duplicate Count	1	0.09%	Blank Count	1153	100.00%	Value	Count	%	Empty field	1153	100.00%
Label	Count	%																										
Row Count	1153	100.00%																										
Null Count	0	0.00%																										
Distinct Count	1	0.09%																										
Unique Count	0	0.00%																										
Duplicate Count	1	0.09%																										
Blank Count	1153	100.00%																										
Value	Count	%																										
Empty field	1153	100.00%																										

Atributo	Other_Notes
Tipo de Dados	String
Descrição	Notas relevantes sobre o indicador
Exemplo	
Preenchido	0%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Todas as linhas encontram-se em branco que serão passadas a “Sem informação”

Estatísticas dos Dados	<p>Column: metadata.Other_notes</p> <ul style="list-style-type: none"> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th><th>Count</th><th>%</th></tr> </thead> <tbody> <tr> <td>Row Count</td><td>1153</td><td>100.00%</td></tr> <tr> <td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr> <td>Distinct Count</td><td>1</td><td>0.09%</td></tr> <tr> <td>Unique Count</td><td>0</td><td>0.00%</td></tr> <tr> <td>Duplicate Count</td><td>1</td><td>0.09%</td></tr> <tr> <td>Blank Count</td><td>1153</td><td>100.00%</td></tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	1	0.09%	Unique Count	0	0.00%	Duplicate Count	1	0.09%	Blank Count	1153	100.00%	
Label	Count	%																					
Row Count	1153	100.00%																					
Null Count	0	0.00%																					
Distinct Count	1	0.09%																					
Unique Count	0	0.00%																					
Duplicate Count	1	0.09%																					
Blank Count	1153	100.00%																					
<p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>1153</td> <td>100.00%</td> </tr> </tbody> </table>	Value	Count	%	Empty field	1153	100.00%																	
Value	Count	%																					
Empty field	1153	100.00%																					

Atributo	Aggregation_method																																																				
Tipo de Dados	String																																																				
Descrição	Descreve como os dados do indicador são agregados																																																				
Exemplo	Weighted average																																																				
Preenchido	57.42%																																																				
Intervalo de Dados	Não mensurável																																																				
Problemas de Qualidade dos Dados	Existem 491 linhas em branco que serão passadas a "Sem informação"																																																				
Estatísticas dos Dados	<p>Column: metadata.Aggregation_method</p> <ul style="list-style-type: none"> Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>1153</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>9</td> <td>0.78%</td> </tr> <tr> <td>Unique Count</td> <td>2</td> <td>0.17%</td> </tr> <tr> <td>Duplicate Count</td> <td>7</td> <td>0.61%</td> </tr> <tr> <td>Blank Count</td> <td>491</td> <td>42.58%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Weighted average</td> <td>581</td> <td>50.39%</td> </tr> <tr> <td>Empty field</td> <td>491</td> <td>42.58%</td> </tr> <tr> <td>Sum</td> <td>59</td> <td>5.12%</td> </tr> <tr> <td>Unweighted average</td> <td>11</td> <td>0.95%</td> </tr> <tr> <td>Median</td> <td>4</td> <td>0.35%</td> </tr> <tr> <td>Linear mixed-effect model estimates</td> <td>3</td> <td>0.26%</td> </tr> <tr> <td>Gap-filled total</td> <td>2</td> <td>0.17%</td> </tr> <tr> <td>See SH.STA.STNT.ME.ZS for...</td> <td>1</td> <td>0.09%</td> </tr> <tr> <td>See SH.STA.OVGH.ME.ZS ...</td> <td>1</td> <td>0.09%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	9	0.78%	Unique Count	2	0.17%	Duplicate Count	7	0.61%	Blank Count	491	42.58%	Value	Count	%	Weighted average	581	50.39%	Empty field	491	42.58%	Sum	59	5.12%	Unweighted average	11	0.95%	Median	4	0.35%	Linear mixed-effect model estimates	3	0.26%	Gap-filled total	2	0.17%	See SH.STA.STNT.ME.ZS for...	1	0.09%	See SH.STA.OVGH.ME.ZS ...	1	0.09%	
Label	Count	%																																																			
Row Count	1153	100.00%																																																			
Null Count	0	0.00%																																																			
Distinct Count	9	0.78%																																																			
Unique Count	2	0.17%																																																			
Duplicate Count	7	0.61%																																																			
Blank Count	491	42.58%																																																			
Value	Count	%																																																			
Weighted average	581	50.39%																																																			
Empty field	491	42.58%																																																			
Sum	59	5.12%																																																			
Unweighted average	11	0.95%																																																			
Median	4	0.35%																																																			
Linear mixed-effect model estimates	3	0.26%																																																			
Gap-filled total	2	0.17%																																																			
See SH.STA.STNT.ME.ZS for...	1	0.09%																																																			
See SH.STA.OVGH.ME.ZS ...	1	0.09%																																																			

Atributo	Limitations_and_exceptions
Tipo de Dados	String

Descrição	Limitações e exceções na medição do indicador																																																						
Exemplo	<p>The Doing Business methodology has limitations that should be considered when interpreting the data. First, the data collected refer to businesses in the economy's largest city and may not represent regulations in other locations of the economy. To address this limitation, subnational indicators are being collected for selected economies. These subnational studies point to significant differences in the speed of reform and the ease of doing business across cities in the same economy. Second, the data often focus on a specific business form - generally a limited liability company of a specified size - and may not represent regulation for other types of businesses such as sole proprietorships. Third, transactions described in a standardized business case refer to a specific set of issues and may not represent the full set of issues a business encounters. Fourth, the time measures involve an element of judgment by the expert respondents. When sources indicate different estimates, the Doing Business time indicators represent the median values of several responses given under the assumptions of the standardized case. Fifth, the methodology assumes that a business has full information on what is required and does not waste time when completing procedures. Please also see: https://www.doingbusiness.org/en/about-us/faq</p>																																																						
Preenchido	35.99%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Existem 738 linhas em branco que serão passadas a "Sem informação"																																																						
Estatísticas dos Dados	<p>Column: metadata.Limitations_and_exceptions</p> <ul style="list-style-type: none"> - Simple Statistics <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>1153</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>77</td> <td>6.68%</td> </tr> <tr> <td>Unique Count</td> <td>23</td> <td>1.99%</td> </tr> <tr> <td>Duplicate Count</td> <td>54</td> <td>4.68%</td> </tr> <tr> <td>Blank Count</td> <td>738</td> <td>64.01%</td> </tr> </tbody> </table> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>738</td> <td>64.01%</td> </tr> <tr> <td>The Women, Business and t...</td> <td>55</td> <td>4.77%</td> </tr> <tr> <td>Collecting valid, reliable, a...</td> <td>47</td> <td>4.08%</td> </tr> <tr> <td>Caution is required when u...</td> <td>24</td> <td>2.08%</td> </tr> <tr> <td>Enrollment indicators are b...</td> <td>21</td> <td>1.82%</td> </tr> <tr> <td>The limited availability of ...</td> <td>19</td> <td>1.65%</td> </tr> <tr> <td>Data on the labor force are...</td> <td>17</td> <td>1.47%</td> </tr> <tr> <td>The limited availability of ...</td> <td>14</td> <td>1.21%</td> </tr> <tr> <td>Data are drawn from labor ...</td> <td>12</td> <td>1.04%</td> </tr> <tr> <td>Data on employment by st...</td> <td>12</td> <td>1.04%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	77	6.68%	Unique Count	23	1.99%	Duplicate Count	54	4.68%	Blank Count	738	64.01%	Value	Count	%	Empty field	738	64.01%	The Women, Business and t...	55	4.77%	Collecting valid, reliable, a...	47	4.08%	Caution is required when u...	24	2.08%	Enrollment indicators are b...	21	1.82%	The limited availability of ...	19	1.65%	Data on the labor force are...	17	1.47%	The limited availability of ...	14	1.21%	Data are drawn from labor ...	12	1.04%	Data on employment by st...	12	1.04%
Label	Count	%																																																					
Row Count	1153	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	77	6.68%																																																					
Unique Count	23	1.99%																																																					
Duplicate Count	54	4.68%																																																					
Blank Count	738	64.01%																																																					
Value	Count	%																																																					
Empty field	738	64.01%																																																					
The Women, Business and t...	55	4.77%																																																					
Collecting valid, reliable, a...	47	4.08%																																																					
Caution is required when u...	24	2.08%																																																					
Enrollment indicators are b...	21	1.82%																																																					
The limited availability of ...	19	1.65%																																																					
Data on the labor force are...	17	1.47%																																																					
The limited availability of ...	14	1.21%																																																					
Data are drawn from labor ...	12	1.04%																																																					
Data on employment by st...	12	1.04%																																																					

Atributo	Notes_from_original_source
Tipo de Dados	String
Descrição	Notas da fonte original dos dados

Exemplo	All surveys were administered using the Enterprise Surveys methodology as outlined in the Methodology page which can be found from www.enterprisesurveys.org .																																																						
Preenchido	9.89%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Existem 1039 linhas em branco que serão passadas a "Sem informação"																																																						
Estatísticas dos Dados	<p>Column: metadata.Notes_from_original_source</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>1153</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>21</td> <td>1.82%</td> </tr> <tr> <td>Unique Count</td> <td>10</td> <td>0.87%</td> </tr> <tr> <td>Duplicate Count</td> <td>11</td> <td>0.95%</td> </tr> <tr> <td>Blank Count</td> <td>1039</td> <td>90.11%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>1039</td> <td>90.11%</td> </tr> <tr> <td>Given the exceptional situ...</td> <td>48</td> <td>4.16%</td> </tr> <tr> <td>This is one of the 35 scored...</td> <td>29</td> <td>2.52%</td> </tr> <tr> <td>Breastfeeding, prolonged ...</td> <td>8</td> <td>0.69%</td> </tr> <tr> <td>For more information please...</td> <td>6</td> <td>0.52%</td> </tr> <tr> <td>This indicator is additional ...</td> <td>3</td> <td>0.26%</td> </tr> <tr> <td>Note that the SMAM takes ...</td> <td>2</td> <td>0.17%</td> </tr> <tr> <td>This indicator is additional ...</td> <td>2</td> <td>0.17%</td> </tr> <tr> <td>All surveys were administer...</td> <td>2</td> <td>0.17%</td> </tr> <tr> <td>This indicator is additional ...</td> <td>2</td> <td>0.17%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	21	1.82%	Unique Count	10	0.87%	Duplicate Count	11	0.95%	Blank Count	1039	90.11%	Value	Count	%	Empty field	1039	90.11%	Given the exceptional situ...	48	4.16%	This is one of the 35 scored...	29	2.52%	Breastfeeding, prolonged ...	8	0.69%	For more information please...	6	0.52%	This indicator is additional ...	3	0.26%	Note that the SMAM takes ...	2	0.17%	This indicator is additional ...	2	0.17%	All surveys were administer...	2	0.17%	This indicator is additional ...	2	0.17%
Label	Count	%																																																					
Row Count	1153	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	21	1.82%																																																					
Unique Count	10	0.87%																																																					
Duplicate Count	11	0.95%																																																					
Blank Count	1039	90.11%																																																					
Value	Count	%																																																					
Empty field	1039	90.11%																																																					
Given the exceptional situ...	48	4.16%																																																					
This is one of the 35 scored...	29	2.52%																																																					
Breastfeeding, prolonged ...	8	0.69%																																																					
For more information please...	6	0.52%																																																					
This indicator is additional ...	3	0.26%																																																					
Note that the SMAM takes ...	2	0.17%																																																					
This indicator is additional ...	2	0.17%																																																					
All surveys were administer...	2	0.17%																																																					
This indicator is additional ...	2	0.17%																																																					

Atributo	General_comments
Tipo de Dados	String
Descrição	Comentários gerais acerca do indicador
Exemplo	Relevance to gender indicator: Women are vastly underrepresented in decision making positions at the top level in the private sector and this indicator monitors progress that has been made.
Preenchido	43.19%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Existem 655 linhas em branco que serão passadas a "Sem informação"

Estatísticas dos Dados	<p>Column: metadata.General_comments</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th><th>Count</th><th>%</th></tr> </thead> <tbody> <tr><td>Row Count</td><td>1153</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>83</td><td>7.20%</td></tr> <tr><td>Unique Count</td><td>49</td><td>4.25%</td></tr> <tr><td>Duplicate Count</td><td>34</td><td>2.95%</td></tr> <tr><td>Blank Count</td><td>655</td><td>56.81%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th><th>Count</th><th>%</th></tr> </thead> <tbody> <tr><td>Empty field</td><td>655</td><td>56.81%</td></tr> <tr><td>2021 variables</td><td>161</td><td>13.96%</td></tr> <tr><td>Time series variables</td><td>90</td><td>7.81%</td></tr> <tr><td>For the reference period, W...</td><td>43</td><td>3.73%</td></tr> <tr><td>National estimates are also...</td><td>23</td><td>1.99%</td></tr> <tr><td>The series for ILD estimates...</td><td>20</td><td>1.73%</td></tr> <tr><td>2014 variables</td><td>12</td><td>1.04%</td></tr> <tr><td>Undernourished children have...</td><td>12</td><td>1.04%</td></tr> <tr><td>2017 variables</td><td>9</td><td>0.78%</td></tr> <tr><td>Data are presented for the ...</td><td>9</td><td>0.78%</td></tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	83	7.20%	Unique Count	49	4.25%	Duplicate Count	34	2.95%	Blank Count	655	56.81%	Value	Count	%	Empty field	655	56.81%	2021 variables	161	13.96%	Time series variables	90	7.81%	For the reference period, W...	43	3.73%	National estimates are also...	23	1.99%	The series for ILD estimates...	20	1.73%	2014 variables	12	1.04%	Undernourished children have...	12	1.04%	2017 variables	9	0.78%	Data are presented for the ...	9	0.78%	<p>Simple Statistics</p>
Label	Count	%																																																						
Row Count	1153	100.00%																																																						
Null Count	0	0.00%																																																						
Distinct Count	83	7.20%																																																						
Unique Count	49	4.25%																																																						
Duplicate Count	34	2.95%																																																						
Blank Count	655	56.81%																																																						
Value	Count	%																																																						
Empty field	655	56.81%																																																						
2021 variables	161	13.96%																																																						
Time series variables	90	7.81%																																																						
For the reference period, W...	43	3.73%																																																						
National estimates are also...	23	1.99%																																																						
The series for ILD estimates...	20	1.73%																																																						
2014 variables	12	1.04%																																																						
Undernourished children have...	12	1.04%																																																						
2017 variables	9	0.78%																																																						
Data are presented for the ...	9	0.78%																																																						
	<p>Value Frequency</p>																																																							

Atributo	Source																																																							
Tipo de Dados	String																																																							
Descrição	Fonte dos dados																																																							
Exemplo	UNICEF-WHO-World Bank Joint Malnutrition Estimates, supplemented with data provided by World Bank staff.																																																							
Preenchido	99.65%																																																							
Intervalo de Dados	Não mensurável																																																							
Problemas de Qualidade dos Dados	Existem 4 linhas em branco que serão passadas a "Sem informação"																																																							
Estatísticas dos Dados	<p>Column: metadata.Source</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Row Count</td><td>1153</td><td>100.00%</td></tr> <tr><td>Null Count</td><td>0</td><td>0.00%</td></tr> <tr><td>Distinct Count</td><td>88</td><td>7.63%</td></tr> <tr><td>Unique Count</td><td>27</td><td>2.34%</td></tr> <tr><td>Duplicate Count</td><td>61</td><td>5.29%</td></tr> <tr><td>Blank Count</td><td>4</td><td>0.35%</td></tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr><td>Global Index database</td><td>272</td><td>23.59%</td></tr> <tr><td>Demographic and Health S...</td><td>194</td><td>16.83%</td></tr> <tr><td>UNESCO Institute for Statistic...</td><td>70</td><td>6.07%</td></tr> <tr><td>World Bank Women, Business...</td><td>55</td><td>4.77%</td></tr> <tr><td>World Bank staff estimates us...</td><td>45</td><td>3.90%</td></tr> <tr><td>Derived based on the data ...</td><td>45</td><td>3.90%</td></tr> <tr><td>Demographic and Health S...</td><td>36</td><td>3.12%</td></tr> <tr><td>World Health Organization, ...</td><td>27</td><td>2.34%</td></tr> <tr><td>International Labour Organizat...</td><td>23</td><td>1.99%</td></tr> <tr><td>UNESCO Institute for Statistic...</td><td>23</td><td>1.99%</td></tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	88	7.63%	Unique Count	27	2.34%	Duplicate Count	61	5.29%	Blank Count	4	0.35%	Value	Count	%	Global Index database	272	23.59%	Demographic and Health S...	194	16.83%	UNESCO Institute for Statistic...	70	6.07%	World Bank Women, Business...	55	4.77%	World Bank staff estimates us...	45	3.90%	Derived based on the data ...	45	3.90%	Demographic and Health S...	36	3.12%	World Health Organization, ...	27	2.34%	International Labour Organizat...	23	1.99%	UNESCO Institute for Statistic...	23	1.99%	<p>Simple Statistics</p>
Label	Count	%																																																						
Row Count	1153	100.00%																																																						
Null Count	0	0.00%																																																						
Distinct Count	88	7.63%																																																						
Unique Count	27	2.34%																																																						
Duplicate Count	61	5.29%																																																						
Blank Count	4	0.35%																																																						
Value	Count	%																																																						
Global Index database	272	23.59%																																																						
Demographic and Health S...	194	16.83%																																																						
UNESCO Institute for Statistic...	70	6.07%																																																						
World Bank Women, Business...	55	4.77%																																																						
World Bank staff estimates us...	45	3.90%																																																						
Derived based on the data ...	45	3.90%																																																						
Demographic and Health S...	36	3.12%																																																						
World Health Organization, ...	27	2.34%																																																						
International Labour Organizat...	23	1.99%																																																						
UNESCO Institute for Statistic...	23	1.99%																																																						

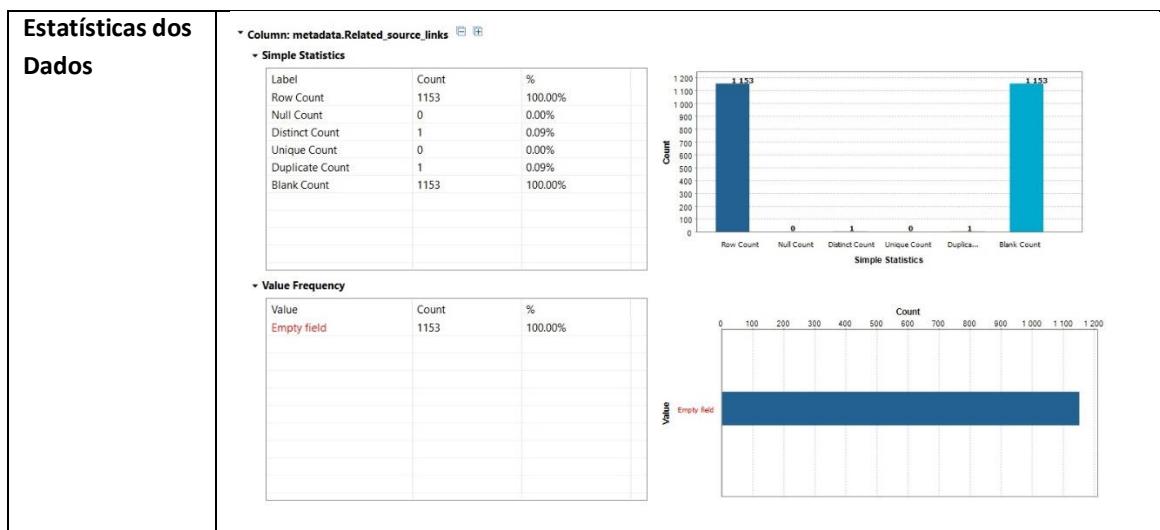
Atributo	Statistical_concept_and_methodology
-----------------	-------------------------------------

Tipo de Dados	String																																																						
Descrição	Conceito estatístico e metodologia usada para a recolha dos dados																																																						
Exemplo	This indicator measures the number of procedures for a small- to medium-size limited liability company to start up and formally operate in each economy's largest business city. To make the data comparable across 190 economies, Doing Business uses a standardized business that is 100% domestically owned, has a start-up capital equivalent to 10 times the income per capita, engages in general industrial or commercial activities and employs between 10 and 50 people one month after the commencement of operations, all of whom are domestic nationals. The starting a business indicators consider two cases of local limited liability companies that are identical in all aspects, except that one company is owned by five married women and the other by five married men.																																																						
Preenchido	49.78%																																																						
Intervalo de Dados	Não mensurável																																																						
Problemas de Qualidade dos Dados	Existem 579 linhas em branco que serão passadas a "Sem informação"																																																						
Estatísticas dos Dados	<p>Column: metadata.Statistical_concept_and_methodology</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>1153</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>175</td> <td>15.18%</td> </tr> <tr> <td>Unique Count</td> <td>89</td> <td>7.72%</td> </tr> <tr> <td>Duplicate Count</td> <td>86</td> <td>7.46%</td> </tr> <tr> <td>Blank Count</td> <td>579</td> <td>50.22%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>579</td> <td>50.22%</td> </tr> <tr> <td>Women, Business and the L...</td> <td>55</td> <td>4.77%</td> </tr> <tr> <td>Data on cause of death are...</td> <td>44</td> <td>3.82%</td> </tr> <tr> <td>Demographic and Health S...</td> <td>39</td> <td>3.38%</td> </tr> <tr> <td>The United Nations Inter-A...</td> <td>20</td> <td>1.73%</td> </tr> <tr> <td>Percentage of female grad...</td> <td>13</td> <td>1.13%</td> </tr> <tr> <td>The status in employment ...</td> <td>12</td> <td>1.04%</td> </tr> <tr> <td>The labor force is the suppl...</td> <td>10</td> <td>0.87%</td> </tr> <tr> <td>Data on menstrual health a...</td> <td>9</td> <td>0.78%</td> </tr> <tr> <td>The International Labour O...</td> <td>9</td> <td>0.78%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	175	15.18%	Unique Count	89	7.72%	Duplicate Count	86	7.46%	Blank Count	579	50.22%	Value	Count	%	Empty field	579	50.22%	Women, Business and the L...	55	4.77%	Data on cause of death are...	44	3.82%	Demographic and Health S...	39	3.38%	The United Nations Inter-A...	20	1.73%	Percentage of female grad...	13	1.13%	The status in employment ...	12	1.04%	The labor force is the suppl...	10	0.87%	Data on menstrual health a...	9	0.78%	The International Labour O...	9	0.78%
Label	Count	%																																																					
Row Count	1153	100.00%																																																					
Null Count	0	0.00%																																																					
Distinct Count	175	15.18%																																																					
Unique Count	89	7.72%																																																					
Duplicate Count	86	7.46%																																																					
Blank Count	579	50.22%																																																					
Value	Count	%																																																					
Empty field	579	50.22%																																																					
Women, Business and the L...	55	4.77%																																																					
Data on cause of death are...	44	3.82%																																																					
Demographic and Health S...	39	3.38%																																																					
The United Nations Inter-A...	20	1.73%																																																					
Percentage of female grad...	13	1.13%																																																					
The status in employment ...	12	1.04%																																																					
The labor force is the suppl...	10	0.87%																																																					
Data on menstrual health a...	9	0.78%																																																					
The International Labour O...	9	0.78%																																																					

Atributo	Development_relevance
Tipo de Dados	String
Descrição	Relevância do indicador para o desenvolvimento
Exemplo	Firms evaluating investment options, governments interested in improving business conditions, and economists seeking to explain economic performance have all grappled with defining and measuring the business environment. The firm-level data from Enterprise Surveys provide a useful tool for benchmarking economies across a large number of indicators measured at the firm level.

	Firms with female top manager measures women's integration as decision makers. Benchmarking firms with female top manager is important to achieving gender equality promotion and empowerment of women. The gender topic provides information about women's entrepreneurship and economic participation in the labor force.																														
Preenchido	46.05%																														
Intervalo de Dados	Não mensurável																														
Problemas de Qualidade dos Dados	Existem 622 linhas em branco que serão passadas a "Sem informação"																														
Estatísticas dos Dados	<p>* Column: metadata.Development_relevance </p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>1153</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>76</td> <td>6.59%</td> </tr> <tr> <td>Unique Count</td> <td>25</td> <td>2.17%</td> </tr> <tr> <td>Duplicate Count</td> <td>51</td> <td>4.42%</td> </tr> <tr> <td>Blank Count</td> <td>622</td> <td>53.95%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>622</td> <td>53.95%</td> </tr> <tr> <td>Ownership of assets, partic... The knowledge and analysi... Measuring the number of d... Violence against women is ... A relative high concentrat... Women's participation in d... Violence against women is ... Gross enrollment ratios ind... Estimates of women in the ...</td> <td>96 55 42 26 24 23 21 18 17</td> <td>8.33% 4.77% 3.64% 2.25% 2.08% 1.99% 1.82% 1.56% 1.47%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	76	6.59%	Unique Count	25	2.17%	Duplicate Count	51	4.42%	Blank Count	622	53.95%	Value	Count	%	Empty field	622	53.95%	Ownership of assets, partic... The knowledge and analysi... Measuring the number of d... Violence against women is ... A relative high concentrat... Women's participation in d... Violence against women is ... Gross enrollment ratios ind... Estimates of women in the ...	96 55 42 26 24 23 21 18 17	8.33% 4.77% 3.64% 2.25% 2.08% 1.99% 1.82% 1.56% 1.47%
Label	Count	%																													
Row Count	1153	100.00%																													
Null Count	0	0.00%																													
Distinct Count	76	6.59%																													
Unique Count	25	2.17%																													
Duplicate Count	51	4.42%																													
Blank Count	622	53.95%																													
Value	Count	%																													
Empty field	622	53.95%																													
Ownership of assets, partic... The knowledge and analysi... Measuring the number of d... Violence against women is ... A relative high concentrat... Women's participation in d... Violence against women is ... Gross enrollment ratios ind... Estimates of women in the ...	96 55 42 26 24 23 21 18 17	8.33% 4.77% 3.64% 2.25% 2.08% 1.99% 1.82% 1.56% 1.47%																													

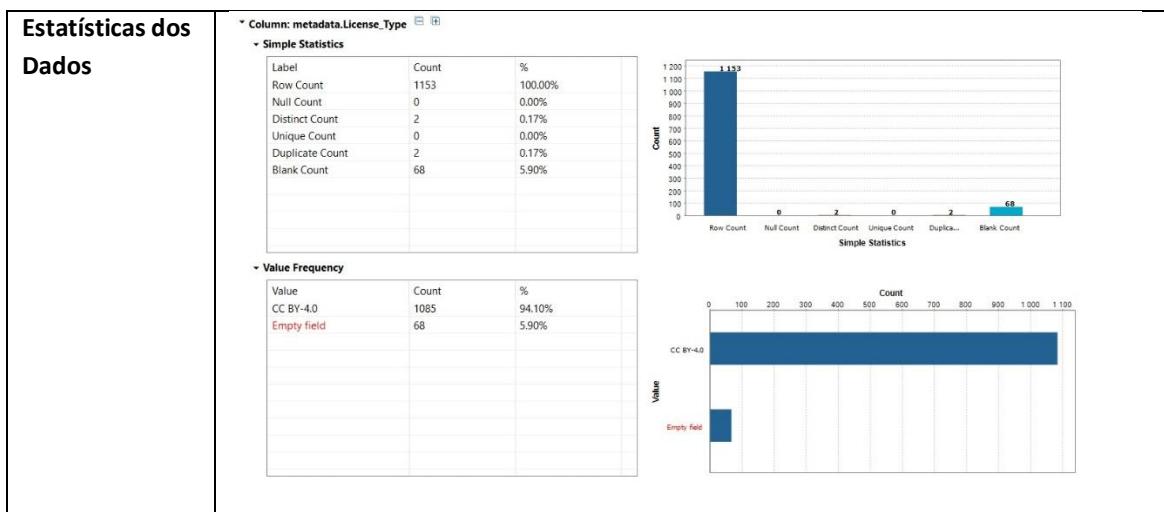
Atributo	Related_source_links
Tipo de Dados	String
Descrição	Links para fontes relacionadas
Exemplo	
Preenchido	0%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Todas as linhas encontram-se em branco que serão passadas a "Sem informação"



Atributo	Other_web_links																											
Tipo de Dados	String																											
Descrição	Outros links relevantes																											
Exemplo																												
Preenchido	0%																											
Intervalo de Dados	Não mensurável																											
Problemas de Qualidade dos Dados	Todas as linhas encontram-se em branco que serão passadas a “Sem informação”																											
Estatísticas dos Dados	<p>* Column: metadata.Other_web.links</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>1153</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>1</td> <td>0.09%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>1</td> <td>0.09%</td> </tr> <tr> <td>Blank Count</td> <td>1153</td> <td>100.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>1153</td> <td>100.00%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	1	0.09%	Unique Count	0	0.00%	Duplicate Count	1	0.09%	Blank Count	1153	100.00%	Value	Count	%	Empty field	1153	100.00%
Label	Count	%																										
Row Count	1153	100.00%																										
Null Count	0	0.00%																										
Distinct Count	1	0.09%																										
Unique Count	0	0.00%																										
Duplicate Count	1	0.09%																										
Blank Count	1153	100.00%																										
Value	Count	%																										
Empty field	1153	100.00%																										

Atributo	Related_indicators																											
Tipo de Dados	String																											
Descrição	Indicadores relacionados																											
Exemplo																												
Preenchido	0%																											
Intervalo de Dados	Não mensurável																											
Problemas de Qualidade dos Dados	Todas as linhas encontram-se em branco que serão passadas a “Sem informação”																											
Estatísticas dos Dados	<p>* Column: metadata.Related_indicators</p> <p>Simple Statistics</p> <table border="1"> <thead> <tr> <th>Label</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Row Count</td> <td>1153</td> <td>100.00%</td> </tr> <tr> <td>Null Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Distinct Count</td> <td>1</td> <td>0.09%</td> </tr> <tr> <td>Unique Count</td> <td>0</td> <td>0.00%</td> </tr> <tr> <td>Duplicate Count</td> <td>1</td> <td>0.09%</td> </tr> <tr> <td>Blank Count</td> <td>1153</td> <td>100.00%</td> </tr> </tbody> </table> <p>Value Frequency</p> <table border="1"> <thead> <tr> <th>Value</th> <th>Count</th> <th>%</th> </tr> </thead> <tbody> <tr> <td>Empty field</td> <td>1153</td> <td>100.00%</td> </tr> </tbody> </table>	Label	Count	%	Row Count	1153	100.00%	Null Count	0	0.00%	Distinct Count	1	0.09%	Unique Count	0	0.00%	Duplicate Count	1	0.09%	Blank Count	1153	100.00%	Value	Count	%	Empty field	1153	100.00%
Label	Count	%																										
Row Count	1153	100.00%																										
Null Count	0	0.00%																										
Distinct Count	1	0.09%																										
Unique Count	0	0.00%																										
Duplicate Count	1	0.09%																										
Blank Count	1153	100.00%																										
Value	Count	%																										
Empty field	1153	100.00%																										

Atributo	License_Type
Tipo de Dados	String
Descrição	Tipo de licença dos dados
Exemplo	CC BY-4.0
Preenchido	94.1%
Intervalo de Dados	Não mensurável
Problemas de Qualidade dos Dados	Existem 68 linhas em branco que serão passadas a “Sem informação”





Auto e Hétero avaliação

- ▶ Pedro Miguel Araújo Pires (a95549) – N + 2
- ▶ Liandro Macedo da Cruz (a100436) – N
- ▶ Catarina Brandão Fernandes (a101441) – N - 1
- ▶ Maria Inês da Costa Maia (a93347) – N – 1

Somatório:

N