



# **ENGENHARIA DE DADOS**

## **MÍNIMO PRODUTO VIÁVEL (MVP)**

**Professor:**

Victor Almeida

**Aluna:**

Julia Bragança dos Santos Ferreira Lima

## OBJETIVO

O objetivo deste trabalho foi construir um pipeline de dados, utilizando a tecnologia da nuvem. O pipeline consistiu na busca, coleta, modelagem, carga e análise dos dados.

A análise dos dados teve como premissa, o Índice de Desenvolvimento humano, que é uma medida resumida do progresso a longo prazo em três dimensões básicas do desenvolvimento humano: renda, educação e saúde. O objetivo da criação do IDH foi oferecer um contraponto a outro indicador utilizado, o Produto Interno Bruto (PIB) per capita, que considera apenas a dimensão econômica do desenvolvimento.

As Faixas do IDH são:

Abaixo de 0,499	Muito Baixo
Entre 0,500 e 0,599	Baixo
Entre 0,600 e 0,699	Médio
Entre 0,700 e 0,799	Alto
Acima de 0,800	Muito alto

O objetivo desse trabalho é verificar quais indicadores podem estar ligados ao IDHM e que tipos de medidas podem ser executadas com o intuito de elevar esse índice.

Quais são as cidades com o maior IDHM? E o que elas têm em comum?

Quais são as cidades com o menor IDHM? E o que elas têm em comum?

O aumento nos impostos arrecadados pode aumentar o IDHM?

Qual o valor médio das despesas municipais das cidades com os maiores e menores IDHM?

O aumento das despesas municipais pode aumentar o IDHM?

## DETALHAMENTO

### 1. Busca pelos dados

Para elaboração deste estudo, foi necessário realizar as buscas pelos dados que poderiam responder a essas perguntas. Para tanto, foi encontrado no repositório de dados abertos da kaggle, um dataset com as informações necessárias para atingir o objetivo deste projeto:

<https://www.kaggle.com/datasets/crisparada/brazilian-cities>

Esse dataset já estava pronto, com os dados retirados de diversos sites, conforme apresentado no catálogo de dados.

Cabe salientar, que os dados são referentes à anos variados, entre 2010 e 2019.

## Brazilian Cities

A collection of 79 attributes from Brazilian Cities



Data Card Code (25) Discussion (3)

### About Dataset

#### Context

Brazil is the world's fifth-largest country by area, with 8.5 million square kilometer, and the fifth most populous, with over 208 million people.

The Federative Republic of Brazil is composed of the union of the 26 states, the Federal District, and the 5,570 municipalities.

This dataset is a compilation of several publicly available information about Brazilian Municipalities.

#### Content

The data was retrieved from several websites, as presented below. I processed and merged all data by city name, resulting in 79 fields for

#### Usability

10.00

#### License

CC BY-SA 4.0

#### Expected update frequency

Annually

#### Tags

Tabular Economics

## 2. Coleta dos dados

A coleta dos dados foi realizada por intermédio do download dos dados e estes foram armazenados na nuvem, para tanto, foi necessário realizar a criação de uma plataforma na nuvem para armazenamento dos dados e posterior tratamento destes.

Neste passo, foi criada uma conta na plataforma da Microsoft Azure, inicialmente foi necessário realizar a criação de um Grupo de Recursos, onde os recursos criados seriam armazenados dentro do Grupo.

Nome	Tipo	Localização
mvpsqldatabase_tst (mvpsqlserver-db/mvpsqldatabase_tst)	Banco de dados SQL	East US
mvpsqlserver-db	SQL Server	East US
mvpsqlstorage1	Conta de armazenamento	East US
mvptestefactory	Data factory (V2)	East US

Após a criação deste grupo de recursos, foi necessário realizar a criação de um storage, onde seriam efetivamente carregados os dados, foi criado então no Azure Blob Storage a criação de um Data Lake.

Foi realizada a criação de um SQL server e de um Banco de Dados na Azure, nesta etapa foi necessário executar a configuração de acesso, o que permitiu que o IP pudesse acessar o recurso e realizar alterações no Banco de Dados.

## 3. Modelagem de Dados e Catálogo de Dados

### a. Modelagem de Dados

Nesta etapa, foi criado um recurso no Azure Data Factory para viabilizar o serviço de ETL do MVP.

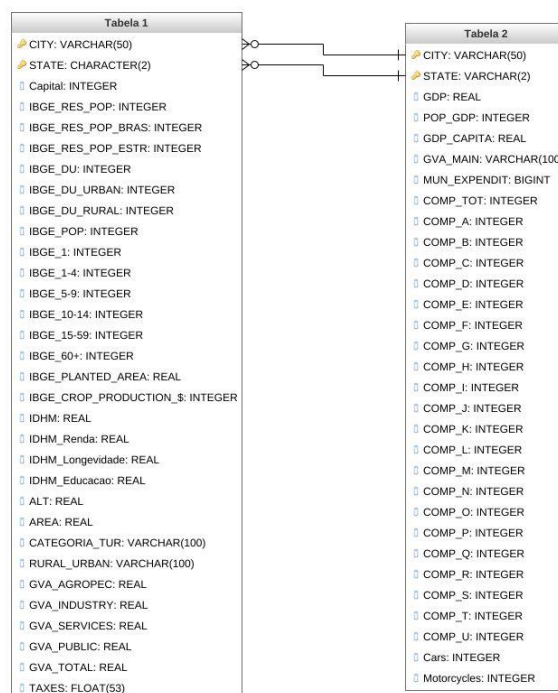
O Azure Data Factory é um serviço na nuvem da Azure para integração e transformação de dados que permite criar fluxos de trabalho orientados a dados para orquestrar a movimentação de dados.

Foi idealizado inicialmente a idéia de ter somente uma tabela, tendo em vista que o dataset que seria utilizado possuía todas as informações necessárias, entretanto, ao se importar o Dataset no Dataflow do Azure Data Factory, foi verificado que os atributos não eram exibidos no select individualmente, pois a quantidade de atributos no dataset excedia o número de 50, o que acarretava em exibir todos os atributos juntos, portanto, foi realizado o particionamento deste dataset em dois.

Para realizar o particionamento dos datasets, foi utilizado o Virtual Studio Code e realizado por intermédio da linguagem Python a divisão do DataSet em dois, conforme pode ser observado na figura a seguir:

```
separando em duas colunas.py X separando as planilhas em duas.py
separando em duas colunas.py > ...
1 from sys import displayhook
2 import pandas as pd
3 tabela = pd.read_csv("BRAZIL_CITIES_REV2022.CSV")
4
5 df= pd.read_csv("BRAZIL_CITIES_REV2022.CSV")
6 df1=df.iloc[:, :39]
7 displayhook(df1)
8 df2=df.iloc[:, 39:].assign(CITY=df["CITY"], STATE=df["STATE"])
9 displayhook(df2)
10 df1.to_csv("primeiras 40 colunas v3")
11 df2.to_csv("resto das colunas v3")
```

Por fim, a modelagem do meu banco de dados está definida foi definida da seguinte forma:



b. Catálogo de Dados

O catálogo de dados foi elaborado em planilha de excel (Catálogo de Dados.csv) que se encontra anexada a este MVP e onde tem cada informação relativa aos dados e também foi elaborada uma planilha com os metadados (Metadados.csv), também anexa a este MVP.

#### 4. Carga

Nesta etapa foi realizada a Extração, transformação e carga (ETL) dos dados no banco de dados. Para executar tal feito, foi utilizado o Azure Data Factory, onde foi possível importar os datasets no armazenamento de Blob da Azure, então foram importadas as duas tabelas, e verificada se o schema foi importado automaticamente, o que em ambos os casos ocorreu sem problemas.

Posteriormente, foi elaborado o Dataflow, onde inicialmente foram executadas duas entradas com cada uma das tabelas e nesta etapa foi necessário realizar o ajuste do tipo de cada coluna, tendo em vista que todas as colunas vieram como strings, então foi necessário alterar para integer (os números inteiros), float (os números com casas decimais). Neste momento foi necessário ter bastante atenção, pois poderia ocorrer situações onde os números inteiros fossem como strings para o query, o que atrapalharia a correta análise de dados.

- Entrada 1:

The screenshot displays the configuration interface for a Dataflow named 'Entrada1' in Azure Data Factory. The interface is in Portuguese and includes a top navigation bar with tabs: 'Configurações de origem', 'Opções de origem', 'Projeção', 'Otimizar', 'Inspeccionar', and 'Visualização de dados'. The 'Configurações de origem' tab is active, showing the following configuration details:

- Nome do fluxo de saída \***: Entrada1
- Descrição**: Importe dados do primeiras\_40colunas
- Tipo de fonte \***: Conjunto de Dados (selected), Embutido
- Conjunto de Dados \***: primeiras\_40colunas
- Opções**:
  - ☒ Permitir o descompasso de esquema
  - ☐ Inferir tipos de coluna em descompasso
  - ☐ Validar esquema
- Ignorar contagem de linhas**: (empty field)
- Amostragem \***: Desabilitar (selected)

Additional actions like 'Saiba mais', 'Redefinir', 'Testar conexão', 'Abrir', and 'Novo' are available for the data set selection.

Publicar tudo Experiência de versão prévia Desligar

dataflow2 ×

Validar Depuração de fluxo de dados

Entrada1 atual select1 join1 sink1

Renomeando Entrada1 para

Configurações de origem Opções de origem **Projeção** Otimizar Inspeccionar Visualização de dados

CITY	abc	string	▼	Especificar o formato	▼
STATE	abc	string	▼	Especificar o formato	▼
CAPITAL	123	integer	▼	Especificar o formato	▼
IBGE_RES_POP	123	integer	▼	Especificar o formato	▼
IBGE_RES_POP_BRAS	123	integer	▼	Especificar o formato	▼
IBGE_RES_POP_ESTR	123	integer	▼	Especificar o formato	▼
IBGE_DU	123	integer	▼	Especificar o formato	▼
IBGE_DU_URBAN	123	integer	▼	Especificar o formato	▼
IBGE_DU_RURAL	123	integer	▼	Especificar o formato	▼
IBGE_POP	123	integer	▼	Especificar o formato	▼
IBGE_1	123	integer	▼	Especificar o formato	▼
IBGE_1-4	123	integer	▼	Especificar o formato	▼
IBGE_5-9	123	integer	▼	Especificar o formato	▼
IBGE 10-14	123	integer	▼	Especificar o formato	▼

- **Entrada 2:**

dataflow2 ×

Validar Depuração de fluxo de dados

Colunas: 40 total

Renomeando Entrada1 para select1 com as colunas 'CITY, STATE,

Junção de Interna em 'select1' e 'select2'

Exportar dados para AzureSqlTable4

Entrada2

Colunas: 45 total

select2

Renomeando Entrada2 para select2 com as colunas 'CITY, STATE,

Configurações de origem Opções de origem **Projeção** Otimizar Inspeccionar Visualização de dados

Nome do fluxo de saída \* Entrada2 [Saiba mais](#)

Descrição Importe dados do resto\_das\_colunas [Redefinir](#)

Tipo de fonte \* Conjunto de Dados Embutido

Conjunto de Dados \* resto\_das\_colunas [Testar conexão](#) [Abrir](#) [+ Novo](#)

Opções ☒ Permitir o descompasso de esquema [?](#) ☐ Inferir tipos de coluna em descompasso [?](#)

Foi realizado em cada uma das tabelas, um select com a finalidade de excluir colunas que não seriam aproveitadas, e conseqüentemente diminuir o tamanho dos dados, além de ajudar na visualização e posterior análise dos dados.

- Select 1:

dataflow2 prévia

✓ Validar ☐ Depuração de fluxo de dados

Selecionar configurações Otimizar Inspeccionar Visualização de dados

<input type="checkbox"/>	Coluna do Entrada1		Nomear como		
<input type="checkbox"/>	abc CITY	→	CITY	+	🗑️
<input type="checkbox"/>	abc STATE	→	STATE	+	🗑️
<input type="checkbox"/>	123 CAPITAL	→	CAPITAL	+	🗑️
<input type="checkbox"/>	123 IBGE_RES_POP	→	IBGE_RES_POP	+	🗑️
<input type="checkbox"/>	123 IBGE_RES_POP_BRAS	→	IBGE_RES_POP_BRAS	+	🗑️
<input type="checkbox"/>	123 IBGE_RES_POP_ESTR	→	IBGE_RES_POP_ESTR	+	🗑️
<input type="checkbox"/>	123 IBGE_DU	→	IBGE_DU	+	🗑️
<input type="checkbox"/>	123 IBGE_DU_URBAN	→	IBGE_DU_URBAN	+	🗑️
<input type="checkbox"/>	123 IBGE_DU_RURAL	→	IBGE_DU_RURAL	+	🗑️

- Select 2:

dataflow2 prévia

✓ Validar ☐ Depuração de fluxo de dados

Selecionar configurações Otimizar Inspeccionar Visualização de dados

<input type="checkbox"/>	Coluna do Entrada2		Nomear como		
<input type="checkbox"/>	abc CITY	→	CITY	+	🗑️
<input type="checkbox"/>	abc STATE	→	STATE	+	🗑️
<input type="checkbox"/>	1,2f GDP	→	GDP	+	🗑️
<input type="checkbox"/>	123 POP_GDP	→	POP_GDP	+	🗑️
<input type="checkbox"/>	1,2f GDP_CAPITA	→	GDP_CAPITA	+	🗑️
<input type="checkbox"/>	abc GVA_MAIN	→	GVA_MAIN	+	🗑️
<input type="checkbox"/>	abc MUN_EXPENDIT	→	MUN_EXPENDIT	+	🗑️
<input type="checkbox"/>	123 COMP_TOT	→	COMP_TOT	+	🗑️
<input type="checkbox"/>	123 COMP_A	→	COMP_A	+	🗑️

Após o select, foi inserido um JOIN onde foi realizada uma junção interna das duas entradas dos dados transformando em uma só tabela final. Neste caso, foi necessário utilizar nas duas tabelas as informações relativas a cidade e estado, os quais serviram de condições do JOIN. Inicialmente foi utilizada apenas a coluna de cidade, entretanto, pelo fato de haver cidades com nomes iguais em estados diferentes, foi necessário utilizar a coluna do estado também.

- JOIN:

Publicar tudo Experiência de versão ☐ prévia

dataflow2 ×

Validar ☒ Depuração de fluxo de dados ☐

Entrada1: Importe dados do primeiras\_40colunas

select1: Renomeando Entrada1 para select1 com as colunas 'CITY, STATE,

join1: Colunas: 63 total

sink1: Exportar dados para AzureSqlTable4

Configurações de junção | Otimizar | Inspeccionar | Visualização de dados

Fluxo esquerdo \*

Fluxo direito \*

Tipo de junção \*

Externa completa | **Interna** | Externa esquerda | Externa direita | Personalizado (cruzado)

Usar correspondência difusa ⓘ ☐

Condições de junção \*

Esquerda: Coluna do select1

Direita: Coluna do select2

abc CITY == abc CITY +

abc STATE == abc STATE +

Por fim, foi executado um sink para servir como coletor de dados e direcionar os dados obtidos desse dataflow para a tabela final.

Foi realizado a etapa de validação de dados e depois os dados foram publicados.

O Dataflow pode ser observado na figura abaixo seguido do código do DataFlow:

Data Factory | Validate all | Publish all

Factory Resources

- Pipelines: 2
- Change Data Capture (preview): 0
- Datasets: 6
- Data flows: 3
  - dataflow2
  - dataflow - teste
  - flowlet11
- Power Query: 0

dataflow2 ×

Validar ☒ Data flow debug ☐

Entrada1

select1

Entrada2

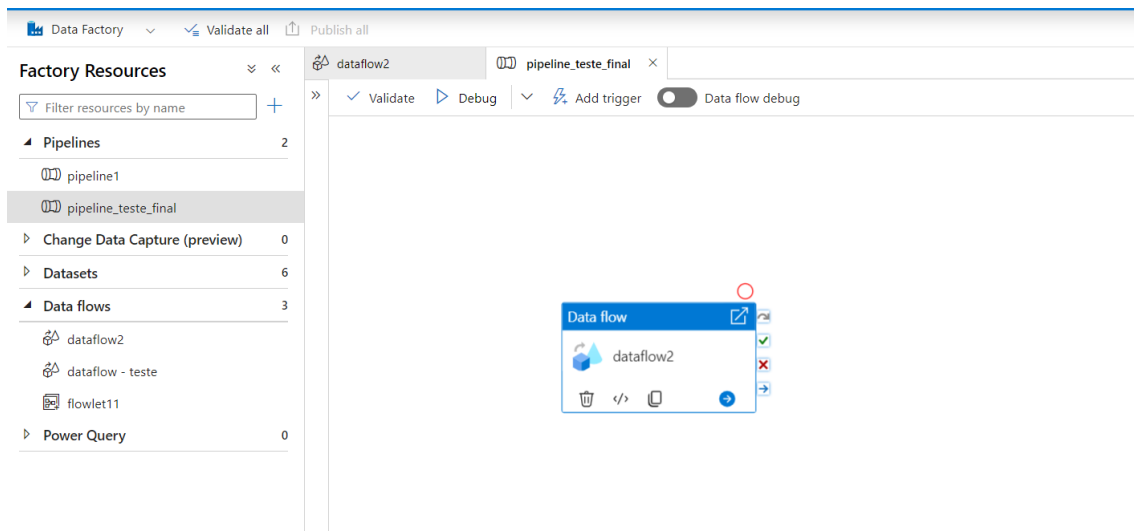
select2

join1

61 Columns

Ainda dentro do Data Factory, o dataflow foi inserido no pipeline, onde foi realizada outra validação e por fim, foi realizada a depuração com êxito.





## 5. Análise

### a. Qualidade de Dados

A qualidade de dados é um fator muito importante em um banco de dados para uma correta análise dos dados e tomadas de decisões.

Foram tomadas algumas medidas na tabela do MVP com intuito de otimizar a qualidade dos dados nela existente dentro do Azure Data Studio.

Um dos problemas que foram observados, foi que valores superiores a 2.147.483.647 em colunas que foram classificadas como integer retornaram com valores nulos, isto porque no Azure Data Factory não foi possível classificar a coluna como bigint, entretanto, este problema foi corrigido dentro do Azure Data Studio, além disso, foram inseridas as restrições de integridade de cada atributo, conforme pode ser observado no ANEXO II.

Por fim, é importante ressaltar que o dataset inicial já estava com uma qualidade bem alta dos dados, sendo necessários apenas pequenos ajustes para otimizar a sua utilização.

### b. Solução do Problema

- 1) Foi realizada a seguinte consulta para verificar as Cidades com o maior índice de Desenvolvimento Humano (Superior a 0.80):

```
select [CITY]
      , [STATE]
      , [IDHM]
      , [IBGE_RES_POP]
      , [MUN_EXPENDIT]
      , (MUN_EXPENDIT / IBGE_RES_POP) AS "TOTAL max"
FROM [dbo].[Tabela_MVP_04]
WHERE [IDHM] > 0.80
and MUN_EXPENDIT >0
ORDER BY "IDHM" desc
```

Retornando com o seguinte resultado:

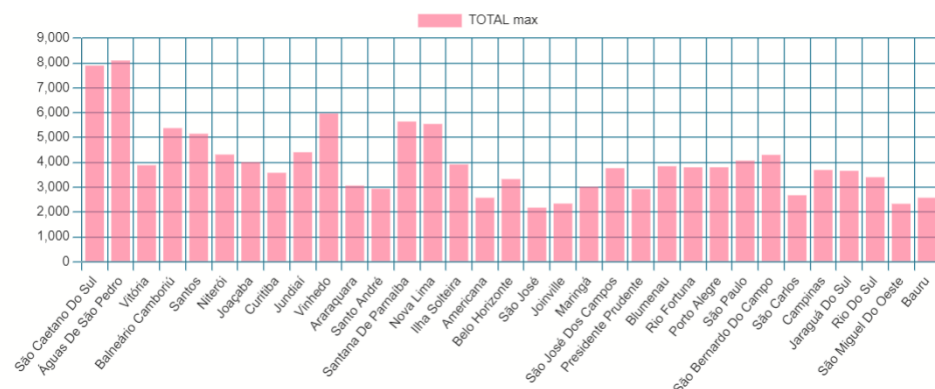
	CITY	STATE	IDHM	IBGE_RES_POP	MUN_EXPENDIT	TOTAL max
1	São Caetano Do Sul	SP	0,862	149263	1178645567	7896
2	Águas De São Pedro	SP	0,854	2707	21934920	8103
3	Vitória	ES	0,845	327801	1273296503	3884
4	Balneário Camboriú	SC	0,845	108089	581366500	5378
5	Santos	SP	0,84	419400	2161093305	5152
6	Niterói	RJ	0,837	487562	2102739559	4312
7	Joaçaba	SC	0,827	27020	108319357	4008
8	Curitiba	PR	0,823	1751907	6269694781	3578
9	Jundiaí	SP	0,822	370126	1630920489	4406
10	Vinhedo	SP	0,817	63611	379854817	5971
11	Araraquara	SP	0,815	208662	639269199	3063
12	Santo André	SP	0,815	676407	1986553463	2936
13	Santana De Parnaíba	SP	0,814	108813	614167631	5644
14	Nova Lima	MG	0,813	80998	449232783	5546
15	Ilha Solteira	SP	0,812	25064	98224256	3918
16	Americana	SP	0,811	210638	541958027	2572
17	Belo Horizonte	MG	0,81	2375151	7899696615	3325
18	São José	SC	0,809	209804	456514062	2175
19	Joinville	SC	0,809	515288	1204303174	2337
20	Maringá	PR	0,808	357077	1077439047	3017

Com a consulta abaixo, foi possível verificar uma média de R\$ 4.001,00 relativo ao Gasto Municipal por pessoa das Cidades com o IDHM acima de 0.80, com a seguinte consulta:

```
SELECT AVG(MUN_EXPENDIT/IBGE_RES_POP)
FROM [dbo].[Tabela_MVP_04]
WHERE [IDHM] > 0.80
```

Com a consulta abaixo, foi possível verificar o gráfico relativo ao Gasto Municipal por pessoa das Cidades com o IDHM acima de 0.80.

```
select [CITY]
, (MUN_EXPENDIT / IBGE_RES_POP) AS "TOTAL max"
FROM [dbo].[Tabela_MVP_04]
WHERE [IDHM] > 0.80
AND MUN_EXPENDIT>0
ORDER BY "IDHM" desc
```



Posteriormente, foram analisados os Gastos Municipais por pessoa das Cidades com o menor índice de Desenvolvimento Humano (Inferiores a 0.50). Conforme consulta a seguir:

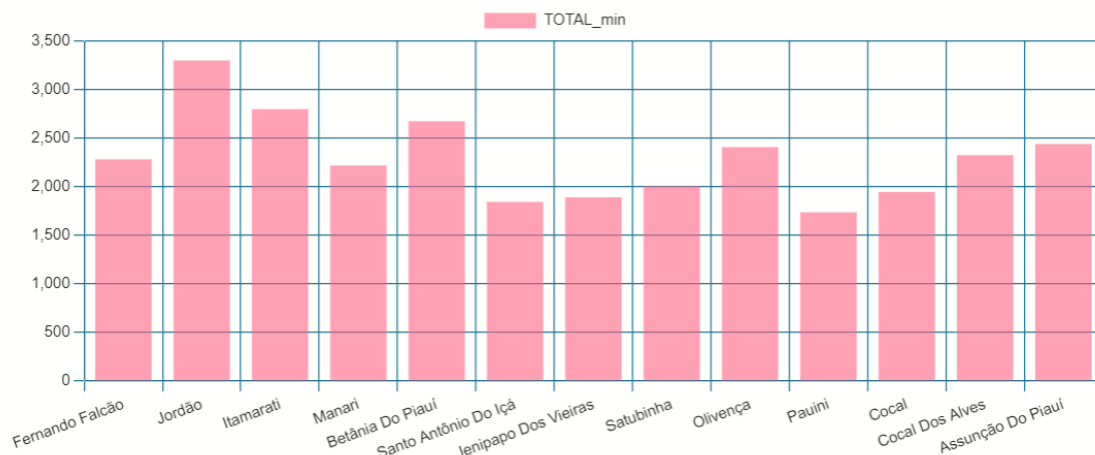
```
select [CITY]
      ,(MUN_EXPENDIT / (IBGE_RES_POP)) AS "TOTAL_min"
FROM [dbo].[Tabela_MVP_04]
WHERE [IDHM] < 0.50
and IBGE_RES_POP>0
AND MUN_EXPENDIT>0
ORDER BY "IDHM"
```

E obtida a tabela resposta a seguir:

	CITY	STATE	IDHM	IBGE_RES_POP	MUN_EXPENDIT	TOTAL_min
1	Fernando Falcão	MA	0,443	9241	21114565	2284
2	Jordão	AC	0,469	6577	21724701	3303
3	Itamarati	AM	0,477	8038	22521795	2801
4	Manari	PE	0,487	18083	40176640	2221
5	Betânia Do Piauí	PI	0,489	6015	16103016	2677
6	Santo Antônio Do Içá	AM	0,49	24481	45201069	1846
7	Jenipapo Dos Vieiras	MA	0,49	15440	29244645	1894
8	Satubinha	MA	0,493	11990	24082803	2008
9	Oliveira	AL	0,493	11047	26624245	2410
10	Pauini	AM	0,496	18166	31589139	1738
11	Cocal	PI	0,497	26036	50733134	1948
12	Cocal Dos Alves	PI	0,498	5572	12966428	2327
13	Assunção Do Piauí	PI	0,499	7503	18325692	2442

Foi então realizada a seguinte consulta para verificar o Gasto Municipal por pessoa das cidades com o IDHM inferior a 0.50:

O qual foi obtido o gráfico a seguir:



Foi então realizada uma consulta para verificar a média do Gasto Municipal por pessoa das Cidades com o IDHM abaixo de 0.50 e obtido o total de R\$2.538,00 por pessoa.

```
SELECT AVG(MUN_EXPENDIT / IBGE_RES_POP) AS AVG_TOTAL_MIN
```

```
FROM [dbo].[Tabela_MVP_04]
WHERE [IDHM] < 0.51
AND MUN_EXPENDIT>0 AND IBGE_POP >0
```

Sendo possível concluir que, aparentemente, há uma correlação entre os gastos municipais e o IDHM, tendo em vista que a média dos gastos municipais é 63% inferior aos gastos municipais nos maiores IDHM.

2) Foi realizada uma análise dos Impostos arrecadados nas Cidades com os maiores índices de desenvolvimento humano e as Cidades com os menores índices de desenvolvimento Humano, conforme gráficos e consultas a seguir:

a) Taxas por pessoas das cidades com o IDHM acima de 0.80

Foi executada a seguinte consulta:

```
select [CITY]
, CAST (TAXES/IBGE_RES_POP as numeric (10,3)) as TAXES_PERSON_Max
FROM [dbo].[Tabela_MVP_04]
WHERE [IDHM] > 0.80
and [IBGE_RES_POP]>0
order by [IDHM] DESC
```



Foi realizada uma consulta da média dos impostos por pessoa:

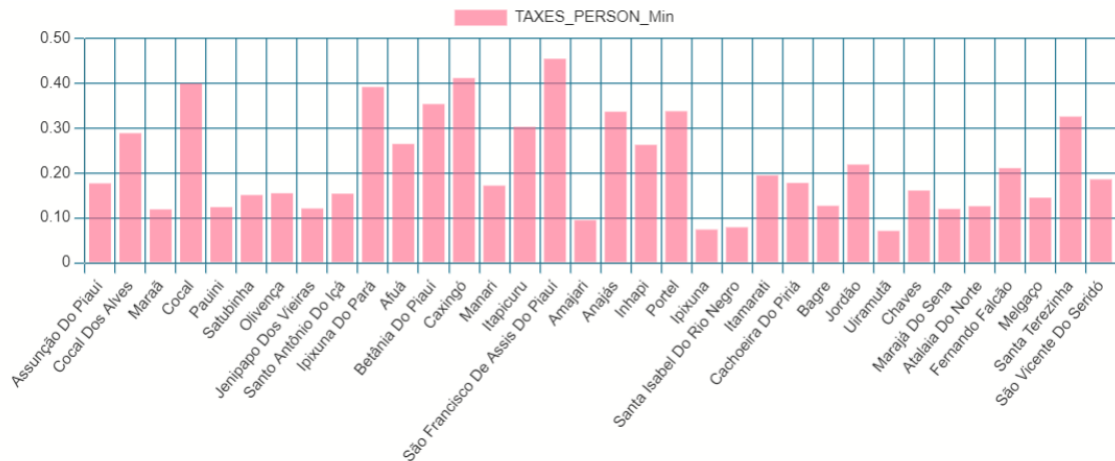
```
SELECT AVG (CAST (TAXES/IBGE_RES_POP as numeric (10,3))) as TAXES_PERSON_Max
FROM [dbo].[Tabela_MVP_04]
WHERE [IDHM] > 0.80
AND IBGE_RES_POP >0
```

Onde foi obtido como resposta que a média dos impostos por pessoa, cujo IDHM é superior a 0.80, é de R\$ 7.621,21 por pessoa.

b) Taxas por pessoas das cidades com o IDHM abaixo de 0.50

Foi executada a seguinte consulta:

```
select [CITY]
, CAST (TAXES/IBGE_RES_POP as numeric (10,3)) as TAXES_PERSON_Min
FROM [dbo].[Tabela_MVP_04]
WHERE [IDHM] < 0.50
and [IBGE_RES_POP]>0
order by [IDHM] DESC
```



Foi realizada uma consulta da média dos impostos por pessoa:

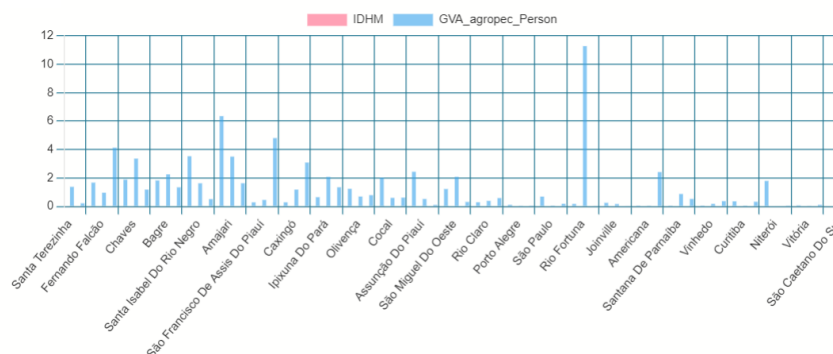
```
SELECT AVG (CAST (TAXES/IBGE_RES_POP as numeric (10,3))) as TAXES_PERSON_Min
FROM [dbo].[Tabela_MVP_04]
WHERE [IDHM] < 0.50
AND IBGE_RES_POP >0
```

Onde foi obtido como resposta que a média dos impostos por pessoa, das cidades com o IDHM abaixo de 0.50, é de R\$ 214,53.

O que se observa que parece existir uma relação entre o IDHM e o aumento de impostos arrecadados por pessoa, tendo em vista que o imposto por pessoa é aproximadamente 35 vezes superior nas cidades com o índice de desenvolvimento acima de 0.80.

- 3) Por fim, foi realizado uma análise acerca da Gestão do Valor Agregado dos setores agropecuário, industrial, serviços públicos e total, conforme consulta a seguir:

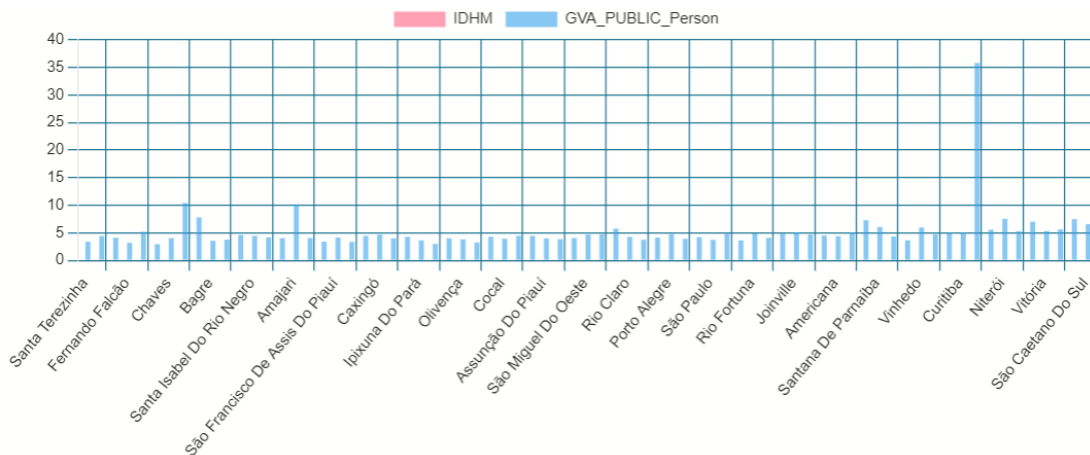
```
select DISTINCT [CITY]
, [IDHM]
, CAST (GVA_AGROPEC/IBGE_RES_POP as numeric (10,3)) as GVA_agropec_Person
FROM [dbo].[Tabela_MVP_04]
WHERE [IDHM] > 0.80
or [IDHM] < 0.50
AND IBGE_RES_POP >0
ORDER BY [IDHM]
```



Onde é possível observar um GVA Agropecuário elevado nas cidades entre Santa Terezinha e Assunção do Piauí que são as cidades com IDHM inferior a

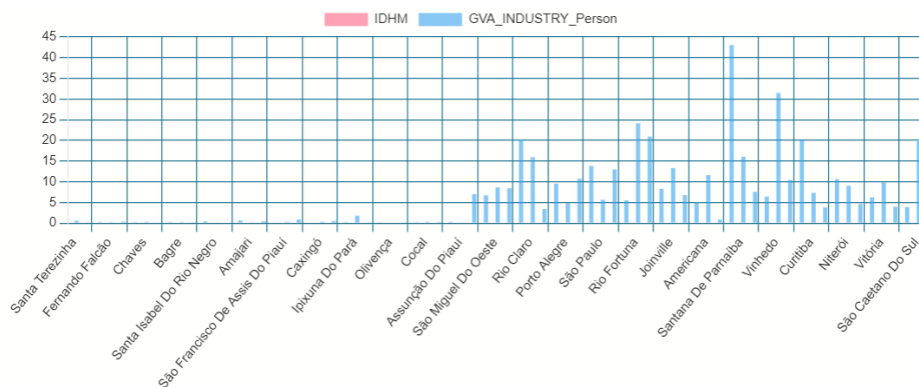
0.50 e um GVA Agropecuário baixo nas cidades entre São Miguel do Oeste e São Caetano do Sul que são cidades com o IDHM acima de 0.80.

```
select DISTINCT [CITY]
, [IDHM]
, CAST (GVA_PUBLIC/IBGE_RES_POP as numeric (10,3)) as GVA_PUBLIC_Person
FROM [dbo].[Tabela_MVP_04]
WHERE [IDHM] > 0.80
or [IDHM] < 0.50
AND IBGE_RES_POP >0
ORDER BY [IDHM]
```



Onde é possível observar um GVA de Serviços Públicos similar nas cidades entre Santa Terezinha e Assunção do Piauí que são as cidades com IDHM inferior a 0.50 e nas cidades entre São Miguel do Oeste e São Caetano do Sul que são cidades com o IDHM acima de 0.80.

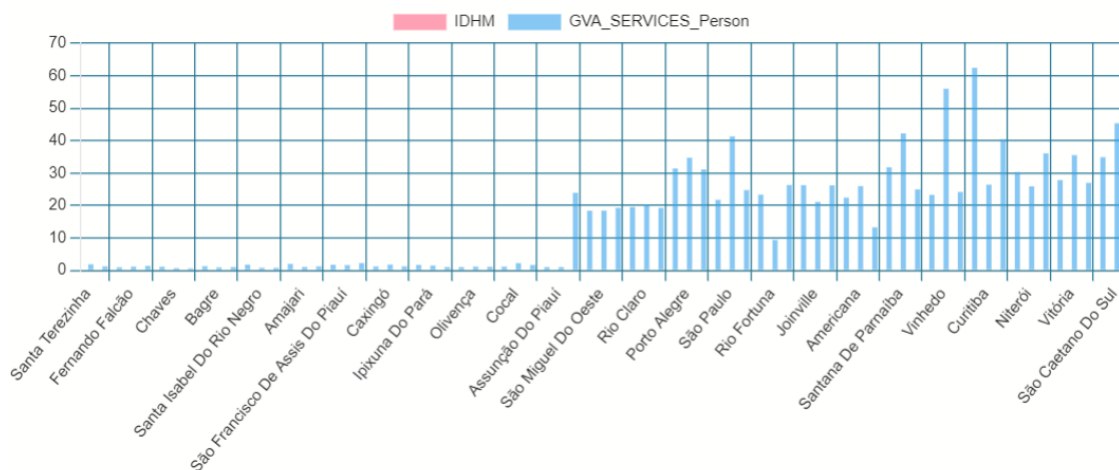
```
select DISTINCT [CITY]
, [IDHM]
, CAST (GVA_INDUSTY/IBGE_RES_POP as numeric (10,3)) as GVA_INDUSTY_Person
FROM [dbo].[Tabela_MVP_04]
WHERE [IDHM] > 0.80
or [IDHM] < 0.50
AND IBGE_RES_POP >0
ORDER BY [IDHM]
```



Onde é possível observar um GVA Industrial inferior nas cidades entre Santa Terezinha e Assunção do Piauí que são as cidades com IDHM inferior a 0.50 se

comparado às cidades entre São Miguel do Oeste e São Caetano do Sul que são cidades com o IDHM acima de 0.80.

```
select DISTINCT [CITY]
      , [IDHM]
      , CAST (GVA_SERVICES/IBGE_RES_POP as numeric (10,3)) as GVA_SERVICES_Person
FROM [dbo].[Tabela_MVP_04]
WHERE [IDHM] > 0.80
or [IDHM] < 0.50
AND IBGE_RES_POP >0
ORDER BY [IDHM]
```



Onde é possível observar um GVA de Serviços enormemente inferior nas cidades entre Santa Terezinha e Assunção do Piauí que são as cidades com IDHM inferior a 0.50 se comparado às cidades entre São Miguel do Oeste e São Caetano do Sul que são cidades com o IDHM acima de 0.80 onde é possível perceber um GVA elevado.

#### 4) CONCLUSÃO:

Por fim, foi possível concluir que baseado nas consultas realizadas, que não há uma relação entre o GVA de Serviços Públicos com o Índice de Desenvolvimento Municipal, tendo em vista que a média de GVA de serviços públicos por pessoa se mantem similar entre os IDHM acima de 0.80 e abaixo de 0.50.

Por outro lado, observou-se que há uma relação entre a arrecadação de impostos e os Gastos Municipais com o aumento do IDHM.

Além disso, foi possível observar que existe uma diferença importante entre o GVA de Serviços e o GVA de indústrias entre as Cidades com o IDHM abaixo de 0.50 e acima de 0.80, o que pode indicar que um investimento neste setor nas Cidades com o IDHM abaixo de 0.50, pode ocasionar em um possível aumento do IDHM dessas Cidades.

## AUTOAVALIAÇÃO

O desenvolvimento do trabalho ocorreu de forma satisfatória e os objetivos foram atingidos, sendo possível responder às perguntas inicialmente propostas e analisar outros fatos para conclusão do problema. Entretanto, poderia ter sido melhor desenvolvido com mais tempo para realizar mais análises, tendo em vista que haviam muitas variáveis e seria necessário realizar diversas consultas para um melhor entendimento de todos os fatores que podem influenciar ou não no IDHM.

Para uma melhora no projeto, verificou-se a possibilidade de realizar essa análise com dados mais atuais e talvez realizar um comparativo com o dataset atualizado de forma que fosse possível verificar quais valores aumentaram ou diminuíram que possam ter impactado no IDHM.

Uma grande dificuldade ocorreu em função do total desconhecimento acerca de todas as plataformas, o que fez com que grande parte do tempo para a elaboração deste trabalho fosse ocupado com o aprendizado da utilização das plataformas, desde as ferramentas de ETL da Azure até as ferramentas de análise, o qual poderia ser utilizado para um aprofundamento ainda maior nas análises.

Dois problemas que ocorreram durante a elaboração do projeto, foi inicialmente ao utilizar o Azure Data Factory, na etapa de realizar as transformações, as colunas não foram identificadas individualmente pelo fato do dataset possuir mais de 50 colunas, para tanto, foi necessário dividir a tabela em duas e carrega-las no repositório e depois importar dois datasets no Azure Data Factory, o que implicou na necessidade da utilização de um Join.

Outro problema que ocorreu, foi quando foi realizada a conexão do Notebook com o Banco de Dados e a coluna relativa aos Gastos Municipais vieram com alguns valores nulos, o que tomou certo tempo para descobrir o que estava ocasionando esse “bug”, ao verificar que era porque o número era grande demais e não deveria ter seu tipo classificado como Integer e sim Big Integer. No Data Factory ao tentar realizar a troca foi observado que não havia essa opção, porém, foi possível realizar a correção no próprio Azure Data Studio, que possui a opção de realizar algumas alterações nas tabelas importadas do Data Factory.

Por fim, não foi possível realizar a etapa de Catálogo de dados da forma esperada, dentro da plataforma da Azure, tendo em vista às dificuldades da utilização de uma solução de governança dentro da plataforma utilizada para a elaboração do projeto (Microsoft Purview). A solução foi realizar o catálogo de dados e os metadados em um arquivo separado que foi anexado ao repositório deste projeto no Github.



# ANEXO I

## 1. Código do Dataflow do Azure Data Factory

```
source(output(  
  Prop_0 as string,  
  CITY as string,  
  STATE as string,  
  CAPITAL as integer,  
  IBGE_RES_POP as integer,  
  IBGE_RES_POP_BRAS as integer,  
  IBGE_RES_POP_ESTR as integer,  
  IBGE_DU as integer,  
  IBGE_DU_URBAN as integer,  
  IBGE_DU_RURAL as integer,  
  IBGE_POP as integer,  
  IBGE_1 as integer,  
  {IBGE_1-4} as integer,  
  {IBGE_5-9} as integer,  
  {IBGE_10-14} as integer,  
  {IBGE_15-59} as integer,  
  {IBGE_60+} as integer,  
  IBGE_PLANTED_AREA as float,  
  {IBGE_CROP_PRODUCTION_$} as integer,  
  {IDHM Ranking 2010} as integer,  
  IDHM as float,  
  IDHM_Renda as float,  
  IDHM_Longevidade as float,  
  IDHM_Educacao as float,  
  LONG as integer,  
  LAT as integer,  
  ALT as float,  
  PAY_TV as integer,  
  FIXED_PHONES as integer,  
  AREA as float,  
  REGIAO_TUR as string,  
  CATEGORIA_TUR as string,  
  ESTIMATED_POP as integer,  
  RURAL_URBAN as string,  
  GVA_AGROPEC as float,  
  GVA_INDUSTRY as float,  
  GVA_SERVICES as float,  
  GVA_PUBLIC as float,  
  GVA_TOTAL as float,  
  TAXES as float  
),  
allowSchemaDrift: true,  
validateSchema: false,  
ignoreNoFilesFound: false) ~> Entrada1  
source(output(  

```

```

    Prop_0 as string,
    CITY as string,
    STATE as string,
    GDP as float,
    POP_GDP as integer,
    GDP_CAPITA as float,
    GVA_MAIN as string,
    MUN_EXPENDIT as string,
    COMP_TOT as integer,
    COMP_A as integer,
    COMP_B as integer,
    COMP_C as integer,
    COMP_D as integer,
    COMP_E as integer,
    COMP_F as integer,
    COMP_G as integer,
    COMP_H as integer,
    COMP_I as integer,
    COMP_J as integer,
    COMP_K as integer,
    COMP_L as integer,
    COMP_M as integer,
    COMP_N as integer,
    COMP_O as integer,
    COMP_P as integer,
    COMP_Q as integer,
    COMP_R as integer,
    COMP_S as integer,
    COMP_T as integer,
    COMP_U as integer,
    HOTELS as integer,
    BEDS as integer,
    Pr_Agencies as integer,
    Pu_Agencies as integer,
    Pr_Bank as integer,
    Pu_Bank as integer,
    Pr_Assets as integer,
    Pu_Assets as integer,
    Cars as integer,
    Motorcycles as integer,
    Wheeled_tractor as integer,
    UBER as integer,
    MAC as integer,
    {WAL-MART} as integer,
    POST_OFFICES as integer
  ),
  allowSchemaDrift: true,
  validateSchema: false,
  ignoreNoFilesFound: false) ~> Entrada2
Entrada1 select(mapColumn(

```

```

    CITY,
    STATE,
    CAPITAL,
    IBGE_RES_POP,
    IBGE_RES_POP_BRAS,
    IBGE_RES_POP_ESTR,
    IBGE_DU,
    IBGE_DU_URBAN,
    IBGE_DU_RURAL,
    IBGE_POP,
    IBGE_1,
    {IBGE_1-4},
    {IBGE_5-9},
    {IBGE_10-14},
    {IBGE_15-59},
    {IBGE_60+},
    IBGE_PLANTED_AREA,
    {IBGE_CROP_PRODUCTION_$},
    IDHM,
    IDHM_Renda,
    IDHM_Longevidade,
    IDHM_Educacao,
    ALT,
    AREA,
    CATEGORIA_TUR,
    RURAL_URBAN,
    GVA_AGROPEC,
    GVA_INDUSTRY,
    GVA_SERVICES,
    GVA_PUBLIC,
    GVA_TOTAL,
    TAXES
  ),
  skipDuplicateMapInputs: true,
  skipDuplicateMapOutputs: true) ~> select1
Entrada2 select(mapColumn(
  CITY,
  STATE,
  GDP,
  POP_GDP,
  GDP_CAPITA,
  GVA_MAIN,
  MUN_EXPENDIT,
  COMP_TOT,
  COMP_A,
  COMP_B,
  COMP_C,
  COMP_D,
  COMP_E,
  COMP_F,

```

```

        COMP_G,
        COMP_H,
        COMP_I,
        COMP_J,
        COMP_K,
        COMP_L,
        COMP_M,
        COMP_N,
        COMP_O,
        COMP_P,
        COMP_Q,
        COMP_R,
        COMP_S,
        COMP_T,
        COMP_U,
        Cars,
        Motorcycles
    ),
    skipDuplicateMapInputs: true,
    skipDuplicateMapOutputs: true) ~> select2
select1, select2 join(select1@CITY == select2@CITY
    && select1@STATE == select2@STATE,
    joinType: 'inner',
    matchType: 'exact',
    ignoreSpaces: false,
    broadcast: 'auto') ~> join1
join1 sink(allowSchemaDrift: true,
    validateSchema: false,
    deletable: false,
    insertable: true,
    updateable: false,
    upsertable: false,
    format: 'table',
    skipDuplicateMapInputs: true,
    skipDuplicateMapOutputs: true,
    errorHandlingOption: 'stopOnFirstError') ~> sink1

```

## 2. Código do Pipeline do Azure Data Factory

```

{
    "name": "pipeline_teste_final",

```

```

"properties": {
  "activities": [
    {
      "name": "dataflow2",
      "type": "ExecuteDataFlow",
      "dependsOn": [],
      "policy": {
        "timeout": "0.12:00:00",
        "retry": 0,
        "retryIntervalInSeconds": 30,
        "secureOutput": false,
        "secureInput": false
      },
      "userProperties": [],
      "typeProperties": {
        "dataflow": {
          "referenceName": "dataflow2",
          "type": "DataFlowReference"
        },
        "compute": {
          "coreCount": 8,
          "computeType": "General"
        },
        "traceLevel": "Fine"
      }
    }
  ],
  "annotations": [],
  "lastPublishTime": "2023-09-20T00:05:50Z"
},
"type": "Microsoft.DataFactory/factories/pipelines"
}

```

## ANEXO II

### Script da tabela no Azure Data Stúdio

```
CREATE TABLE [dbo].[Tabela_MVP_04] (  
    [CITY] NVARCHAR (50) NOT NULL,  
    [STATE] CHAR (2) NOT NULL,  
    [CAPITAL] INT NULL,  
    [IBGE_RES_POP] INT NULL,  
    [IBGE_RES_POP_BRAS] INT NULL,  
    [IBGE_RES_POP_ESTR] INT NULL,  
    [IBGE_DU] INT NULL,  
    [IBGE_DU_URBAN] INT NULL,  
    [IBGE_DU_RURAL] INT NULL,  
    [IBGE_POP] INT NOT NULL,  
    [IBGE_1] INT NULL,  
    [IBGE_1-4] INT NULL,  
    [IBGE_5-9] INT NULL,  
    [IBGE_10-14] INT NULL,  
    [IBGE_15-59] INT NULL,  
    [IBGE_60+] INT NULL,  
    [IBGE_PLANTED_AREA] REAL NULL,  
    [IBGE_CROP_PRODUCTION_$] INT NULL,  
    [IDHM] REAL NULL,  
    [IDHM_Renda] REAL NULL,  
    [IDHM_Longevidade] REAL NULL,  
    [IDHM_Educacao] REAL NULL,  
    [ALT] REAL NULL,  
    [AREA] REAL NULL,  
    [CATEGORIA_TUR] NVARCHAR (MAX) NULL,  
    [RURAL_URBAN] NVARCHAR (MAX) NULL,  
    [GVA_AGROPEC] REAL NULL,  
    [GVA_INDUSTRY] REAL NULL,  
    [GVA_SERVICES] REAL NULL,  
    [GVA_PUBLIC] REAL NULL,  
    [GVA_TOTAL] REAL NULL,  
    [TAXES] FLOAT (53) NULL,  
    [GDP] REAL NULL,  
    [POP_GDP] INT NULL,  
    [GDP_CAPITA] REAL NULL,  
    [GVA_MAIN] NVARCHAR (MAX) NULL,  
    [MUN_EXPENDIT] BIGINT NULL,  
    [COMP_TOT] INT NULL,  
    [COMP_A] INT NULL,  
    [COMP_B] INT NULL,  
    [COMP_C] INT NULL,  
    [COMP_D] INT NULL,  
    [COMP_E] INT NULL,  
    [COMP_F] INT NULL,  
    [COMP_G] INT NULL,  
    [COMP_H] INT NULL,  
    [COMP_I] INT NULL,  
    [COMP_J] INT NULL,  
    [COMP_K] INT NULL,  
    [COMP_L] INT NULL,  
    [COMP_M] INT NULL,  
    [COMP_N] INT NULL,  
    [COMP_O] INT NULL,
```

```

[COMP_P]                INT                NULL,
[COMP_Q]                INT                NULL,
[COMP_R]                INT                NULL,
[COMP_S]                INT                NULL,
[COMP_T]                INT                NULL,
[COMP_U]                INT                NULL,
[Cars]                  INT                NULL,
[Motorcycles]           INT                NULL
);

```

GO

```

EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Outras Atividades de Serviços.', @level0type = N'SHEMA', @level0name =
N'dbo', @level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type =
N'COLUMN', @level2name = N'COMP_S';

```

GO

```

EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número total de
companias', @level0type = N'SHEMA', @level0name = N'dbo', @level1type = N'TABLE',
@level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'COMP_TOT';

```

GO

```

EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Administração Pública, Defesa e Segurança Social.', @level0type =
N'SHEMA', @level0name = N'dbo', @level1type = N'TABLE', @level1name =
N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'COMP_O';

```

GO

```

EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Informação e Comunicação.', @level0type = N'SHEMA', @level0name = N'dbo',
@level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN',
@level2name = N'COMP_J';

```

GO

```

EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'População de
Estrangeiros residentes', @level0type = N'SHEMA', @level0name = N'dbo', @level1type
= N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =
N'IBGE_RES_POP ESTR';

```

GO

```

EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Population',
@level0type = N'SHEMA', @level0name = N'dbo', @level1type = N'TABLE', @level1name =
N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'POP_GDP';

```

GO

```

EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Construção.', @level0type = N'SHEMA', @level0name = N'dbo', @level1type =
N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =
N'COMP_F';

```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Alojamento e Alimentação.', @level0type = N'SCHEMA', @level0name = N'dbo',
@level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN',
@level2name = N'COMP_I';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'IDH Renda taxa',
@level0type = N'SCHEMA', @level0name = N'dbo', @level1type = N'TABLE', @level1name =
N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'IDHM_Renda';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Comércio; Reparação de Veículos Automotores e Motocicletas.', @level0type =
N'SCHEMA', @level0name = N'dbo', @level1type = N'TABLE', @level1name =
N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'COMP_G';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Transporte, Armazenamento e Correio.', @level0type = N'SCHEMA', @level0name
= N'dbo', @level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type =
N'COLUMN', @level2name = N'COMP_H';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Gestão de valor
adicionado bruto - Serviços', @level0type = N'SCHEMA', @level0name = N'dbo',
@level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN',
@level2name = N'GVA_SERVICES';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Eletricidade e Gás', @level0type = N'SCHEMA', @level0name = N'dbo',
@level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN',
@level2name = N'COMP_D';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Atividades Financeiras, de Seguros e Serviços Relacionados.', @level0type =
N'SCHEMA', @level0name = N'dbo', @level1type = N'TABLE', @level1name =
N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'COMP_K';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Área da cidade',
@level0type = N'SCHEMA', @level0name = N'dbo', @level1type = N'TABLE', @level1name =
N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'AREA';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Nome do Estado
do Brasil', @level0type = N'SCHEMA', @level0name = N'dbo', @level1type = N'TABLE',
@level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'STATE';
```



```
GO
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Despesas
municipais - in reais', @level0type = N'SHEMA', @level0name = N'dbo', @level1type =
N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =
N'MUN_EXPENDIT';
```

```
GO
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Atividades de Água, Saneamento, Gestão de Resíduos e Descontaminação.',
@level0type = N'SHEMA', @level0name = N'dbo', @level1type = N'TABLE', @level1name =
N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'COMP_E';
```

```
GO
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Educação.', @level0type = N'SHEMA', @level0name = N'dbo', @level1type =
N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =
N'COMP_P';
```

```
GO
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Moradores em
domicílios particulares - de 15 a 69 anos de idade', @level0type = N'SHEMA',
@level0name = N'dbo', @level1type = N'TABLE', @level1name = N'Tabela_MVP_04',
@level2type = N'COLUMN', @level2name = N'IBGE_15-59';
```

```
GO
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Indústrias de Transformação', @level0type = N'SHEMA', @level0name =
N'dbo', @level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type =
N'COLUMN', @level2name = N'COMP_C';
```

```
GO
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Gestão de valor
adicionado bruto - Agropecuaria', @level0type = N'SHEMA', @level0name = N'dbo',
@level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN',
@level2name = N'GVA_AGROPEC';
```

```
GO
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Atividades Imobiliárias.', @level0type = N'SHEMA', @level0name = N'dbo',
@level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN',
@level2name = N'COMP_L';
```

```
GO
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Moradores em
domicílios particulares - de 1 a 4 anos de idade', @level0type = N'SHEMA',
@level0name = N'dbo', @level1type = N'TABLE', @level1name = N'Tabela_MVP_04',
@level2type = N'COLUMN', @level2name = N'IBGE_1-4';
```

```
GO
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Moradores em
domicílios particulares - de 10 a 14 anos de idade', @level0type = N'SHEMA',
```

```
@level0name = N'dbo', @level1type = N'TABLE', @level1name = N'Tabela_MVP_04',  
@level2type = N'COLUMN', @level2name = N'IBGE_10-14';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'IDH taxa da  
longevidade', @level0type = N'SHEMA', @level0name = N'dbo', @level1type = N'TABLE',  
@level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =  
N'IDHM_Longevidade';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Moradores em  
domicílios particulares - de 4 a 8 anos de idade', @level0type = N'SHEMA',  
@level0name = N'dbo', @level1type = N'TABLE', @level1name = N'Tabela_MVP_04',  
@level2type = N'COLUMN', @level2name = N'IBGE_5-9';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'População de  
Brasileiros Residentes', @level0type = N'SHEMA', @level0name = N'dbo', @level1type =  
N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =  
N'IBGE_RES_POP_BRAS';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Total Gestão de  
valor adicionado bruto', @level0type = N'SHEMA', @level0name = N'dbo', @level1type =  
N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =  
N'GVA_TOTAL';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Numero total de  
carros', @level0type = N'SHEMA', @level0name = N'dbo', @level1type = N'TABLE',  
@level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'Cars';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de  
Empresas: Indústrias Extrativas', @level0type = N'SHEMA', @level0name = N'dbo',  
@level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN',  
@level2name = N'COMP_B';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de  
Empresas: Atividades Profissionais, Científicas e Técnicas.', @level0type =  
N'SHEMA', @level0name = N'dbo', @level1type = N'TABLE', @level1name =  
N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'COMP_M';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Moradores em  
domicílios particulares', @level0type = N'SHEMA', @level0name = N'dbo', @level1type  
= N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =  
N'IBGE_POP';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Moradores em  
domicílios particulares - acima de 60 anos de idade', @level0type = N'SHEMA',  
@level0name = N'dbo', @level1type = N'TABLE', @level1name = N'Tabela_MVP_04',  
@level2type = N'COLUMN', @level2name = N'IBGE_60+';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Produto Interno  
Bruto', @level0type = N'SHEMA', @level0name = N'dbo', @level1type = N'TABLE',  
@level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'GDP';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Atividade com o  
GVA principal', @level0type = N'SHEMA', @level0name = N'dbo', @level1type =  
N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =  
N'GVA_MAIN';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Moradores em  
domicílios particulares - até 1 ano de idade', @level0type = N'SHEMA', @level0name =  
N'dbo', @level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type =  
N'COLUMN', @level2name = N'IBGE_1';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'População  
Residente', @level0type = N'SHEMA', @level0name = N'dbo', @level1type = N'TABLE',  
@level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =  
N'IBGE_RES_POP';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de  
Empresas: Saúde Humana e Serviços Sociais.', @level0type = N'SHEMA', @level0name =  
N'dbo', @level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type =  
N'COLUMN', @level2name = N'COMP_Q';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de  
Empresas: Serviços Domésticos.', @level0type = N'SHEMA', @level0name = N'dbo',  
@level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN',  
@level2name = N'COMP_T';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Área plantada ou  
destinada a colheita em hectares', @level0type = N'SHEMA', @level0name = N'dbo',  
@level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN',  
@level2name = N'IBGE_PLANTED_AREA';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'IDH taxa da  
educação', @level0type = N'SHEMA', @level0name = N'dbo', @level1type = N'TABLE',
```

```
@level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =  
N'IDHM_Educacao';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Total de  
domicilios rurais', @level0type = N'SHEMA', @level0name = N'dbo', @level1type =  
N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =  
N'IBGE_DU_RURAL';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Categoria de  
Turismo', @level0type = N'SHEMA', @level0name = N'dbo', @level1type = N'TABLE',  
@level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =  
N'CATEGORIA_TUR';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Nome do  
Município do Brasil', @level0type = N'SHEMA', @level0name = N'dbo', @level1type =  
N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =  
N'CITY';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Produto Interno  
Bruto per capita', @level0type = N'SHEMA', @level0name = N'dbo', @level1type =  
N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =  
N'GDP_CAPITA';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Total de  
domolicios urbanos', @level0type = N'SHEMA', @level0name = N'dbo', @level1type =  
N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =  
N'IBGE_DU_URBAN';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Gestão de valor  
adicionado bruto - Serviços Públicos', @level0type = N'SHEMA', @level0name = N'dbo',  
@level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN',  
@level2name = N'GVA_PUBLIC';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Produção  
agrícola', @level0type = N'SHEMA', @level0name = N'dbo', @level1type = N'TABLE',  
@level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =  
N'IBGE_CROP_PRODUCTION_$';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de  
Empresas: Instituições Internacionais e Outras Instituições Extraterritoriais.',  
@level0type = N'SHEMA', @level0name = N'dbo', @level1type = N'TABLE', @level1name =  
N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'COMP_U';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Agricultura, Pecuária, Silvicultura, Pesca e Aquicultura', @level0type =
N'SCHEMA', @level0name = N'dbo', @level1type = N'TABLE', @level1name =
N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'COMP_A';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Atividades Administrativas e Serviços Complementares.', @level0type =
N'SCHEMA', @level0name = N'dbo', @level1type = N'TABLE', @level1name =
N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'COMP_N';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Impostos',
@level0type = N'SCHEMA', @level0name = N'dbo', @level1type = N'TABLE', @level1name =
N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'TAXES';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Taxa do IDH',
@level0type = N'SCHEMA', @level0name = N'dbo', @level1type = N'TABLE', @level1name =
N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name = N'IDHM';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Total de
unidades domesticas', @level0type = N'SCHEMA', @level0name = N'dbo', @level1type =
N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =
N'IBGE_DU';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'1 se for capital
de algum estado', @level0type = N'SCHEMA', @level0name = N'dbo', @level1type =
N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =
N'CAPITAL';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Elevação da
cidade (metros)', @level0type = N'SCHEMA', @level0name = N'dbo', @level1type =
N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =
N'ALT';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número Total de
Motocicletas', @level0type = N'SCHEMA', @level0name = N'dbo', @level1type = N'TABLE',
@level1name = N'Tabela_MVP_04', @level2type = N'COLUMN', @level2name =
N'Motorcycles';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Número de
Empresas: Artes, Cultura, Esporte e Recreação.', @level0type = N'SCHEMA', @level0name
= N'dbo', @level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type =
N'COLUMN', @level2name = N'COMP_R';
```

GO

```
EXECUTE sp_addextendedproperty @name = N'MS_Description', @value = N'Gestão de valor  
adicionado bruto - Industria', @level0type = N'SHEMA', @level0name = N'dbo',  
@level1type = N'TABLE', @level1name = N'Tabela_MVP_04', @level2type = N'COLUMN',  
@level2name = N'GVA_INDUSTRY';
```