

Figura 1 – Hans Peter Luhn



Fonte: (IEEE Spectrum, 2018)

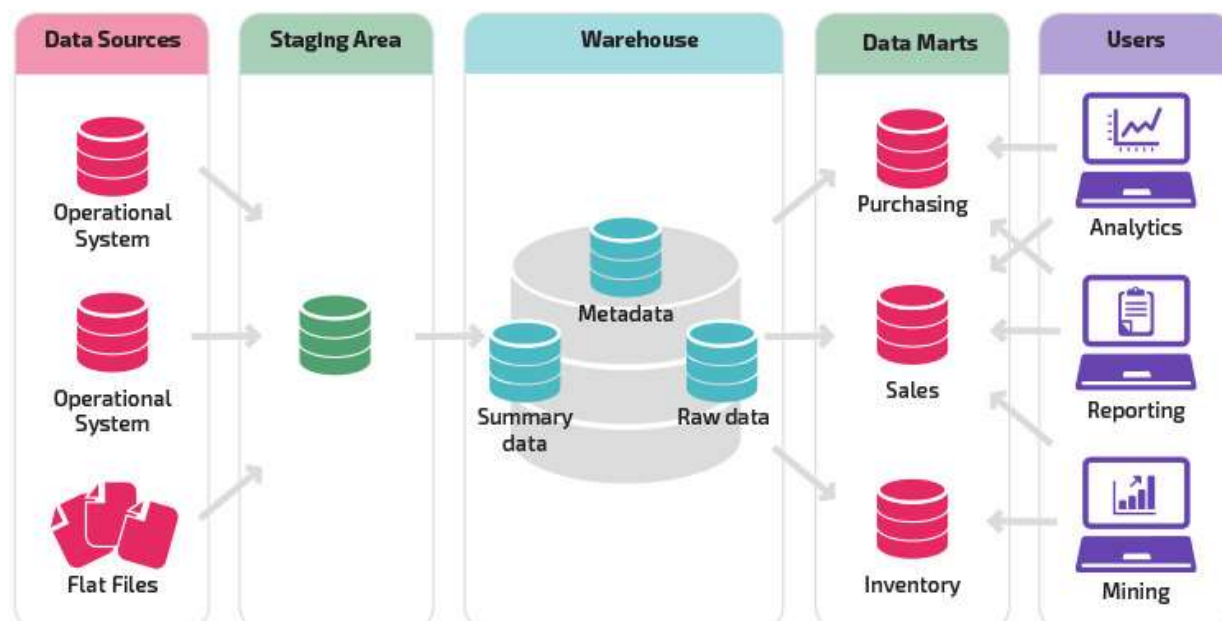
A partir dos anos 60, houve o surgimento de novas formas de armazenamento como os DBMS (*Database Management Systems*), tendo uma evolução no modo de gerenciar grandes volumes de dados, no final da década de 70, nasce o modelo relacional no DBMS. A partir desse momento, todas as informações eram apresentadas e armazenadas em formato digital, fazendo com que fosse possível a concretização do Business Intelligence nas próximas décadas.

Também nessa mesma época, os CPD (Centros de Processamento de Dados), estavam se consolidando, se transformando no meio-termo da tecnologia da informação com os negócios. Os CPD eram focados totalmente em dados, diferente da TI que o centro focava em software, hardware e redes.

Com o surgimento do conceito de sistemas de informações executivas (EIS) há uma grande disseminação do assunto, sendo um dos maiores aliados aos sistemas de BI. Segundo Turban (2009, p. 27), "Esse conceito expandiu o suporte computadorizado aos gerentes e executivos de nível superior. Alguns dos recursos introduzidos foram sistemas de geração de relatórios dinâmicos multidimensionais [...]".

Na década de 80, alguns fabricantes de softwares voltados ao campo do BI começaram a ganhar terreno. Softwares como *MicroStrategy*, *Business Objects* e

Figura 2 - Arquitetura do ambiente de BI



Fonte: Panoly (2019).

4.2 Montagem do ambiente – Fontes de Dados (*Data Source*).

O primeiro passo na aplicação dos processos de Business Intelligence é definir quais serão as bases de dados utilizadas para o processo e quais dados serão extraídos delas. No caso do presente trabalho, foram utilizadas as bases de micro dados do censo escolar do INEP, disponíveis no Portal Brasileiro de Dados Abertos no link: <http://dados.gov.br/dataset/microdados-do-censo-escolar> e no próprio site do INEP no link: <http://inep.gov.br/web/guest/microdados>. Para a melhor delimitação do trabalho, foram utilizados os censos dos anos de 2015 a 2018.

Os arquivos estão em formato CSV (*Comma-separated Values*) que é um tipo de arquivo onde seus dados estão separados por algum delimitador, no caso das bases do INEP é utilizado o delimitador *Pipe* (|). Eles são divididos em Turmas, Escolas, Matrículas (Centro-Oeste, Nordeste, Norte, Sudeste e Sul), e Docentes (Centro-Oeste, Nordeste, Norte, Sudeste e Sul), onde se encontra as informações das turmas, das escolas, dos alunos e dos docentes envolvidos nos censos de cada ano, respectivamente.

Além dos dados principais, faz-se necessário o uso de tabelas auxiliares para auxiliar na definição dos dados do INEP, já que são utilizados campos com os códigos dos Países, Unidades da Federação (UF), Municípios, Distritos, Mesorregiões e

Microrregiões. Para o primeiro, o INEP disponibiliza em sua base, ao fazer download, uma tabela (Figura 3) que contém os códigos dos países descritos no censo, já que alunos estrangeiros também são envolvidos no censo escolar.

Figura 3 - Tabela de códigos dos países

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Anexo 4 - Tabela de Países			Legenda:													
3	(referente à variável CO_PAIS_ORIGEM das tabelas de matrícula e de docente)			Não coletado de 2015 em diante													
5	Cód.	Código ISO ALPHA-3	Nome do País ou Área	Coleta por ano ("s"=sim,"n"=não)													
6				07	08	09	10	11	12	13	14	15	16	17	18		
8	4	AFG	AFEGANISTÃO	s	s	s	s	s	s	s	s	s	s	s	s	s	s
9	8	ALB	ALBÂNIA	s	s	s	s	s	s	s	s	s	s	s	s	s	s
10	12	DZA	ARGÉLIA	s	s	s	s	s	s	s	s	s	s	s	s	s	s
11	16	ASM	SAMOA AMERICANA	s	s	s	s	s	s	s	s	s	s	s	s	s	s
12	20	AND	ANDORRA	s	s	s	s	s	s	s	s	s	s	s	s	s	s
13	24	AGO	ANGOLA	s	s	s	s	s	s	s	s	s	s	s	s	s	s
14	28	ATG	ANTÍGUA E BARBUDA	s	s	s	s	s	s	s	s	s	s	s	s	s	s
15	31	AZE	AZERBAIJÃO	s	s	s	s	s	s	s	s	s	s	s	s	s	s
16	32	ARG	ARGENTINA	s	s	s	s	s	s	s	s	s	s	s	s	s	s
17	36	AUS	AUSTRÁLIA	s	s	s	s	s	s	s	s	s	s	s	s	s	s
18	40	AUT	ÁUSTRIA	s	s	s	s	s	s	s	s	s	s	s	s	s	s
19	44	BHS	BAHAMAS	s	s	s	s	s	s	s	s	s	s	s	s	s	s
20	48	BHR	BAHREIN	s	s	s	s	s	s	s	s	s	s	s	s	s	s
21	50	BGD	BANGLADESH	s	s	s	s	s	s	s	s	s	s	s	s	s	s
22	51	ARM	ARMÊNIA	s	s	s	s	s	s	s	s	s	s	s	s	s	s
23	52	BRB	BARBADOS	s	s	s	s	s	s	s	s	s	s	s	s	s	s
24	56	BEL	BÉLGICA	s	s	s	s	s	s	s	s	s	s	s	s	s	s
25	60	BMU	BERMUDAS	s	s	s	s	s	s	s	s	s	s	s	s	s	s
26	64	BTN	BUTÃO	s	s	s	s	s	s	s	s	s	s	s	s	s	s
27	68	BOL	BOLÍVIA (ESTADO PLURINACIONAL DA)	s	s	s	s	s	s	s	s	s	s	s	s	s	s
28	70	BIH	BÓSNIA E HERZEGOVINA	s	s	s	s	s	s	s	s	s	s	s	s	s	s
29	72	BWA	BOTSUANA	s	s	s	s	s	s	s	s	s	s	s	s	s	s
30	76	BRA	BRASIL	s	s	s	s	s	s	s	s	s	s	s	s	s	s
31	84	BLZ	REIJE	s	s	s	s	s	s	s	s	s	s	s	s	s	s

Fonte: Adaptado de INEP (2019).

Para as UF, Municípios, Distritos, Mesorregiões e Microrregiões, foram utilizadas as bases de códigos *Geodata* disponíveis no site *GitHub* no link: <https://github.com/paulofreitas/geodata-br/tree/master/data/pt>. O *GitHub* é um site para a criação de repositórios públicos e privados com o intuito de compartilhar informações e códigos e o repositório *Geodata* tem como propósito prover informações precisas e atualizadas acerca dos dados geográficos do Brasil. Essas informações são um compilado das informações disponíveis na SIDRA (Sistema IBGE de Recuperação Automática) formados pelo IBGE (Instituto Brasileiro de Geografia e Estatística).

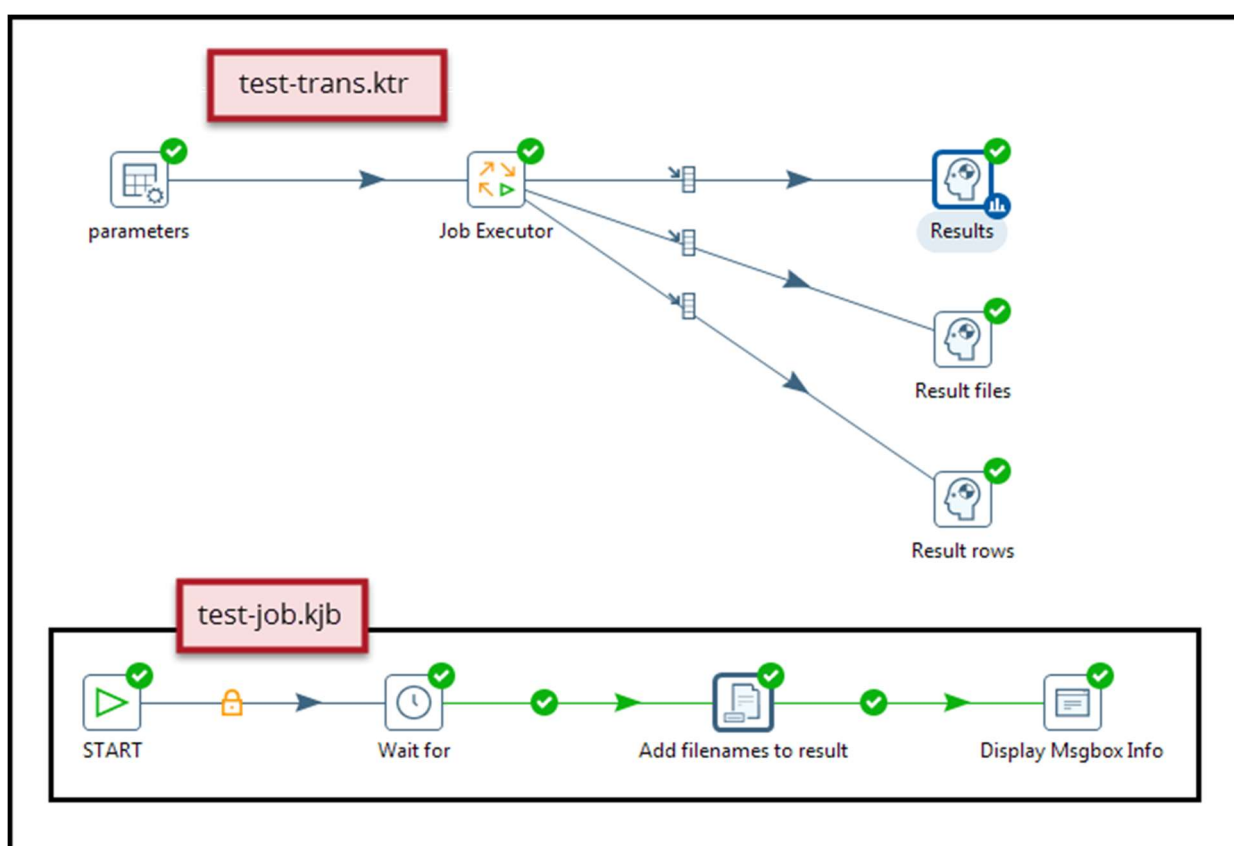
4.3 Montagem do ambiente – Área de *Staging*

Inmon (2005, p. 29) define em um dos seus postulados sobre *Data Warehouse* a não volatilidade, ou seja, os dados dentro do mesmo não podem sofrer alterações. Isso significa que se faz necessária uma fase intermediária antes de carregar os dados

no DW, para isso tem-se a *Staging Area* ou *Data Stage*. Com todos os dados já na máquina é iniciada a montagem dos processos de ETL para fazer a carga no Banco de Dados de *Staging*.

Será utilizado o *Pentaho Data Integrator* (PDI) versão 5.0.1 para iniciar os processos de ETL, separando as cargas por assunto. O PDI utiliza duas nomenclaturas como *Job* e *Transformation*, o primeiro é a menor ação possível que o programa possa fazer como ler o arquivo ou fazer inserção, e o segundo é um conjunto de outros *Jobs* para fazer uma execução única e contínua. Como na imagem abaixo:

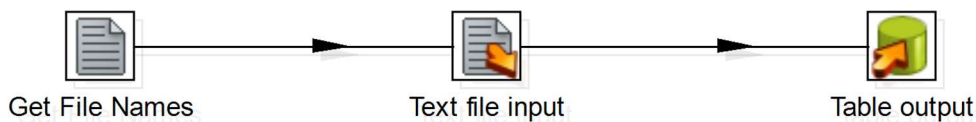
Figura 4 - Exemplo de Transformation e Job



Fonte: Pentaho (20-?).

A carga dos arquivos no BD dos arquivos principais (turmas, matrícula, escolas, docentes) e das bases de códigos das UF, Municípios, Distritos, Mesorregiões e Microrregiões são compostas por três passos, em que o PDI encontra os arquivos, prepara-os para a inserção e grava-os no BD, como pode ser visto na imagem abaixo:

Figura 5 - Visão da ETL das bases principais



Fonte: Autores (2019).

Os passos são descritos abaixo:

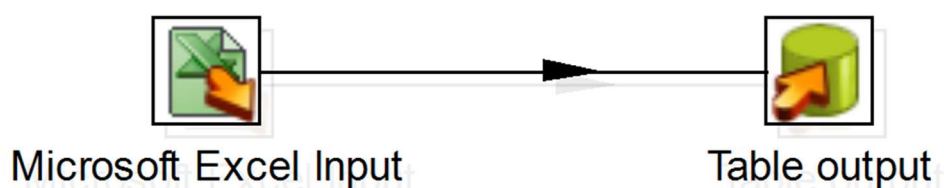
Get File Names: Esse *step* procura nomes de arquivos ou pastas. Ele é recomendado para quando se tem uma grande massa de dados em que todos precisam ser gravados. Os padrões dos nomes são adquiridos conforme uma expressão regular.

Text File Input: Aqui o Pentaho prepara um ou mais arquivos de textos para a inserção, nele são configuradas diversas opções como os delimitadores do texto, linha de título, formato e colunas adicionais para serem adicionadas no momento da carga.

Table Output: Realiza a carga dos dados estruturados em uma tabela no banco de dados. A tabela não precisa ser criada com antecedência, já que o PDI prepara um comando SQL automaticamente para a mesma ser criada.

Já para a carga das tabelas contendo o código dos países, foi utilizado um padrão de carga diferente, já que o arquivo que possui esse dado está em um formato diferente das outras bases, como pode ser visto na imagem abaixo:

Figura 6 - Visão geral da ETL de auxiliares



Fonte: Autores (2019).

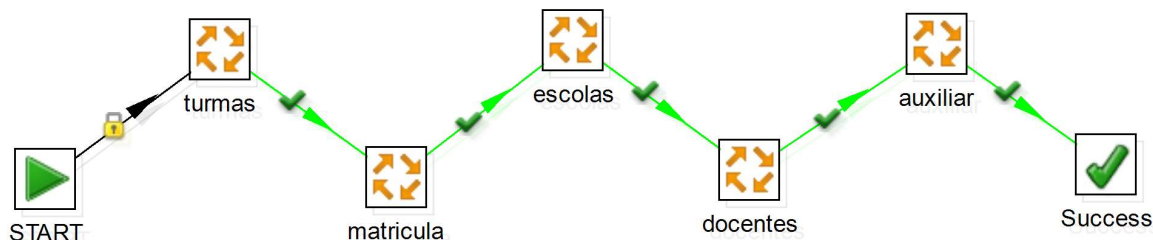
Os passos são descritos abaixo:

Microsoft Excel Input: Esse *step* procura nomes de arquivos do tipo XLS (formato utilizado nas versões de 97 até 2003) e/ou XLSX (utilizado na versão de 2007 em diante). Nele podem-se configurar opções como, especificar de qual linha e/ou coluna deve-se iniciar a análise, se os títulos das colunas estão na primeira linha (*Header*), além de especificar campos adicionais no momento da carga.

Table Output: Como descrito nas cargas principais, esse passo realiza a carga dos dados estruturados em uma tabela no banco de dados. A tabela não precisa ser criada com antecedência, já que o PDI prepara um comando SQL automaticamente para a mesma ser criada.

Após definir cada uma das ETLs, será usado uma *Transformation* para unir todos os outros *Jobs*, como pode ser visto na imagem abaixo:

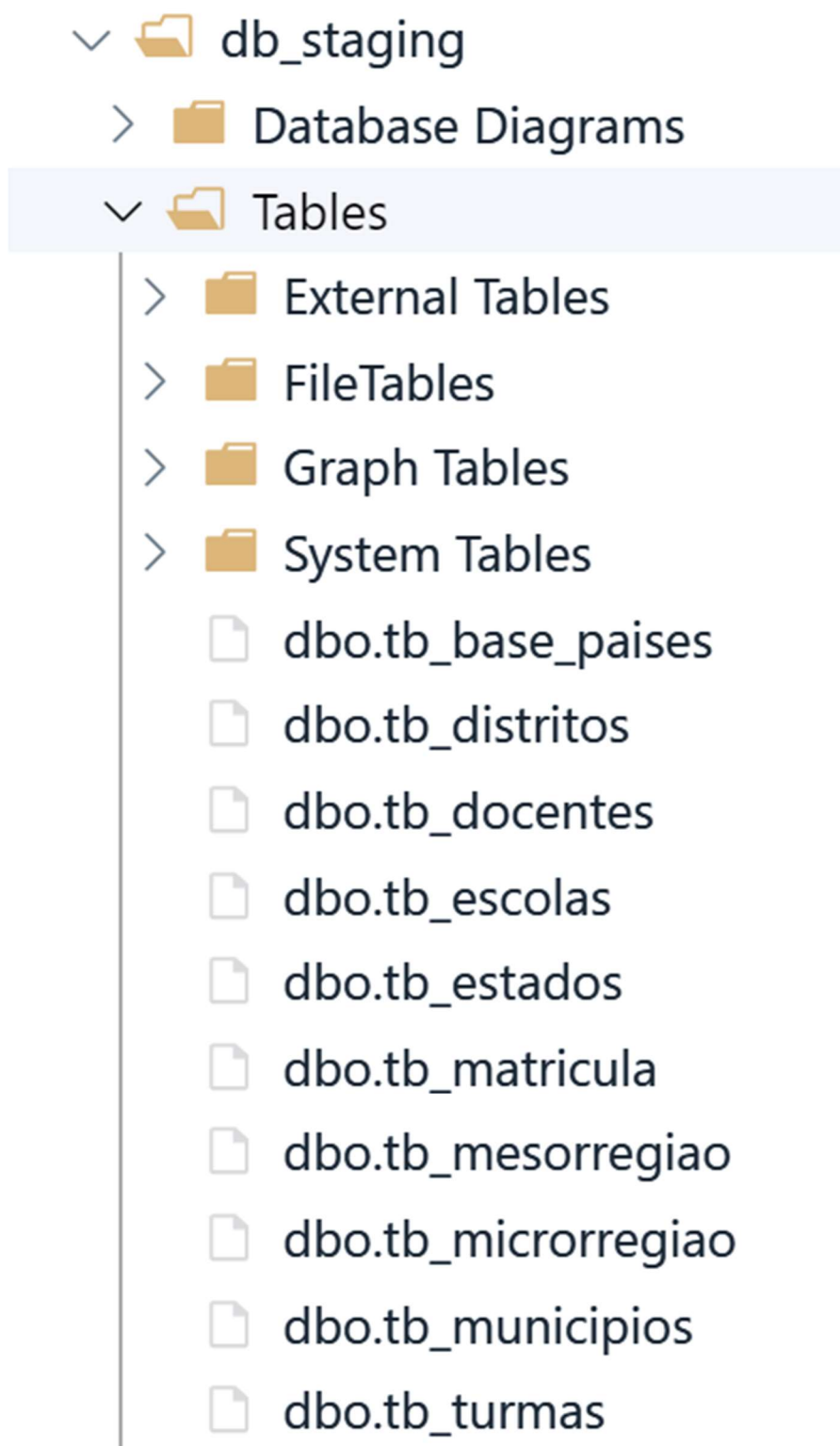
Figura 7 - Visão geral da ETL Staging



Fonte: Autores (2019).

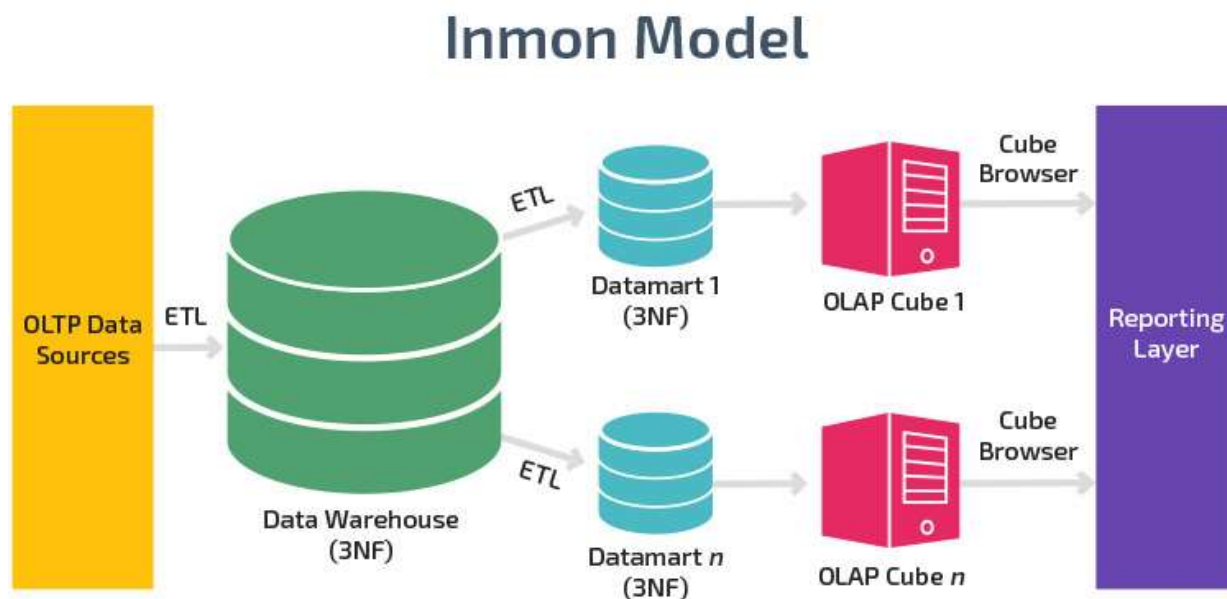
Com todo o fluxo executado, o banco de dados *Staging* foi finalizado na forma da imagem abaixo:

Figura 8 - Visão do Banco Staging



Fonte: Autores (2019).

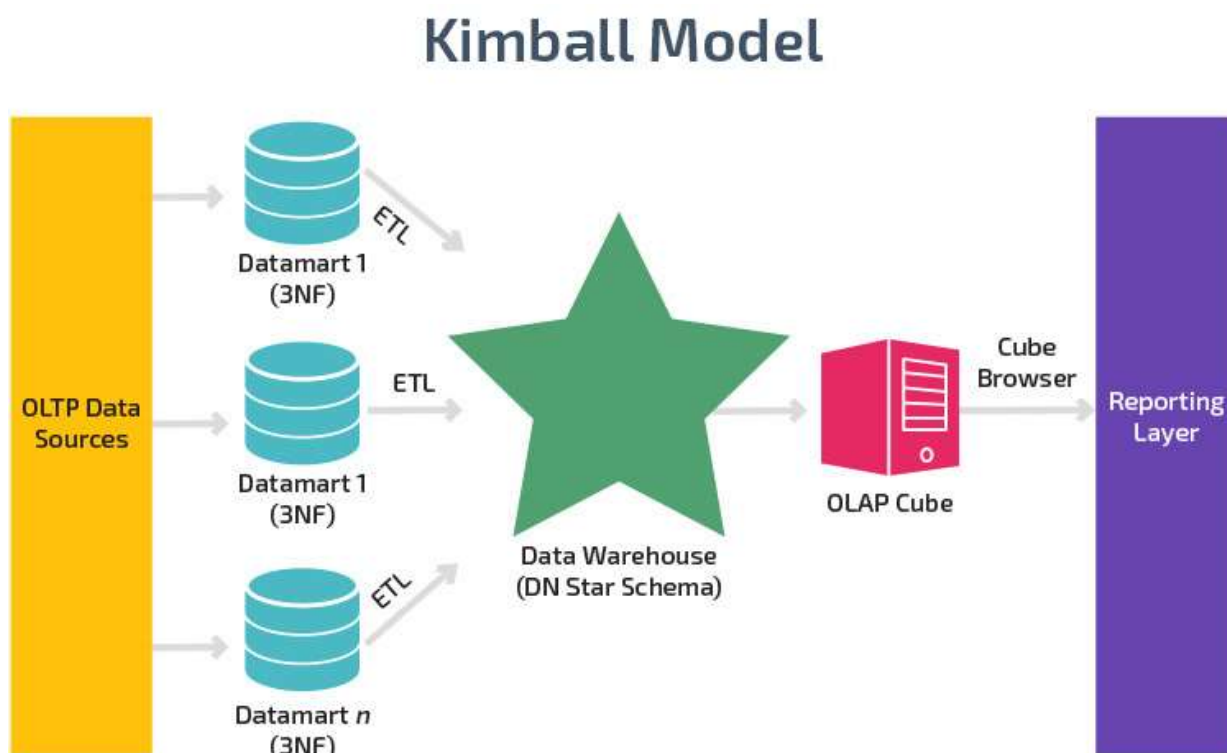
Figura 9 - Modelo Inmon



Fonte: Panoly (2019).

Em contrapartida, Kimball utiliza a abordagem *bottom-up* em que é feita primeiramente a criação de *Data Marts* em cada área de interesse para depois se criar um grande *Data Warehouse* que é unicamente uma junção de todos esses *Marts* (PANOLY, 2019). Tal como Kimball (2013) afirma: “O *Data Warehouse* não é nada mais do que uma junção de diversos *Data Marts*”. Seu diagrama é mostrado abaixo:

Figura 10 - Modelo Kimball



Fonte: Panoly (2019).

Abaixo é feita uma comparação entre as abordagens Inmon e Kimball:

Tabela 1 - Comparação entre as abordagens do Inmon e Kimball

	Inmon	Kimball
Manutenção	Fácil	Difícil - muitas vezes redundante e sujeito a revisão
Tempo	Maior tempo para iniciar	Menor tempo para iniciar
Conhecimento preciso	Time especialista	Time generalista
Tempo de construção do Data Warehouse	Demorado	Rápido
Custo para implantar	Custos iniciais altos, com menores custos subsequentes de desenvolvimento do projeto	Custos iniciais pequenos, com cada projeto subsequente custando-o mais
Persistência dos dados	Alta taxa de mudança dos dados	Relativamente estável

Fonte: Autores (2019)

Para a realização desse trabalho foi escolhida a abordagem Inmon porque o projeto não terá uso de *Data Marts*, assim sendo, será criado unicamente o *Data Warehouse* para armazenar os dados.

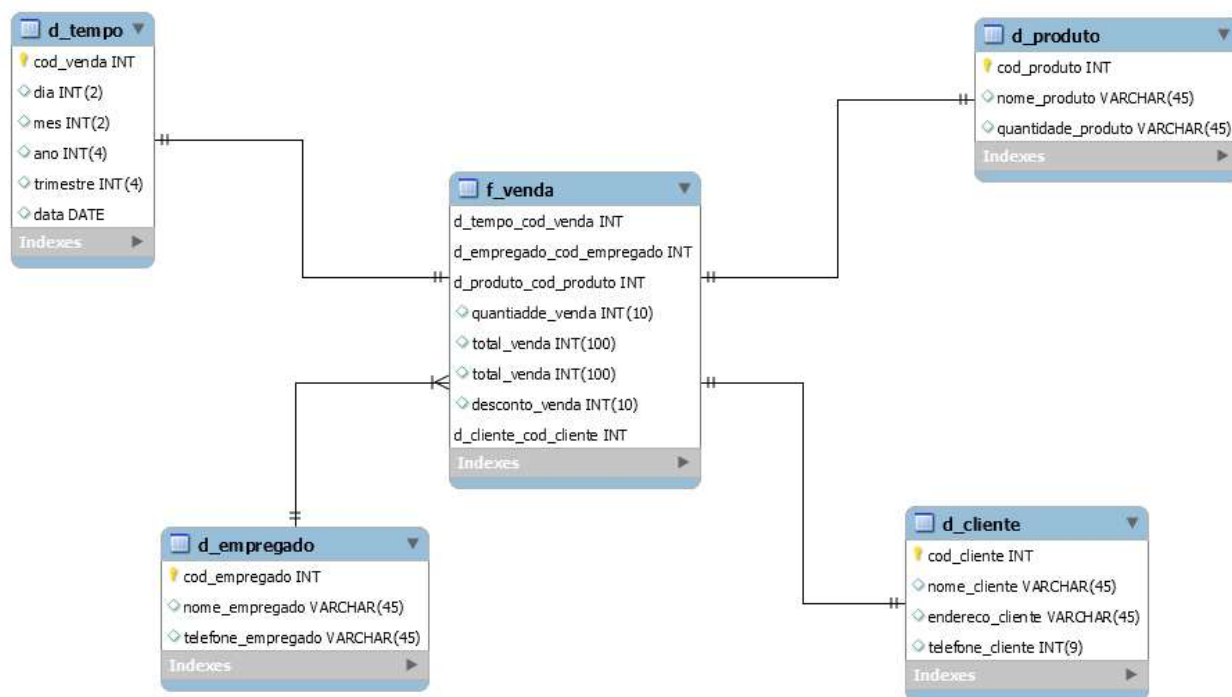
4.4.3 Modelos Estrela e Floco de Neve (*Star Schema and Snow-Flake Schema*)

Tendo definida a estrutura, inicia-se o desenvolvimento do modelo do DW.

Em um modelo de dados multidimensional pode ser utilizado dois tipos de modelos, que são: tipo Estrela (*Star Schema*) ou tipo Floco de Neve (*Snow-Flake Schema*).

O modelo Estrela é o mais básico e mais comum para a arquitetura do *Data Warehouse*. No seu desenho, a tabela fato (F_VENDA, Figura 11) assume o centro da arquitetura seguido pelas tabelas de dimensões, que em volta dela, definem a quantidade de pontas da Estrela (CARVALHAES e ALVES, 2015). Possui como vantagem uma visualização simplificada dos dados, além de mais agilidade nas análises.

Figura 11 - Exemplo de modelo Estrela

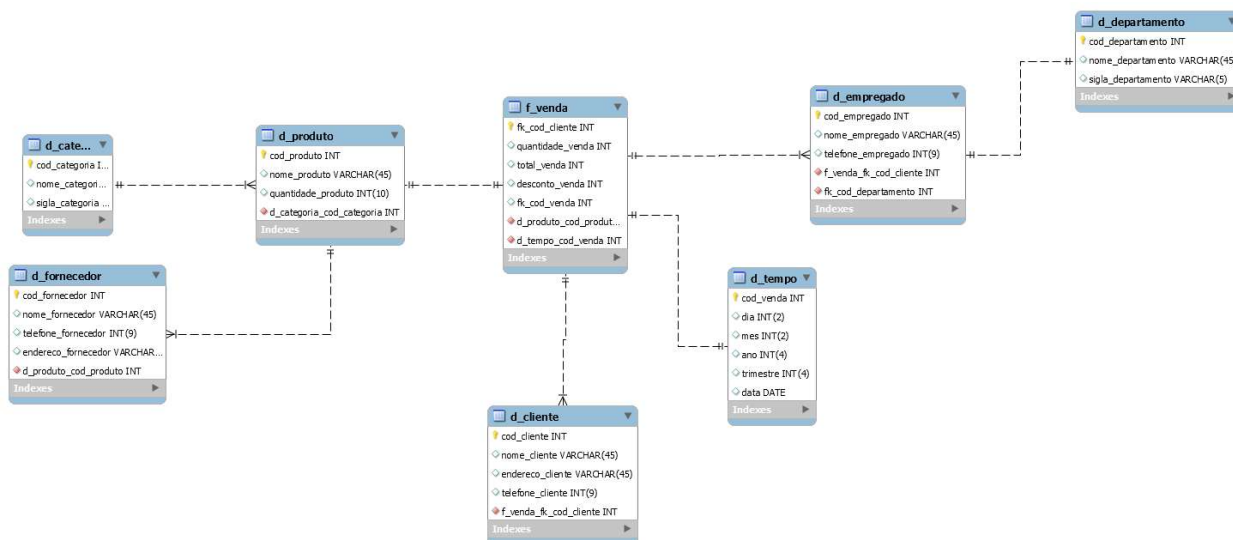


Fonte: Autores (2019).

O modelo Floco de Neve é um modelo específico que, partindo do modelo Estrela, as dimensões que possuem hierarquia são decompostas em outras tabelas (CARVALHAES e ALVES, 2015). Nesse modelo tem-se uma redução de redundância nas tabelas de dimensões e uma menor quantidade de memória utilizada. Um exemplo seria a dimensão chamada Data, em que ela poderá ser decomposta em

outras tabelas como dia, mês, ano, trimestre, etc. Assim, essas “sub-dimensões” vão compor a dimensão principal. Seu diagrama é mostrado abaixo:

Figura 12 - Exemplo de modelo Floco de Neve



Fonte: Autores (2019).

Abaixo é feita uma comparação entre os modelos Estrela e Floco de Neve:

Tabela 2 - Comparação entre Star Schema e Snow Flake Schema

	<i>Star Schema</i>	<i>Snow Flake</i>
Manutenção	Possui dados redundantes que dificultam manter ou alterar	Sem redundância, portanto, facilita manter e alterar
Facilidade de Uso	Consultas menos complexas e de fácil entendimento	As consultas são mais complexas o que torna difícil de entender
Performance nas consultas	Menos <i>foreign keys</i> o que torna a consulta mais rápida	Mais <i>foreign keys</i> o que torna a consulta mais lenta
Espaço	Não tem tabelas normalizadas o que aumenta o espaço	Tem tabelas normalizadas
Bom para:	Bom para <i>Data Mart</i> com relacionamentos simples (1:1 ou 1:N)	Bom para uso em <i>Data Warehouse</i> para simplificar as relações complexas (N:N)

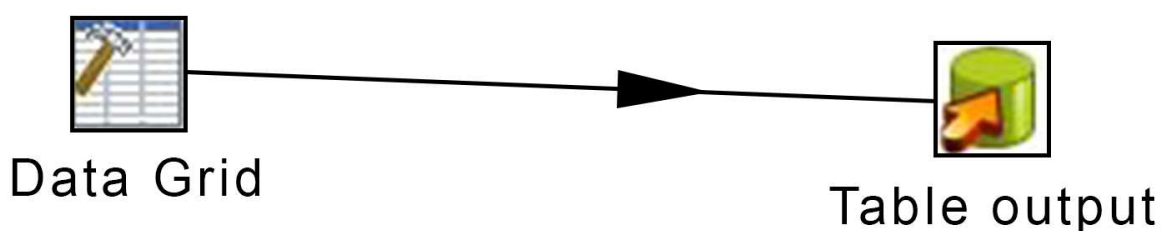
Fonte: Autores (2019).

Conforme definido pelos dois autores, todo projeto necessita obrigatoriamente de uma dimensão de tempo, em contrapartida, esse 'tempo' pode ser descrito de formas diferentes por cada projeto, com base nas necessidades das análises. Ele pode ser definido tanto como informações separadas (ano ou mês ou dia), uma data, formada por ano, mês, e dia, ou até uma informação mais complexa inserindo trimestre, dia da semana, hora, etc.

Para o presente trabalho, a dimensão de tempo será identificada pelos anos referentes a cada análise (2015 até 2018), um identificador para cada uma delas e seu indicador nulo que será explicado adiante.

Seu diagrama de carga é mostrado na imagem abaixo:

Figura 13 - Visão geral da ETL Ano



Fonte: Autores (2019).

Os seguintes passos foram utilizados:

Data Grid: Neste *step* será criada uma tabela com um conjunto constante de dados, informando os nomes dos campos, seus tipos, e seus respectivos dados. Aqui está sendo carregado cada um dos anos da análise e um indicador para cada um deles.

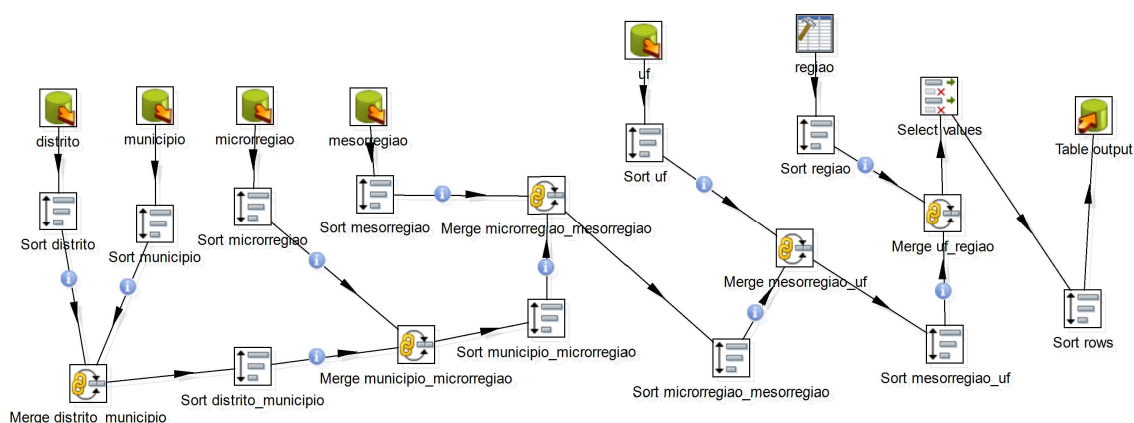
Table Output: Conforme explicado na parte do *Staging*, esse passo realiza a carga dos dados estruturados em uma tabela no banco de dados. A tabela não precisa ser criada com antecedência, já que o PDI prepara um comando SQL automaticamente para a mesma ser criada. Aqui os dados estão sendo gravados na tabela D_TEMPO do banco de dados.

Indicador nulo da dimensão:

pela combinação das informações de distrito, município, microrregião, mesorregião, UF e região, junto de um identificador único para cada uma dessas combinações, além dos identificadores nulos.

Seu diagrama de carga é mostrado na imagem abaixo:

Figura 14 - Diagrama da ETL Localidade Distrito



Fonte: Autores (2019).

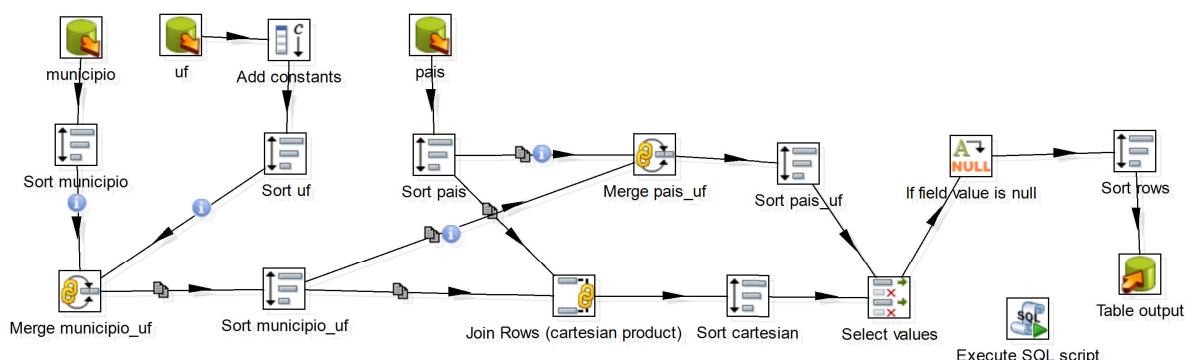
Os seguintes passos foram utilizados:

Table Input (na imagem acima com os nomes: distrito, município, microrregião, mesorregião, uf): Este *step* permite utilizar os dados já existentes em alguma tabela para fazer outras operações, como a inserção em outro banco, por exemplo. Os dados nesse passo são adquiridos por meio de um comando SQL, mas o *Pentaho* possui uma interface gráfica para selecionar esses dados, sem necessidade do comando, se assim o usuário preferir.

Data Grid: Neste *step* será criada uma tabela com um conjunto constante de dados, informando os nomes dos campos, seus tipos, e seus respectivos dados. Aqui está sendo carregado cada uma das regiões da análise.

Sort rows (na imagem acima com os nomes: Sort distrito, Sort município, Sort microrregião, Sort mesorregião, Sort região, Sort distrito_município, Sort município_microrregião, Sort microrregião_mesorregião, Sort mesorregião_uf): Esse passo possibilita a ordenação de um conjunto de dados com base em uma coluna informada. Seu uso é semelhante ao comando *order by* do SQL. Nele podem ser

Figura 15 - Diagrama da ETL Localidade Município



Fonte: Autores (2019).

Os seguintes passos foram utilizados:

Table Input (na imagem acima com os nomes: municipio, uf, pais): Como descrito na carga anterior, este *step* permite utilizar os dados já existentes em alguma tabela para fazer outras operações, como a inserção em outro banco, por exemplo. Os dados nesse passo são adquiridos por meio de um comando SQL, mas o *Pentaho* possui uma interface gráfica para selecionar esses dados, sem necessidade do comando, se assim o usuário preferir.

Sort rows (na imagem acima com os nomes: Sort municipio, Sort uf, Sort pais, Sort municipio_uf, Sort pais_uf, Sort cartesian): Como dito na carga anterior, esse passo possibilita a ordenação de um conjunto de dados com base em uma coluna informada. Seu uso é semelhante ao comando *order by* do SQL. Nele podem ser configuradas outras opções como ordenação ascendente ou descendente e diferenciação de maiúsculas e minúsculas.

Merge Join (na imagem acima com os nomes: Merge município_uf, Merge pais_uf): Explicado na carga anterior, possui um funcionamento semelhante ao comando *JOIN* do SQL. Esse *step* une dois fluxos de informação com base em uma coluna compartilhada (*key*), além de ser possível a configuração da forma de união (retornar apenas os dados que se relacionam ou também aqueles que não se relacionam). Requer o uso do *step Sort rows* antes deste para ordenação da coluna escolhida.

(-1, -1, -1); quando o aluno tem como país de origem o Brasil, mas não possui informações referentes ao seu município e sua UF (-1, -1, 76).

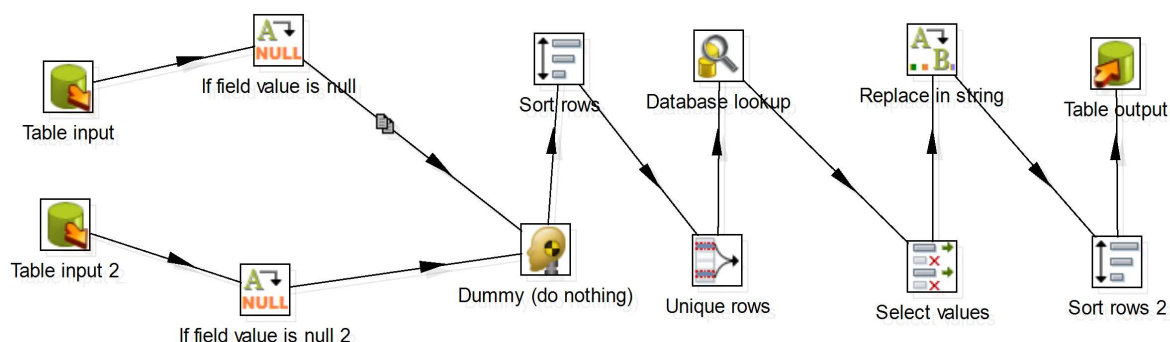
-2: Caso o aluno for estrangeiro. Esse indicador foi criado em uma combinação: Quando o aluno tiver como país de origem qualquer valor diferente de 76 (que faz referência ao Brasil) e suas informações de UF e município não estiverem disponíveis, por exemplo, se o aluno for natural dos Estados Unidos: (-2, -2, 840).

-3: Caso o aluno for naturalizado/nascido no exterior. Esse indicador foi criado em uma combinação: Quando o aluno for naturalizado/nascido no exterior e suas informações de município e UF não estiverem disponíveis (-3, -3, 76).

4.4.5.4 Dimensão Escola

Para a carga da dimensão das escolas, a ordem de ações consiste na extração dos dados dos dois tipos de escolas (públicas e privadas), transformação dos dados inserindo os indicadores de nulo dos dois tipos de escolas, remover as informações duplicadas e posterior inserção delas no *Data Warehouse*, como segue:

Figura 16 - Diagrama da ETL Escola



Fonte: Autores (2019).

Os passos estão descritos a seguir:

Table Input: Este *step* permite utilizar os dados já existentes em alguma tabela para fazer outras operações, como a inserção em outro banco, por exemplo. Os dados nesse passo são adquiridos por meio de um comando SQL, mas o *Pentaho* possui uma interface gráfica para selecionar esses dados, sem necessidade do comando, se assim o usuário preferir. A utilização de dois passos será explicada adiante.

informações referentes às escolas pelo o seu significado segundo o dicionário de dados do INEP. Por exemplo, a coluna IN_AGUA_INEXISTENTE indica se a escola possui ou não abastecimento de água, no fluxo, os dados existentes nessa coluna são 0 para 'Não' e 1 para 'Sim', assim, esse *step* faz essa substituição do valor numérico pelo seu valor de significado. Ao finalizar, os dados são mais uma vez ordenados e inseridos na tabela da dimensão escolar.

Indicadores nulos da dimensão:

Como citado anteriormente, essa dimensão possuirá os seguintes indicadores de informação nula:

-1: Caso alguma informação estiver nula.

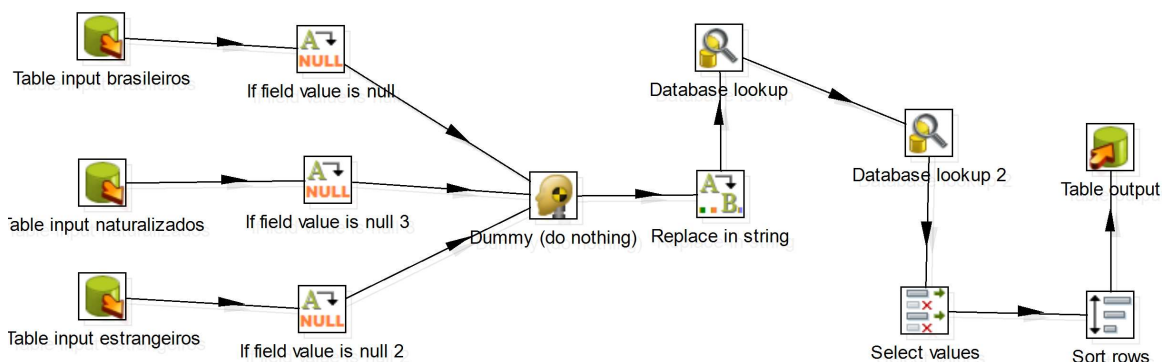
-2: Inserido nas colunas referentes aos mantenedores privados e na categoria de escola privada, quando a escola for pública, já que essas duas informações não se aplicam a uma escola que é pública.

4.4.5.5 Fato Aluno

Agora na última tabela do *Data Warehouse*, tem-se a fato aluno que é gerada após todas as dimensões estiverem prontas no BD.

Na sua carga, o processo é semelhante à carga da dimensão escolas, com o diferencial da substituição dos anos da análise por seus respectivos códigos definidos na dimensão ano no momento da carga.

Figura 17 - Diagrama da ETL Aluno



Fonte: Autores (2019).

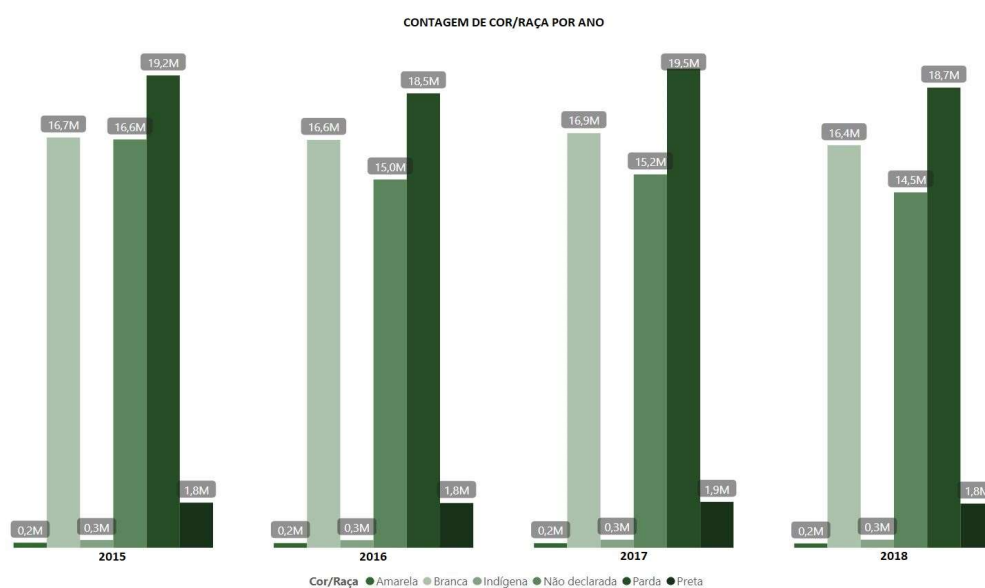
5 RESULTADOS DA ANÁLISE

Ao finalizar todas as ETLs para o *Data Warehouse* e geração dos indicadores pela ferramenta *Power BI* foi possível realizar a análise desses indicadores. Foram gerados quinze indicadores apresentados abaixo, a partir da leitura, interpretação de como eles poderiam ser úteis como indicadores e da bibliografia consultada constante no capítulo 3 deste trabalho.

Este trabalho não tem a pretensão de cobrir todo o assunto, mas pode ser útil para indicar caminhos para outras análises. É importante lembrar que essas análises são afetas a um recorte temporal, a saber, os anos de 2015 a 2018 da Educação Básica dos estados, municípios e distritos brasileiros.

1. Qual o total de alunos por cada Cor/Raça definida pelo Censo Escolar entre os anos da análise?

Figura 18 - Contagem de cor/raça por ano

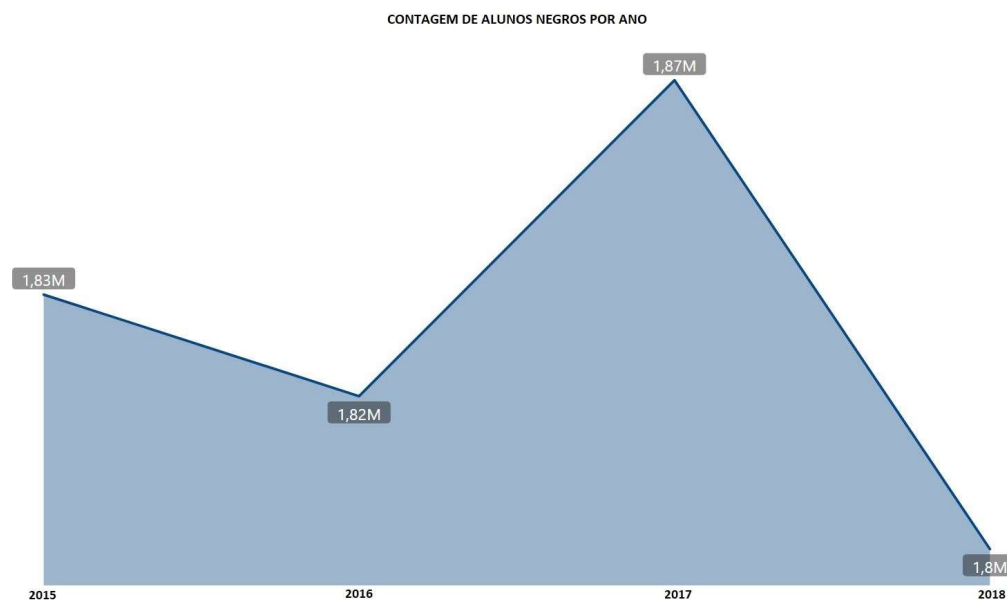


Fonte: Autores (2019).

Segundo o gráfico da figura acima a cor/raça de maior quantidade da base é dos alunos que se consideraram pardos, mantendo quase 20 milhões de alunos entre todos os anos da análise, logo após tem-se a Cor/Raça branca, atingindo quase 17 milhões, além dos que preferiram não se declarar, com a sua menor quantidade em 2018 onde estiveram com 14 milhões. Os negros se mantêm entre 1,8 e 1,9 milhões, sendo a quarta maior Cor/Raça na base do INEP.

2. Qual o total de alunos que se declararam negros entre os anos da análise?

Figura 19 - Contagem de alunos negros por ano

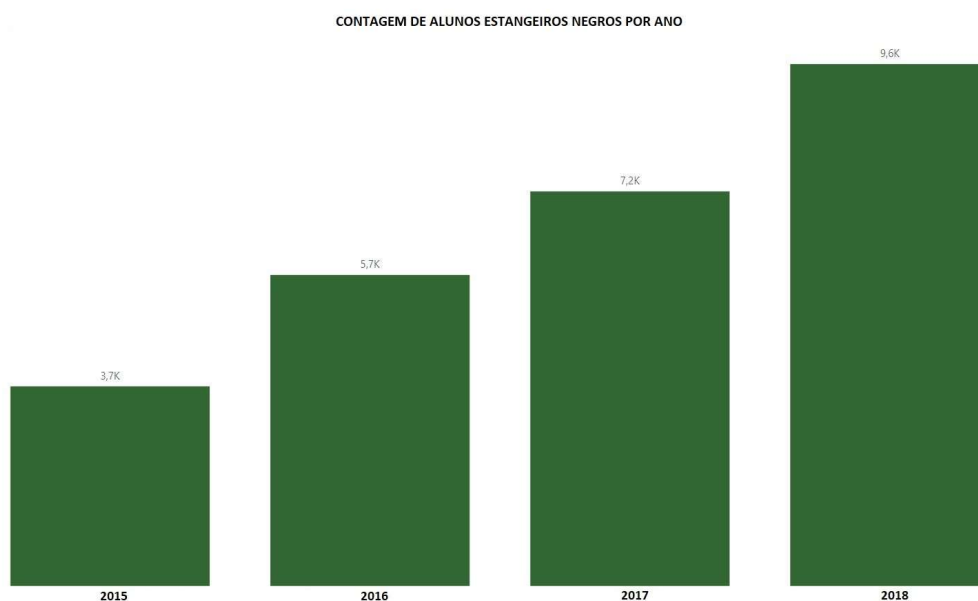


Fonte: Autores (2019).

Segundo o gráfico para o segundo indicador, a quantidade de alunos negros na base do INEP não passou de 1,9 milhões. Sua menor quantidade foi em 2018 onde se teve apenas 1,8 milhões de alunos que se declararam negros e seu maior pico foi em 2017 tendo 1,87 alunos com auto declaração negra.

3. Qual o total de alunos estrangeiros que se declararam negros entre os anos da análise?

Figura 20 - Contagem de alunos estrangeiros negros por ano

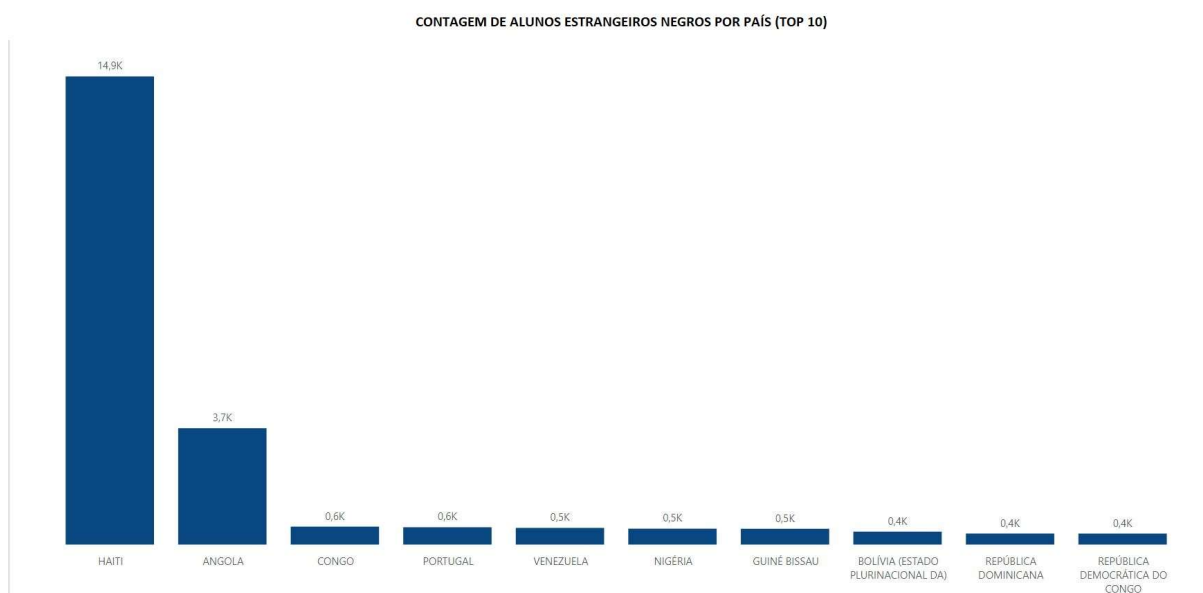


Fonte: Autores (2019).

Para o gráfico referente ao terceiro indicador, é possível notar um aumento desde o início da análise em 2015, de 3,7 mil alunos, até o último ano onde chegou a marca de quase 10 mil alunos estrangeiros segundo a base do INEP.

4. Qual o país que possui a maior quantidade de alunos estrangeiros negros no Brasil entre os anos da análise?

Figura 21 - Contagem de alunos estrangeiros negros por país (Top 10)



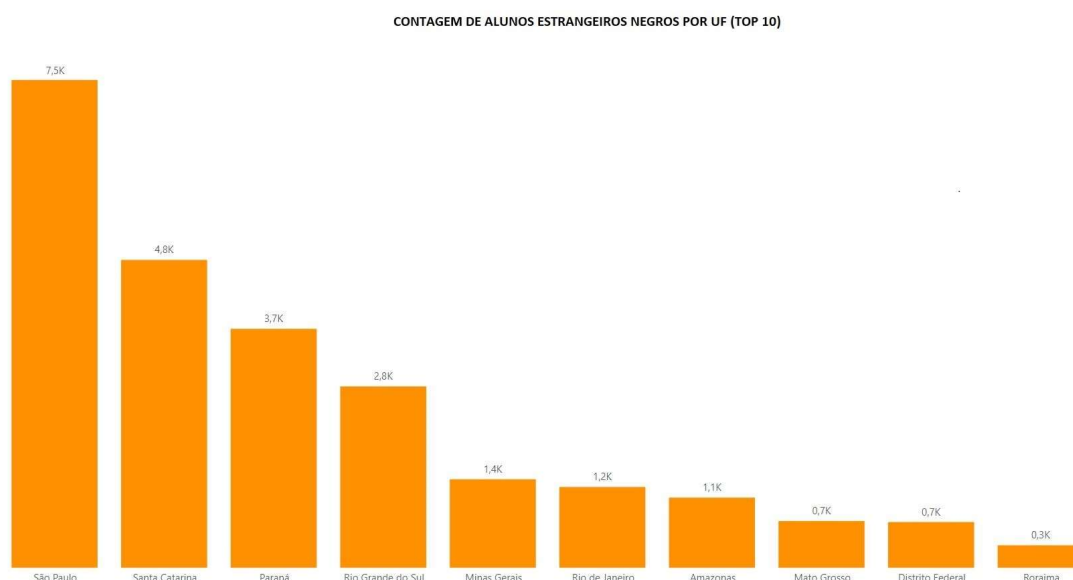
Fonte: Autores (2019).

Para o quarto indicador, por questões de visualização, a análise foi reduzida para mostrar apenas os dez primeiros países que mais possuem alunos no Brasil, na educação básica segundo a base do INEP. O Haiti se mostra o país com a maior quantidade, chegando a quase 15 mil alunos, seguido por Angola e Congo. Portugal é o quarto país e os Estados Unidos estão abaixo desses 10 primeiros.

Não faz parte deste trabalho a correlação das missões de paz no Haiti com o fato deste país ser aquele com mais estrangeiros negros na Educação Básica, porém trata-se de um possível estudo futuro.

5. Qual a UF que possui a maior concentração de alunos estrangeiros negros?

Figura 22 - Contagem de alunos estrangeiros negros por UF (Top 10)

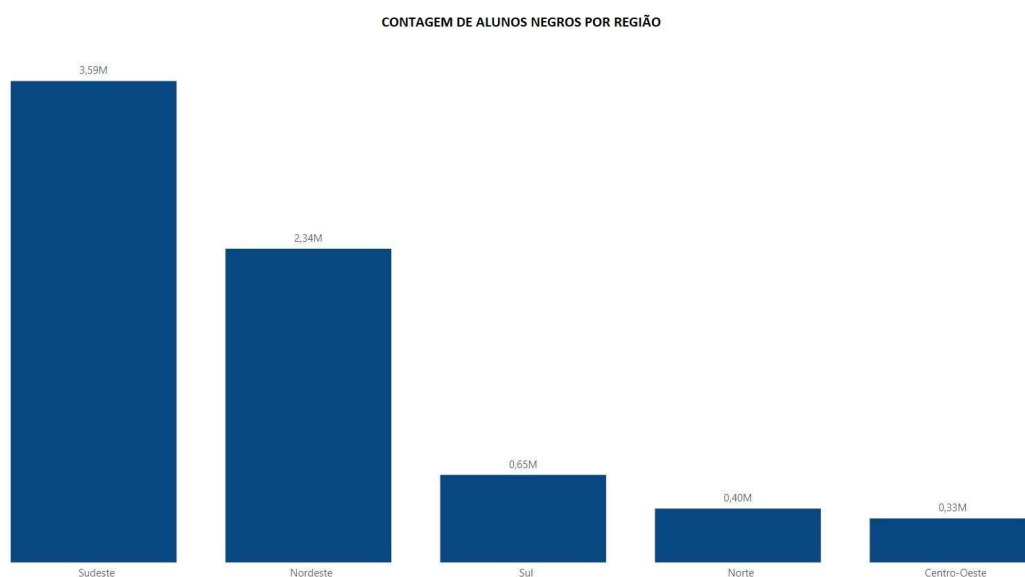


Fonte: Autores (2019).

Para o gráfico acima, por questões de visualização, o gráfico foi reduzido para mostrar apenas os 10 primeiros. A UF onde mais se concentram os alunos estrangeiros negros é o estado de São Paulo, seguido por Santa Catarina e Paraná. O Distrito Federal é o nono maior, chegando a quase mil registros.

6. Qual o total de alunos negros por região, UF e município entre os anos da análise?

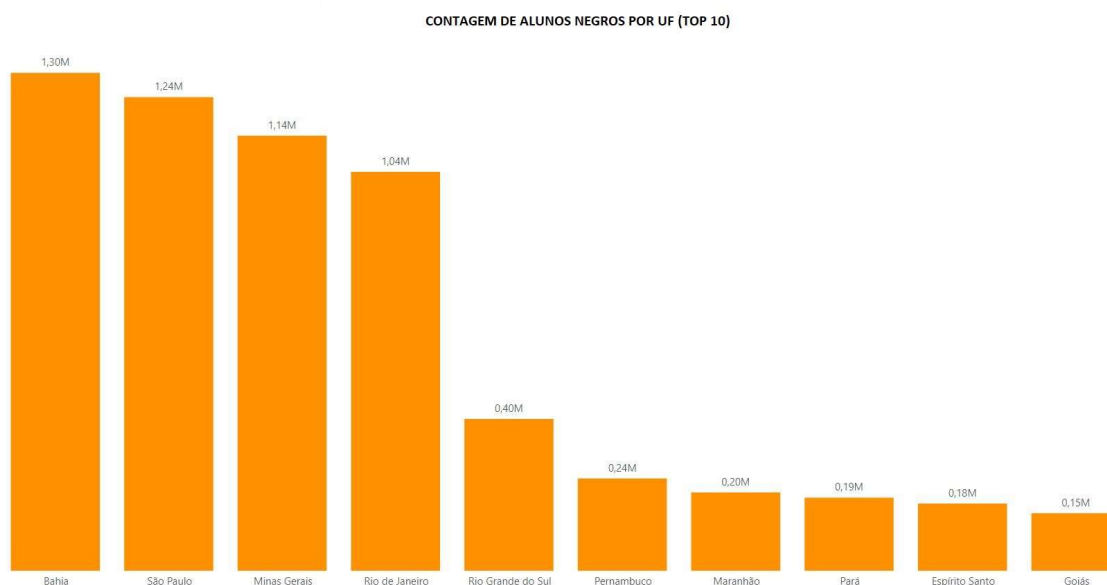
Figura 23 - Contagem de alunos negros por região



Fonte: Autores (2019).

Segundo o gráfico acima, o maior registro de alunos se concentra na região sudeste onde tem-se 3,59 milhões de dados, seguido logo após pelo nordeste com 2,34 milhões de registros, o sul vem depois com 65 mil resultados, após o norte com 40 mil resultados, e por último o centro-oeste com 33 mil resultados.

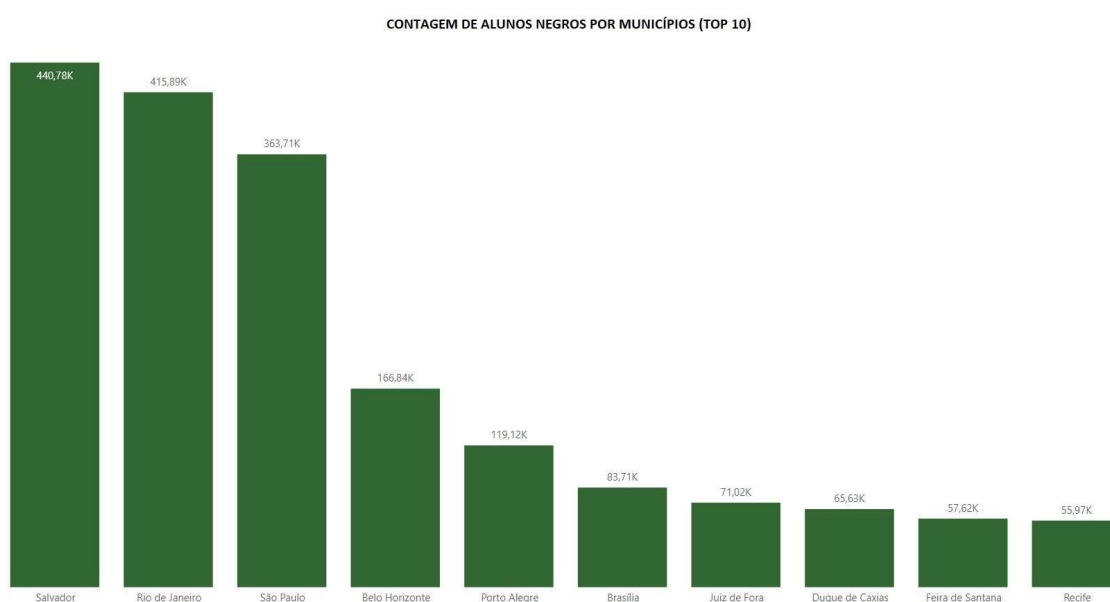
Figura 24 - Contagem de alunos negros por UF (Top 10)



Fonte: Autores (2019).

Segundo o gráfico acima (por questões de visualização foi reduzido para mostrar apenas os 10 primeiros), a maior quantidade de registros dos alunos negros se concentra no estado da Bahia, com 1,3 milhões de resultados. Logo após vem São Paulo com 1,24 milhões e Minas Gerais com 1,14 milhões, fechando os três primeiros. O Distrito Federal não fica entre os 10 primeiros nessa análise.

Figura 25 - Contagem de alunos negros por município (Top 10)



Fonte: Autores (2019).

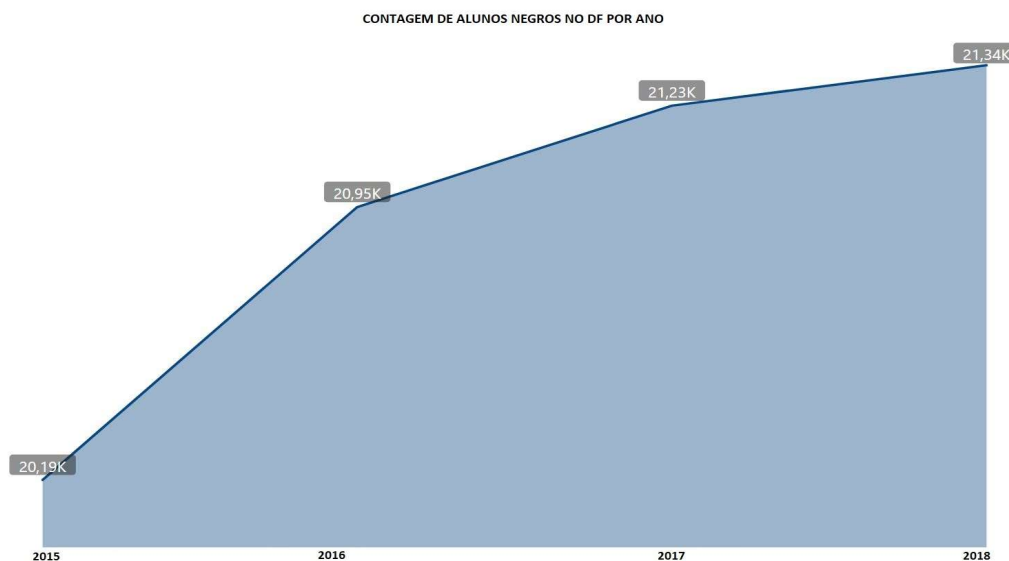
Segundo o gráfico acima (que por questões de visualização, foi reduzido para mostrar apenas os 10 primeiros), o município com a maior concentração de alunos negros é o município de Salvador na Bahia com 440 mil alunos, seguido pelo município do Rio de Janeiro com 415 mil resultados e após o município de São Paulo com 363 mil resultados. Brasília aparece na análise como o sexto maior município com 83 mil alunos (na base do INEP, Brasília, mesmo sendo oficialmente um distrito, possui um código de município para manter a padronização).

- Qual é a diferença de alunos negros entre as regiões nordeste e sudeste nos anos da análise?

A figura 23 do indicador anterior também responde este, demonstrando a diferença de quase 1,25 milhões entre a região sudeste e nordeste.

- Qual é a quantidade de alunos negros no Distrito Federal entre os anos da análise?

Figura 26 - Contagem de alunos negros no DF por ano

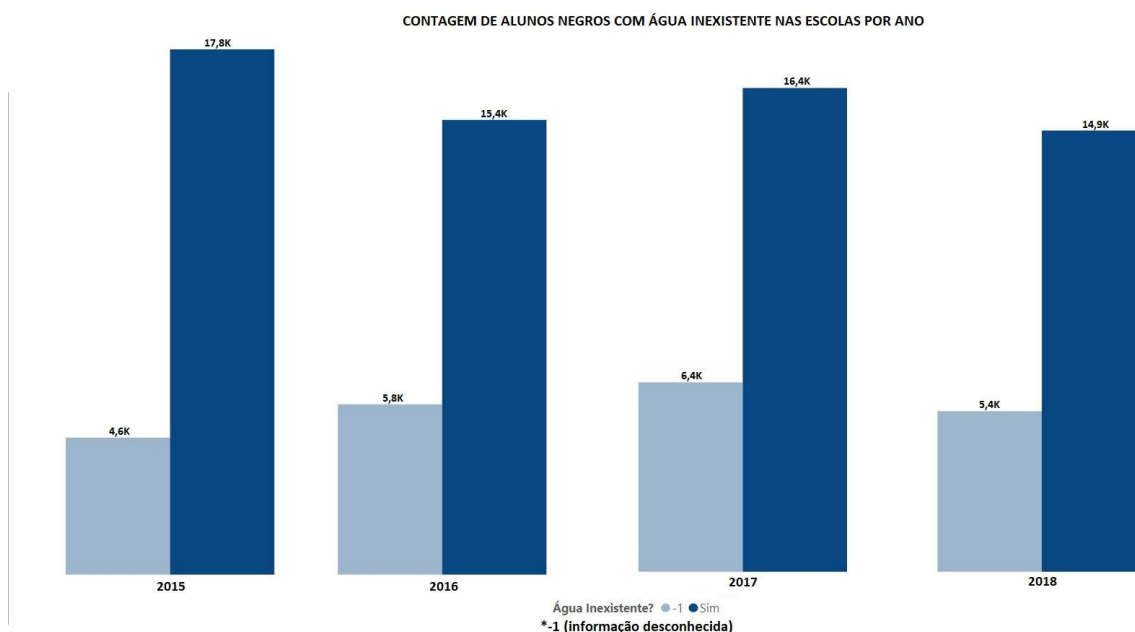


Fonte: Autores (2019).

Em análise ao gráfico, é possível observar que, em média, cerca de 25.675 estudantes negros estão anualmente matriculados em escolas de ensino básico. Segundo dados da Secretária de Estado da Educação do DF, cerca de 450 mil estudantes são atendidos pela Secretária de Educação do Distrito Federal, envolvendo Escolas de Educação Básica, Escolas Parque, Centros Interescolares de Línguas, Centros de Ensino Profissionalizante, além de um Centro de Ensino Médio Integrado (SECRETARIA DE ESTADO DA EDUCAÇÃO, 2018). Em dados levantados pela Codeplan (2014), a população negra do DF corresponde a 57%, porém esta realidade não é transmitida no gráfico, pois a avaliação é feita por auto declaração, sendo vários alunos não se declarando como negros ou nem sequer declaram uma cor/raça.

9. Qual a quantidade de alunos negros que estudam em escolas sem água, energia, esgoto e alimentação entre os anos da análise?

Figura 27 - Contagem de alunos negros com água inexistente nas escolas por ano



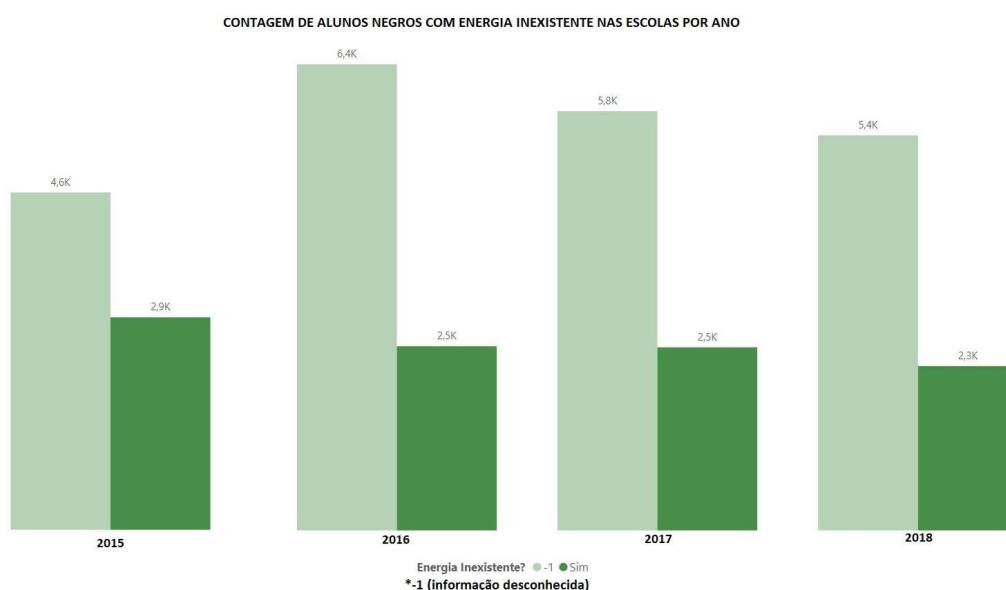
Fonte: Autores (2019).

Significado dos códigos da análise:

-1: Informação desconhecida.

A partir de uma análise do gráfico, é possível identificar que uma parcela de 17,8 mil alunos negros no ano de 2015 não tem acesso a água em suas escolas. Esta quantidade pode estar ligada a fatores geográficos e geopolíticos, pois segundo uma pesquisa realizada pelo jornal O Globo, casos como este, onde há falta de um dos princípios de saneamento básico, são extremados, afetando principalmente comunidades pobres e rurais (VASCONCELLOS, RIBEIRO e LINS, 2014). Também consta no gráfico a quantidade de dados inexistentes na base do INEP (representado por “-1”), onde em 2017 teve-se a maior quantidade de dados faltantes, 6,4 mil.

Figura 28 - Contagem de alunos negros com energia inexistente nas escolas por ano



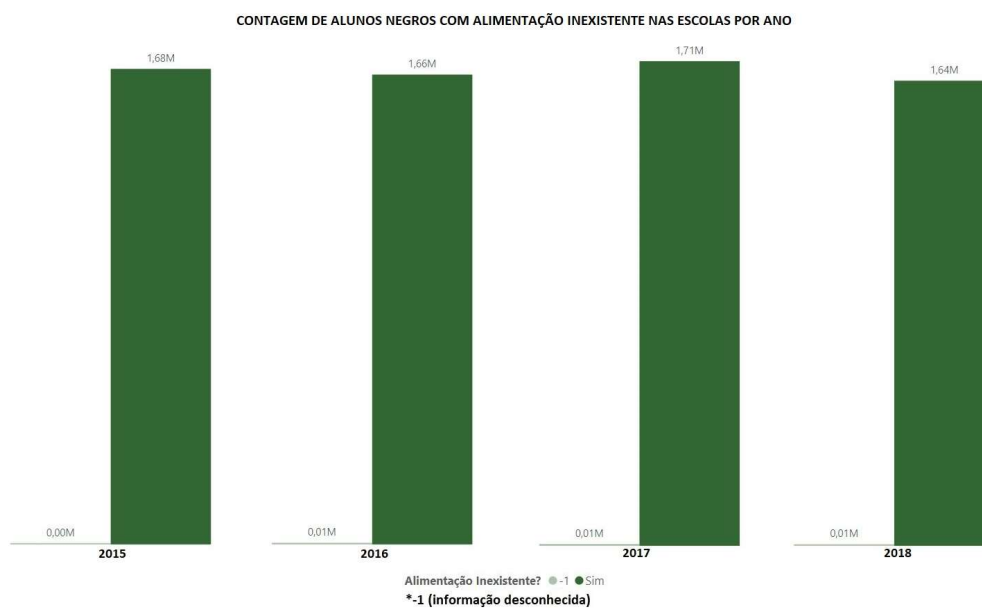
Fonte: Autores (2019).

Significado dos códigos da análise:

-1: Informação desconhecida.

Com base no gráfico, é possível detectar que a quantidade de 2,9 mil alunos negros no ano de 2015 que frequentam escolas com déficit de energia elétrica. Este número pode ser considerado baixo em relação à quantidade total de alunos negros, pois a partir de programas de desenvolvimento como o Programa Luz para Todos que asseguram a prioridade de desenvolvimento de medidas para o melhoramento do acesso à energia em escolas (PLANO DE DESENVOLVIMENTO DA EDUCAÇÃO, [200-?]). As informações inexistentes passam as informações existentes em todos os anos, chegando ao seu maior pico em 2017 com 6,4 mil de registros.

Figura 29 - Contagem de alunos negros com alimentação inexistente nas escolas por ano



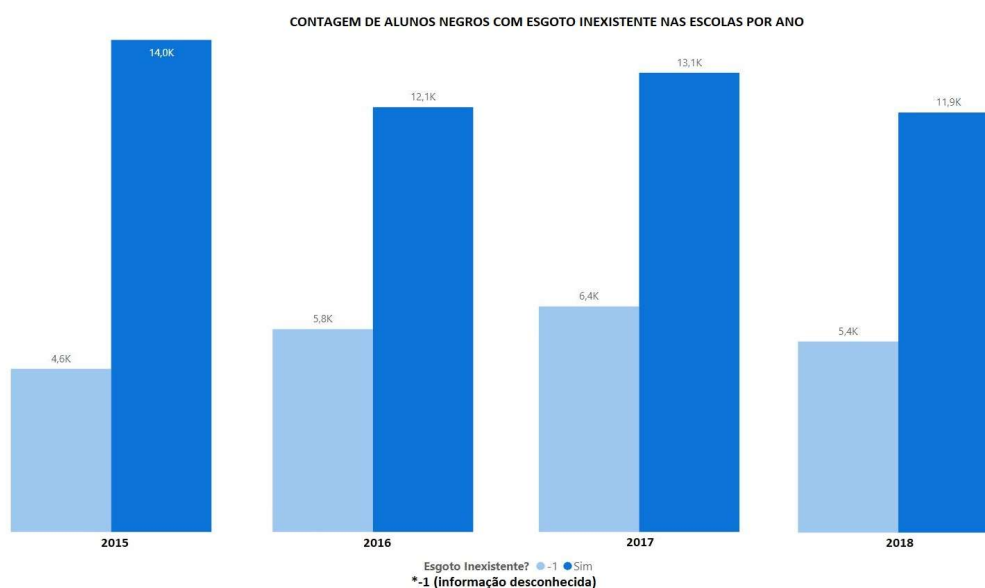
Fonte: Autores (2019).

Significado dos códigos da análise:

-1: Informação desconhecida.

Segundo o gráfico acima, tem-se a informação que em média, 1,6 milhões de alunos negros estudam em escolas sem alimentação, tendo o seu maior pico em 2017, onde teve-se 1,71 milhões de registros. As informações inexistentes não chegam a quase mil registros, se mostrando a menor inexistência entre os outros gráficos.

Figura 30 - Contagem de alunos negros com esgoto inexistente nas escolas por ano



Fonte: Autores (2019).

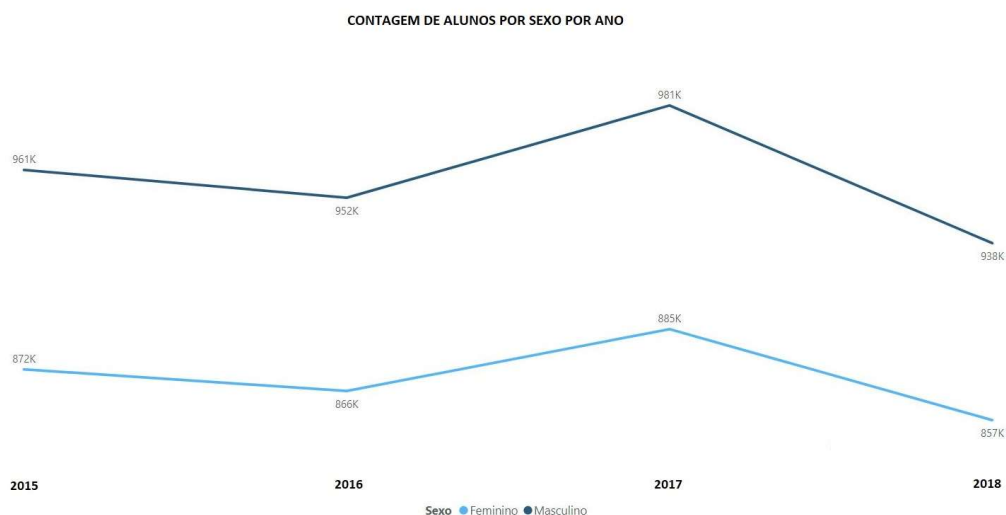
Significado dos códigos da análise:

-1: Informação desconhecida.

Segundo o gráfico acima, em média, 12,7 mil alunos estudam em escolas sem esgoto, com o seu maior pico em 2015 com quase 14 mil registros de alunos. Sobre as informações inexistentes, sua maior quantidade foi em 2017, onde teve-se 6,4 mil registros faltantes.

10. Qual a quantidade de alunos negros por sexo entre os anos da análise?

Figura 31 - Contagem de alunos por sexo por ano

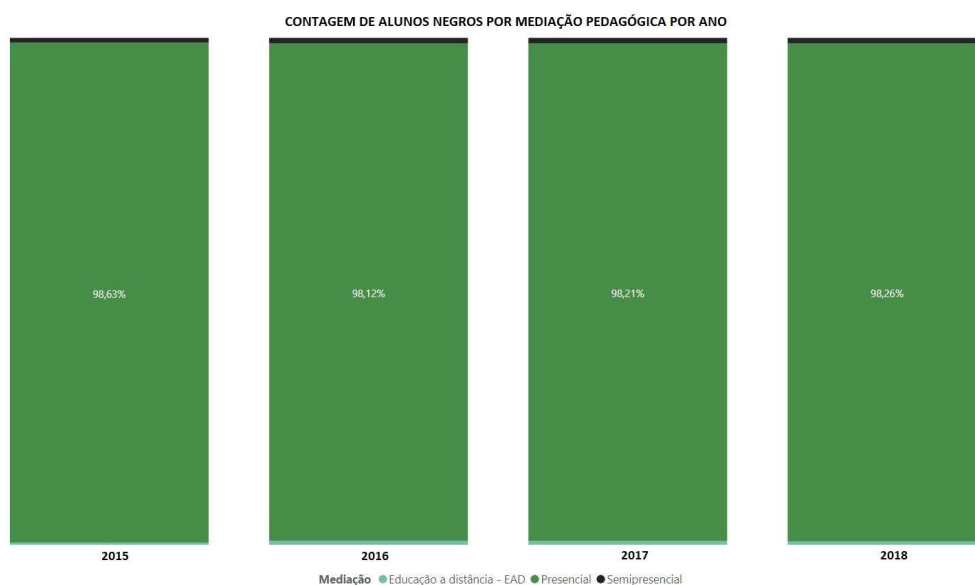


Fonte: Autores (2019).

Segundo o gráfico apresentado, percebe-se uma maior quantidade de alunos negros envolvidos na educação básica, segundo os dados da base do INEP. O maior pico foi em 2017, onde teve-se 981 mil alunos, contra 885 mil alunas no mesmo ano. Percebe-se também que nos dois sexos apresentados, essa quantidade foi variando, onde em 2015 teve-se uma quantidade maior que em 2016, que teve uma quantidade menor que em 2017, que teve uma quantidade maior em 2018.

11. Qual a quantidade de alunos negros nos módulos de ensino presencial, semipresencial e a distância entre os anos da análise?

Figura 32 - Contagem de alunos negros por mediação pedagógica por ano

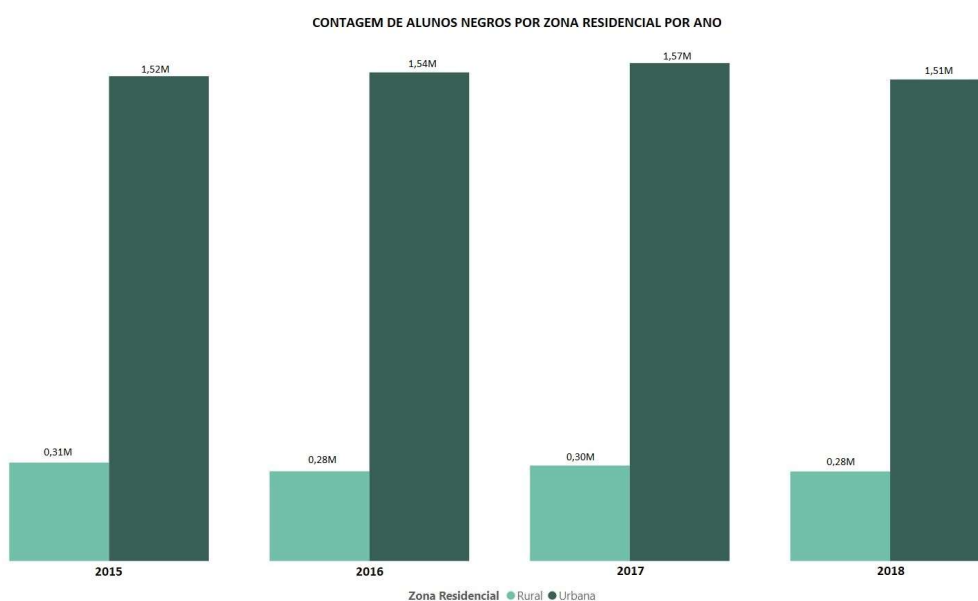


Fonte: Autores (2019).

Segundo o gráfico apresentado, percebe-se a maior inserção dos negros na educação presencial com quase 100% dos dados da base demonstrando isso.

12. Qual a quantidade de alunos negros que moram em zona urbana ou rural entre os anos da análise?

Figura 33 - Contagem de alunos negros por zona residencial por ano

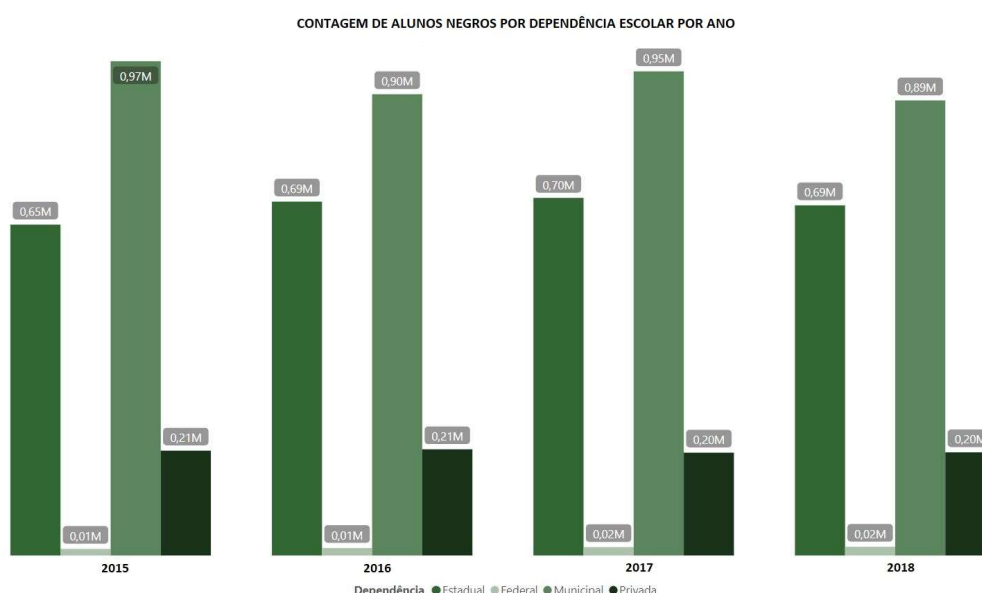


Fonte: Autores (2019).

Partindo de uma análise do gráfico, pode-se observar que grande parte da população de alunos negros se concentra em áreas urbanas, sendo mais exatos 83,96% desta amostra, e apenas 16,04% da população negra concentra-se em zonas rurais. Isso pode ser por conta de uma tendência nacional de concentração de população urbana. Segundo uma pesquisa do IBGE, cerca de 84,72% da população brasileira está reunida em centros urbanos e apenas 15,28% está localizada em zonas rurais (IBGE, 2015).

13. Qual a quantidade de alunos negros que estudam em escolas públicas e privadas?

Figura 34 - Contagem de alunos negros por dependência escolar por ano



Fonte: Autores (2019).

Segundo o gráfico há uma maior concentração de alunos negros em escolas com dependência municipal, chegando até 52,72% no ano de 2015. Logo após tem-se a modalidade estadual, com o seu pico atingindo 70 mil registros no ano de 2017. As escolas privadas não passaram de 21 mil registros em todos os anos da análise.

14. Qual a quantidade de alunos negros que estudam em escolas urbanas e rurais?

Figura 35 - Contagem de alunos negros por localização escolar por ano



Fonte: Autores (2019).

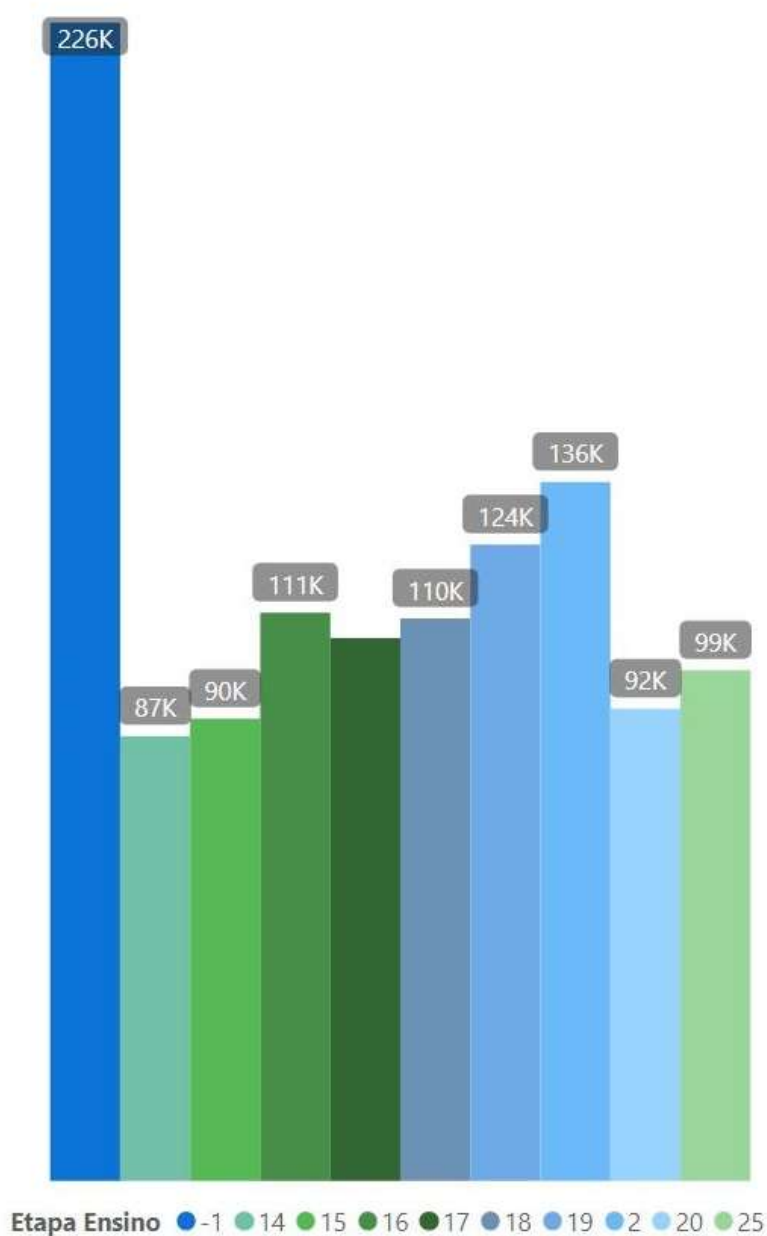
Segundo os dados apresentados no gráfico, percebe-se uma maior quantidade de alunos negros envolvidos nas escolas de localização urbana, tendo seu maior pico em 2017, onde tem-se 1,67 milhões de alunos nas escolas urbanas, comparado com o maior pico das escolas rurais em 2015, com 20 mil alunos.

15. Qual a quantidade de alunos negros em cada etapa de ensino definida no censo entre os anos da análise?

Por questões de performance, não foram alterados os códigos referentes a cada uma das etapas de ensino, mas em cada um dos gráficos será descrito o significado dos mesmos segundo o dicionário de dados na base do INEP.

Figura 36 - Contagem de alunos negros por etapa de ensino em 2015 (Top 10)

CONTAGEM DE ALUNOS NEGROS POR ETAPA DE ENSINO EM 2015 (TOP 10)



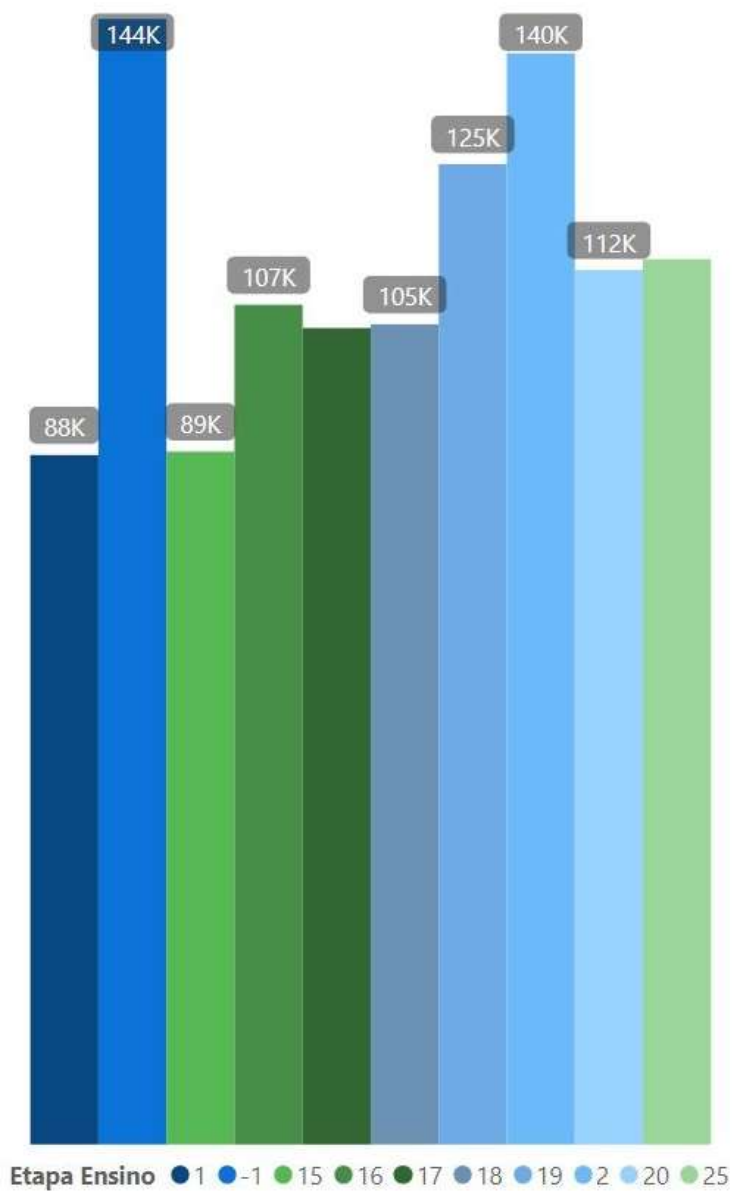
Fonte: Autores (2019).

Significado dos códigos da análise:

- 1: Informação desconhecida;
- 14: Ensino Fundamental de 9 anos - 1º Ano;
- 15: Ensino Fundamental de 9 anos - 2º Ano;
- 16: Ensino Fundamental de 9 anos - 3º Ano;

Figura 37 - Contagem de alunos negros por etapa de ensino em 2016 (Top 10)

CONTAGEM DE ALUNOS NEGROS POR ETAPA DE ENSINO EM 2016 (TOP 10)



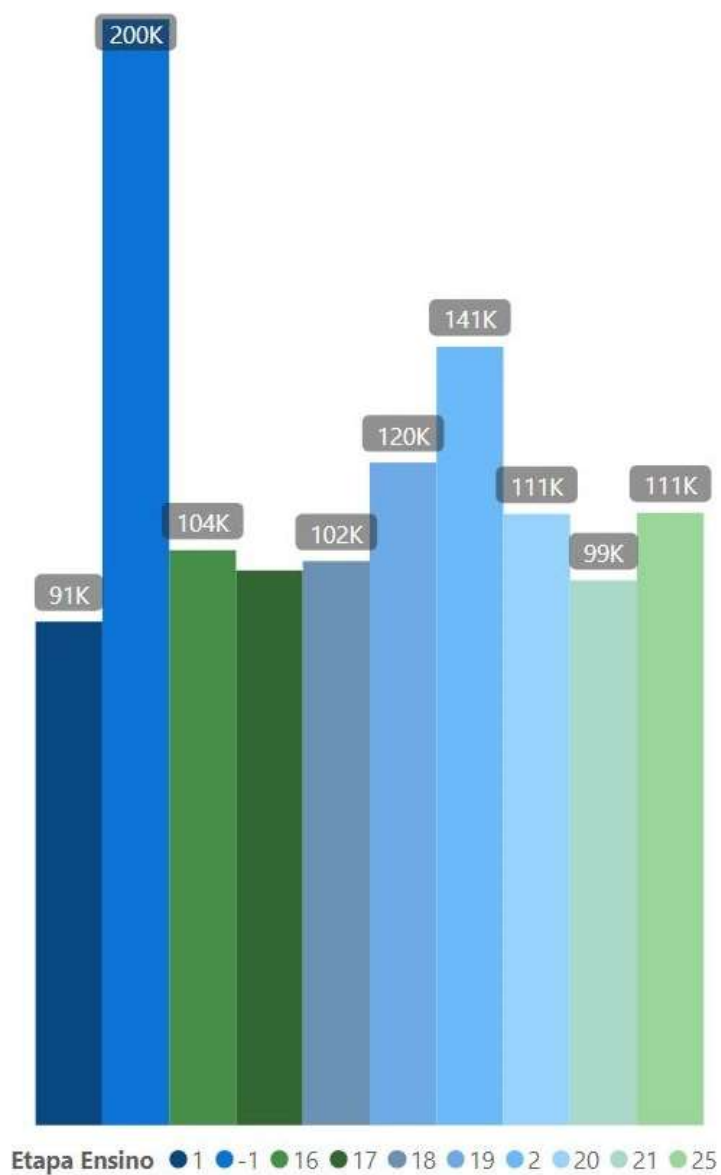
Fonte: Autores (2019).

Significado dos códigos da análise:

- 1: Educação Infantil – Creche;
- 1: Informação desconhecida;
- 15: Ensino Fundamental de 9 anos - 2º Ano;
- 16: Ensino Fundamental de 9 anos - 3º Ano;
- 17: Ensino Fundamental de 9 anos - 4º Ano;

Figura 38 - Contagem de alunos negros por etapas de ensino em 2017 (Top 10)

CONTAGEM DE ALUNOS NEGROS POR ETAPA DE ENSINO EM 2017 (TOP 10)



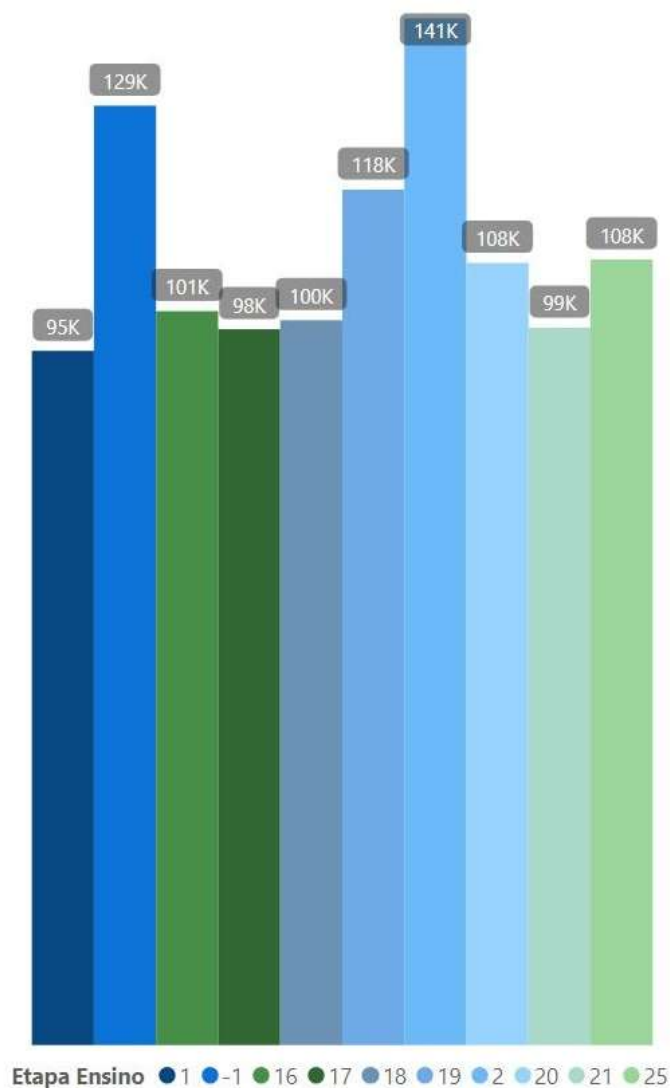
Fonte: Autores (2019).

Significado dos códigos da análise:

- 1: Educação Infantil – Creche;
- 1: Informação desconhecida;
- 16: Ensino Fundamental de 9 anos - 3º Ano;
- 17: Ensino Fundamental de 9 anos - 4º Ano;

Figura 39 - Contagem de alunos negros por etapa de ensino em 2018 (Top 10)

CONTAGEM DE ALUNOS NEGROS POR ETAPA DE ENSINO EM 2018 (TOP 10)



Fonte: Autores (2019).

Significado dos códigos da análise:

- 1: Educação Infantil – Creche;
- 1: Informação desconhecida;
- 16: Ensino Fundamental de 9 anos - 3º Ano;
- 17: Ensino Fundamental de 9 anos - 4º Ano;
- 18: Ensino Fundamental de 9 anos - 5º Ano;
- 19: Ensino Fundamental de 9 anos - 6º Ano;