

CENTRO UNIVERSITÁRIO SENAC

CURSO DE TECNOLOGIA EM BANCO DE DADOS

PROJETO INTEGRADOR V: APOIO DECISÓRIO AOS NEGÓCIOS

AVALIAÇÃO DO ENSINO BÁSICO NOS MUNICÍPIOS BRASILEIROS

Integrantes do grupo:

Cesar Inui

Douglas Bazo de Castro

Felipe Alves Nilo

Michel Ferreira da Silva

Polyanna Ynaí Miranda Simões

EAD – ENSINO À DISTÂNCIA

2025

Integrantes do grupo:

Cesar Inui

Douglas Bazo de Castro

Felipe Alves Nilo

Michel Ferreira da Silva

Polyanna Ynaí Miranda Simões

## AVALIAÇÃO DO ENSINO BÁSICO NOS MUNICÍPIOS BRASILEIROS

Prof. Dr. Gustavo Moreira Calixto

## TRABALHO PARA APROVAÇÃO EM DISCIPLINA

EAD – ENSINO À DISTÂNCIA

2025

## RESUMO

O investimento em educação é um motor fundamental do desenvolvimento econômico. Diante dos desafios no setor educacional brasileiro, é crucial avaliar como os recursos públicos são aplicados. Este estudo propõe um sistema de informação que integre dados sobre infraestrutura escolar, investimentos e desempenho educacional nos municípios brasileiros. Por meio de processos de ETL (Extração, Transformação e Carga), os dados serão consolidados em um banco único e informações serão produzidas e visualizadas em um painel interativo. A ferramenta pretende apoiar análises comparativas e embasar políticas públicas voltadas à melhoria da educação básica no Brasil.

Palavras-chave: avaliação educacional, bancos de dados, educação básica, gastos públicos, infraestrutura de escolas.

## ABSTRACT

Investment in education is a key driver of economic development. Given the challenges facing Brazil's education sector, it is crucial to assess how public funds are allocated. This study proposes an information system that integrates data on school infrastructure, investments, and educational outcomes across Brazilian municipalities. Using ETL (Extract, Transform, Load) processes, the data will be consolidated into a unified database and information will be produced and presented through an interactive dashboard. The tool aims to support comparative analyses and inform public policies focused on improving basic education in Brazil.

Keywords: educational assessment, databases, basic education, public spending, school infrastructure.

## LISTA DE ABREVIACÕES

BI – *Business Intelligence*

ETL – *Extract, Transform, Load* / Processo de Extração, Transformação e Carga

FUNDEB – Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação

FUNDEF – Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério

IDEB – Índice de Desenvolvimento da Educação Básica

INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

LDB – Lei das Diretrizes e Bases da Educação Nacional (Lei nº 9.394/1996)

MEC – Ministério da Educação

SAEB – Sistema Nacional de Avaliação da Educação Básica

SQL – *Structured Query Language* / Linguagem de Consulta Estruturada

TI – Tecnologia da Informação

## LISTA DE ILUSTRAÇÕES

Figura 1: Tabela sobre a origem dos recursos do FUNDEF .....	09
Figura 2: Proficiências médias em Matemática e em Língua Portuguesa no Rio de Janeiro .....	14
Figura 3: Proficiência média em Matemática na rede privada do Rio de Janeiro .....	14
Figura 4: Proficiência média em Matemática na rede pública do Rio de Janeiro .....	15
Figura 5: Investimento por aluno no Rio de Janeiro em 2024 .....	15
Figura 6: Investimento por aluno em Angra dos Reis em 2024 .....	16
Figura 7: Painel de matrículas no Rio de Janeiro .....	17
Figura 8: Painel de docentes no Rio de Janeiro .....	17
Figura 9: Painel de escolas no Rio de Janeiro .....	18

## SUMÁRIO

1. INTRODUÇÃO: VISÃO GERAL DA SOLUÇÃO PROPOSTA .....	08
1.1 Contexto e Motivação .....	10
1.2 Problema .....	10
1.3 Objetivos .....	11
2. DEFINIÇÃO DAS ATIVIDADES DE APOIO DECISÓRIO AOS NEGÓCIOS .....	12
2.1 Coleta e Armazenamento de Informações Relevantes .....	12
2.2 Definição de Métricas e Indicadores de Desempenho .....	13
2.3 Análise Exploratória de Dados .....	14
2.4 Desenvolvimento de <i>Dashboards</i> e Relatórios gerenciais .....	16
3. DEFINIÇÃO DO PROCESSO DE EXTRAÇÃO, TRANSFORMAÇÃO, CARGA .	19
3.1 Extração .....	19
3.2 Transformação .....	20
3.3 Carga .....	21
4. CONSIDERAÇÕES FINAIS .....	22
REFERÊNCIAS .....	23

## **1. INTRODUÇÃO: VISÃO GERAL DA SOLUÇÃO PROPOSTA**

É consenso que um dos pilares do desenvolvimento econômico de um país passa pelo investimento em educação. Assim, é importante avaliarmos se os gastos governamentais em educação no Brasil são adequados aos enormes desafios dessa área.

Embora existam diversos estudos sobre a eficiência na gestão dos recursos educacionais, poucos integram dados financeiros, de infraestrutura e de desempenho escolar de forma sistemática e acessível aos formuladores de políticas.

Neste contexto, sistemas de informação (bancos de dados) que possam auxiliar a tomada de decisão dos gestores públicos, considerando gastos, infraestrutura das escolas e resultados de avaliações educacionais periódicas, são fundamentais.

A educação básica, conforme o artigo 4º da LDB, é um direito do cidadão e uma obrigação do Estado, que deve garantir uma oferta de qualidade (CURY, 2008). Um dos deveres do Estado em relação à educação é assegurar a igualdade de condições para o acesso e a permanência na educação básica (CARREIRA, 2013). A rápida democratização das escolas públicas brasileiras no final do século XX tornou o sistema educacional excessivamente flexível, prejudicando sua qualidade (BASTOS, 2017). Segundo Peruchin (2018), a importância da educação em conjunto com outras estratégias sociais é primordial para o pleno desenvolvimento da sociedade.

No Brasil, desde a década de 1990, várias iniciativas foram tomadas objetivando a avaliação do desempenho dos estudantes, como a criação do Sistema Nacional de Avaliação da Educação Básica (SAEB), do Exame Nacional do Ensino Médio (ENEM) e do Exame Nacional de Desempenho dos Estudantes (ENADE), este último voltado aos estudantes do ensino superior.

Esteves (2020) aponta que a sanção da Lei nº 9.448/1997 reestruturou o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), tornando-o uma autarquia vinculada ao Ministério da Educação (MEC). O INEP é a



entidade de estado responsável por conduzir pesquisas, estudos e avaliações sobre o sistema educacional, que possibilitam traçar novas estratégias e prospectos para desenvolver o sistema educacional brasileiro (TOMASELLI, 2022).

O Censo Escolar é o principal instrumento de coleta de informações da educação básica no Brasil. É realizado anualmente pelo INEP e a declaração é obrigatória para todas as escolas públicas e privadas. Esta pesquisa abrange vários aspectos das escolas, entre os quais, infraestrutura física, recursos humanos, fluxo dos alunos, etc..

Os recursos destinados à educação tem sido ampliados desde os anos 90. Em 1996, a LDB estabeleceu as porcentagens mínimas de investimentos que (i) a União, (ii) os Estados e o Distrito Federal e (iii) os Municípios deveriam direcionar para a manutenção e o desenvolvimento do ensino público.

Assim, em 1996, com a Emenda Constitucional nº 14, foi criado o Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério – FUNDEF, para um período de dez anos. Posteriormente, em 2006, com a Emenda Constitucional nº 53, foi criado o FUNDEB, substituindo o FUNDEF. A tabela abaixo resume a origem dos recursos constitutivos do FUNDEF:

Unidade Federativa	Legislação
União	30% dos 18% da arrecadação dos impostos da União destinados à Educação; Complementação para os Estados que não atingirem o mínimo igual à média nacional;
Estados Distrito Federal	15% do ICMS devido ao Distrito Federal e aos Estados (art. 155, inciso II da Constituição Federal);
	15% do FPE (art. 159, inciso I, alínea a da Constituição Federal e Lei 5.172 de 25 de outubro de 1966);
	15% do IPI-X (art.159, inciso II da Constituição Federal e lei complementar nº 61, 26/12/1989)
	15% da renúncia fiscal proveniente da Lei Kandir – 87/1996
Municípios	15% do ICMS (art. 158, inciso IV da Constituição Federal);
	15% do FPM (art. 159, inciso I, alínea b da Constituição Federal e lei nº 5.172 de 25/12/1966).
	15% da renúncia fiscal proveniente da Lei Kandir – 87/1996

Figura 1. Tabela sobre a origem dos recursos do FUNDEF.

Basicamente, o FUNDEB incorporou percentuais de arrecadação de outros três impostos: Imposto sobre a Propriedade de Veículos Automotores – IPVA, Imposto sobre Transmissão Causa Mortis e Doação – ITCMD e o Imposto sobre a Propriedade Territorial Rural – ITR.

Os Sistemas de Gerenciamento de Banco de Dados são essenciais para as operações de TI das instituições públicas e privadas contemporâneas, à medida que lidam com o aumento da quantidade de dados a serem administrados (CARR, 2023). Segundo Silva (2016), a Linguagem de Consulta Estruturada SQL (*Structured Query Language*) é a linguagem de banco de dados mais amplamente utilizada.

## **1.1 Contexto e Motivação**

Este projeto “Avaliação do Ensino Básico nos Municípios Brasileiros” está sendo desenvolvido em um contexto específico, conforme descrito anteriormente, voltado para a pesquisa e para o auxílio à tomada de decisão de gestores públicos, agregando várias fontes de dados distintas e dispersas (avaliações educacionais, infraestrutura das escolas e recursos financeiros disponibilizados).

O projeto surge como uma iniciativa dos estudantes do SENAC, que são da área de tecnologia, com o objetivo de promover a conscientização da importância do gasto eficiente de recursos públicos limitados e de fornecer ferramentas relevantes para os pesquisadores e gestores públicos que lidam com a educação básica no Brasil.

## **1.2 Problema**

Os pesquisadores e os gestores públicos enfrentam diversas dificuldades em suas atividades, sejam financeiras, de gestão de pessoas ou de materiais. A dispersão de fontes de dados de naturezas distintas sobre a educação é mais um desses problemas/restrições enfrentados por esses atores, dada a complexidade dos inúmeros aspectos envolvendo o tema. Isso gera a necessidade de desenvolver habilidades de implementação de bancos de dados que permitam a produção de painéis (*dashboard*) e relatórios contendo informações sobre a educação no Brasil.

### 1.3 Objetivos

Inicialmente, por meio de um processo de Extração, Transformação e Carga (ETL), pretendemos reunir dados quantitativos dispersos e de naturezas distintas sobre educação básica nos municípios brasileiros (avaliações educacionais, infraestrutura das escolas e recursos financeiros disponibilizados). Com isso, será possível produzir painéis (*dashboard*) e relatórios contendo informações sobre a educação básica no Brasil, assim como, eventualmente, construir uma ferramenta para avaliação da eficiência dos gastos públicos com ensino básico nos municípios.

**Objetivo Geral:** Desenvolver uma ferramenta de análise que correlacione gastos públicos, infraestrutura e desempenho educacional para auxiliar na tomada de decisão.

**Objetivos Específicos:**

- Coletar e armazenar dados do FUNDEB, SAEB e Censo Escolar.
- Criar métricas de eficiência (ex.: custo por aluno vs. IDEB).
- Gerar *dashboards* interativos para visualização de resultados.

Ao integrar dados educacionais dispersos, este projeto fortalece a transparência na gestão pública e apoia o desenvolvimento de políticas educacionais fundamentadas em evidências.

## 2. DEFINIÇÃO DAS ATIVIDADES DE APOIO DECISÓRIO AOS NEGÓCIOS

A educação básica é um pilar fundamental para o desenvolvimento socioeconômico de um país. No Brasil, apesar dos investimentos públicos em educação, persistem desafios como desigualdades regionais, infraestrutura inadequada e variações no desempenho escolar. A dispersão de dados sobre financiamento (FUNDEB), avaliações (SAEB/IDEB) e infraestrutura escolar dificulta a análise integrada por gestores públicos.

Este projeto propõe um sistema de apoio à decisão que consolida e analisa dados educacionais, permitindo avaliar a eficiência dos gastos públicos por município. A motivação surge da necessidade de:

- integrar fontes de dados dispersas (financeiros, desempenho, infraestrutura);
- oferecer *insights* para otimizar recursos e políticas educacionais.

A seguir, são apresentadas as atividades de apoio decisório aos negócios.

### 2.1 Coleta e Armazenamento de Informações Relevantes

Este projeto passa pela construção de bancos de dados nacionais relacionados ao ensino básico nos municípios, abrangendo três aspectos:

- I. Avaliação de desempenho dos estudantes: serão usados dados oriundos do SAEB, que tem o Índice de Desenvolvimento da Educação Básica – IDEB como principal indicador, por reunir dois conceitos importantes: fluxo escolar (aprovação e reprovação) e médias de desempenho nas avaliações do INEP. Para maiores detalhes, acessar:

<https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/saeb/resultados>

- II. Infraestrutura das escolas: serão usados dados do Censo Escolar da Educação Básica, realizado pelo INEP, que coleta informações das instituições de ensino. Para maiores detalhes, acessar:

<https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-escolar>

- III. Financiamento da educação nos municípios: serão usados dados sobre os recursos transferidos aos municípios pelo Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação – FUNDEB, criado em 2006 e mantido conjuntamente pela União, pelos Estados e pelos Municípios. Para maiores detalhes, acessar:

<https://www.gov.br/fnde/pt-br/acesso-a-informacao/acoes-e-programas/financiamento/fundeb>

#### **Fontes de Dados:**

- **FUNDEB**: Dados de repasses financeiros por município.
- **SAEB/IDEB**: Resultados de avaliações de desempenho.
- **Censo Escolar**: Infraestrutura, número de alunos e professores.

**Armazenamento**: Banco de dados relacional (MySQL) com tabelas normalizadas (3FN) para evitar redundâncias.

## **2.2 Definição de Métricas e Indicadores de Desempenho**

A extração de informações do banco de dados por meio de métricas e indicadores de desempenho são fundamentais para uma descrição inicial da realidade da educação básica nos municípios brasileiros. Por exemplo, poderíamos usar as seguintes estatísticas:

- **Eficiência Financeira**: (Recursos Recebidos / Número de Alunos) vs. IDEB.

- **Infraestrutura Adequada:** % de escolas com biblioteca, laboratórios, etc.
- **Correlação Gastos-Desempenho:** Identificar municípios com melhor custo-benefício.

## 2.3 Análise Exploratória de Dados

Além da definição de métricas e indicadores de desempenho sobre a educação básica no Brasil, o resultado da extração de informações do banco de dados pode ser apresentado na forma de tabelas e gráficos.

Por exemplo, a partir dos *websites* indicados na subseção 2.1, é possível construir gráficos com as proficiências médias em Língua Portuguesa e Matemática dos alunos do 5º ano do Ensino Fundamental do Rio de Janeiro no período 2011-2023 (Painel SAEB):

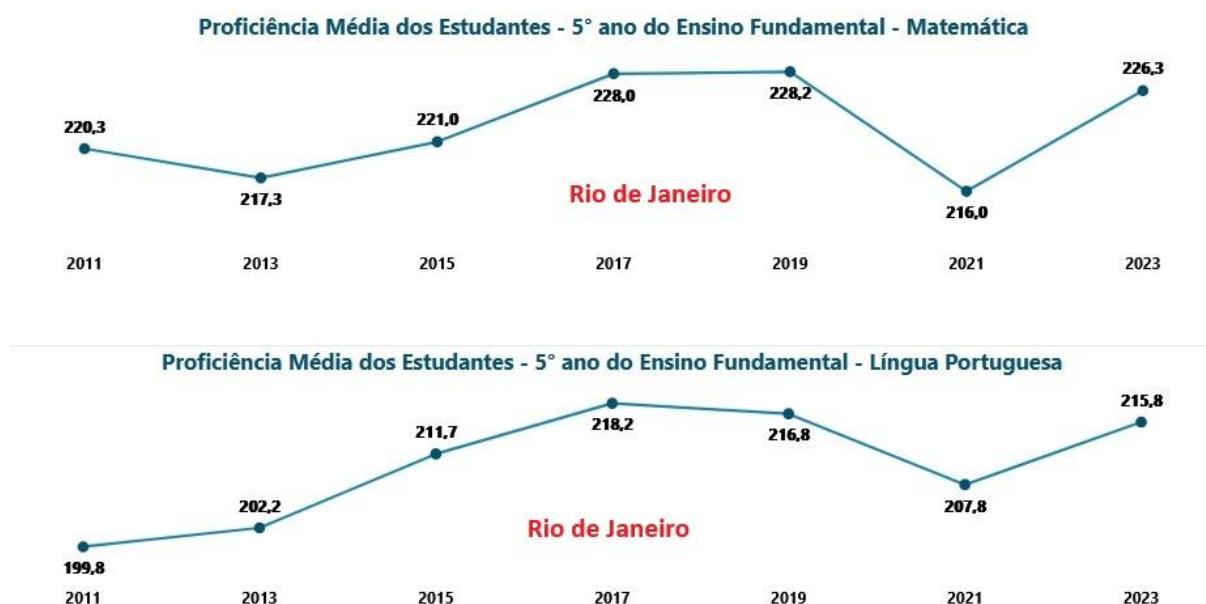


Figura 2. Proficiências médias em Matemática e em Língua Portuguesa no Rio de Janeiro.

Também podemos ter essa informação de desempenho de acordo com a rede escolar: pública e privada, por exemplo, em Matemática:



Figura 3. Proficiência média em Matemática na rede privada do Rio de Janeiro.



Figura 4. Proficiência média em matemática na rede pública do Rio de Janeiro.

Em todos esses gráficos, é possível notar a queda de rendimento no ano de 2021, durante a pandemia de Covid-19.

Informações relativas a investimentos por aluno também podem ser obtidas pra cada bimestre do ano. Por exemplo, em 2024, no Rio de Janeiro e Angra dos Reis, as tabelas abaixo mostram os investimentos por aluno com recursos do FUNDEB (ensino médio, educação superior e educação profissional não são financiados por este fundo):

#### ► Rio de Janeiro – 2024



## Indicadores de Investimento por Aluno

[<< Visualizar Anos Anteriores](#) | [Imprimir](#) | [Visualizar Próximos Anos >>](#)

Código	Indicador	2024					
		1º	2º	3º	4º	5º	6º
4.1	Investimento educacional por aluno da educação infantil	<u>R\$768,57</u>	<u>R\$1.570,46</u>	<u>R\$2.417,03</u>	<u>R\$3.567,93</u>	<u>R\$7.232,08</u>	<u>R\$6.010,09</u>
4.2	Investimento educacional por aluno do ensino fundamental	<u>R\$781,11</u>	<u>R\$4.775,66</u>	<u>R\$8.060,60</u>	<u>R\$10.983,84</u>	<u>R\$12.694,17</u>	<u>R\$18.061,86</u>
4.3	Investimento educacional por aluno do ensino médio	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>
4.4	Investimento educacional por aluno da educação superior	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>
4.5	Investimento educacional por aluno da educação de jovens e adultos	<u>R\$623,34</u>	<u>R\$978,72</u>	<u>R\$1.662,02</u>	<u>R\$2.596,26</u>	<u>R\$7.488,14</u>	<u>R\$5.527,54</u>
4.6	Investimento educacional por aluno da educação especial	<u>R\$687,70</u>	<u>R\$4.032,55</u>	<u>R\$6.634,64</u>	<u>R\$10.176,69</u>	<u>R\$10.550,40</u>	<u>R\$16.071,15</u>
4.7	Investimento educacional por aluno da educação profissional	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>
4.8	Investimento educacional por aluno da educação básica	<u>R\$757,68</u>	<u>R\$3.955,28</u>	<u>R\$6.669,47</u>	<u>R\$9.203,75</u>	<u>R\$11.590,76</u>	<u>R\$15.237,70</u>

Figura 5. Investimento por aluno no Rio de Janeiro em 2024.

## ► Angra dos Reis – 2024

## Indicadores de Investimento por Aluno

[<< Visualizar Anos Anteriores](#) | [Imprimir](#) | [Visualizar Próximos Anos >>](#)

Código	Indicador	2024					
		1º	2º	3º	4º	5º	6º
4.1	Investimento educacional por aluno da educação infantil	<u>R\$958,36</u>	<u>R\$3.082,43</u>	<u>R\$4.347,66</u>	<u>R\$6.872,73</u>	<u>R\$9.932,26</u>	<u>R\$13.821,80</u>
4.2	Investimento educacional por aluno do ensino fundamental	<u>R\$1.562,58</u>	<u>R\$4.469,86</u>	<u>R\$6.603,36</u>	<u>R\$8.930,29</u>	<u>R\$11.431,54</u>	<u>R\$15.740,05</u>
4.3	Investimento educacional por aluno do ensino médio	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>
4.4	Investimento educacional por aluno da educação superior	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>
4.5	Investimento educacional por aluno da educação de jovens e adultos	<u>R\$366,91</u>	<u>R\$1.337,61</u>	<u>R\$6.194,66</u>	<u>R\$9.960,00</u>	<u>R\$12.901,80</u>	<u>R\$14.949,66</u>
4.6	Investimento educacional por aluno da educação especial	<u>R\$9.322,50</u>	<u>R\$4.461,88</u>	<u>R\$7.866,13</u>	<u>R\$10.377,45</u>	<u>R\$13.191,35</u>	<u>R\$9.274,23</u>
4.7	Investimento educacional por aluno da educação profissional	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>	<u>R\$0,00</u>
4.8	Investimento educacional por aluno da educação básica	<u>R\$1.880,73</u>	<u>R\$5.809,71</u>	<u>R\$8.813,65</u>	<u>R\$12.237,95</u>	<u>R\$15.975,18</u>	<u>R\$23.031,53</u>

Figura 6. Investimento por aluno em Angra dos Reis em 2024.



Além dessas informações, podem ser utilizadas outras técnicas, como:

- Agregação de dados por região/município.
- Identificação de *outliers* (ex.: altos gastos com baixo IDEB).
- Análise temporal (evolução dos indicadores).

## 2.4 Desenvolvimento de *Dashboards* e Relatórios Gerenciais

*Dashboards* (painéis) são muito úteis por fornecerem uma visão mais ampla sobre o objeto de estudo, apresentando métricas, indicadores, tabelas e gráficos simultaneamente, facilitando assim a visualização de tendências/padrões.

**Ferramentas:** Power BI ou Metabase.

**Visualizações:**

- Mapa georreferenciado com indicadores por município.
- Gráficos de dispersão (gastos x desempenho).
- Relatórios PDF automatizados para gestores.

Por exemplo, a partir dos *websites* indicados na subseção 2.1, é possível construir alguns painéis relacionados a matrículas, docentes e escolas do Rio de Janeiro, apresentados abaixo:

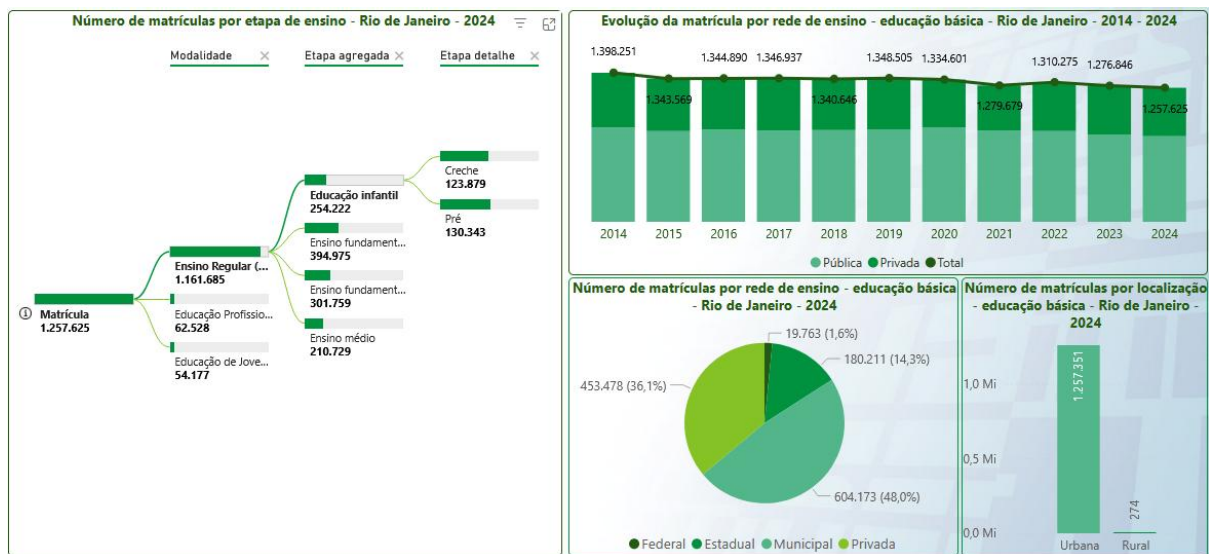


Figura 7. Painel de matrículas no Rio de Janeiro.

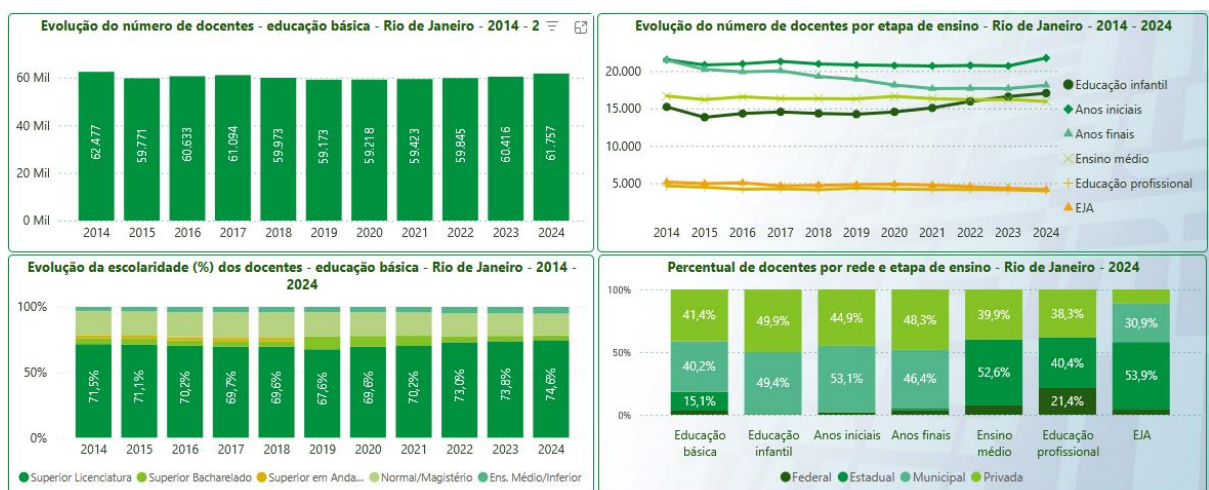


Figura 8. Painel de docentes no Rio de Janeiro.

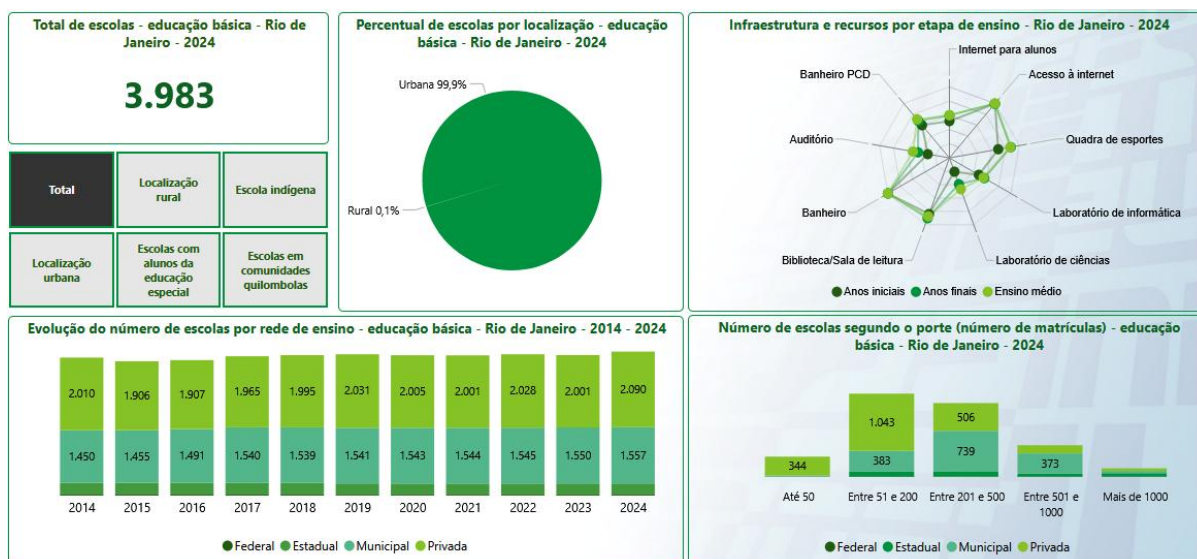


Figura 9. Painel de escolas no Rio de Janeiro.

### **3. DEFINIÇÃO DO PROCESSO DE EXTRAÇÃO, TRANSFORMAÇÃO, CARGA**

O principal objetivo do ETL é a extração, o refinamento, correção e padronização de dados derivados de diferentes fontes, para que sejam compatibilizados com o modelo de dados do *Data Warehouse*.

A periodicidade da execução dos processos ETL deverá ser alinhada, preferencialmente, com a frequência com que os dados da fonte são atualizadas. Geralmente, o acionamento do processo é realizado nos períodos noturnos, madrugada ou nos finais de semana para não interferir eventuais análises realizadas no banco durante o horário comercial.

Nesta parte serão descritas as três principais fases do processo ETL para a execução do projeto.

#### **3.1 Extração**

A extração da fonte de dados poderá ser realizada de duas maneiras. A primeira é acessar os dados do portal URL manualmente, realizar o download periodicamente e disponibilizar o arquivo numa pasta compartilhada para que a aplicação do ETL possa ler os dados, tratar e carregar no banco de dados.

Eventualmente, outra opção é automatizar o processo do ETL para que ele possa buscar os dados na URL e baixar os dados em períodos predeterminados. A desvantagem é que a URL de origem disponha de recursos que impeça a obtenção de dados por meio de “aplicações robô” por meio do recursos de “CAPTCHA”, por exemplo.

Dependendo do software do ETL, ele não dispõe de recursos para baixar dados em URL. No caso do projeto em particular, os arquivos de entrada seriam no formato de arquivos externos para serem carregados e não de “banco para banco”, conforme ocorre em alguns casos.

Essa é a grande vantagem da utilização do software ETL: possui a versatilidade para ler uma ampla diversidade de formatos e tipos de dados (*input*) e também converter e armazenar em vários formatos de bancos ou criar arquivos externos (*output*).

### 3.2 Transformação

A transformação é a etapa mais crítica do ETL. É a fase onde os dados obtidos são formatados adequadamente para serem armazenados no banco de dados para posterior análise em BI.

As operações de transformação descritas abaixo são baseadas no software “*ETL FME*”, porém conceitualmente são aplicáveis em qualquer outro software ETL:

- *DateTimeFormatter* – padronização do formato de data para “aaaa-mm-dd”;
- *Attribute Remover* – eliminação de atributos ou campos desnecessários à análise. Essa triagem dos dados é importante para não sobrecarregar o *Data Warehouse* e afetar a performance do banco;
- *AttributeCreator* – criação de novos atributos, resultantes da concatenação de outros campos;
- *StringReplacer* – tratamento de caracteres especiais;
- *StringCaseChanger* – padronização de caracteres maiúsculas e minúsculas;
- *NullAttributeMapper* – tratamento e padronização para campos que conterem valores nulos;
- *AttributeRenamer* – conversão dos nomes de atributos de origem (*input*) para atributos de destino (*output*) do banco;
- *AttributeTester*, *AttributeFilter*, *AttributeRangeMapper* – verificação de valores para ver se estão em conformidade com o banco. Usado geralmente para tabelas-tipo (pai) quando se exige algum tipo de classe de atributos para tabela-filho. Também pode ser aplicado para descartar dados desnecessários ao escopo do projeto;
- *Counter* – criador de valores para Primary-Key de 1 a N quantidades, caso a tabela de destino do banco não tenha o auto-incremento;
- *FeatureMerger* – executa o *join* entre duas tabelas, utilizadas em casos em que se integra dados extraídos de diferentes fontes. Também é essencial para padronização de dados em casos onde se utiliza tabelas-pai e filho.
- *StatisticsCalculator* – usado para calcular, converter e padronizar dados referentes a unidades de medida como distâncias, quantidades, pesos, etc.;
- *VertexCreator* – converter dados alfanuméricos de localização como coordenadas geográficas em dados gráficos para mapas digitais.

Essas operações podem ser contextualizadas por:

- Limpeza: tratamento de valores nulos ou duplicados.
- Padronização: unificação de nomes de municípios e códigos IBGE.
- Cálculo de métricas: "Gasto por aluno = Total Recebido / Número de Matrículas", por exemplo.

### **3.3 Carga**

O estágio final do processo ETL é o armazenamento dos dados em *Data Warehouse* para serem disponibilizados para consulta e visualização aos analistas de BI.

A carga dos dados podem ser feitas em modo incremental, com a preservação de dados já existentes e adição de novos dados, ou substituindo e atualizando completamente os dados já existentes pelos novos.

#### 4. CONSIDERAÇÕES FINAIS

Este projeto demonstra como a integração de dados educacionais pode transformar informações dispersas em *insights* acionáveis. A solução proposta permite:

- **Transparência:** visualização clara da alocação de recursos;
- **Eficiência:** identificação de boas práticas e áreas críticas;
- **Tomada de Decisão:** baseada em evidências para gestores públicos.

Futuros desdobramentos podem incluir modelos preditivos (ex.: impacto de aumentos de verba no IDEB) e expansão para outras etapas educacionais.

## REFERÊNCIAS

BASTOS, Manoel. Os desafios da educação brasileira. Revista Científica Multidisciplinar Núcleo do Conhecimento, v. 39, p. 39-46, 2017. DOI: 10.32749/nucleodoconhecimento.com.br/educacao/educacao-brasileira.

CARR, Caroline et al. Oracle Database: Uma análise aprofundada do SGBD que domina o Mundo Corporativo. Research, Society and Development, v. 12, 2023, artigo e0812541323. DOI: 10.33448/rsd-v12i5.41323.

CARREIRA, Denise (Coord.). \*Informe Brasil: gênero e educação\*. São Paulo: Ação Educativa, 2013.

CURY, Roberto. A educação básica como direito. Cadernos de Pesquisa, v. 38, 2008. DOI: 10.1590/S0100-15742008000200002.

ESTEVES, Thiago; SOUZA, José dos Santos. O papel do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) na reforma gerencial dos sistemas públicos de ensino no Brasil. Educere et Educare, v. 15, 2020. DOI: 10.17648/educare.v15i35.24136.

PERUCHIN, Débora. A educação no Brasil: desenvolvimento e desigualdades. Revista Triângulo, v. 11, p. 199, 2018. DOI: 10.18554/rt.v0i0.2587.

SILVA, Yasin; ALMEIDA, Isadora; QUEIROZ, Michell. SQL: From Traditional Databases to Big Data. In: ACM International Conference Proceeding Series, 2016, p. 413-418. DOI: 10.1145/2839509.2844560.

TOMASELLI, Guilherme Borchardt; SILVA SANTOS, Maria; FRIGO, Luciana; GASPARINI, Isabela. Ferramenta de Visualização de Dados do Censo da Educação Superior do INEP. In: Workshop de Computação Aplicada à Gestão do Ensino Superior (WCGE), 2022, Online. Anais do Workshop de Computação Aplicada à



Gestão do Ensino Superior. Porto Alegre: Sociedade Brasileira de Computação, 2022. p. 227-234. DOI: 10.5753/wcge.2022.223061.

MACHADO, Felipe Nery Rodrigues. Banco de Dados. 4. ed. São Paulo: Editora Érica, 2020.