

# EduRadar

Lucas Eduardo de Oliveira Mariz  
lucasmaliz2002@gmail.com

Fonte dos Dados: <https://www.gov.br/inep/pt-br>

Aplicação OLAP: [ProjetoPowerBI EduRadar](#)

Github: [https://github.com/LucasMariz673/ProjetoSAD\\_EduRadar.git](https://github.com/LucasMariz673/ProjetoSAD_EduRadar.git)

# ETAPA 1 - PLANEJAMENTO

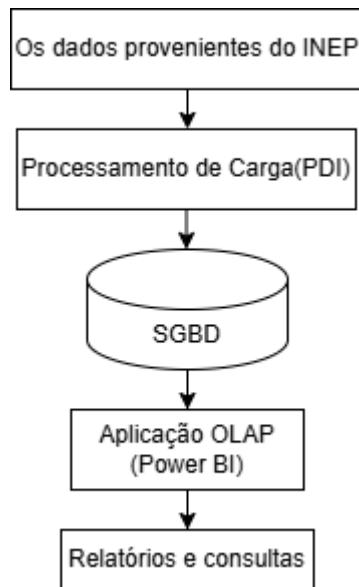
# 1. Contextualização

Os dados se tornaram ferramentas essenciais para orientar decisões estratégicas na sociedade. Quando se trata de educação, eles impacta diretamente o desenvolvimento social e econômico do país. No entanto, muitas vezes esses dados são subutilizados ou de difícil acesso para análises aprofundadas. Diante disso, este projeto tem como objetivo criar um data mart educacional com base nos dados disponibilizados pelo INEP. A proposta é consolidar informações por etapa de ensino, rede escolar e recorte geográfico, permitindo a geração de insights relevantes, como a identificação dos locais com maiores desafios educacionais e a comparação do desempenho ao longo dos anos. Essas análises são fundamentais para apoiar a gestão da educação básica e orientar políticas públicas mais eficazes em todo o território brasileiro.

## 2. Escopo/objetivo do Data Mart

Como objetivo principal, este data mart busca gerar insights que apoiem a gestão da educação básica no Brasil, possibilitando a criação de políticas públicas mais eficazes e baseadas em dados. Para isso, foram analisadas informações disponibilizadas pelo INEP, referentes às redes públicas e privadas de ensino nos anos de 2007 a 2023, com recortes por etapa de ensino e região geográfica.

### 3. Arquitetura Tecnológica



## 4. Processo

Para a construção do Data Mart, será adotado um processo organizado em etapas. Primeiramente, serão analisados os dados educacionais disponibilizados pelo INEP, com foco no IDEB e em seus principais componentes: os resultados do SAEB (desempenho dos estudantes em Língua Portuguesa e Matemática) e o índice de aprovação (indicador de fluxo escolar). Em seguida, serão extraídos os dados relevantes, abrangendo diferentes etapas de ensino e recortes geográficos. Na fase de transformação, os dados passarão por limpeza, padronização de formatos e verificação de consistência, garantindo uma base estruturada. Após isso, os dados serão carregados no Data Mart e estarão prontos para serem explorados por meio de ferramentas OLAP, permitindo a geração de análises históricas, comparativas e indicadores que auxiliem na avaliação e melhoria da qualidade da educação pública.

## 4. Abordagem

Para este projeto, optamos pela abordagem Bottom-Up, proposta por Ralph Kimball, por ser adequada à criação de soluções analíticas voltadas a áreas específicas. A estrutura do Data Mart será baseada no modelo dimensional em estrela (Star Schema), que favorece a simplicidade nas consultas e a eficiência no uso de sistemas gerenciadores de banco de dados relacionais. A modelagem contempla uma tabela fato central, que consolida métricas como notas do SAEB, índices de aprovação e valores do IDEB, conectada a diversas tabelas dimensão, como tempo, rede de ensino, localização e etapa escolar. Essa arquitetura facilita a análise dos dados e a geração de relatórios direcionados para a avaliação da qualidade da educação nas diferentes regiões do país.

## 5. Usuários

O Data Mart desenvolvido poderá atender a diferentes perfis de usuários dentro do setor educacional. Gestores públicos poderão utilizá-lo como base para decisões estratégicas voltadas à melhoria da qualidade do ensino. Além disso, analistas de dados e pesquisadores terão à disposição um ambiente estruturado para realizar estudos comparativos, acompanhar a evolução de indicadores como o IDEB, SAEB e taxa de aprovação, e identificar padrões regionais. A ferramenta também pode ser útil para órgãos de controle e planejamento, ao oferecer uma visão clara sobre o desempenho educacional nas diferentes redes e territórios do país.



# ETAPA 2 - LEVANTAMENTO DAS NECESSIDADES

## 6. Consultas de Apoio à Decisão

Nesta fase, foram selecionados indicadores estratégicos para o monitoramento da qualidade educacional, a partir dos dados fornecidos pelo INEP sobre índice de aprovação, SAEB e IDEB. As consultas de apoio à decisão podem contemplar questões como:

- Qual foi a disciplina do SAEB de 2009 a 2019 no Sul que apresentou maior valor?
- Como evoluíram as notas do SAEB nas escolas públicas do Norte entre 2007 e 2023?
- Quais estados registraram maior IDEB dos Anos Iniciais em escolas públicas e particulares no último ciclo avaliativo?
- Como se compara o desempenho geral do IDEB nos Anos Finais entre as escolas públicas e privadas da região Nordeste?

## 7. Indicadores do <negócio>

Índice de aprovação escolar, para medir a proporção de alunos aprovados. Auxilia gestores a acompanhar o desempenho acadêmico geral e identificar possíveis necessidades de intervenção nas escolas públicas e privadas. (Fonte: INEP)

Notas médias do SAEB , para avaliar o nível de proficiência dos estudantes em disciplinas como Língua Portuguesa e Matemática. Permite analisar o progresso da aprendizagem em diferentes estados e regiões, identificando desigualdades e orientando políticas educacionais. (Fonte: INEP)

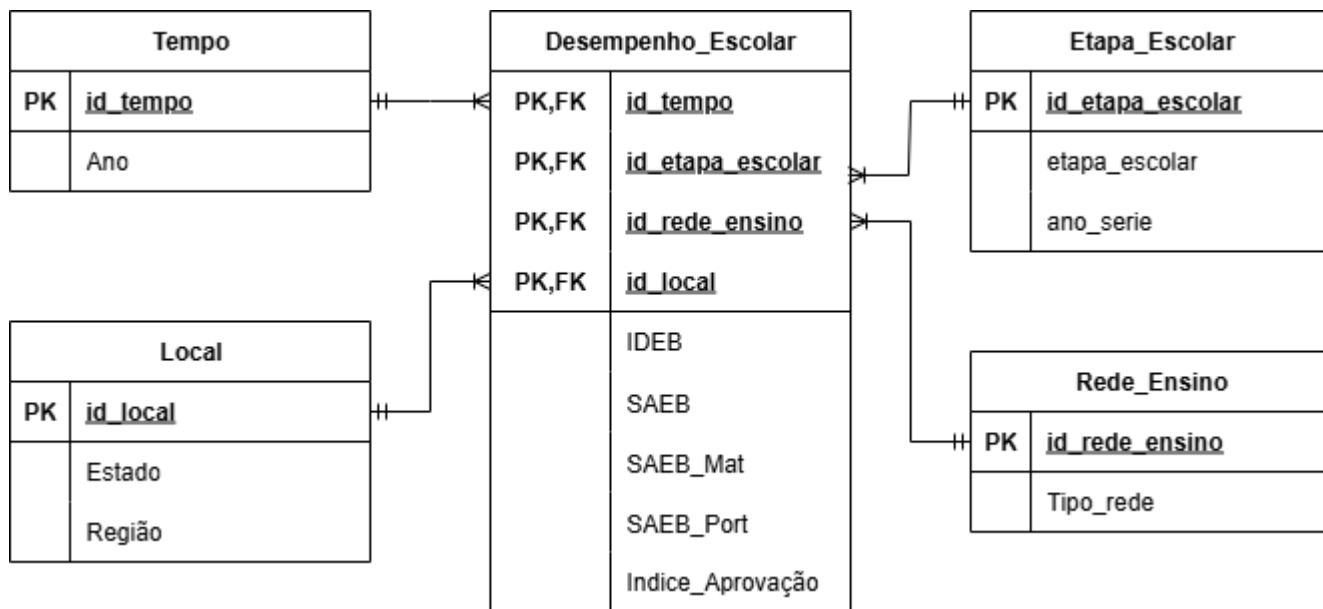
Índice de Desenvolvimento da Educação Básica (IDEB), para mensurar a qualidade da educação nas etapas do ensino fundamental e médio, combinando taxas de aprovação e desempenho em avaliações. Serve para monitorar a evolução da qualidade do ensino e definir metas de melhoria. (Fonte: INEP)

# ETAPA 3 - MODELAGEM

## 9. Modelo Relacional

Para gerar o modelo relacional, foram definidas quatro tabelas de dimensões principais. A dim\_local contempla os recortes geográficos, com atributos como estado e região, permitindo análises territoriais. A dim\_tempo organiza a dimensão temporal, com o atributo ano, possibilitando o acompanhamento da evolução dos indicadores ao longo do tempo. A dim\_etapa descreve o nível de ensino, incluindo atributos como ano/série e etapa de ensino. Já a dim\_rede\_ensino classifica os dados conforme o tipo de rede (pública ou privada), oferecendo comparações entre as redes de ensino. No centro do modelo, a tabela de fatos fato\_desempenho\_escolar consolida as principais métricas: IDEB, nota do SAEB e taxa de aprovação, fornecendo a base para relatórios e análises sobre a qualidade da educação em todo o território nacional.

## 9. Modelo Relacional



# 10. Modelo Dimensional

Este Modelo Dimensional foi desenvolvido com foco na gestão da educação básica, com o objetivo de apoiar a formulação de políticas públicas e o monitoramento da qualidade do ensino. A área de negócios abrange o acompanhamento de indicadores educacionais nacionais, como o IDEB, o SAEB e a taxa de aprovação, permitindo análises comparativas entre regiões, estados e redes de ensino. O processo central é a avaliação do desempenho educacional, em que cada fato registrado representa a combinação de métricas educacionais por rede, etapa e unidade geográfica. A granularidade do modelo está definida no nível de estado e região, o que possibilita análises agregadas, mantendo o foco em tendências educacionais e disparidades regionais.

# 10. Modelo Dimensional

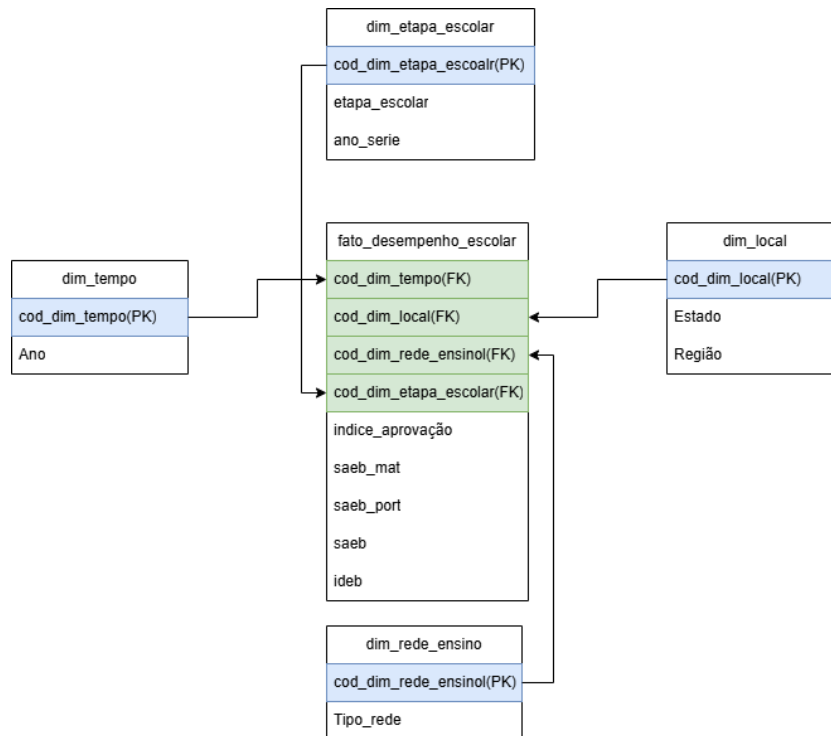
As hierarquias dentro das dimensões estão bem definidas. Na dim\_tempo, a hierarquia é simples, composta pelo atributo ano, permitindo análises históricas. Na dim\_local, a hierarquia segue região > estado, viabilizando comparações por recortes geográficos. Já na dim\_etapa, a hierarquia é formada por etapa de ensino (anos iniciais, anos finais, ensino médio) e série/ano (1º ao 9º ano do ensino fundamental e 1ª a 4ª série do ensino médio). Por fim, a dim\_rede\_ensino categoriza o tipo de rede em pública ou privada. Essas estruturas possibilitam uma análise flexível e detalhada dos indicadores educacionais, oferecendo suporte a uma gestão orientada por dados e baseada em evidências.



# 10.Modelo Dimensional

A tabela `fato_desempenho_escolar` contém três métricas principais relacionadas ao desempenho educacional. A primeira é o IDEB (Índice de Desenvolvimento da Educação Básica), um indicador que combina o desempenho dos alunos nas avaliações do SAEB com a taxa de aprovação. Esta métrica é do tipo não aditiva, pois representa um índice e não pode ser somada diretamente. Mais três métricas são a `saeb_port` e `saeb_mat`, que correspondem à nota padronizada obtida pelos estudantes nas áreas de língua portuguesa e matemática, e `saeb`, que é a nota média padronizada. Assim como o IDEB, são métricas não aditivas. Por fim, a `indice_aprovação`, expressa em percentual, que indica a proporção de alunos aprovados em determinada etapa e rede de ensino. Esta métrica também é não aditiva, sendo utilizada para acompanhamento e análise de fluxo escolar. Essas métricas, em conjunto, possibilitam uma análise abrangente da qualidade da educação nas diferentes regiões e redes de ensino do país.

# 10. Modelo Dimensional



# 10.Modelo Dimensional

	cod_dim_tempo	cod_dim_local	cod_dim_etapa_escolar	cod_dim_rede_ensino	indice_aprovação	saeb_mat	saeb_port	saeb	ideb
1	3	13	10	1	70,1	184,20	178,28	4,73	4,1
2	3	8	5	1	84,4	209,84	203,81	5,68	5,6
3	2	2	1	2	90,2	201,90	195,62	5,38	5,3
4	3	20	5	2	75	248,11	243,31	7,13	7,1
5	2	13	4	2	69,8	221,69	218,25	6,17	6,1
6	1	13	1	1	78,5	227,21	220,91	6,32	6,3
7	2	5	2	2	98,6	219,98	214,85	6,07	6,1
8	3	2	7	2	93,5	251,89	241,63	7,17	7,1
9	3	1	9	1	86,2	251,66	244,60	7,22	7,2
10	1	17	7	1	92,7	201,84	193,89	5,35	5,0

# 10. Modelo Dimensional do Data Mart (lógico)

## **Tabela Fato:**

Quantidade de linhas da fato = 7.020 linhas

Quantidade de colunas = 9 (4 chaves + 5 métricas)

Bytes por linha = 9 colunas \* 4 bytes = 36 bytes por linha

$36 * 7.020 = 252.720 \text{ bytes} = 252,7 \text{ KB}$

**Tabelas Dimensões** (estimado em 25% do tamanho da fato):

25% de 252,7 KB = 63,2 KB

## **Total estimado:**

$252,7 \text{ KB (fato)} + 63,2 \text{ KB (dimensões)} = 315,9 \text{ KB}$

# ETAPA 4 - PROJETO FÍSICO DO BD

# 11. Modelo Relacional do Data Mart (físico)

```
CREATE TABLE tempo (  
  id_tempo SERIAL PRIMARY KEY,  
  ano INT);
```

```
CREATE TABLE rede (  
  id_rede_ensino SERIAL PRIMARY KEY,  
  tipo_rede TEXT);
```

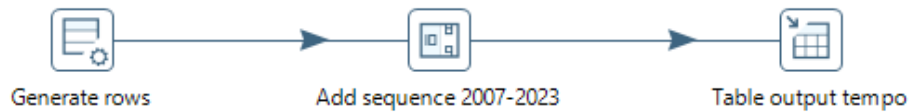
```
CREATE TABLE local (  
  id_local SERIAL PRIMARY KEY,  
  estado TEXT,  
  regioao TEXT);
```

```
CREATE TABLE etapa_escolar (  
  id_etapa_escolar SERIAL PRIMARY KEY,  
  etapa_escolar TEXT,  
  ano_serie TEXT);
```

```
CREATE TABLE desempenho_escolar (  
  id_local INT,  
  id_tempo INT,  
  id_rede_ensino INT,  
  id_etapa_escolar INT,  
  ideb NUMERIC(5,2),  
  indice_aprovacao NUMERIC(5,2),  
  saeb_port NUMERIC(5,2),  
  saeb_mat NUMERIC(5,2),  
  saeb NUMERIC(5,2),  
  CONSTRAINT pk_fato_educacao PRIMARY KEY (id_local, id_tempo,  
  id_rede_ensino, id_etapa_escolar),  
  CONSTRAINT fk_fato_local FOREIGN KEY (id_local) REFERENCES  
  local (id_local),  
  CONSTRAINT fk_fato_tempo FOREIGN KEY (id_tempo)  
  REFERENCES tempo (id_tempo),  
  CONSTRAINT fk_fato_rede FOREIGN KEY (id_rede_ensino)  
  REFERENCES rede (id_rede_ensino),  
  CONSTRAINT fk_fato_etapa FOREIGN KEY (id_etapa_escolar)  
  REFERENCES etapa_escolar (id_etapa_escolar));
```

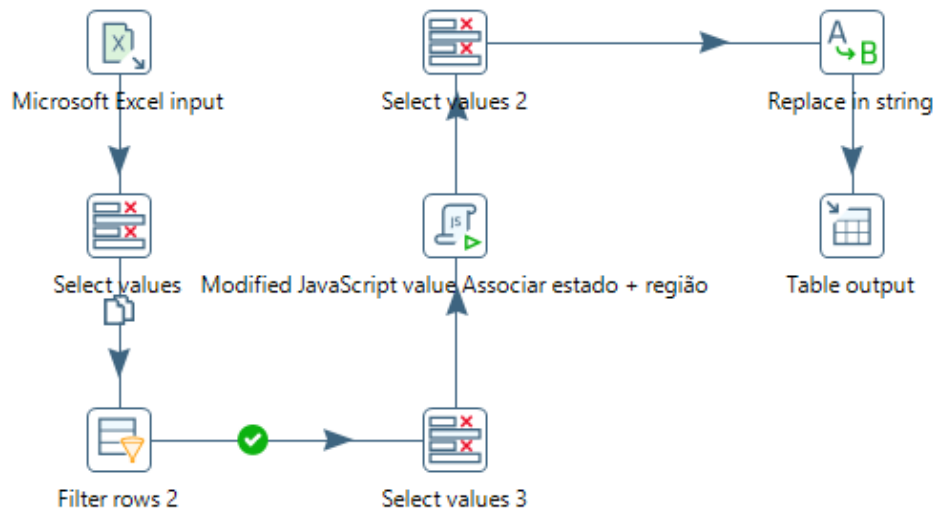
# ETAPA 5 - EXTRAÇÃO, TRANSFORMAÇÃO E CARGA

## 12.Plano de Carga da Dimensão Tempo

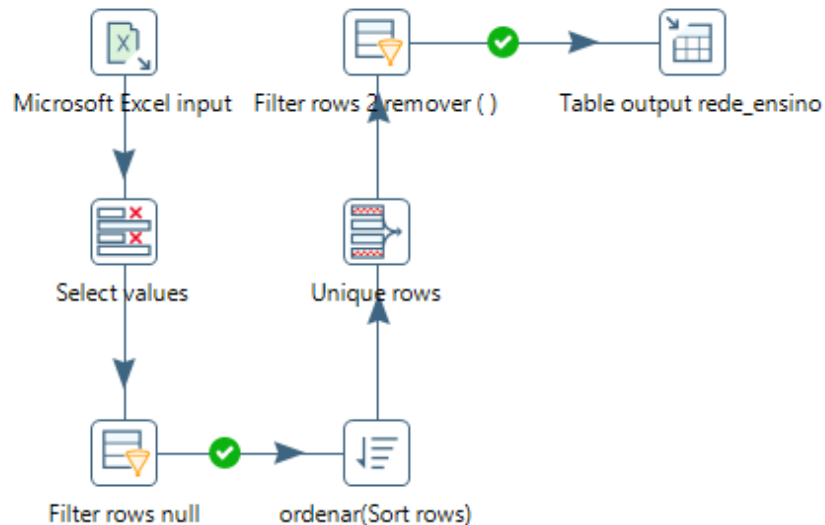




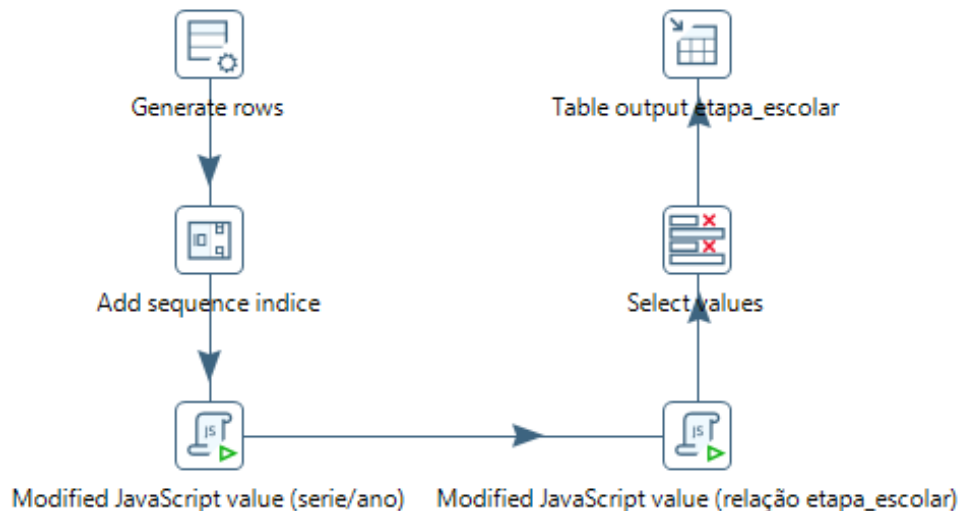
# 13. Plano de Carga da Dimensão Local



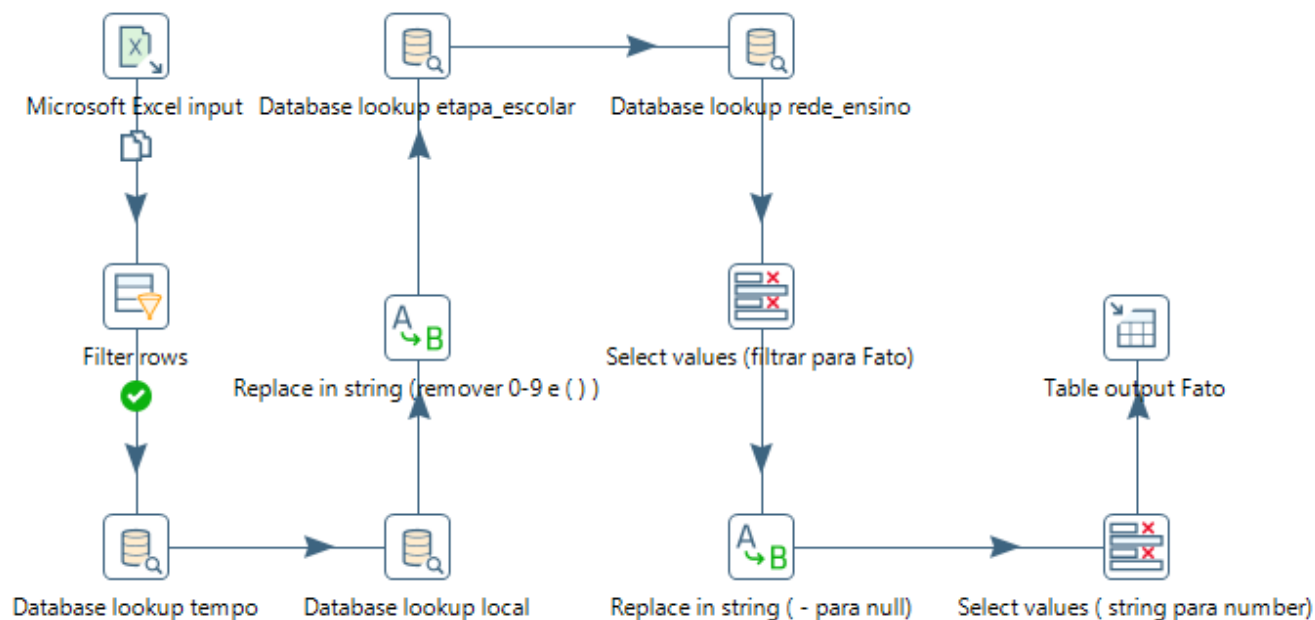
# 14.Plano de Carga da Dimensão Rede Ensino



## 14. Plano de Carga da Dimensão Etapa Escolar



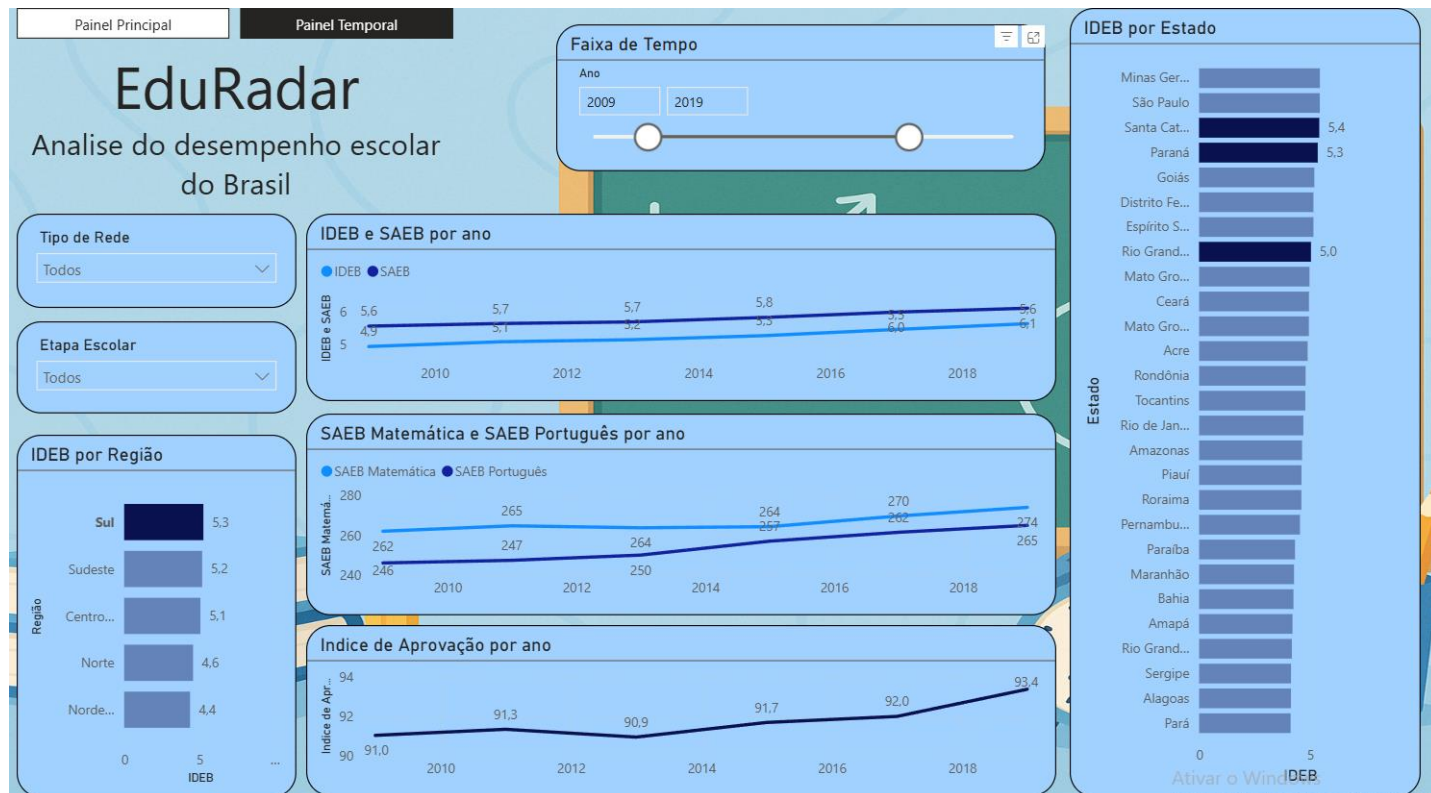
# 15. Plano de Carga da Fato



# ETAPA 6 - APLICAÇÃO OLAP e PAINEL DE BORDO

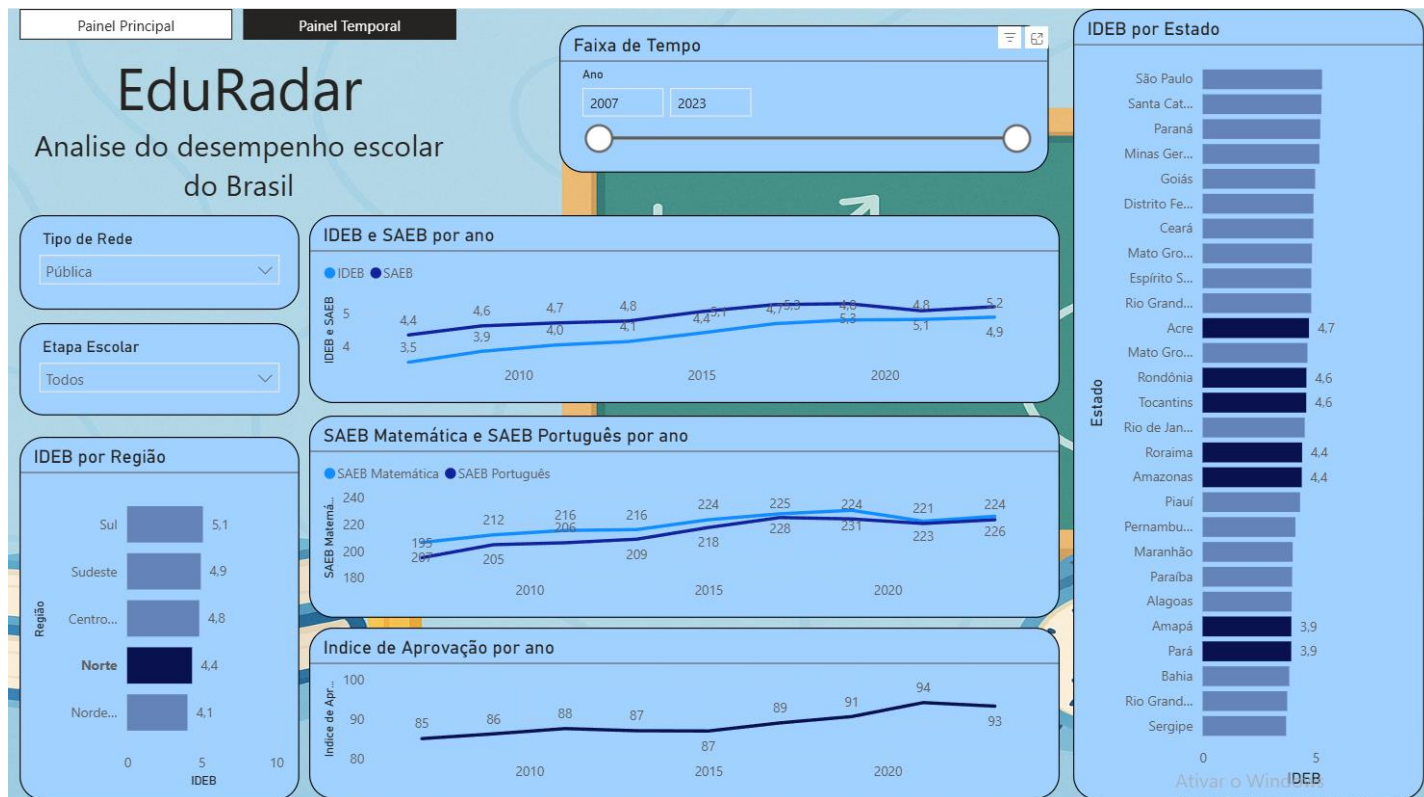
# 16.Consulta OLAP 1

Qual foi a disciplina do SAEB de 2009 a 2019 no Sul que apresentou maior valor?



# 16.Consulta OLAP 2

Como evoluíram as notas do SAEB nas escolas públicas do Norte entre 2007 e 2023?



# 16.Consulta OLAP 3

Quais estados registraram maior IDEB dos Anos Iniciais em escolas públicas no último ciclo avaliativo?





# Referências

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