

Perfil dos beneficiários do PROUNI

2015 à 2019

Emanoel Rainey Curvelo Manço

eraineycm8@gmail.com

Links

- ❖ Dataset: <https://www.kaggle.com/datasets/lfarhat/brasil-students-scholarship-prouni-20052019>
- ❖ Git: https://github.com/eraineycm8/profile_PROUNI_beneficiaries

Etapa 1 - Planejamento



PROUNI – Contexto geral

- ❖ O PROUNI é um programa que oferece bolsas de estudo, integrais e parciais, em cursos de graduação em instituições privadas. Podem participar estudantes sem diploma de nível superior que tenham realizado o ENEM, tendo obtido no mínimo 450 pontos na média das provas e não tenham zerado a redação, além de atender critérios de renda familiar.
- ❖ Para a bolsa integral, a renda familiar mensal per capita deve ser de até 1,5 salário mínimo, enquanto para a parcial, é de até 3 salários mínimos. É necessário ter cursado o ensino médio em escola pública ou privada como bolsista integral, parcial ou de forma mista, ou se enquadrar como pessoa com deficiência. Professores da rede pública de ensino também podem se candidatar para cursos de licenciatura e pedagogia, sem limite de renda.

Objetivos do Data Mart

- ❖ O objetivo do Data Mart a ser desenvolvido é fornecer uma análise detalhada do perfil dos beneficiários do PROUNI entre os anos de 2015 e 2019. Este Data Mart possibilitará extrair informações relevantes sobre a distribuição dos bolsistas por região, faixa etária, tipo de bolsa (integral ou parcial) e cursos frequentados ao longo desse período. A análise permitirá identificar tendências, variações e características específicas dos estudantes contemplados pelo programa, e outros pontos de vista sobre o programa em si. Com a análise dos dados espera-se criar questionamentos e insights para avaliar e aumentar a eficácia do programa, em busca de entender a distribuição de recursos e beneficiários, os estudantes de baixa renda em busca de educação superior.

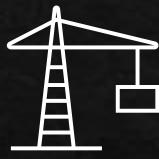
Arquitetura Tecnológica

Fonte de Dados



CSV

ETL



Pentaho

Data Mart



Mysql

Aplicação OLAP



Power BI

Processo

- ❖ Planejamento
- ❖ Levantamento das Necessidades
- ❖ Modelagem dimensional
- ❖ Projeto Físico do Banco de Dados
- ❖ Projeto ETL
- ❖ Aplicação OLAP

Processo

- ❖ Para criar este Data Mart, foi adotada a estratégia Bottom-up, a qual se desenvolve de forma progressiva e permite a entrega imediata de resultados. Essa abordagem foi escolhida por ser específica para a construção de um Data Mart e por sua ênfase na análise detalhada de partes para compor o todo.
- ❖ Será utilizado o modelo "Star Schema" por ser uma extensão incremental da estratégia Bottom-up, focando na montagem do Data Mart. Essa técnica concentra-se na criação de tabelas de dimensões e fatos, alinhada à natureza descentralizada de um Data Warehouse, facilitando a análise e a compreensão dos dados em um contexto mais segmentado e analítico.

Etapa 2 - Levantamento das Necessidades



Consultas de Apoio à Decisão

- ❖ Crescimento de beneficiários do PROUNI
- ❖ Quantidade de estudantes por região e ano
- ❖ Comparação de estudantes entre regiões por ano
- ❖ Distribuição dos estudantes por idade e tipo da bolsa
- ❖ Distribuição dos estudantes por ano e idade

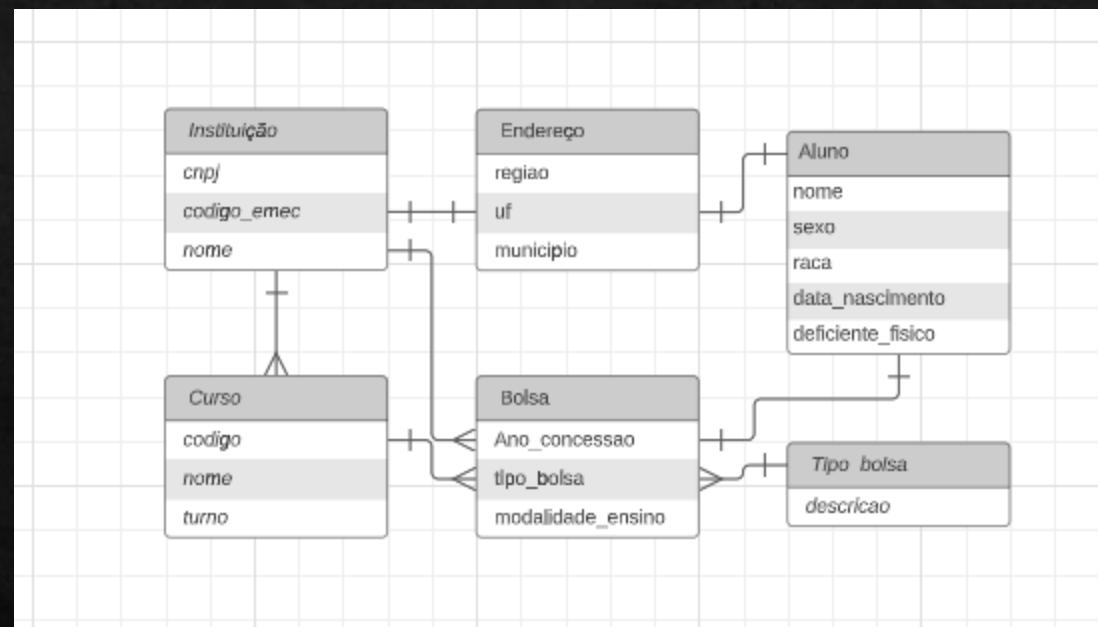
Indicadores de análise

- ❖ Quantidade de estudantes por região
- ❖ Quantidade de estudantes por estado
- ❖ Quantidade de estudantes por faixa etária
- ❖ Quantidade de estudantes por ano

Etapa 3 - Modelagem



Modelo Relacional



Modelo Dimensional

A – Área de Negócios

Educação

B – Processo

Análise dos beneficiários do PROUNI no Brasil entre 2015 a 2019

C – Granularidade

Estado x Região x Ano x Faixa etária

Modelo Dimensional

D – Atributos e Hierarquia das Dimensões

(DIM_Local: região, estado, cidade)

(DIM_Tempo: ano)

(DIM_Perfil_Beneficiário: faixa_etaria, sexo, condicao_fisica, grupo_racial)

(DIM_Perfil_Bolsa: tp_bolsa, modalidade, turno)

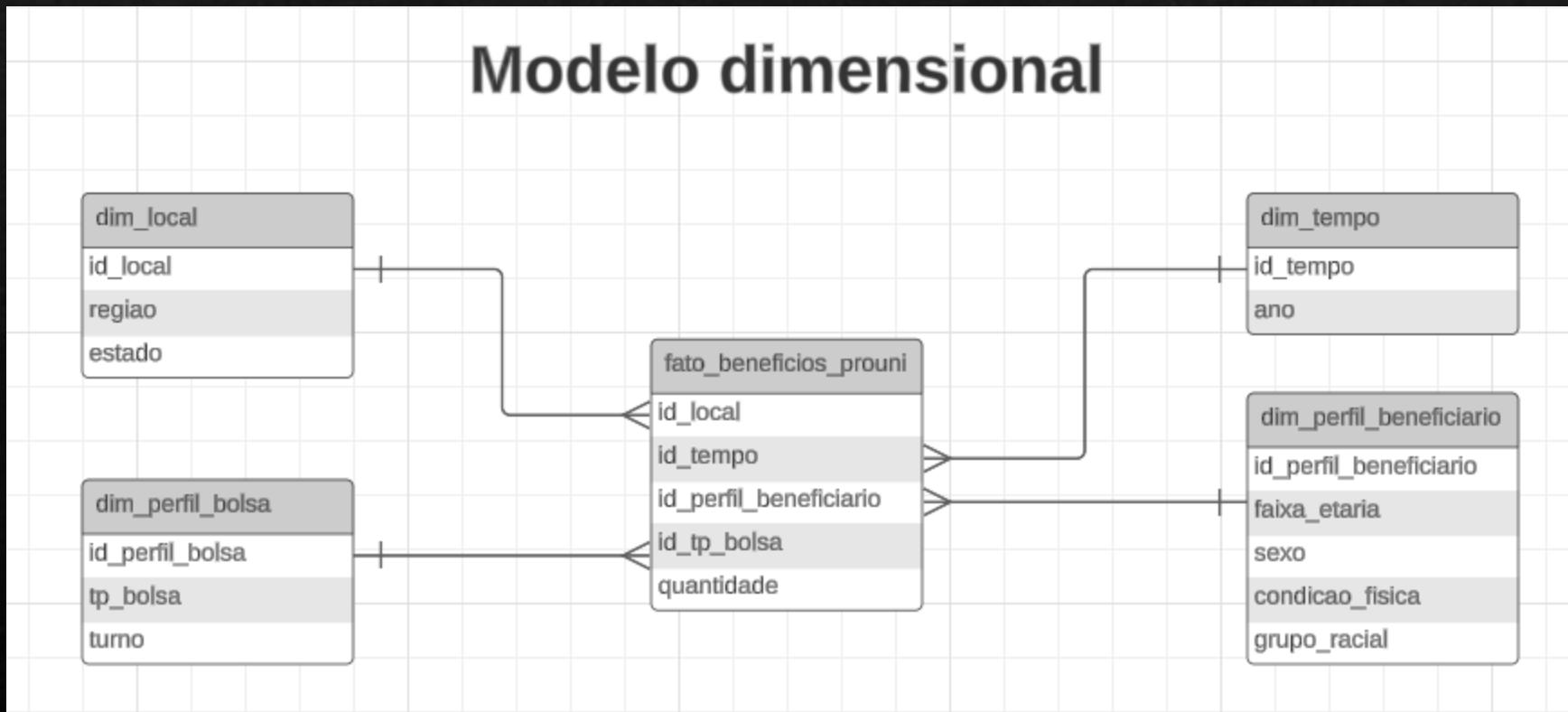
Modelo Dimensional

E – Métricas da Fato

- ❖ Quantidade (aditiva)

Modelo Dimensional

F – Esquema Estrela



dim_tempo	
id_tempo	ano
1	2015
2	2016
3	2017
4	2018
5	2019

Modelo Dimensional

G – Inserção de 10 “Fatos”

fato_beneficios_prouni				
id_tempo	id_local	id_perfil_bolsa	id_perfilBeneficiario	quantidade
1	1	6		64
1	1	8		110
1	1	8		151
1	1	9		63
1	1	9		64
1	1	9		95
1	1	9		111
1	1	9		151
1	1	9		169
1	1	13		64

dim_perfil_bolsa			
id_perfil_bolsa	tp_bolsa	modalidade	turno
1	Bolsa 25% (Complementar)	Educação a distância	A Distância
2	Bolsa 25% (Complementar)	Presencial	Integral
3	Bolsa 25% (Complementar)	Presencial	Matutino
4	Bolsa 25% (Complementar)	Presencial	Noturno
5	Bolsa 25% (Complementar)	Presencial	Vespertino
6	Bolsa 100% (Integral)	Educação a distância	A Distância
7	Bolsa 100% (Integral)	Presencial	Integral
8	Bolsa 100% (Integral)	Presencial	Matutino
9	Bolsa 100% (Integral)	Presencial	Noturno
10	Bolsa 100% (Integral)	Presencial	Vespertino

dim_perfil_beneficiario				
id_perfil_beneficiario	faixa_etaria	sexo	condicao_fisica	grupo_racial
	1 Sub 16	Feminino	Pessoa sem deficiência	Amarela
	2 De 16 a 18 anos	Feminino	Pessoa sem deficiência	Amarela
	3 De 19 a 21 anos	Feminino	Pessoa sem deficiência	Amarela
	4 De 22 a 25 anos	Feminino	Pessoa sem deficiência	Amarela
	5 De 26 a 30 anos	Feminino	Pessoa sem deficiência	Amarela
	6 De 31 a 35 anos	Feminino	Pessoa sem deficiência	Amarela
	7 De 36 a 40 anos	Feminino	Pessoa sem deficiência	Amarela
	8 De 41 a 50 anos	Feminino	Pessoa sem deficiência	Amarela
	9 51+ anos	Feminino	Pessoa sem deficiência	Amarela
	10 De 19 a 21 anos	Feminino	Pessoa com deficiência	Amarela

dim_local			
id_local	regiao	estado	cidade
2	Centro-Oeste	Distrito Federal	Brasília
3	Centro-Oeste	Goiás	Abadia De Goias
4	Centro-Oeste	Goiás	Abadiania
5	Centro-Oeste	Goiás	Acreuna
6	Centro-Oeste	Goiás	Adelandia
7	Centro-Oeste	Goiás	Agua Fria De Goias
8	Centro-Oeste	Goiás	Agua Limpa
9	Centro-Oeste	Goiás	Aguas Lindas De Goias
10	Centro-Oeste	Goiás	Alexania
11	Centro-Oeste	Goiás	Aloandia

Modelo Dimensional

H – Estimativa de Espaço

- ❖ Tendo em vista 5600 locais, 15 perfis de bolsa, 543 perfis de beneficiários, 5 tempos, temos:
- ❖ $5600 \times 15 \times 543 \times 5 = 228.060.000$ registros.
- ❖ Sendo esses campos chaves cada com 4 bytes e mais um valor numérico também de 4 bytes temos:
- ❖ $228.060.000 \times ((4 \times 4) + (1 \times 4)) = 14.595.840.000$ bytes

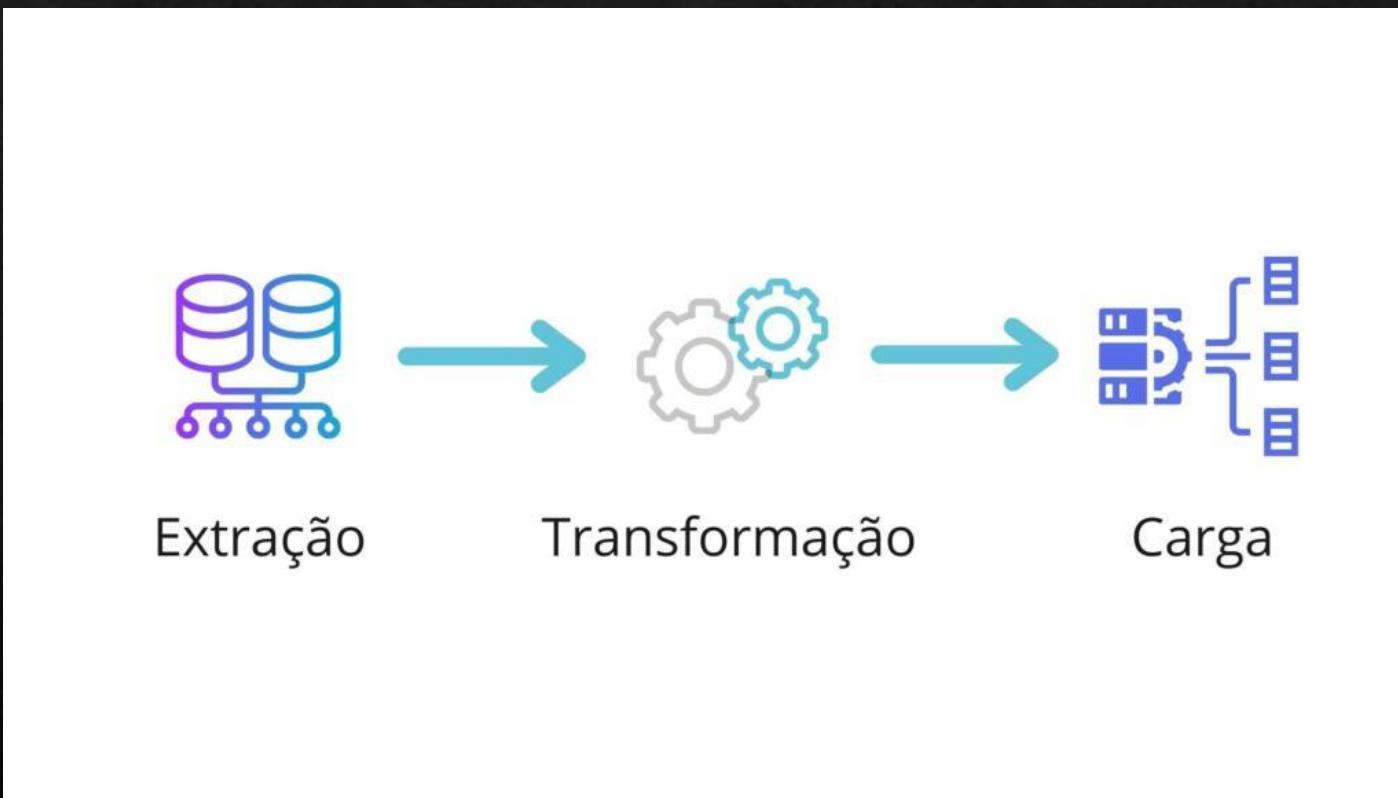
Etapa 4 - Projeto Físico do BD

Modelo Relacional do DataMart

I – DML

```
1  CREATE SCHEMA profile_prouni_beneficiaries;
2
3  USE profile_prouni_beneficiaries;
4
5  CREATE TABLE dim_tempo (
6      id_tempo INT AUTO_INCREMENT PRIMARY KEY,
7      ano INT
8  );
9
10 CREATE TABLE dim_local (
11     id_local INT AUTO_INCREMENT PRIMARY KEY,
12     regiao VARCHAR(100),
13     estado VARCHAR(100),
14     cidade VARCHAR(100)
15 );
16
17 CREATE TABLE dim_perfil_bolsa (
18     id_perfil_bolsa INT AUTO_INCREMENT PRIMARY KEY,
19     tp_bolsa VARCHAR(100),
20     modalidade VARCHAR(100),
21     turno VARCHAR(100)
22 );
23
24
25 CREATE TABLE dim_perfil_beneficiario (
26     id_perfil_beneficiario INT AUTO_INCREMENT PRIMARY KEY,
27     faixa_etaria VARCHAR(50),
28     sexo VARCHAR(50),
29     condicao_fisica VARCHAR(100),
30     grupo_racial VARCHAR(50)
31 );
32
33
34 CREATE TABLE fato_beneficios_prouni (
35     id_tempo INT,
36     id_local INT,
37     id_perfil_bolsa INT,
38     id_perfil_beneficiario INT,
39     -- Metricas abaixo
40     quantidade INT,
41     PRIMARY KEY (id_tempo, id_local, id_perfil_bolsa, id_perfil_beneficiario),
42     CONSTRAINT fk_tempo FOREIGN KEY (id_tempo) REFERENCES dim_tempo(id_tempo),
43     CONSTRAINT fk_local FOREIGN KEY (id_local) REFERENCES dim_local(id_local),
44     CONSTRAINT fk_perfil_bolsa FOREIGN KEY (id_perfil_bolsa) REFERENCES dim_perfil_bolsa(id_perfil_bolsa),
45     CONSTRAINT fk_perfil_beneficiario FOREIGN KEY (id_perfil_beneficiario) REFERENCES dim_perfil_beneficiario(id_perfil_beneficiario)
46 );
47
```

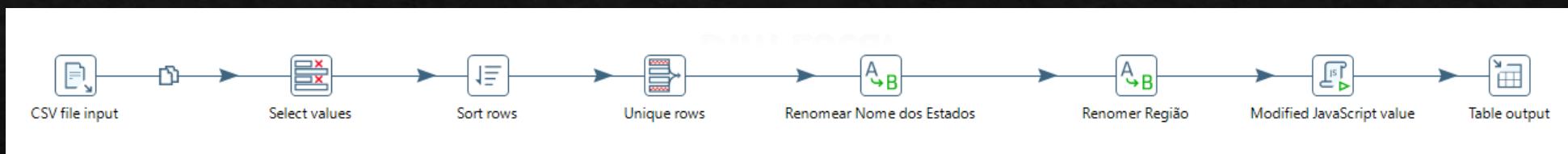
Etapa 5 - Extração, Transformação e Carga



Extração, Transformação e Carga

J – Plano de cargas

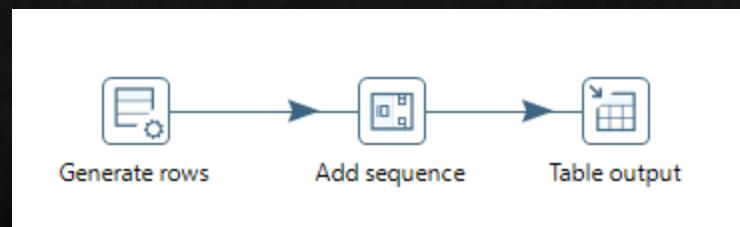
Dim Local



Extração, Transformação e Carga

J – Plano de cargas

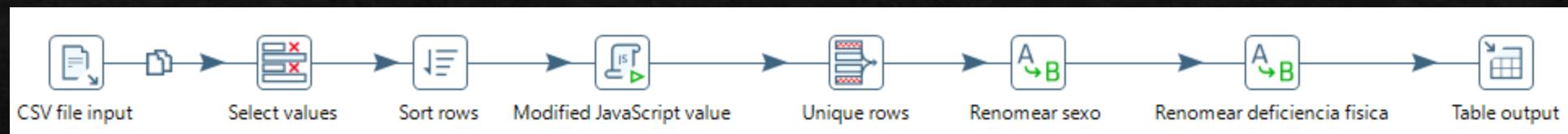
Dim Tempo



Extração, Transformação e Carga

J – Plano de cargas

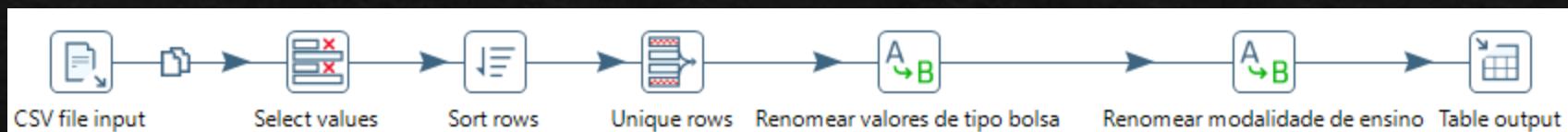
Dim Perfil do Beneficiário



Extração, Transformação e Carga

J – Plano de cargas

Dim Perfil da Bolsa



Extração, Transformação e Carga

J – Plano de cargas

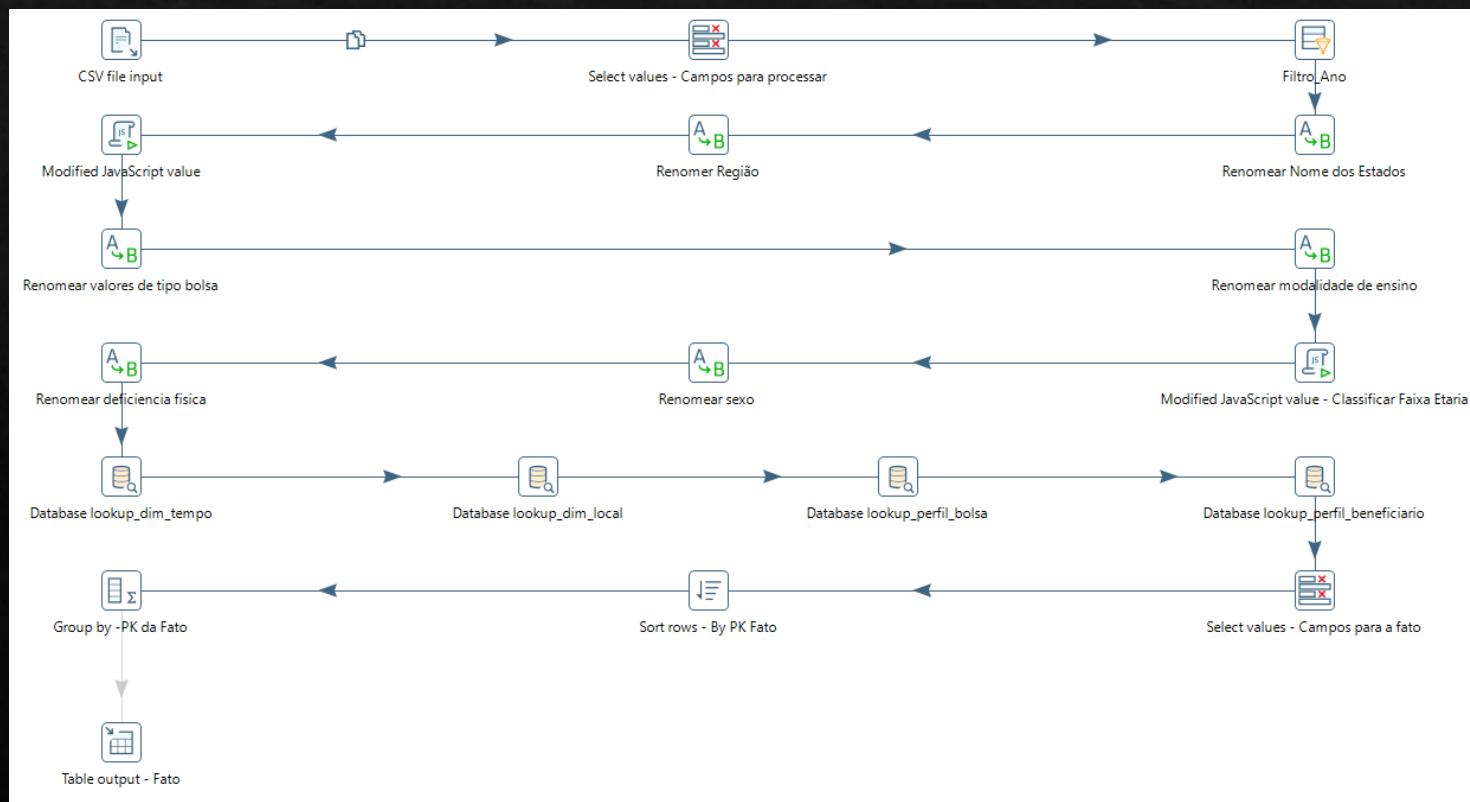
Dim Perfil da Bolsa



Extração, Transformação e Carga

J – Plano de cargas

Fato Benefícios Prouni

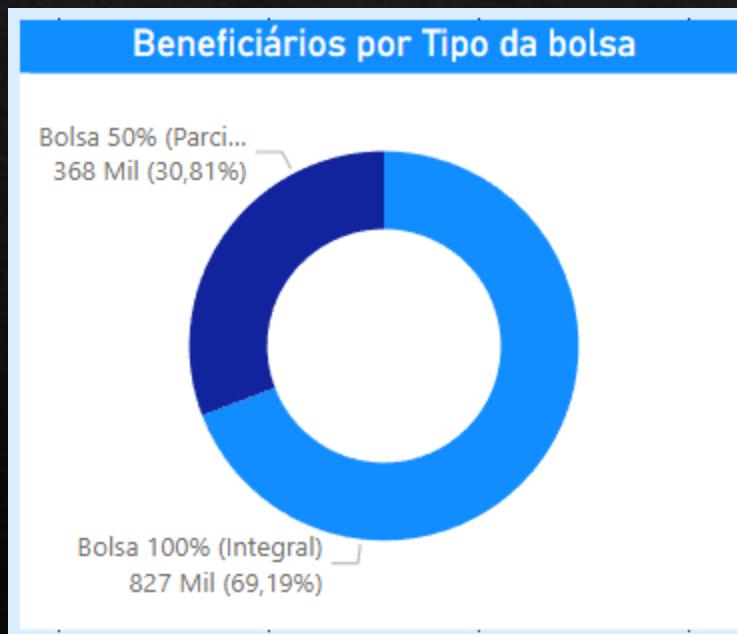


Etapa 6 - Aplicação OLAP



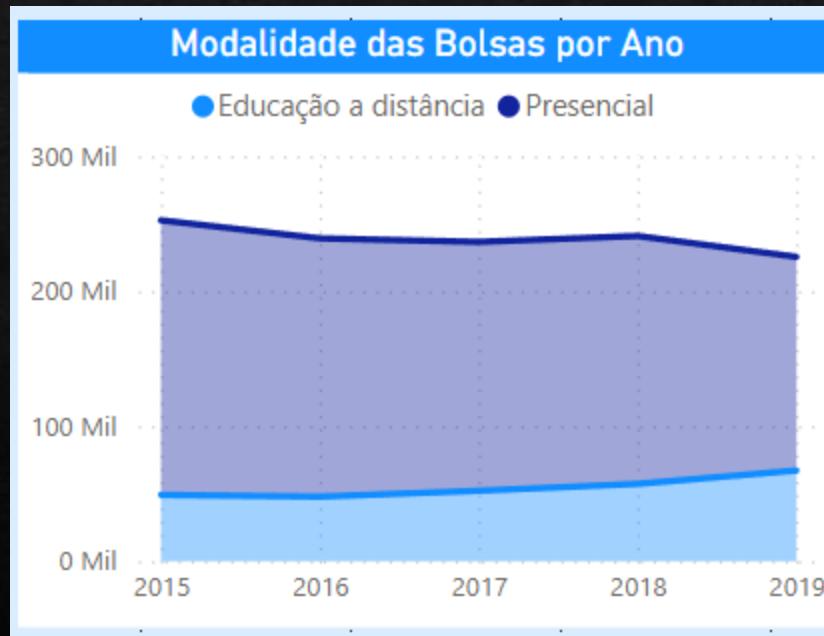
Aplicação OLAP

K- Consultas OLAP - Beneficiários por Tipo da bolsa



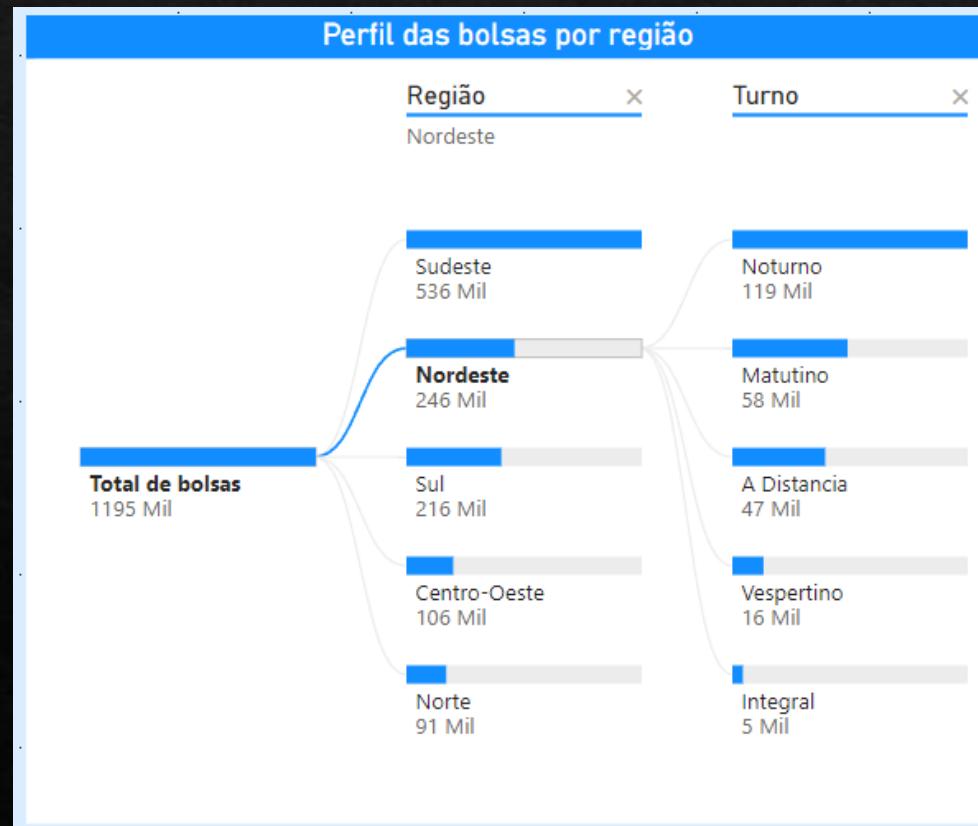
Aplicação OLAP

K– Consultas OLAP – Modalidade das Bolsas por Ano



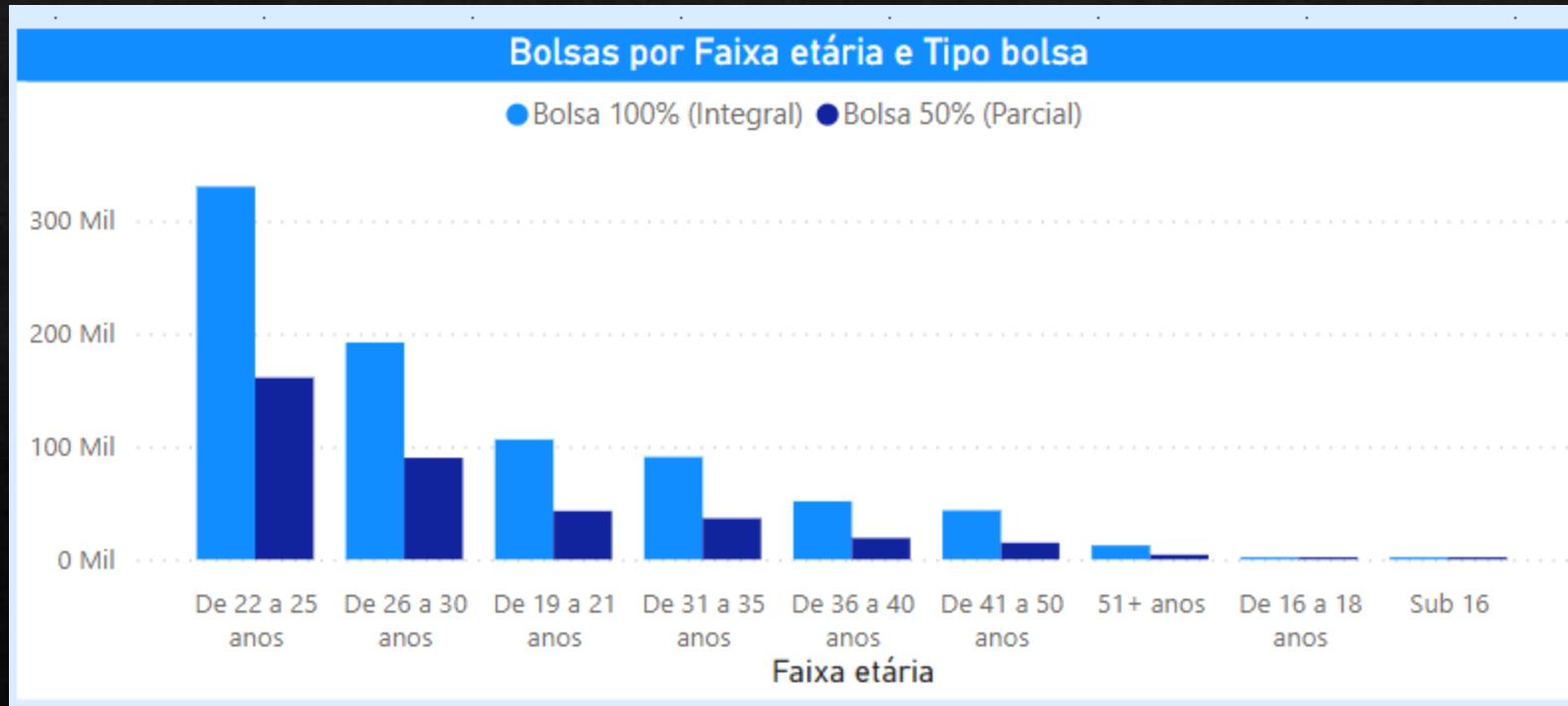
Aplicação OLAP

K– Consultas OLAP – Perfil das bolsas por região



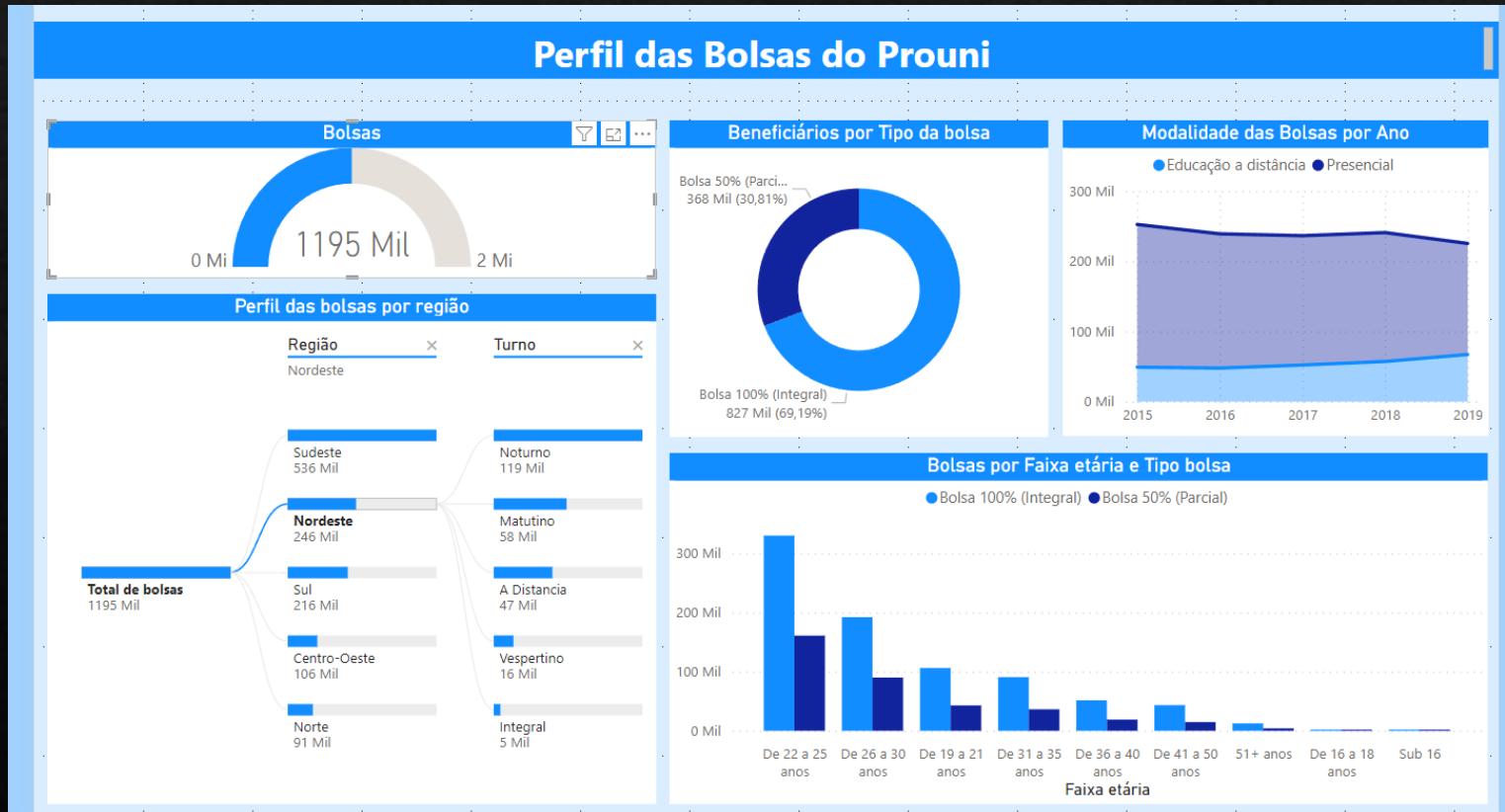
Aplicação OLAP

K– Consultas OLAP – Bolsas por Faixa Etária e Tipo Bolsa



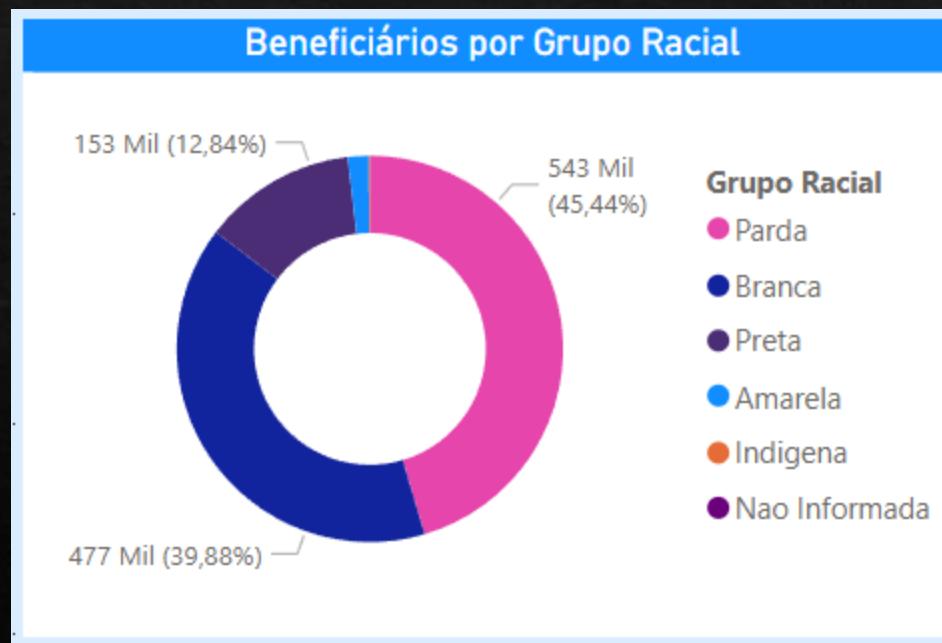
Aplicação OLAP

K– Página de Consultas OLAP - Perfil das Bolsas do Prouni



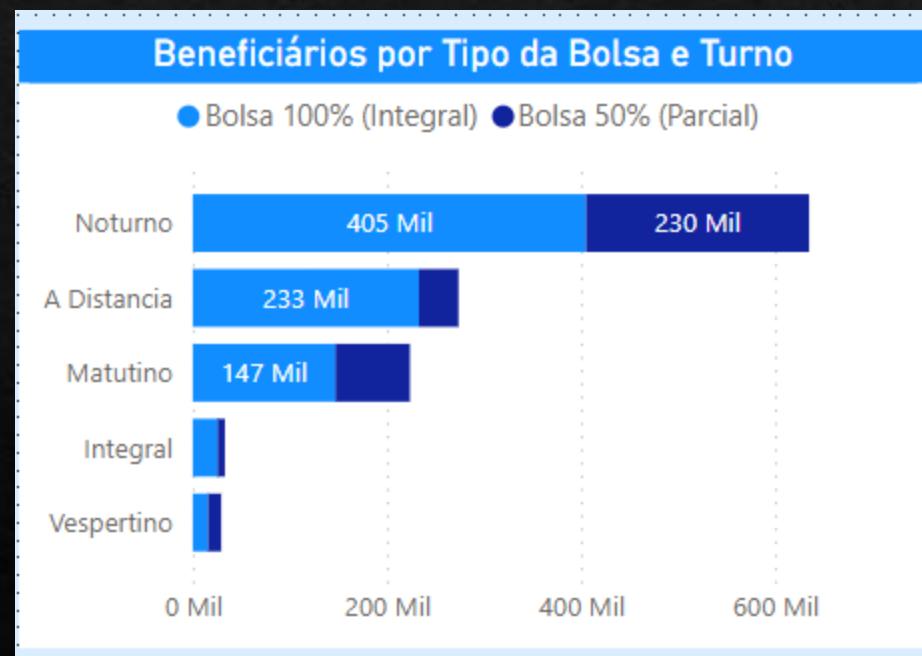
Aplicação OLAP

K- Consultas OLAP - Beneficiários por Grupo Racial



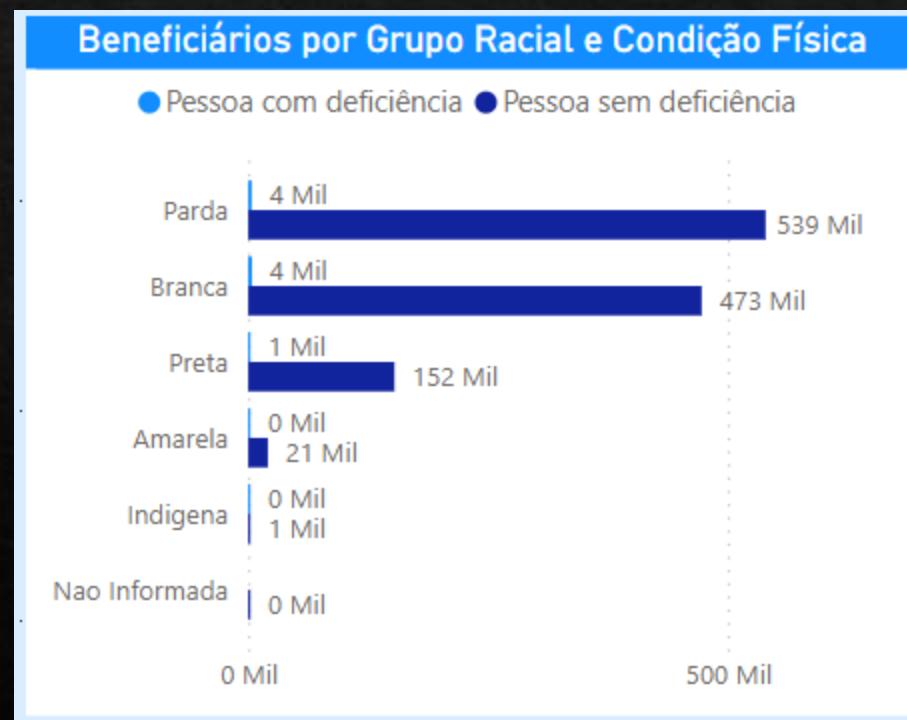
Aplicação OLAP

K- Consultas OLAP - Beneficiários por Tipo da Bolsa e Turno



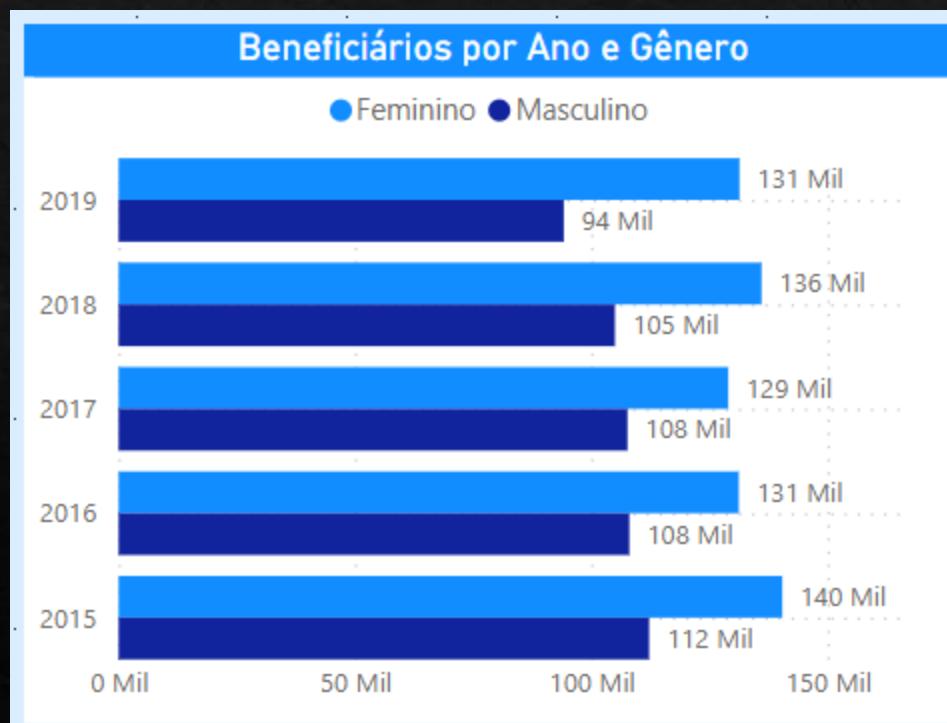
Aplicação OLAP

K- Consultas OLAP - Beneficiários por Grupo Racial e Condição Física



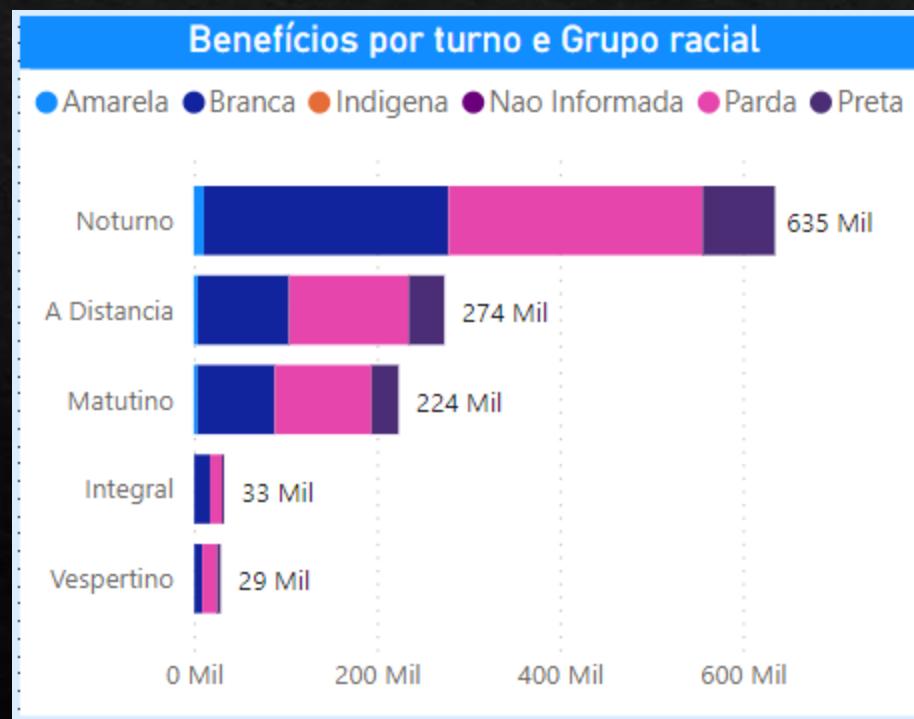
Aplicação OLAP

K- Consultas OLAP - Beneficiários por Ano e Gênero



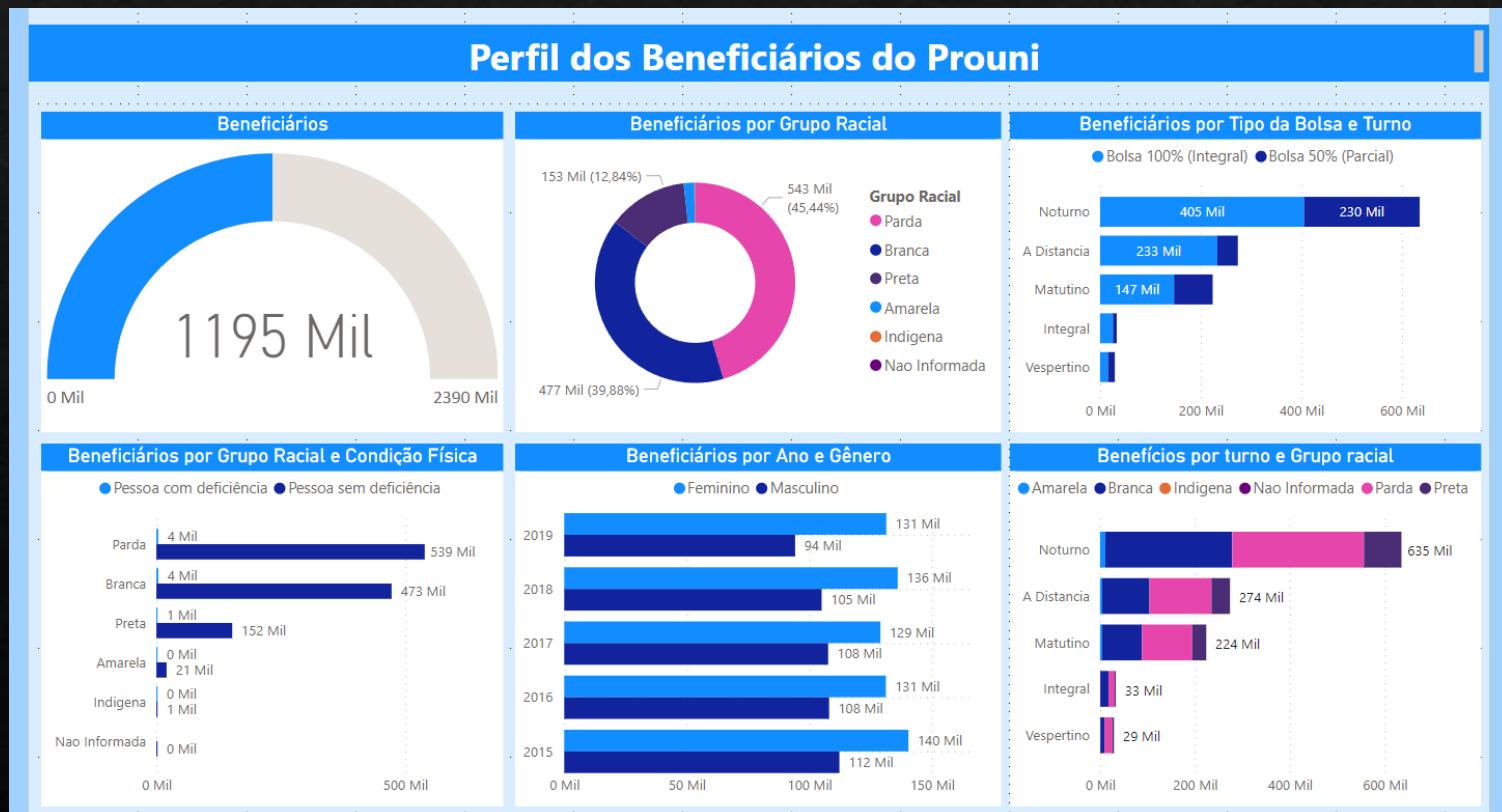
Aplicação OLAP

K- Consultas OLAP - Beneficiários por Turno e Grupo Racial



Aplicação OLAP

K– Página de Consultas OLAP – Perfil dos Beneficiários do Prouni



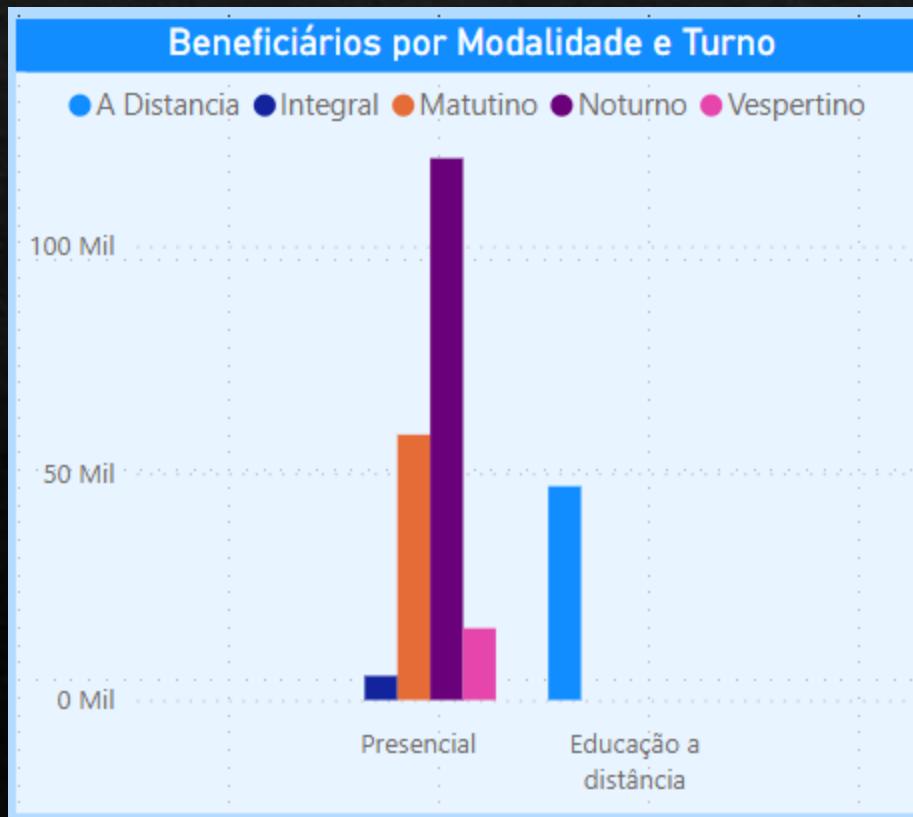
Aplicação OLAP

K– Consultas OLAP - Beneficiários por ano



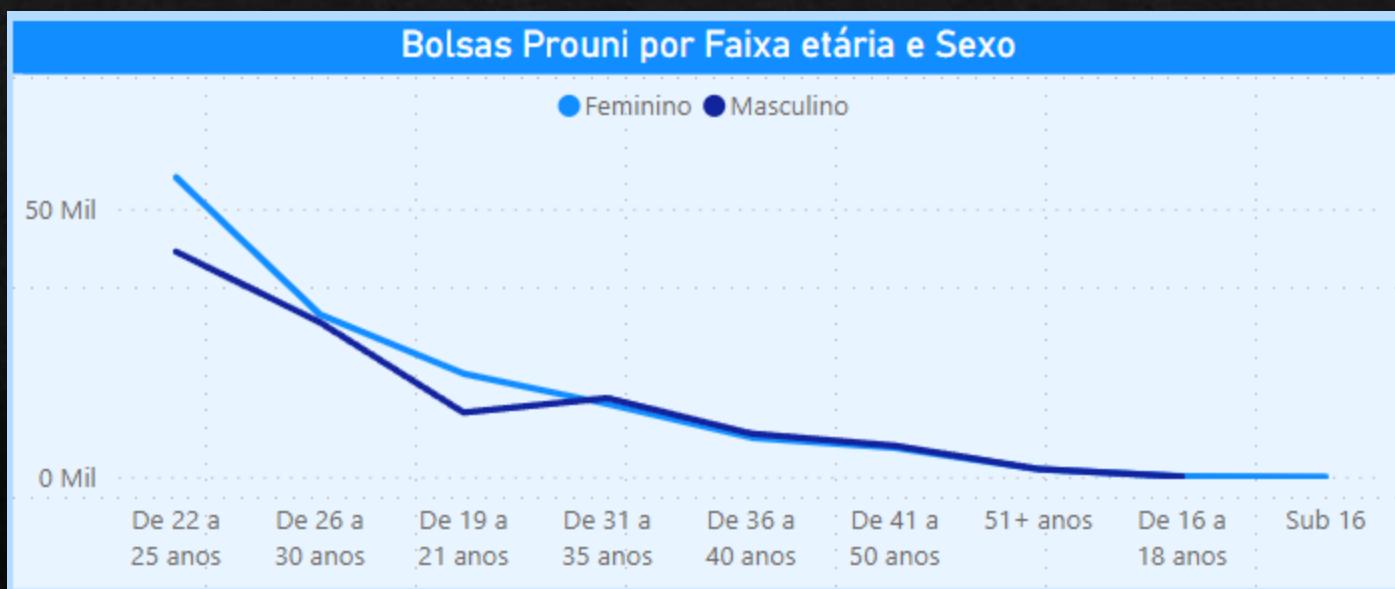
Aplicação OLAP

K- Consultas OLAP - Beneficiários por ano



Aplicação OLAP

K- Consultas OLAP - Beneficiários por ano



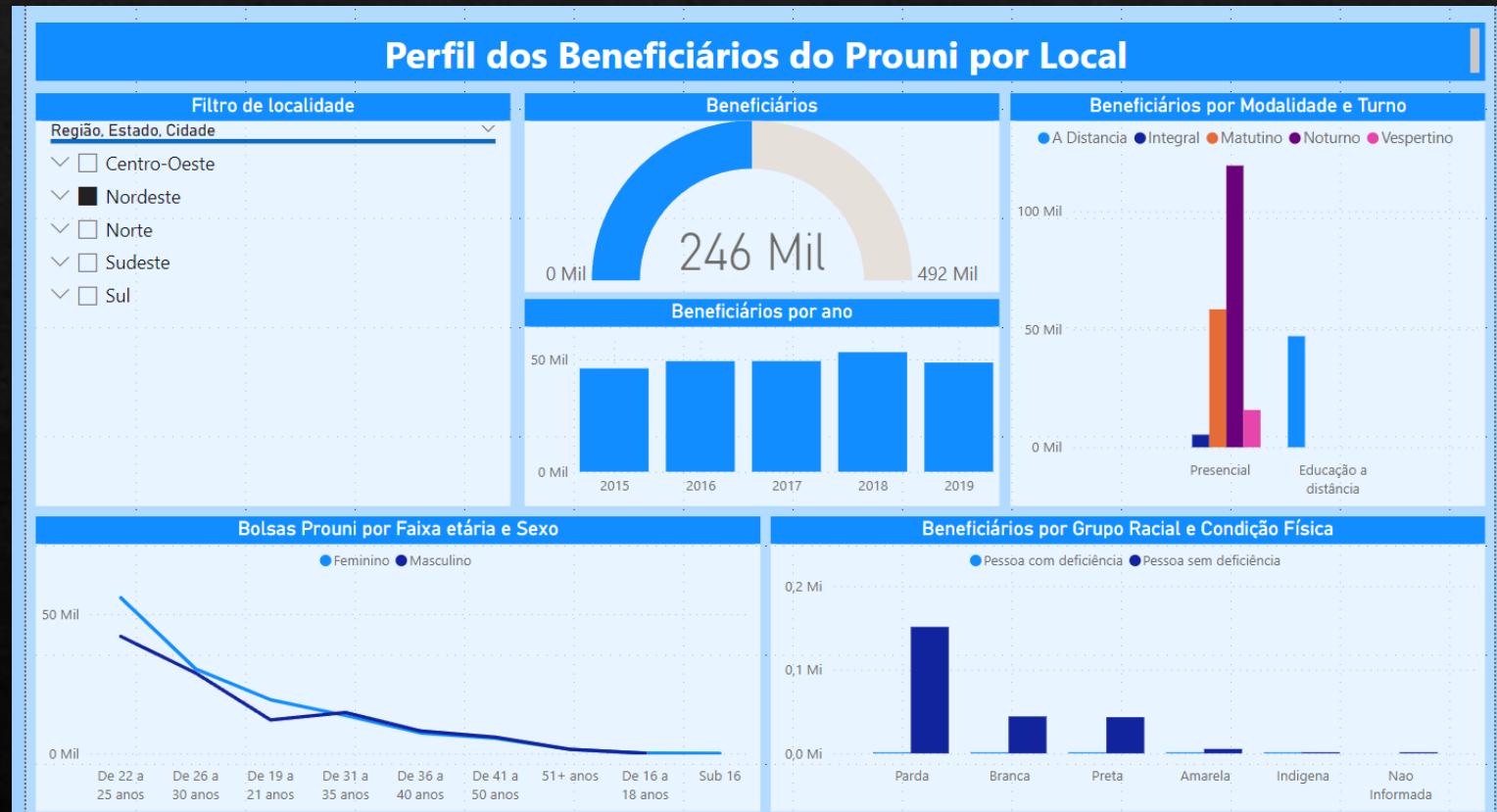
Aplicação OLAP

K- Consultas OLAP - Beneficiários por ano



Aplicação OLAP

K– Página de Consultas OLAP – Perfil dos Beneficiários Prouni por Local



Referências

Base de dados aberta - **Brasil Students Scholarship - Prouni - (2005-2019)**

<https://www.kaggle.com/datasets/lfarhat/brasil-students-scholarship-prouni-20052019>

Informações sobre o programa - **PROUNI**

<https://acessounico.mec.gov.br/prouni>