

Projeto Data Warehouse (BI) - projeto_dw

PUC Minas [Disciplina: Business Intelligence]

Este projeto visa o desenvolvimento de pipelines de extração, transformação e carga em Data Warehouse e o desenvolvimento de dashboards. O tema escolhido foi **Mobilidade Urbana** de Belo Horizonte, com enfoque em dados sobre **Acidentes de Trânsito** ocorridos nos perímetros da cidade.

- github.com/djmop/projeto_dw

Contribuidores

1. Daniel Nunes
2. João Dias
3. Maycon Felipe
4. Octávio Martins
5. Pedro Damasceno

Conjuntos de dados

Para o desenvolvimento do Data Warehouse, foram utilizados quatro (4) conjuntos de dados sobre acidentes de trânsito em Belo Horizonte, tendo sido adquiridos no [Portal Brasileiro de Dados Abertos](#). No momento desta redação, os dados compreendem o período de 2011 a 2021. Seguem os conjuntos:

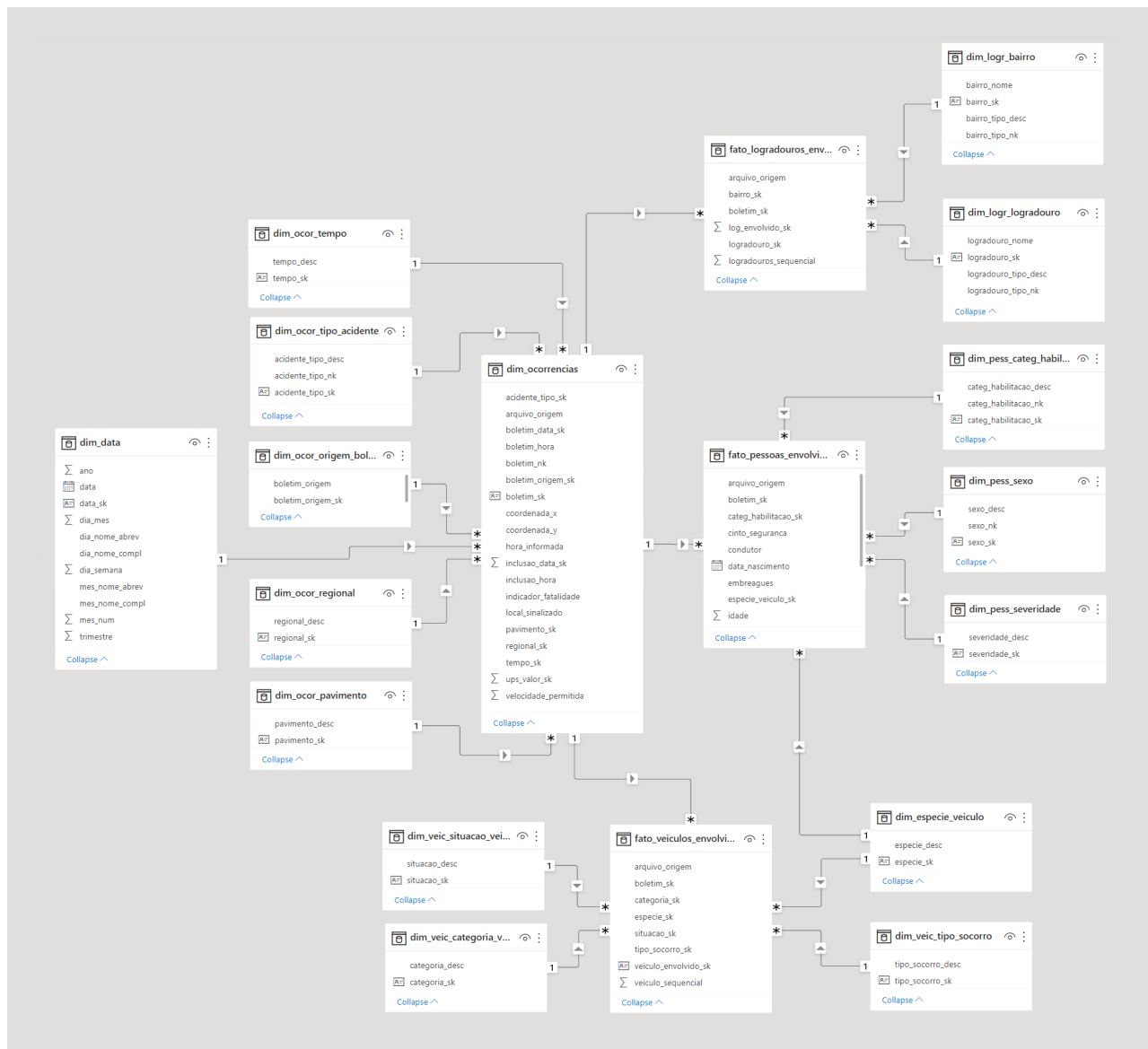
1. **Relação das pessoas envolvidas nos acidentes de trânsito com vítima** ([link](#))
2. **Relação dos logradouros dos locais de acidentes de trânsito com vítima** ([link](#))
3. **Relação dos veículos envolvidos nos acidentes de trânsito com vítima** ([link](#))
4. **Relação de ocorrências de acidentes de trânsito com vítima** ([link](#))

Data Warehouse

Os scripts de R contidos neste repositório realizam o download, extração e transformação dos dados. Na transformação, os dados são organizados em um esquema do tipo *snowflake*, contemplado pelas seguintes tabelas principais:

1. **Fato: Logradouros Envolvidos**
2. **Fato: Veículos Envolvidos**
3. **Fato: Pessoas Envolvidas**
4. **Fato/Dimensão: Ocorrências**

Modelo dimensional (*snowflake*)



Instalação

Configuração do projeto em R e RStudio

1. Instale base da [linguagem R](#) baixando-a do repositório CRAN. Recomenda-se a instalação direta em `c:/`. A versão utilizada neste projeto é a `R-4.2.1`.
2. Realize o download da IDE [RStudio](#) para configuração do projeto. Também recomenda-se a instalação direta em `c:/`.
3. Clone o repositório do projeto

```
git clone https://github.com/djmop/projeto_dw.git
```

4. Abra o projeto a partir do arquivo `projeto.dw.Rproj` situado na pasta raiz do projeto.

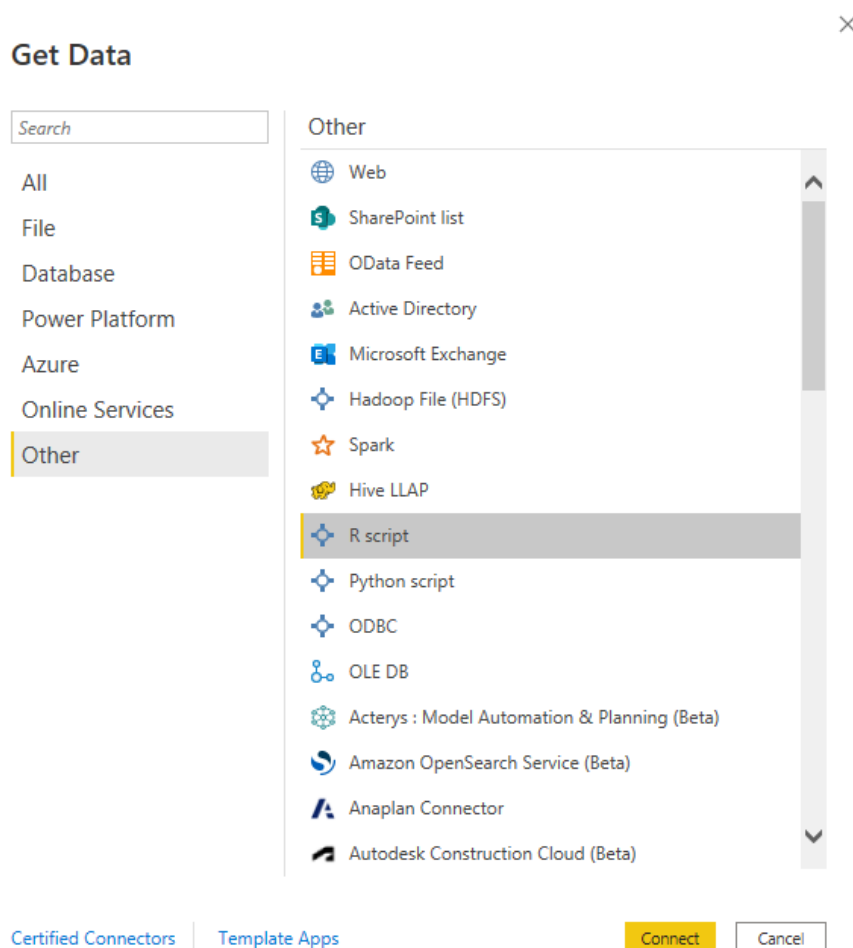
5. Instale o pacote `{renv}` via console do R e configure o ambiente do projeto.

```
# Instala {renv}
install.packages('renv')

# Configuração do ambiente
renv::restore()
```

Carregando dados no Power BI

1. Instale o [Microsoft Power BI](#).
2. Abra o **Power BI** e conecte a importação de dados via **R Script**.



*Selecione **R script** na importação de dados do **Power BI***

3. Insira o seguinte script no modal R script, alterando `<PATH_TO_PROJECT>` com o caminho do projeto em sua máquina. Clique em `OK` e aguarde o processamento. O log de processamento pode ser acompanhado a partir do arquivo situado em `log/global.log`.

```
# Parameters
project_root = '<PATH_TO_PROJECT>/projeto.dw/'

# Load main script
rscript = base::file.path(project_root, 'src/main.R')
source(rscript, encoding = 'UTF-8')

# Run
run_etl(project_root, download = TRUE, log = TRUE)
```

Exemplo:

R script

Script

```
# Parameters
project_root = 'E:/My drive/PUC Minas/6o_semestre/BUSINESS_INTEL/projeto.dw/'

# Load main script
rscript = base::file.path(project_root, 'src/main.R')
source(rscript, encoding = 'UTF-8')

# Run
run_etl(project_root, download = FALSE, log = FALSE)
```

The script will run with the following R installation C:\R\R-4.2.1.

To configure your settings and change which R installation you want to run, go to Options and settings.

OK

Cancel

*Insira o script no campo disponibilizado pela conexão **R script***