

Nome: Tiago Ferreira Saraiva

RGM: 02494649-4

Documentação:

Objetivo: Carregar um banco de dados, extrair, limpar e transformar as informações, carregando-as em um serviço de nuvem, criar uma conexão local com o serviço de nuvem para manipular as tabelas de dados de forma mais ágil, utilizar os dados para criar um painel interativo em Power BI que possa apresentar os dados de forma dinâmica, com informações acessíveis ao usuário.

O projeto envolve a extração dos dados, que estão em formato sqlite3 e a análise e transformação para uso após tratamento utilizando Jupyter Notebook. Com a transformação das tabelas é realizada a ingestão dos dados ao serviço de nuvem Google Cloud, na solução BigQuery.

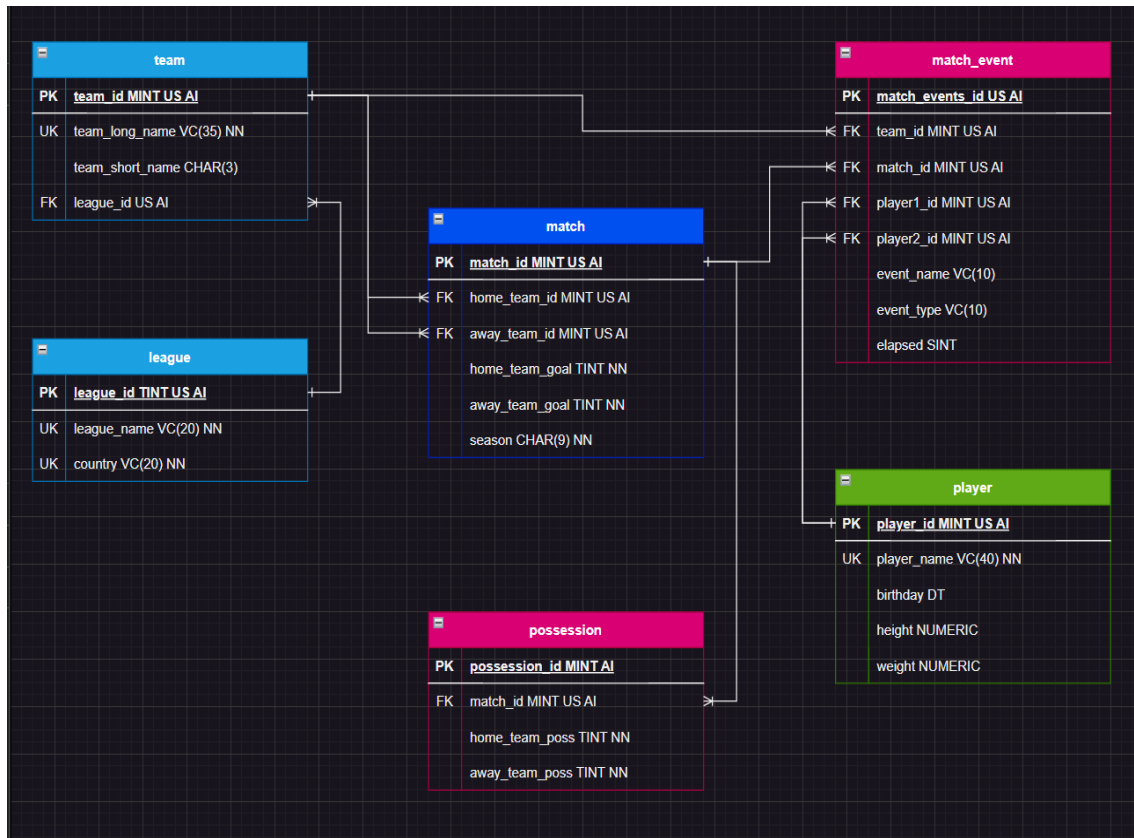
Após a ingestão dos dados é feito um acesso local para manipulação das tabelas utilizando SQL Server, conectando com o DBeaver -- qualquer outra ferramenta de preferência. A conexão permite que sejam realizados procedimentos nas tabelas para criação ou exclusão de tabelas, inclusão ou retirada de dados.

Com os dados tratados e disponíveis em serviço de nuvem é o utilizado o Power BI para criação de um dashboard interativo onde o usuário pode consultar informações sobre os dados, de forma acessível.

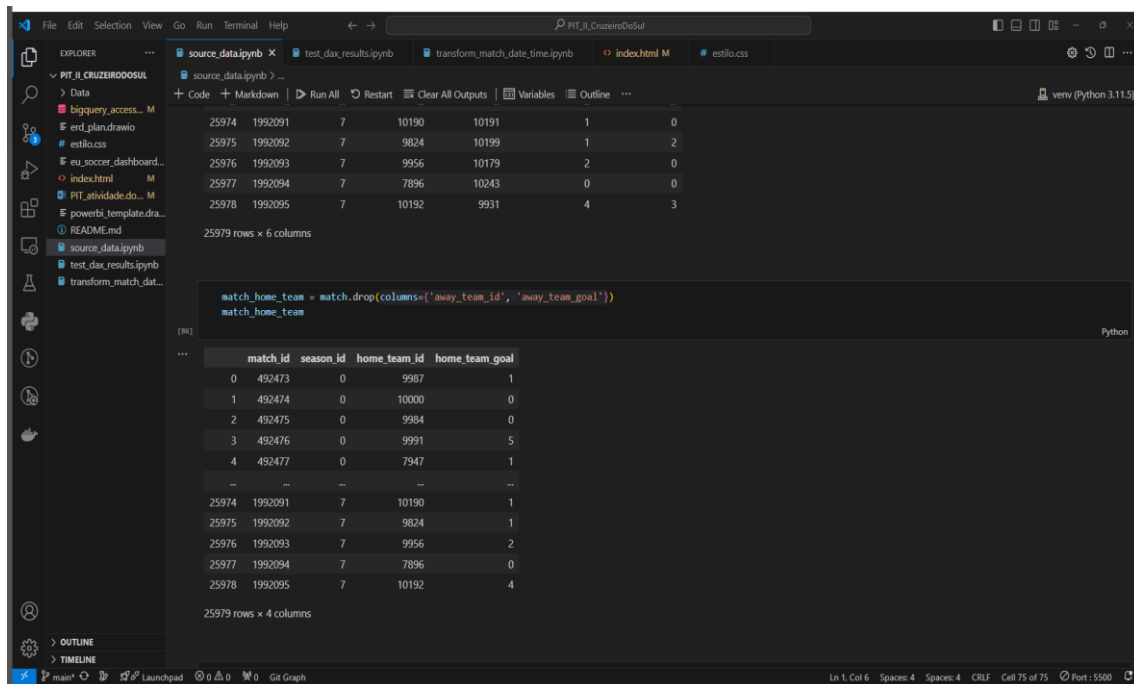
Os dados estão disponíveis na plataforma Kaggle, conforme site abaixo, e possuem alguns problemas, nem todas as ligas têm informações completas para utilização. As tabelas devem ficar conforme o ERD, anexo e disponível no Github.

Foram solucionados alguns problemas nos dados, também foram realizados alguns testes para utilização no Power BI, alguns eficazes, outros ineficazes. A construção do Power BI trouxe vários desafios e a projeção inicial dos painéis foi alterada devido limitações impostas.

Diagrama Entidade Relacionamento – ERD:



Extração, limpeza e transformação dos dados:



```

source_data.ipynb
test_dax_results.ipynb
transform_match_date_time.ipynb
index.html
estilo.css

source_data.ipynb
test_dax_results.ipynb
transform_match_date_time.ipynb

25974 1992091 7 10190 10191 1 0
25975 1992092 7 9824 10199 1 2
25976 1992093 7 9956 10179 2 0
25977 1992094 7 7896 10243 0 0
25978 1992095 7 10192 9931 4 3

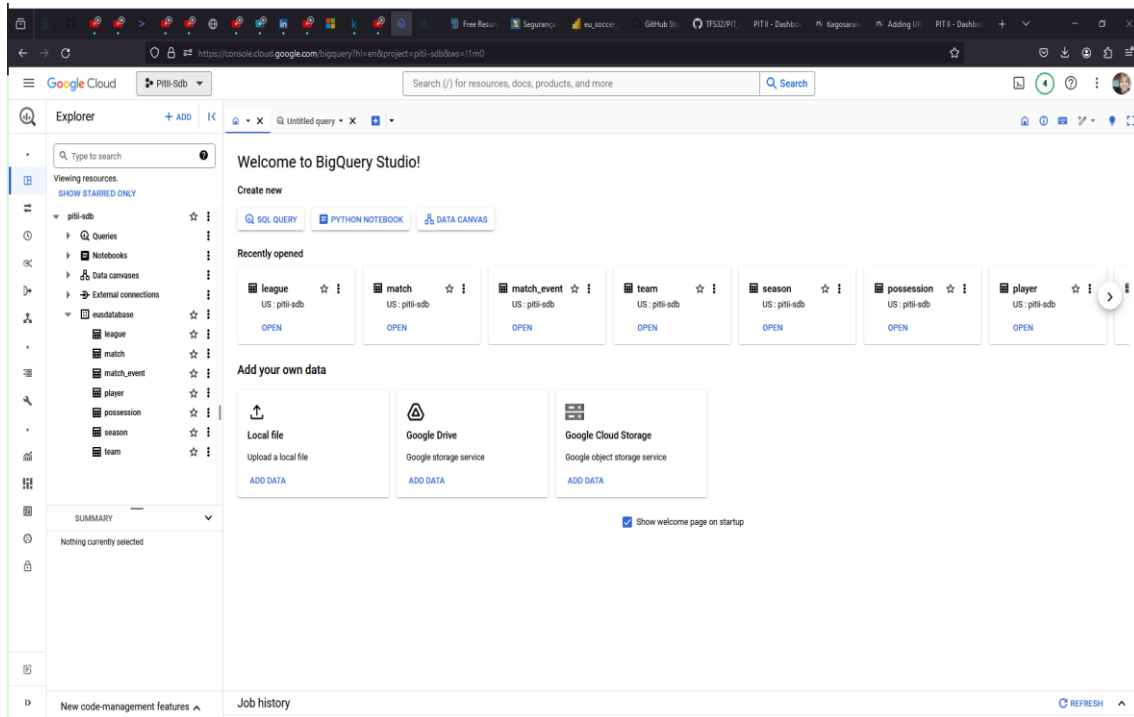
25979 rows x 6 columns

match_home_team = match.drop(columns=['away_team_id', 'away_team_goal'])
match_home_team

[86]
...
   match_id  season_id  home_team_id  home_team_goal
0    492473         0         9987             1
1    492474         0        10000             0
2    492475         0         9984             0
3    492476         0         9991             5
4    492477         0         7947             1
...      ...      ...      ...
25974 1992091         7        10190             1
25975 1992092         7         9824             1
25976 1992093         7         9956             2
25977 1992094         7         7896             0
25978 1992095         7        10192             4

25979 rows x 4 columns
    
```

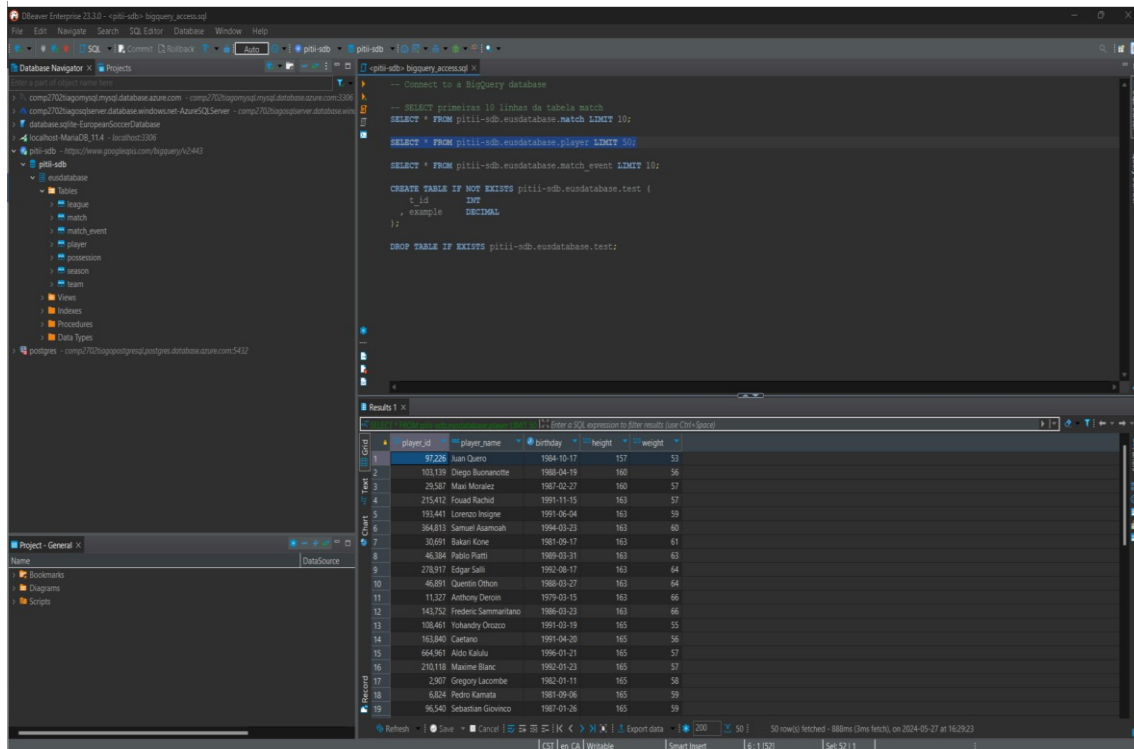
Google Cloud – BigQuery:



The screenshot shows the Google Cloud BigQuery Studio interface. The top navigation bar includes the Google Cloud logo, a search bar, and various icons. The main content area is divided into several sections:

- Explorer:** A sidebar on the left showing a tree view of resources. It includes sections for "Viewing resources" (SHOW STARRED ONLY), "Recently opened" (listing files like league, match, match_event, team, season, possession, player), and "Add your own data" (with options for Local file, Google Drive, and Google Cloud Storage).
- Welcome to BigQuery Studio!**: A central panel with a "Create new" section (SQL QUERY, PYTHON NOTEBOOK, DATA CANVAS) and a "Recently opened" section (listing files like league, match, match_event, team, season, possession, player).
- Job history:** A section at the bottom right showing a list of jobs.

Conexão framework local:



The screenshot shows the DBeaver Enterprise 23.3.0 interface. The left sidebar displays the "Database Navigator" with a tree view of databases. The main window shows a SQL query editor with the following code:

```
-- Connect to a BigQuery database
-- SELECT primeira Jo linha da tabela match
SELECT * FROM pitil-sdb.eudatabase.match LIMIT 10;

SELECT * FROM pitil-sdb.eudatabase.player LIMIT 50;

SELECT * FROM pitil-sdb.eudatabase.match_event LIMIT 10;

CREATE TABLE IF NOT EXISTS pitil-sdb.eudatabase.test (
  id INT,
  example DECIMAL
);

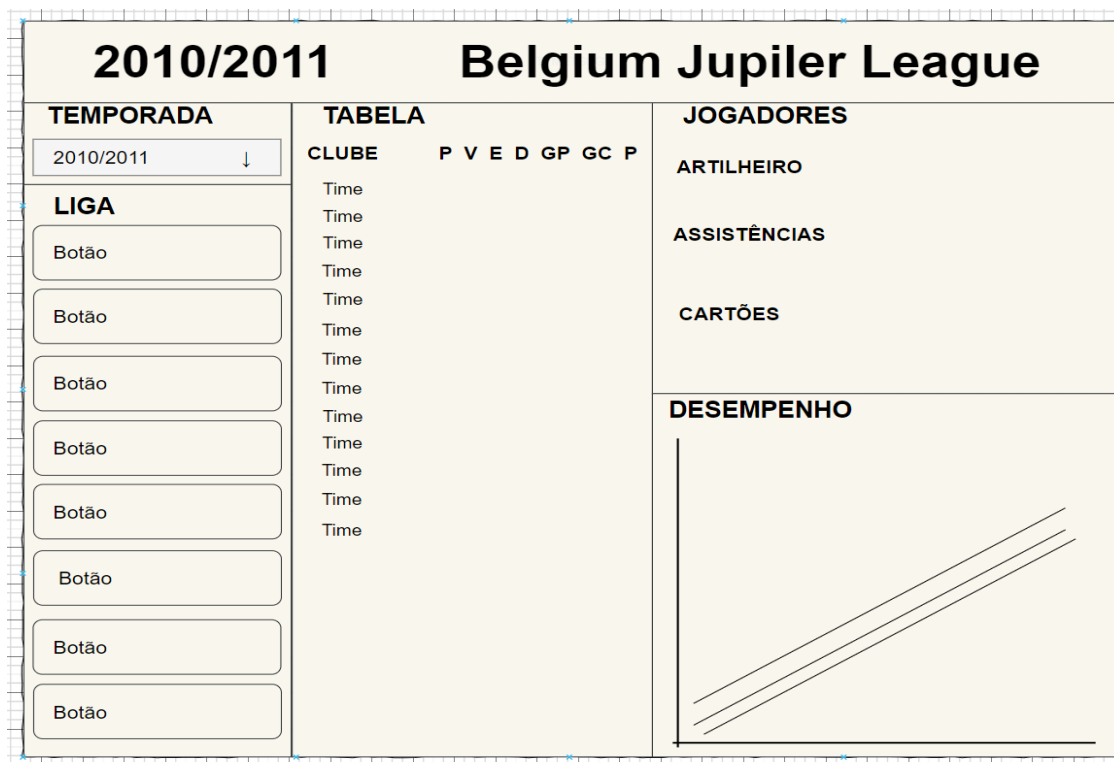
DROP TABLE IF EXISTS pitil-sdb.eudatabase.test;
```

The bottom panel shows the "Results" tab with a table of data. The table has columns: player_id, player_name, birthday, height, and weight. The data is as follows:

| player_id | player_name | birthday | height | weight |
|-----------|---------------------|------------|--------|--------|
| 97226 | Juan Quiso | 1984-10-17 | 157 | 53 |
| 103139 | Diego Buonanotte | 1988-04-19 | 160 | 56 |
| 29587 | Masi Morales | 1987-02-27 | 160 | 57 |
| 215412 | Fouad Rachid | 1991-11-15 | 163 | 57 |
| 193441 | Lorenzo Insigne | 1991-06-04 | 163 | 59 |
| 394813 | Serghei Aramocan | 1994-03-23 | 163 | 60 |
| 38691 | Brian Kane | 1981-09-17 | 163 | 61 |
| 46394 | Pablo Patti | 1989-03-31 | 163 | 63 |
| 278917 | Edgar Salli | 1992-08-17 | 163 | 64 |
| 46891 | Quentin Othou | 1988-03-27 | 163 | 64 |
| 11327 | Anthony Deroin | 1979-03-15 | 163 | 66 |
| 143752 | Frederic Sammartino | 1986-03-23 | 163 | 66 |
| 106461 | Kohandry Orozco | 1991-03-19 | 165 | 55 |
| 163340 | Cristiano | 1991-08-20 | 165 | 56 |
| 664961 | Aldo Kakulu | 1996-01-21 | 165 | 57 |
| 210118 | Maxime Blanc | 1992-01-23 | 165 | 57 |
| 2907 | Gregory Lacombe | 1982-01-11 | 165 | 58 |
| 6824 | Pedro Kamata | 1981-09-06 | 165 | 59 |
| 96540 | Sebastian Giovinco | 1987-01-26 | 165 | 59 |

Power BI:

Painel 1:



Painel 2:



Tabela de Construção (Desenvolvimento Ágil)

| ID | História do Usuário | Estimativa em Pontos | Prioridade |
|----|--|----------------------|------------|
| 0 | Escolha do banco de dados para uso no projeto | 4 | A |
| 1 | Análise dos dados componentes do banco, normalização e projeção do diagrama de entidade e relacionamento | 8 | B |
| 4 | Extração e transformação do banco de dados | 16 | A |
| 2 | Escolha da aplicação de nuvem e carregamento dos dados para nuvem | 4 | C |
| 3 | Conexão local para manipulação dos arquivos em nuvem | 2 | D |
| 5 | Painel de apresentação desenvolvido em Power BI, após recebimento dos dados tratados. | 20 | A |
| 6 | Criação de uma página da web para acoplar o painel criado em Power BI | 5 | C |

ID: 0

Título: Banco de dados

Requerente: Product Owner

| |
|--|
| Ação: Toda a equipe escolhe as fontes do banco de dados para ser utilizado no projeto. |
| Comentários: Observar quais serão as fontes, qual tipo de arquivo será recebido, como será o tratamento. |
| Critérios de aceitação: Confiabilidade e integridade. |
| Regras de negócio: |
| Requisito não funcional: |
| Prioridade: <input checked="" type="checkbox"/> A <input type="checkbox"/> B <input type="checkbox"/> C <input type="checkbox"/> D |
| Pontos de história: 4 |

| |
|---|
| ID: 1 |
| Título: Diagrama Entidade Relacionamento (ERD) |
| Requerente: Product Owner |
| Ação: Ambas equipes definem como será a construção do relacionamento entre as entidades para que seja realizada a transformação de acordo com os critérios de construção da visualização. |
| Comentários: Normalizar as tabelas, evitando redundância. |
| Critérios de aceitação: Design simples que atenda aos requisitos de construção do painel ao final. |
| Regras de negócio: |
| Requisito não funcional: |
| Prioridade: <input type="checkbox"/> A <input checked="" type="checkbox"/> B |

☐ C

☐ D

Pontos de história: 8

ID: 2

Título: Serviço de Nuvem

Requerente: Product Owner

Ação: A equipe de engenharia de dados escolhe o serviço de nuvem adequado para hospedar o banco de dados resultante do tratamento dos dados.

Comentários: Atenção para os critérios de segurança e replicabilidade.

Critérios de aceitação: Conformidade com o código de segurança e com o backup.

Regras de negócio:

Requisito não funcional:

Prioridade:

☐ A

☐ B

☒ C

☐ D

Pontos de história: 4

ID: 3

Título: Conexão Nuvem

Requerente: Product Owner

Ação: Equipe de engenharia de dados cria o fluxo de informação para que os dados em nuvem sejam manipulados diretamente em framework local.

Comentários:

Critérios de aceitação: Funcionamento.



| |
|--|
| Regras de negócio: |
| Requisito não funcional: |
| Prioridade: <input type="checkbox"/> A <input type="checkbox"/> B <input type="checkbox"/> C <input checked="" type="checkbox"/> D |
| Pontos de história: 4 |

| |
|---|
| ID: 4 |
| Título: ETL |
| Requerente: Product Owner |
| Ação: Equipe de engenharia de dados realiza a extração e transformação dos dados conforme os critérios estabelecidos, verificando a integridade das informações e tratando os problemas que o banco apresentou. |
| Comentários: |
| Critérios de aceitação: Garantir que os critérios do ERD sejam atendidos. |
| Regras de negócio: |
| Requisito não funcional: |
| Prioridade: <input checked="" type="checkbox"/> A <input type="checkbox"/> B <input type="checkbox"/> C <input type="checkbox"/> D |
| Pontos de história: 16 |

| |
|------------------|
| ID: 5 |
| Título: Power BI |

| |
|--|
| Requerente: Product Owner |
| Ação: A equipe de visualização desenha os painéis de Power BI conforme o projeto. |
| Comentários: Tentar atender aos critérios do planejamento. |
| Critérios de aceitação: Utilizar todas as informações que o banco possui. |
| Regras de negócio: |
| Requisito não funcional: |
| Prioridade: <input checked="" type="checkbox"/> A <input type="checkbox"/> B <input type="checkbox"/> C <input type="checkbox"/> D |
| Pontos de história: 20 |

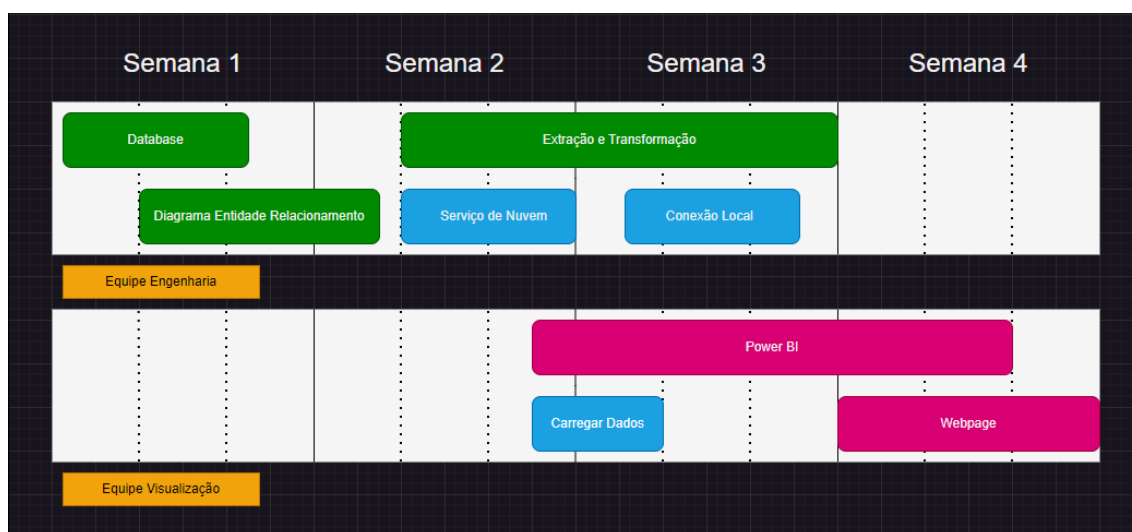
| |
|--|
| ID: 6 |
| Título: Website |
| Requerente: Product Owner |
| Ação: Equipe de visualização desenvolve um site para receber o Power BI facilitando o acesso dos usuários internos a painel. |
| Comentários: |
| Critérios de aceitação: |
| Regras de negócio: |
| Requisito não funcional: |
| Prioridade: <input type="checkbox"/> A <input type="checkbox"/> B <input checked="" type="checkbox"/> C <input type="checkbox"/> D |

Pontos de história: 5

Fluxo de apresentação



Fluxo de Trabalho



TESTES

1. Teste da limpeza dos dados.
2. Teste de conexão ao banco de dados com o acesso local.
3. Teste de dados carregados no Power BI
4. Teste do painel no Power BI
5. Teste de conexão ao site e acoplamento do Power BI.

1. Teste da limpeza dos dados.

Teste para verificar a integridade dos dados.

- 1.1. Tamanho das tabelas, verificando se linhas e colunas correspondem a situação real

Próximo passo é uma breve checagem nos dados, verificando integridade.

```
# Verificar se o tamanho das tabelas
print("Tables size\n"
      "League:", league.shape, "\n"
      "Team:", team.shape, "\n"
      "Player:", player.shape, "\n"
      "Match:", match.shape, "\n"
      "Possession:", possession_set.shape, "\n"
      "Match Event:", match_event.shape
      )
```

[27]

```
... Tables size
League: (11, 2)
Team: (299, 4)
Player: (11060, 5)
Match: (25979, 6)
Possession: (34815, 3)
Match Event: (883000, 7)
```

- 1.2. Verificar o tipo de dado de cada tabela



```
# Verificar qual o tipo de dados de cada tabela  
league.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 11 entries, 0 to 10  
Data columns (total 2 columns):  
#   Column      Non-Null Count  Dtype  
---  ---  
0   league_name  11 non-null    object  
1   country      11 non-null    object  
dtypes: object(2)  
memory usage: 308.0+ bytes
```

```
team.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
Index: 299 entries, 0 to 25815  
Data columns (total 4 columns):  
#   Column          Non-Null Count  Dtype  
---  ---  
0   team_id          299 non-null    int64  
1   team_long_name   299 non-null    object  
2   team_short_name  299 non-null    object  
3   league_id        299 non-null    int64  
dtypes: int64(2), object(2)  
memory usage: 11.7+ KB
```



```
player.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 11060 entries, 0 to 11059  
Data columns (total 5 columns):  
#   Column          Non-Null Count  Dtype  
---  ---  
0   player_id       11060 non-null  int64  
1   player_name     11060 non-null  object  
2   birthday        11060 non-null  object  
3   height          11060 non-null  int64  
4   weight          11060 non-null  int64  
dtypes: int64(3), object(2)  
memory usage: 432.2+ KB
```

```
match.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 25979 entries, 0 to 25978  
Data columns (total 6 columns):  
#   Column              Non-Null Count  Dtype  
---  ---  
0   match_id            25979 non-null  int64  
1   season              25979 non-null  object  
2   home_team_id        25979 non-null  int64  
3   away_team_id        25979 non-null  int64  
4   home_team_goal      25979 non-null  int64  
5   away_team_goal      25979 non-null  int64  
dtypes: int64(5), object(1)  
memory usage: 1.2+ MB
```

```
possession_set.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 34815 entries, 0 to 34814
Data columns (total 3 columns):
#   Column          Non-Null Count  Dtype
---  ---
0   away_team_poss   34796 non-null   object
1   home_team_poss   34796 non-null   object
2   match_id         34815 non-null   int64
dtypes: int64(1), object(2)
memory usage: 816.1+ KB
```

```
match_event.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 883000 entries, 0 to 882999
Data columns (total 7 columns):
#   Column          Non-Null Count  Dtype
---  ---
0   event_name       883000 non-null   object
1   event_type       718097 non-null   object
2   player1          857319 non-null   object
3   player2          207966 non-null   object
4   elapsed          883000 non-null   int64
5   team             881841 non-null   object
6   match_id         883000 non-null   int64
dtypes: int64(2), object(5)
memory usage: 47.2+ MB
```

1.3. Verificar a tabela 'possession' pois suas colunas deveriam ser int64

```
# A tabela "possession" deveria ter todas as colunas como integer, porém são strings
# Há linhas em branco, que não deveriam existir, serão tratados em outro arquivo
print(possession_set['home_team_poss'].isnull().sum().max())
print(possession_set['away_team_poss'].isnull().sum().max())
```

```
19
19
```



```
# Verificando se ainda existem dados em branco
print(possession['home_team_poss'].isnull().sum().max())
print(possession['away_team_poss'].isnull().sum().max())
```

```
0
0
```

```
# Visualizando como ficou a situação após a interpolação, nas linhas 186 e 31068
possession[(possession['match_id'] == 489170) | (possession['match_id'] == 1051814)]
```

| | away_team_poss | home_team_poss | match_id |
|-------|----------------|----------------|----------|
| 185 | 58.0 | 42.0 | 489170 |
| 186 | 55.5 | 44.5 | 489170 |
| 187 | 53.0 | 47.0 | 489170 |
| 188 | 51.0 | 49.0 | 489170 |
| 31067 | 32.0 | 68.0 | 1051814 |

1.4. Verificando ligas com dados faltantes em colunas da tabela 'match_event'



```
# Com o resultado verifica-se que há dados faltantes em quase todas as ligas  
sample['league_id'].unique()
```

```
array([ 1,  2,  3,  4,  5,  9, 10], dtype=int64)
```

league

| | league_name | country |
|----|--------------------------|-------------|
| 0 | Belgium Jupiler League | Belgium |
| 1 | England Premier League | England |
| 2 | France Ligue 1 | France |
| 3 | Germany 1. Bundesliga | Germany |
| 4 | Italy Serie A | Italy |
| 5 | Netherlands Eredivisie | Netherlands |
| 6 | Poland Ekstraklasa | Poland |
| 7 | Portugal Liga ZON Sagres | Portugal |
| 8 | Scotland Premier League | Scotland |
| 9 | Spain LIGA BBVA | Spain |
| 10 | Switzerland Super League | Switzerland |

Resultados do teste: Há inconsistência nos dados de eventos de partida, que não são passíveis de saneamento.

2. Teste de conexão ao banco de dados com o acesso local.

Teste para verificar se a conexão local do DBeaver junto ao Google Cloud.

2.1. Criação de uma tabela por meio de SQL query no DBeaver, replicando no Google Cloud

Google Cloud BigQuery Explorer

match

Schema

| Field name | Type | Mode | Key | Collation | Default Value | Policy Tags | Description |
|----------------|---------|----------|-----|-----------|---------------|-------------|-------------|
| match_id | INTEGER | NULLABLE | - | - | - | - | - |
| season | STRING | NULLABLE | - | - | - | - | - |
| home_team_id | INTEGER | NULLABLE | - | - | - | - | - |
| away_team_id | INTEGER | NULLABLE | - | - | - | - | - |
| home_team_goal | INTEGER | NULLABLE | - | - | - | - | - |
| away_team_goal | INTEGER | NULLABLE | - | - | - | - | - |

test

Schema

| Field name | Type | Mode | Key | Collation | Default Value | Policy Tags | Description |
|------------|---------|----------|-----|-----------|---------------|-------------|-------------|
| t_id | INTEGER | NULLABLE | - | - | - | - | - |
| example | NUMERIC | NULLABLE | - | - | - | - | - |

Resultados do teste: OK

3. Teste de dados carregados no Power BI

Verificar se os dados carregados no Power BI correspondem ao tamanho das tabelas criadas no Jupyter Notebook.

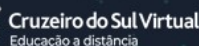
3.1. Comparar as tabelas mensuradas no Jupyter Notebook com as tabelas carregadas no Power BI

Próximo passo é uma breve checagem nos dados, verificando integridade.

```
# Verificar se o tamanho das tabelas
print("Tables size\n"
      "League:", league.shape, "\n"
      "Team:", team.shape, "\n"
      "Player:", player.shape, "\n"
      "Match:", match.shape, "\n"
      "Possession:", possession_set.shape, "\n"
      "Match Event:", match_event.shape
      )
```

[27]

```
... Tables size
League: (11, 2)
Team: (299, 4)
Player: (11060, 5)
Match: (25979, 6)
Possession: (34815, 3)
Match Event: (883000, 7)
```



Name

league

Mark as date table

Calendars

Manage relationships

Relationships

New measure

measure

Quick measure

measure column

New column

column

New table

table

Structure

×

✓

league_id

league_name

country

0

Belgium Jupiler League

Belgium

1

England Premier League

England

2

France Ligue 1

France

3

Germany 1. Bundesliga

Germany

4

Italy Serie A

Italy

5

Netherlands Eredivisie

Netherlands

6

Poland Ekstraklasa

Poland

7

Portugal Liga ZON Sagres

Portugal

8

Scotland Premier League

Scotland

9

Spain LIGA BBVA

Spain

10

Switzerland Super League

Switzerland

table: league (11 rows)

Name

match

Mark as date table

Calendars

Manage relationships

Relationships

New measure

measure

Quick measure

measure column

New column

column

New table

table

Structure

×

✓

match_id

season

home_team_id

away_team_id

home_team_goal

away_team_goal

selected_match_id

home_match_id

away_match_id

1979837

2015/2016

10001

9991

1

1

1979837

1979837

1979896

2015/2016

8635

8203

1

1

1979896

1979896

1979901

2015/2016

1773

274581

1

1

1979901

1979901

1979902

2015/2016

10001

8571

1

1

1979902

1979902

1979926

2015/2016

8571

8635

1

1

1979926

1979926

1979932

2015/2016

9986

8573

1

1

1979932

1979932

1979941

2015/2016

9994

9997

1

1

1979941

1979941

1979945

2015/2016

9985

8571

1

1

1979945

1979945

1979949

2015/2016

9994

8635

1

1

1979949

1979949

1979968

2015/2016

8635

8573

1

1

1979968

1979968

1979970

2015/2016

8203

9994

1

1

1979970

1979970

1979981

2015/2016

9994

10000

1

1

1979981

1979981

1980002

2015/2016

8203

9987

1

1

1980002

1980002

1980004

2015/2016

274581

9994

1

1

1980004

1980004

1980022

2015/2016

8571

10001

1

1

1980022

1980022

1980028

2015/2016

9986

8571

1

1

1980028

1980028

1980050

2015/2016

9991

1773

1

1

1980050

1980050

1979851

2015/2016

9997

8573

1

1

1979851

1979851

1979853

2015/2016

8635

9991

1

1

1979853

1979853

1980060

2015/2016

9986

9991

1

1

1980060

1980060

1980063

2015/2016

9997

9994

1

1

1980063

1980063

1979857

2015/2016

274581

9985

1

1

1979857

1979857

1979875

2015/2016

10000

9991

1

1

1979875

1979875

1979879

2015/2016

9997

9986

1

1

1979879

1979879

1979885

2015/2016

9986

8635

1

1

1979885

1979885

1988804

2015/2016

8650

8466

1

1

1988804

1988804

1988820

2015/2016

9825

8586

1

1

1988820

1988820

1988829

2015/2016

8654

8668

1

1

1988829

1988829

1988843

2015/2016

8197

10260

1

1

1988843

1988843

1988846

2015/2016

9850

9825

1

1

1988846

1988846

1988849

2015/2016

8654

8659

1

1

1988849

1988849

1988852

2015/2016

8668

9826

1

1

1988852

1988852

1988855

2015/2016

8466

10252

1

1

1988855

1988855

1988859

2015/2016

8659

8586

1

1

1988859

1988859

1988866

2015/2016

9850

8668

1

1

1988866

1988866

1988874

2015/2016

10261

10252

1

1

1988874

1988874

1988880

2015/2016

10252

8654

1

1

1988880

1988880

1988887

2015/2016

8668

8586

1

1

1988887

1988887

table: match (25,979 rows)



| | | | | | | |
|-----------------------------------|-------------|--------------------|----------------------|-------------|----------------------|-----------|
| Name | match_event | Mark as date table | Manage relationships | New measure | Quick measure column | New table |
| Structure | Calendars | Relationships | Calculations | | | |
| event_name | event_type | player1 | player2 | elapsed | team | match_id |
| cross | cross | | | 90 | 8654 | 489132 |
| cross | cross | | | 90 | 9825 | 489133 |
| cross | cross | | | 90 | 8528 | 489149 |
| cross | cross | | | 90 | 8559 | 489054 |
| cross | cross | | | 90 | 8586 | 489284 |
| cross | cross | | | 90 | 8658 | 658944 |
| cross | cross | | | 90 | 9879 | 658972 |
| cross | cross | | | 90 | 8528 | 659095 |
| cross | cross | | | 90 | 8655 | 659103 |
| cross | cross | | | 90 | 8462 | 658593 |
| cross | cross | | | 90 | 8667 | 659155 |
| cross | cross | | | 90 | 8472 | 659161 |
| cross | cross | | | 90 | 8472 | 659170 |
| cross | cross | | | 90 | 10252 | 659374 |
| cross | cross | | | 90 | 10252 | 659374 |
| cross | cross | | | 90 | 8658 | 658678 |
| cross | cross | | | 90 | 8650 | 839928 |
| cross | cross | | | 90 | 8472 | 839960 |
| cross | cross | | | 90 | 10261 | 839970 |
| cross | cross | | | 90 | 10261 | 839827 |
| cross | cross | | | 90 | 8658 | 840009 |
| cross | cross | | | 90 | 9825 | 840060 |
| cross | cross | | | 90 | 8654 | 840072 |
| cross | cross | | | 90 | 9879 | 840092 |
| cross | cross | | | 90 | 8559 | 840099 |
| cross | cross | | | 90 | 8559 | 840099 |
| cross | cross | | | 90 | 8602 | 840099 |
| cross | cross | | | 90 | 9879 | 840120 |
| cross | cross | | | 90 | 9879 | 840120 |
| cross | cross | | | 90 | 8655 | 840127 |
| cross | cross | | | 90 | 10194 | 840137 |
| cross | cross | | | 90 | 8528 | 840164 |
| cross | cross | | | 90 | 8658 | 840171 |
| cross | cross | | | 90 | 8472 | 840181 |
| cross | cross | | | 90 | 10252 | 840201 |
| cross | cross | | | 90 | 8586 | 840206 |
| cross | cross | | | 90 | 8455 | 840229 |
| cross | cross | | | 90 | 10261 | 839847 |
| Table: match_event (883,000 rows) | | | | | | |

| | | | | | | |
|-----------------------------|----------------------|--------------------|----------------------|-------------|----------------------|-----------|
| Name | player | Mark as date table | Manage relationships | New measure | Quick measure column | New table |
| Structure | Calendars | Relationships | Calculations | | | |
| player_id | player_name | birthday | height | weight | | |
| 352887 | Abdoulaye Doucoure | January 1, 1993 | 183 | 75 | | |
| 148827 | Abdoulwahid Sissoko | March 20, 1990 | 183 | 75 | | |
| 29904 | Abel Gomez | February 20, 1982 | 183 | 75 | | |
| 449241 | Adnan Januzaj | February 5, 1995 | 183 | 75 | | |
| 41399 | Ahmed Apimah Barusso | December 26, 1984 | 183 | 75 | | |
| 441166 | Ahmet Ozcan | March 7, 1995 | 183 | 75 | | |
| 32845 | Alan Martin | January 1, 1989 | 183 | 75 | | |
| 575364 | Alberto Grassi | March 7, 1995 | 183 | 75 | | |
| 41729 | Alessandro Parisi | April 15, 1977 | 183 | 75 | | |
| 37479 | Alexis | August 4, 1985 | 183 | 75 | | |
| 629594 | Alfonso Pedraza | April 9, 1996 | 183 | 75 | | |
| 25816 | Almen Abdi | October 21, 1986 | 183 | 75 | | |
| 183327 | Alvaro Gimenez | May 19, 1991 | 183 | 75 | | |
| 213489 | Alvaro Vazquez | April 27, 1991 | 183 | 75 | | |
| 668307 | Amadou Diawara | July 17, 1997 | 183 | 75 | | |
| 170316 | Amara Baby | February 23, 1989 | 183 | 75 | | |
| 612863 | Andre Anguissa | November 16, 1995 | 183 | 75 | | |
| 163839 | Andre Claro | March 31, 1991 | 183 | 75 | | |
| 41404 | Andrea Lazzari | December 3, 1984 | 183 | 75 | | |
| 33986 | Andres Palop | October 22, 1973 | 183 | 75 | | |
| 32419 | Andrew Considine | April 1, 1987 | 183 | 75 | | |
| 17312 | Anthar Yahia | March 21, 1982 | 183 | 75 | | |
| 148331 | Anthony Moris | April 29, 1990 | 183 | 75 | | |
| 39485 | Antonio Bocchetti | June 11, 1980 | 183 | 75 | | |
| 317932 | Baptiste Aloi | June 29, 1994 | 183 | 75 | | |
| 348408 | Birama Toure | June 6, 1992 | 183 | 75 | | |
| 40274 | Birkir Bjarnason | May 27, 1988 | 183 | 75 | | |
| 171046 | Bob Schepers | March 30, 1992 | 183 | 75 | | |
| 569038 | Bobo | March 20, 1982 | 183 | 75 | | |
| 33895 | Borja Neru | January 20, 1974 | 183 | 75 | | |
| 301237 | Brad McKay | March 26, 1993 | 183 | 75 | | |
| 23593 | Bradley Orr | November 1, 1982 | 183 | 75 | | |
| 6803 | Brecht Verbrugghe | April 29, 1982 | 183 | 75 | | |
| 575118 | Bruno Santos | February 7, 1993 | 183 | 75 | | |
| 46386 | Bruno Teles | May 1, 1986 | 183 | 75 | | |
| 110138 | Calvin Jong-a-Pin | July 18, 1986 | 183 | 75 | | |
| 210186 | Chistopher Mfuyi,25 | July 3, 1989 | 183 | 75 | | |
| 94219 | Clair Malonga | July 11, 1987 | 182 | 75 | | |
| Table: player (11,060 rows) | | | | | | |

Name

Mark as date table

Calendars

Manage relationships

Relationships

New measure

measure

Quick measure

measure

New column

Calculations

New table

Structure

away_team_poss

home_team_poss

match_id

| | | |
|----|----|--------|
| 48 | 52 | 489045 |
| 48 | 52 | 489046 |
| 48 | 52 | 489137 |
| 48 | 52 | 489139 |
| 48 | 52 | 489140 |
| 48 | 52 | 489151 |
| 48 | 52 | 489152 |
| 48 | 52 | 489153 |
| 48 | 52 | 489159 |
| 48 | 52 | 489160 |
| 48 | 52 | 489168 |
| 48 | 52 | 489168 |
| 48 | 52 | 489171 |
| 48 | 52 | 489188 |
| 48 | 52 | 489194 |
| 48 | 52 | 489198 |
| 48 | 52 | 489200 |
| 48 | 52 | 489201 |
| 48 | 52 | 489203 |
| 48 | 52 | 489214 |
| 48 | 52 | 489216 |
| 48 | 52 | 489235 |
| 48 | 52 | 489246 |
| 48 | 52 | 489252 |
| 48 | 52 | 489253 |
| 48 | 52 | 489255 |
| 48 | 52 | 489271 |
| 48 | 52 | 489273 |
| 48 | 52 | 489273 |
| 48 | 52 | 489277 |
| 48 | 52 | 489283 |
| 48 | 52 | 489291 |
| 48 | 52 | 489291 |
| 48 | 52 | 489295 |
| 48 | 52 | 489298 |
| 48 | 52 | 489301 |
| 48 | 52 | 489301 |
| 48 | 52 | 489303 |

Table: possession (34,815 rows)

Name

Mark as date table

Calendars

Manage relationships

Relationships

New measure

measure

Quick measure

measure

New column

Calculations

New table

Structure

team_id

team_long_name

team_short_name

league_id

| | | | |
|--------|----------------------|-----|---|
| 9987 | KRC Genk | GEN | 0 |
| 10000 | SV Zulte-Waregem | ZUL | 0 |
| 9984 | KSV Cercle Brugge | CEB | 0 |
| 9991 | KAA Gent | GEN | 0 |
| 7947 | FCV Dender EH | DEN | 0 |
| 8203 | KV Mechelen | MEC | 0 |
| 9999 | KSV Roeselare | ROS | 0 |
| 4049 | Tubize | TUB | 0 |
| 10001 | KVC Westerlo | WES | 0 |
| 8342 | Club Brugge KV | CLB | 0 |
| 9985 | Standard de Liège | STL | 0 |
| 9994 | Sporting Lokeren | LOK | 0 |
| 9996 | Royal Excel Mouscron | MOU | 0 |
| 9993 | Beerschot AC | BAC | 0 |
| 8635 | RSC Anderlecht | AND | 0 |
| 9998 | RAEC Mons | MON | 0 |
| 9986 | Sporting Charleroi | CHA | 0 |
| 8571 | KV Kortrijk | KOR | 0 |
| 9997 | Sint-Truidense VV | STT | 0 |
| 9989 | Lierse SK | LIE | 0 |
| 6351 | KAS Eupen | EUP | 0 |
| 1773 | Oud-Heverlee Leuven | O-H | 0 |
| 8475 | Waasland-Beveren | WAA | 0 |
| 8573 | KV Oostende | OOS | 0 |
| 274581 | Royal Excel Mouscron | MOP | 0 |
| 10260 | Manchester United | MUN | 1 |
| 9825 | Arsenal | ARS | 1 |
| 8472 | Sunderland | SUN | 1 |
| 8654 | West Ham United | WHU | 1 |
| 10252 | Aston Villa | AVL | 1 |
| 8668 | Everton | EVE | 1 |
| 8549 | Middlesbrough | MID | 1 |
| 8559 | Bolton Wanderers | BOL | 1 |
| 8667 | Hull City | HUL | 1 |
| 8455 | Chelsea | CHE | 1 |
| 10194 | Stoke City | STK | 1 |
| 8650 | Liverpool | LIV | 1 |
| 10261 | Newcastle United | NEW | 1 |

Table: team (299 rows)

Resultados dos testes: OK

4. Teste do painel no Power BI

Verificar se os slicers e measures utilizados no Power BI correspondem ao comportamento desejado.

4.1. Verificar o slicer de ligas:

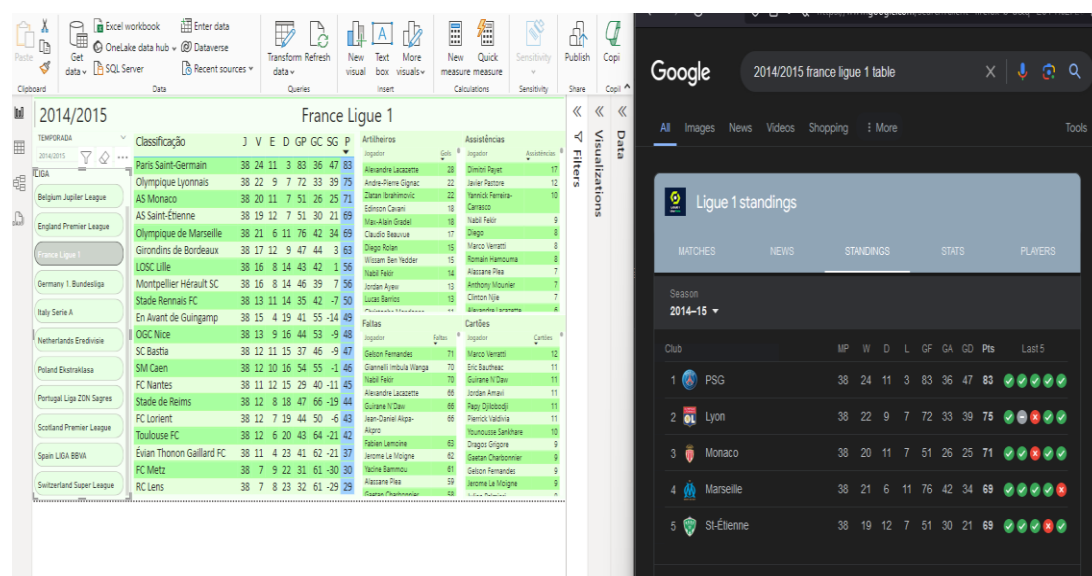


TEMPORADA: 2015/2016

LIGA: Belgium Jupiler League, England Premier League, France Ligue 1, Germany 1. Bundesliga, Italy Serie A, Netherlands Eredivisie, Poland Ekstraklasa, Portugal Liga ZON Sagres, Scotland Premier League, Spain LIGA BBVA, Switzerland Super League

| Classificação | J | V | E | D | GP | GC | SG | P |
|------------------------|----|----|----|----|-----|----|-----|----|
| Paris Saint-Germain | 38 | 30 | 6 | 2 | 102 | 19 | 83 | 96 |
| AS Monaco | 38 | 17 | 14 | 7 | 57 | 50 | 7 | 65 |
| Olympique Lyonnais | 38 | 19 | 8 | 11 | 67 | 43 | 24 | 65 |
| OGC Nice | 38 | 18 | 9 | 11 | 58 | 41 | 17 | 63 |
| LOSC Lille | 38 | 15 | 15 | 8 | 39 | 27 | 12 | 60 |
| AS Saint-Étienne | 38 | 17 | 7 | 14 | 42 | 37 | 5 | 58 |
| SM Caen | 38 | 16 | 6 | 16 | 39 | 52 | -13 | 54 |
| Stade Rennais FC | 38 | 13 | 13 | 12 | 52 | 54 | -2 | 52 |
| Angers SCO | 38 | 13 | 11 | 14 | 40 | 38 | 2 | 50 |
| Girondins de Bordeaux | 38 | 12 | 14 | 12 | 50 | 57 | -7 | 50 |
| SC Bastia | 38 | 14 | 8 | 16 | 36 | 42 | -6 | 50 |
| Montpellier Hérault SC | 38 | 14 | 7 | 17 | 49 | 47 | 2 | 49 |
| FC Nantes | 38 | 12 | 12 | 14 | 33 | 44 | -11 | 48 |
| Olympique de Marseille | 38 | 10 | 18 | 10 | 48 | 42 | 6 | 48 |
| FC Lorient | 38 | 11 | 13 | 14 | 47 | 58 | -11 | 46 |
| En Avant de Guingamp | 38 | 11 | 11 | 16 | 47 | 56 | -9 | 44 |
| Toulouse FC | 38 | 9 | 13 | 16 | 45 | 55 | -10 | 40 |
| Stade de Reims | 38 | 10 | 9 | 19 | 44 | 57 | -13 | 39 |
| GFC Ajaccio | 38 | 8 | 13 | 17 | 37 | 58 | -21 | 37 |
| ES Troyes AC | 38 | 3 | 9 | 26 | 28 | 83 | -55 | 18 |

4.2. Verificar o slicer de temporada:

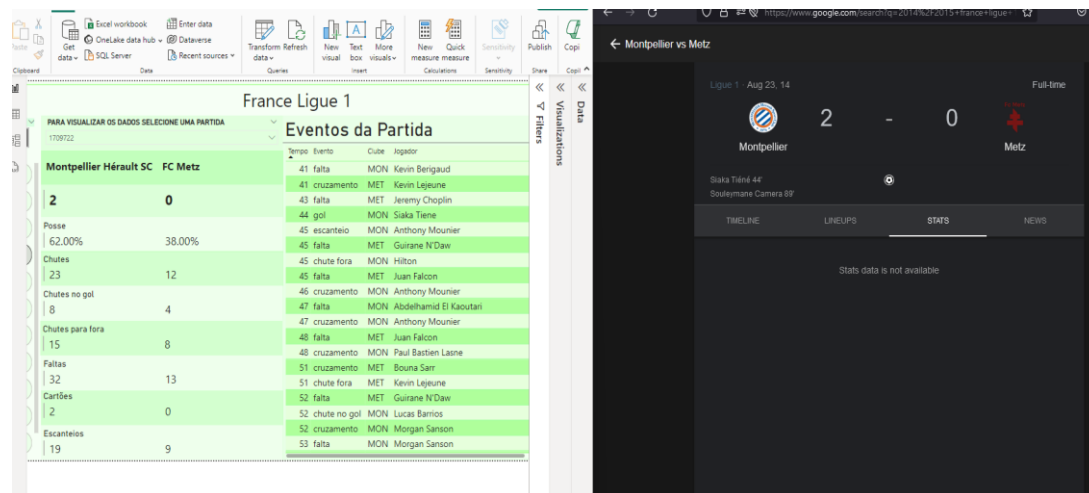


TEMPORADA: 2014/2015

LIGA: Belgium Jupiler League, England Premier League, France Ligue 1, Germany 1. Bundesliga, Italy Serie A, Netherlands Eredivisie, Poland Ekstraklasa, Portugal Liga ZON Sagres, Scotland Premier League, Spain LIGA BBVA, Switzerland Super League

| Classificação | J | V | E | D | GP | GC | SG | P |
|--------------------------|----|----|----|----|----|----|-----|----|
| Paris Saint-Germain | 38 | 24 | 11 | 3 | 83 | 36 | 47 | 83 |
| Olympique Lyonnais | 38 | 22 | 9 | 7 | 72 | 33 | 39 | 75 |
| AS Monaco | 38 | 20 | 11 | 7 | 51 | 26 | 25 | 71 |
| AS Saint-Étienne | 38 | 19 | 12 | 7 | 51 | 30 | 21 | 69 |
| Olympique de Marseille | 38 | 12 | 6 | 11 | 76 | 42 | 34 | 69 |
| Girondins de Bordeaux | 38 | 17 | 12 | 9 | 47 | 44 | 3 | 63 |
| LOSC Lille | 38 | 16 | 8 | 14 | 43 | 42 | 1 | 56 |
| Montpellier Hérault SC | 38 | 16 | 8 | 14 | 46 | 39 | 7 | 56 |
| Stade Rennais FC | 38 | 13 | 11 | 14 | 35 | 42 | -7 | 50 |
| En Avant de Guingamp | 38 | 15 | 4 | 19 | 41 | 55 | -14 | 49 |
| OGC Nice | 38 | 13 | 9 | 16 | 44 | 53 | -9 | 48 |
| SC Bastia | 38 | 12 | 11 | 15 | 37 | 46 | -9 | 47 |
| SM Caen | 38 | 12 | 10 | 16 | 54 | 55 | -1 | 46 |
| FC Nantes | 38 | 11 | 12 | 15 | 29 | 40 | -11 | 45 |
| Stade de Reims | 38 | 12 | 8 | 18 | 47 | 66 | -19 | 44 |
| FC Lorient | 38 | 12 | 7 | 19 | 44 | 50 | -6 | 43 |
| Toulouse FC | 38 | 12 | 6 | 20 | 43 | 64 | -21 | 42 |
| Évian Thonon Gaillard FC | 38 | 11 | 4 | 23 | 41 | 62 | -21 | 37 |
| FC Metz | 38 | 7 | 9 | 22 | 31 | 61 | -30 | 30 |
| RC Lens | 38 | 7 | 8 | 23 | 32 | 61 | -29 | 29 |

4.3. Verificar o slicer de partida no painel clube:



Resultados dos testes: slicer de temporada só funcionou em um dos painéis e foi sincronizado. Slicer de clube não funcionou e foi excluído. Não foi possível criar string para substituir o id de partida no slicer de partida.


5. Teste de conexão e acoplamento do Power BI

Verificar se o DNS do site está funcionando, verificar se o Power BI dentro do site está funcionando.

5.1. Certificar que o site está com o DNS corretamente configurado:

[→](#)
[↺](#)


<https://www.whatsmydns.net/#CNAME/www.tiagoaraiva.works>

























whatsmydns.net
Global DNS Propagation Checker

BEST BUY

Ultimate PC Gaming Sale



22%
Lenovo Legion...
\$2,699.99

CNAME
 

 Search

| | | |
|---|--|-----------------------------------|
|  | San Jose CA, United States Corporate West | tfs32.github.io ✓ |
|  | Holtsville NY, United States OpenDNS | tfs32.github.io ✓ |
|  | Miami FL, United States AT&T | tfs32.github.io ✓ |
|  | Reston VA, United States Sprint | tfs32.github.io ✓ |
|  | New York NY, United States Speakeasy | tfs32.github.io ✓ |
|  | Kelowna, Canada Shaw | tfs32.github.io ✓ |
|  | Culiacan, Mexico Megacable | tfs32.github.io ✓ |
|  | Santa Cruz do Sul, Brazil Claro | tfs32.github.io ✓ |
|  | Paterna de Rivera, Spain ServHosting | tfs32.github.io ✓ |
|  | Manchester, United Kingdom Ancor B | tfs32.github.io ✓ |
|  | Paris, France France Telecom | tfs32.github.io ✓ |
|  | Diemen, Netherlands Tele2 Nederland | tfs32.github.io ✓ |
|  | Oberhausen, Germany Deutsche Telekom | tfs32.github.io ✓ |
|  | Zizers, Switzerland Oskar Emmenegger | tfs32.github.io ✓ |
|  | Sassuolo, Italy Telecom Italia | tfs32.github.io ✓ |
|  | Cullinan, South Africa Liquid | tfs32.github.io ✓ |
|  | Antalya, Turkey Tsknet Yastim | tfs32.github.io ✓ |
|  | Saint Petersburg, Russia Prometei | tfs32.github.io ✓ |
|  | Rawalpindi, Pakistan CMPak | tfs32.github.io ✓ |

DNS Propagation Checker

whatsmydns.net lets you instantly perform a DNS lookup to check a domain name's current IP address and DNS record information against multiple nameservers located in different parts of the world.



5.2. Verificar o Power BI dentro do site:

Dashboard

Para acessar utilize seu e-mail institucional da Universidade Cruzeiro do Sul.
nome.sobrenome@cs.cruzeirodosul.edu.br

[illegible]

Resultados dos testes: OK

Codificação:

Na Tabela a seguir insira as informações referentes ao desenvolvimento do código do processo.

| | |
|---|---|
| Linguagem | Python, SQL, DAX |
| Banco de Dados | https://www.kaggle.com/datasets/hugomathien/soccer/code?datasetId=63 |
| Hospedagem | Git Hub |
| Plataforma | Power BI |
| Modo de Codificação | (X) Tradicional () Low-code |
| Link do repositório no GitHub com os códigos abertos | https://github.com/TFS32/PIT_II_CruzeiroDoSul |
| Link da solução em funcionamento | https://tiagosaraiva.works/ |
| Link do vídeo narrado (no mínimo 5 min) | https://www.youtube.com/watch?v=7F1IT4sS_Lw |

Testes da Solução

Escolha 5 colegas para testar sua aplicação, disponibilize o *link* de acesso ou os recursos necessários para que testem como usuários. Preencha a Tabela a seguir com as informações obtidas:

| | |
|---|--------------------------------|
| Nome: Diego Goncalves Silva | Data do teste: 27/05/24 |
| O que testou e funcionou: Todas as funcionalidades, botões das listas. | |
| O que testou e não funcionou – O que deve ser corrigido: N/A | |
| Funcionalidade não testada (faltou ou não foi implementada): N/A | |

| | |
|---|----------------------------------|
| Nome: Rodrigo Waltmann Nascimento | Data do teste: 28/05/2024 |
| O que testou e funcionou: - | |
| O que testou e não funcionou – O que deve ser corrigido: relatou que não conseguiu acessar o site. | |
| Funcionalidade não testada (faltou ou não foi implementada): - | |

| | |
|---|----------------------------------|
| Nome: Rafael da Silva Camargo | Data do teste: 28/05/2024 |
| O que testou e funcionou: acessei o site, acessei Power BI, fiz seleções usando as opções de botões, ok. | |
| O que testou e não funcionou – O que deve ser corrigido: os dados de time fora de casa apresentam os mesmos dados de time da casa. Painel do Power BI tem delay. Os dados de gol ou cartões apresentam as vezes informações do jogador em branco. O botão de selecionar temporada no painel do clube não funciona. | |
| Funcionalidade não testada (faltou ou não foi implementada): | |

| | |
|--|---------------------------|
| Nome: Giovanna Albuquerque Riello | Data do teste: 28/05/2024 |
| O que testou e funcionou: site funcionando, Power BI funcionando, é possível fazer seleções. | |
| O que testou e não funcionou – O que deve ser corrigido: o site é muito simples não tem aparência legal. | |
| Funcionalidade não testada (faltou ou não foi implementada): não há identidade visual no projeto. | |

| | |
|--|---------------------------|
| Nome: Fernando Alves de Oliveira | Data do teste: 30/05/2024 |
| O que testou e funcionou: site ok, acesso ao Power BI ok. | |
| O que testou e não funcionou – O que deve ser corrigido: as vezes aparecem dados em branco nas tabelas com nomes de jogadores. Quando o dado de ocorrências na partida é zero, aparece como (blank). | |
| Funcionalidade não testada (faltou ou não foi implementada): | |

Vídeo da Solução atualizada

Após levantar os *feedbacks* e executar as correções necessárias e pertinentes, grave um vídeo de **até 5 minutos** apresentando as modificações realizadas no sistema.

| | |
|-------------------|---|
| Link para o vídeo | https://youtu.be/ei_MSyh7IY |
|-------------------|---|