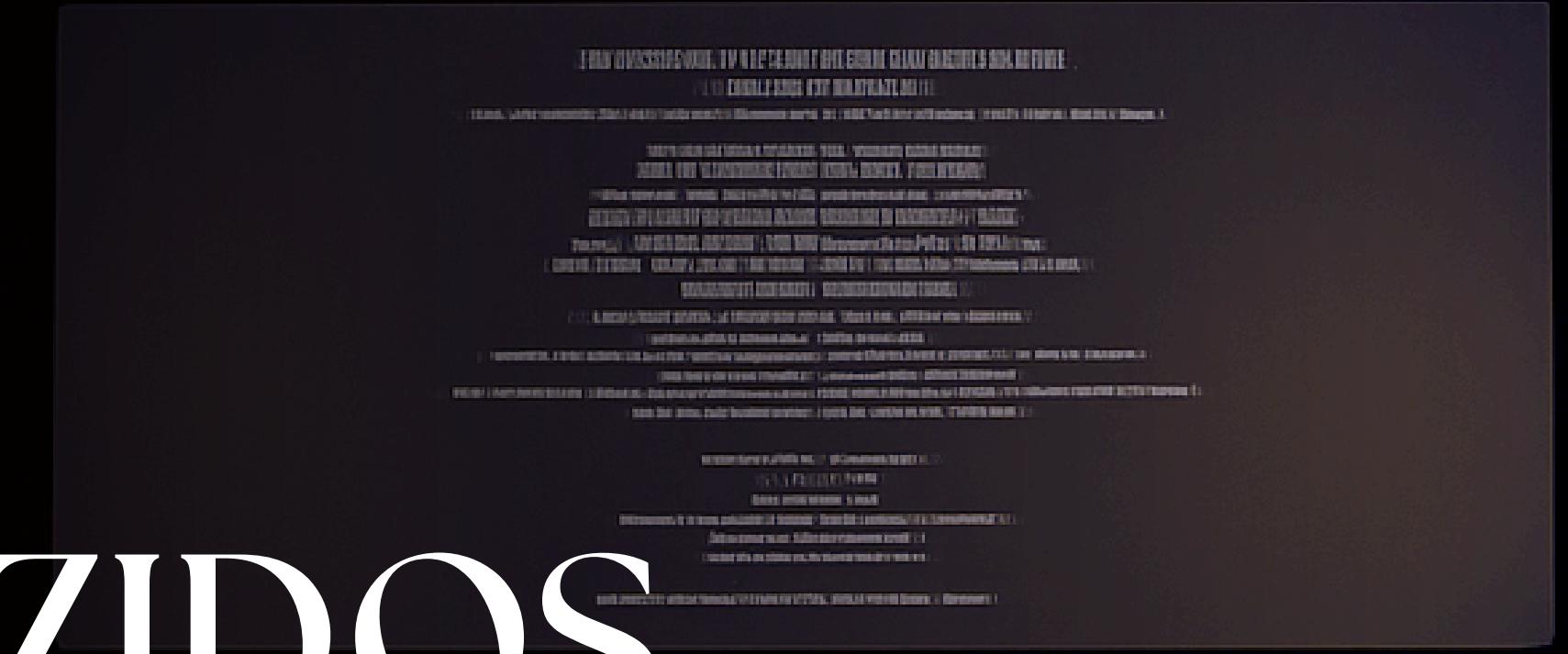


# FILMES PRODUZIDOS NOS ÚLTIMOS 150 ANOS



# O DATASET UTILIZADO

**Letterboxd (Movies Dataset)**

Data Card    Code (5)    Discussion (1)    Suggestions (0)

**Letterboxd**



Data obtained using a program from the site [letterboxd.com](http://letterboxd.com).



950 mil filmes com os dados sobre:

- descrição
- cartazes
- elencos
- equipe técnica
- datas de lançamento

# VISÃO COMERCIAL



Intenção de promover o mercado cinematográfico, com uma análise do impacto que o público tem na produção dos filmes, atingindo:

- Studios de Produção de filmes
- Aplicativos de Streaming

# TRATAMENTO DOS DADOS

Parameters

Name	Description
authorization	string   (string   null) (query) Bearer aluno123

Request body required

Edit Value | Schema

```
{  
    "input_dir": "C:/Users/thiag/Downloads/countries",  
    "pattern": "*.csv",  
    "sep": ";",  
    "encoding": "utf-8-sig",  
    "chunksize": 100000,  
    "enforce_cols": [  
        "id",  
        "country",  
        "genre",  
        "actor_name",  
        "actor_role",  
        "crew_name",  
        "crew_role",  
        "language",  
        "language_type",  
        "link",  
        "releases_country",  
        "releases_date",  
        "releases_type",  
        "releases_rating",  
        "theme",  
        "studio",  
        ""]  
}
```

Execute

Responses

Curl

```
curl -X 'POST' \  
  'http://127.0.0.1:8000/api/v1/process?authorization=Bearer%20aluno123' \  
  -H 'accept: application/json' \  
  -H 'Content-Type: application/json' \  
  -d '{  
    "input_dir": "C:/Users/thiag/Downloads/countries",  
    "pattern": "*.csv",  
    "sep": ";",  
    "encoding": "utf-8-sig",  
    "chunksize": 100000,  
    "enforce_cols": [  
        "id",  
        "country",  
        "genre",  
        "actor_name",  
        "actor_role",  
        "crew_name",  
        "crew_role",  
        "language",  
        "language_type",  
        "link",  
        "releases_country",  
        "releases_date",  
        "releases_type",  
        "releases_rating",  
        "theme",  
        "studio",  
        ""]  
  }'
```

Request URL

```
http://127.0.0.1:8000/api/v1/process?authorization=Bearer%20aluno123
```

Server response

Code Details

Utilizamos API em com python para o tratamento dos dados

# BIG QUERY

Consulta sem título

Executar Salvar Fazer o download Compartilhar :

```
1 SELECT
2   studio.id_filmes, studio.filmes,
3   releases.country_filmes, date_filmes, type_filmes, rating_filmes,
4   language.language, language_type,
5   genero.genero_filmes,
6   actor.coutrie_filmes,
7
8 FROM
9   applied-radar-474411-h3.bigquery.studio AS studio
10 JOIN
```

Consulta concluída

Resultados da consulta

Salvar resultados Abrir em

Linha	id_filmes	studio_filmes	country_filmes	date_filmes	type_filmes
1	1163241	Media Castle	China	2019-10-19	Theatrical
2	1163241	Media Castle	South Korea	2019-08-14	Theatrical
3	1072144	Media Complex Japan	Germany	2011-08-26	Theatrical

Criamos uma tabela com todos as pastas separadas dos dados.

# QUERYs

```
SELECT
    genero_filmes,
    AVG(rating_filmes) AS media_avaliacao
FROM
    `applied-radar-474411-h3.bigquery.resultadofinal`
GROUP BY
    genero_filmes
ORDER BY
    media_avaliacao DESC;
```

Calcula a média de avaliação de filmes por gênero.

# QUERYs

```
WITH filmes_por_ano AS (
    SELECT
        date_filmes,
        COUNT(*) AS total_filmes
    FROM
        `applied-radar-474411-h3.bigquery.resultadofinal`
    GROUP BY
        ano
)
SELECT
    ano,
    total_filmes,
    LAG(total_filmes) OVER (ORDER BY ano) AS total_filmes_ano_anterior,
    SAFE_DIVIDE(total_filmes - LAG(total_filmes) OVER (ORDER BY ano), LAG(total_filmes) OVER (ORDER BY ano)
* 100 AS evolucao_percentual
FROM
    filmes_por_ano
ORDER BY
    ano;
```

Conta o número de filmes por ano e calcular a evolução do número de filmes, comparando o percentual entre dois anos consecutivos.

# QUERYs

```
WITH total_filmes AS (
    SELECT COUNT(*) AS total
    FROM `applied-radar-474411-h3.bigquery.resultadofinal`
)
SELECT
    pais,
    COUNT(*) AS numero_filmes,
    (COUNT(*) / (SELECT total FROM total_filmes)) * 100 AS percentual_filmes_por_pais
FROM
    `applied-radar-474411-h3.bigquery.resultadofinal`
GROUP BY
    pais
ORDER BY
    percentual_filmes_por_pais DESC;
```

Calcula a porcentagem de filmes de cada país em relação ao total de filmes.

# QUERYs

```
SELECT
    genero,
    pais,
    AVG(nota) AS media_avaliacao
FROM
    `applied-radar-474411-h3.bigquery.resultadofinal`
GROUP BY
    genero, pais
ORDER BY
    media_avaliacao DESC;
```

Calcula a porcentagem de filmes de cada país em relação ao total de filmes.

# POWER BI

