



UNIVERSIDADE
DE ÉVORA

Desenho do Modelo Dimensional

Licenciatura em Engenharia Informática

Integração e Processamento Analítico de Informação

2023-2024



Alunos:

Bárbara Loureiro (48469), Gustavo Gomes (48392).

Docentes:

Carlos Pampulim Caldeira

1ª Fase do Trabalho - Desenho do Modelo Dimensional, desenvolvido no âmbito da unidade curricular de Integração e Processamento Analítico de Informação da Licenciatura em Engenharia Informática.

Évora, 25 de março de 2024

~~Conteúdo~~

1	Introdução	1
2	Escolha do Processo de Negócio	2
3	Declaração do Grão do Processo de Negócio	3
4	Escolha das Tabelas de Dimensão	4
4.1	Dimensão Filme	4
4.2	Dimensão Realizador	4
4.3	Dimensão Ator	5
4.4	Dimensão Género	5
4.5	Dimensão Data	6
5	Identificação dos Factos	7
6	Conclusão	8
7	Anexos	9

1 Introdução

Um modelo dimensional de dados é utilizado em data warehousing para organizar e estruturar os dados de uma forma que facilite a análise e a obtenção de insights. Ele é especialmente útil para lidar com conjuntos de dados complexos e grandes volumes de informações.

Num modelo dimensional, os dados são organizados em duas categorias principais:

- **Dimensões:** São as categorias pelas quais os dados podem ser analisados. Elas representam os aspectos pelos quais se analisam os dados e geralmente contêm informações descritivas.
- **Factos:** São as medidas quantitativas ou métricas que representam os dados a analisar. Eles contêm informações numéricas que podem ser agregadas.

A relação entre dimensão e facto é estabelecida através de chaves estrangeiras. Cada registo numa tabela de factos está associado a registos nas tabelas de dimensões. Isso permite analisar os dados de diferentes perspectivas, combinando informações descritivas (dimensões) com medidas quantitativas (factos). Os modelos são frequentemente representados em esquemas de estrela, os factos estão no centro, rodeados pelas dimensões, criando uma estrutura que se assemelha a uma estrela quando visualizada graficamente.

A primeira fase do trabalho consistiu em elaborar um desenho do modelo dimensional de dados, a escolha do processo de negócio, a declaração do grão do processo de negócio, a escolha das tabelas de dimensão, a identificação dos factos e a identificação das fontes de dados.

O modelo dimensional desenvolvido destina-se à análise do desempenho no cinema, usando como referência as plataformas de avaliação do IMDb e do Rotten Tomatoes. O objetivo é fornecer uma estrutura analítica abrangente para explorar o sucesso de filmes e séries, considerando aspectos como elenco, realizadores, género e avaliações.

Para as amostras de dados, utilizadas como exemplo no nosso Google Sheets (em anexos), foram geradas aleatoriamente no Google Colab, em documentos .xlsx, por scripts de Python.

2 Escolha do Processo de Negócio

O processo de negócio selecionado para este modelo dimensional é o acompanhamento do desempenho de filmes e séries. Inclui, não apenas a análise das avaliações e críticas por parte dos utilizadores e dos profissionais especializados, mas também relaciona o impacto que o elenco e os realizadores têm no resultado final da produção.

Permite-nos identificar alguns aspectos impulsionadores de sucesso no mundo do entretenimento audiovisual, que podem orientar decisões estratégicas na indústria, desde a seleção de elenco e realizador até à estratégia de marketing e distribuição.

3 Declaração do Grão do Processo de Negócio

O grão escolhido para este modelo é "a avaliação de cada filme ou série por plataforma de avaliação (IMDb, Rotten Tomatoes) e ano de lançamento". Esse nível de granularidade permite análises detalhadas do desempenho individual de cada produção, ao mesmo tempo que possibilita agregações para insights mais amplos.

4 Escolha das Tabelas de Dimensão

As dimensões **selecionadas** para o modelo foram escolhidas para fornecer uma visão abrangente dos elementos-chave que **influenciam** o desempenho de filmes e séries. Elas incluem:

4.1 Dimensão Filme

Chave e Atributos: Chave do filme, título, ano de lançamento, gênero, duração, classificação indicativa, idioma original, realizador principal, elenco principal, avaliação IMDb e avaliação Rotten Tomatoes.

Finalidade: Facilita análises detalhadas sobre o desempenho de **filmes e séries**, permitindo a identificação de padrões e tendências.

Coluna	Descrição
Chave do Filme	Identificador único para cada filme ou série
Título do Filme	Nome do filme ou série
Ano de Lançamento	Ano em que o filme ou série foi lançado
Gênero	Gêneros do filme ou série
Duração	Duração do filme em minutos ou temporadas/episódios para séries
Classificação Indicativa	Faixa etária recomendada para o público
Idioma Original	Idioma original do filme ou série
Realizador Principal	Nome do realizador principal do filme ou série
Elenco Principal	Nomes dos atores principais do elenco
Avaliação do IMDb	Avaliação do filme/série no site da IMDb
Avaliação do Rotten Tomatoes	Avaliação do filme/série no site da Rotten Tomatoes

Figura 1: Tabela Dimensão Filme

4.2 Dimensão Realizador

Chave e Atributos: Chave do realizador, nome completo, data de nascimento, sexo, nacionalidade, número de filmes produzidos, data de falecimento e ano de início de carreira.

Finalidade: Permite avaliar o impacto dos realizadores no sucesso das produções e explorar a relação entre a experiência do realizador e a qualidade dos filmes.

Coluna	Descrição
Chave do Realizador	Identificador único para cada realizador
Nome do Realizador	Nome completo do realizador
Data de Nascimento	Data de nascimento do realizador
Sexo	Sexo do realizador
Nacionalidade	Nacionalidade do realizador
Número de Filmes	Total de filmes produzidos pelo realizador
Data de Falecimento	Data de falecimento do realizador
Ano de Início de Carreira	Ano em que o diretor iniciou sua realizador

Figura 2: Tabela Dimensão Realizador

4.3 Dimensão Ator

Chave e Atributos: Chave do ator, nome completo, nome artístico, data de nascimento, sexo, nacionalidade, número de filmes em que participou, data de falecimento e ano de início de carreira.

Finalidade: Oferece insights sobre a influência do elenco no desempenho das produções e possibilita análises sobre a popularidade e o sucesso de atores específicos.

Coluna	Descrição
Chave do Ator	Identificador único para cada ator
Nome do Ator	Nome completo do ator
Nome Artístico	Nome artístico do ator
Data de Nascimento	Data de nascimento do ator
Sexo	Sexo do ator
Nacionalidade	Nacionalidade do ator
Número de Filmes	Total de filmes em que o ator participou
Data de Falecimento	Data de falecimento do ator
Ano de Início de Carreira	Ano em que o ator iniciou sua carreira no cinema/televisão

Figura 3: Tabela Dimensão Realizador

4.4 Dimensão Género

Chave e Atributos: Chave do género, Género do filme ou série e descrição.

Finalidade: Permite categorizar filmes e séries de acordo com seus géneros, facilitando análises relacionadas aos gostos do público e tendências do mercado cinematográfico.

Coluna	Descrição
Chave do Género	Identificador único para cada género
Ação	Filmes dinâmicos com destaque em cenas de luta, perseguições e explosões.
Aventura	Narrativas centradas em explorações e jornadas, muitas vezes em locais exóticos ou perigosos.
Comédia	Obras focadas no humor, com situações cómicas e diálogos engraçados.
Drama	Histórias com desenvolvimento intenso de personagens e enredos emocionais.
Fantasia	Filmes que incorporam elementos mágicos, mitológicos ou sobrenaturais nas suas narrativas.
Ficção Científica	Enredos baseados em avanços científicos e tecnológicos, muitas vezes ambientados no futuro.
Horror	Filmes projetados para assustar e criar suspense, frequentemente com presença de entidades sobrenaturais.
Mistério	Tramas que envolvem a solução de um crime ou o desvendar de segredos.
Romance	Histórias de amor e relacionamentos íntimos entre personagens.
Suspense	Filmes que mantêm os espectadores na expectativa e tensão, muitas vezes através de reviravoltas na trama.
Terror	Subgénero do horror, projetado para provocar terror e medo extremo.
Documentário	Obras que apresentam factos reais, pessoas e eventos, muitas vezes com um propósito informativo ou educacional.
Animação	Filmes que utilizam técnicas de animação para criar os personagens e ambientes.
Musical	Filmes que incorporam canções e coreografias como parte essencial do enredo.
Biografia	Histórias baseadas na vida de pessoas reais, retratando eventos históricos ou conquistas significativas.

Figura 4: Tabela Dimensão Género

4.5 Dimensão Data

Chave e Atributos: Chave da Data, Data Completa, Ano, ... ~~(ver imagem correspondente)~~

Finalidade: Fornece informações temporais cruciais para análises de tendências e padrões ao longo do tempo em relação ao desempenho de filmes e séries.

Coluna	Descrição
Chave da Data	Identificador único para cada data
Data Completa	Data no formato AAAA-MM-DD
Ano	Ano da data
Semestre	Semestre (1 ou 2)
Trimestre	Trimestre (1 a 4)
Mês	Mês (1 a 12)
Nome do Mês	Nome do mês (Janeiro a Dezembro)
Dia	Dia do mês (1 a 31)
Dia da Semana	Dia da semana (Segunda a Domingo)
Nome do Dia	Nome do dia da semana
Semana do Ano	Semana do ano (1 a 52)
Dia do Ano	Dia do ano (1 a 365/366)
Feriado	Indica se é um feriado
É Final de Semana	Indica se é final de semana
É Dia Útil	Indica se é dia útil
Quinzena	Quinzena do mês (1 ou 2)
Época do Ano	Época do ano (Primavera, Verão, Outono, Inverno)
Ano Fiscal	Ano fiscal
Trimestre Fiscal	Trimestre fiscal (1 a 4)
Mês Fiscal	Mês fiscal (1 a 12)
Semana Fiscal	Semana fiscal (1 a 52)
Início da Semana	Data de início da semana
Fim da Semana	Data de fim da semana
Início do Mês	Data de início do mês
Fim do Mês	Data de fim do mês
Início do Trimestre	Data de início do trimestre
Fim do Trimestre	Data de fim do trimestre
Início do Ano	Data de início do ano
Fim do Ano	Data de fim do ano

Figura 5: Tabela Dimensão Data

A **inclusão das dimensões de género e data é essencial** para garantir uma análise completa e abrangente do desempenho no cinema, permitindo uma compreensão mais profunda das tendências e ~~fatores~~ que influenciam o sucesso das produções.

5 Identificação dos Factos

Neste caso, os factos são e podem ser observados nas figuras 6 e 7:

- **Avaliação do IMDb:** A avaliação média atribuída pelos utilizadores do IMDb para cada filme.
- **Avaliação do Rotten Tomatoes:** A avaliação média atribuída pelos críticos ou público no Rotten Tomatoes para cada filme.

Chave do Filme	Título do Filme	Ano de Lançamento	Gênero	Duração	Classificação Indicativa	Idioma Original
1	Descoberta Final	2007	Horror	82	16 anos	Francês
2	Aventura Mágica	2015	Romance	134	Livre	Italiano
3	Descoberta Invisível	2011	Ação	134	12 anos	Inglês

Realizador Principal	Elenco Principal	Avaliação do IMDB	Avaliação do Rotten Tomatoes
Alex Johnson	Alex Smith, Alex Jones,	3	51
Jordan Williams	Taylor Wilson, Jamie Jo	5	59
Jordan Williams	Jordan Miller, Dakota Jo	4	60

Figura 6 e 7: Exemplo de Amostra de Dados
Filme

6 Conclusão

O desenho do modelo dimensional desenvolvido oferece uma estrutura para a análise do desempenho no cinema, permitindo aos utilizadores extrair insights valiosos sobre as características que influenciam o sucesso das produções. Este desenho servirá como base para as fases seguintes da unidade curricular. Graças a ele, temos as bases para começar a desenvolver um esquema em estrela (em anexo) e consequentemente as seguintes etapas do projeto.

7 Anexos

Google Sheets da Fase 1:

<https://docs.google.com/spreadsheets/d/1FTZ17RKPnha5AOt6Ve4XmpAtS6T7ngGySBAXnFkl65E/edit?usp=sharing>

Esquema em Estrela:

