



Integração e Processamento Analítico de Informação

Modelo Dimensional do Projeto Final

Miguel Grilo 58387

Jorge Couto 58656

Tabela de Conteúdos

Conteúdo

Tabela de Conteúdos.....	i
Lista de Figuras	ii
Resumo.....	iii
Introdução.....	1
Metodologia.....	2
Temas	2
Ferramentas.....	2
Aplicações e/ou Procedimentos	2
Caso de Estudo.....	3
Dimensões.....	3
Dimensão da Data.....	3
Dimensão da Hora.....	4
Dimensão do Piloto	4
Dimensão das Equipas/Marcas	5
Dimensão da Competição	5
Dimensão dos Circuitos.....	6
Dimensão das Corridas	6
Dimensão dos Pitstops	7
Esquema em Estrela	7
Tabela de Fatos.....	8
Grão Mais Atómico Possível.....	8
Conclusões [e Perspetivas Futuras].....	9
Referências	10

Lista de Figuras

Figura 1 - Dimensão da Data	3
Figura 2 - Dimensão da Hora	4
Figura 3 - Dimensão do Piloto	4
Figura 4 - Dimensão das Equipas/Marcas.....	5
Figura 5 - Dimensão da Competição	5
Figura 6 - Dimensão dos Circuitos	6
Figura 7 - Dimensão das Corridas	6
Figura 8 - Dimensão dos Pitstops.....	7
Figura 9 - Esquema em Estrela	7
Figura 10 - Tabela de Fatos	8
Figura 11 - Grão Mais Atómico Possível Parte1.....	8
Figura 12 - Grão Mais Atómico Possível Parte2.....	8

Resumo

Este projeto teve como objetivo criar um modelo dimensional para analisar o desempenho de pilotos e/ou equipas/marcas de Fórmula 1, permitindo uma visão detalhada ao longo dos anos. O modelo dimensional organiza os dados para facilitar o estudo de diversos fatores, como a data, hora, condições atmosféricas, circuito e pitstops realizados durante cada corrida e ainda as características dos pilotos e equipas/marcas, com recurso a um Esquema em Estrela. A Tabela de Fatos apresenta-nos alguns indicadores como o Número de Troféus Ganhos pelo Piloto e/ou Equipa/Marca, Pontos Totais do Piloto e Maior Posição do Piloto e/ou Equipa/Marca, que nos permitem comparar desempenhos. Este projeto aplica conhecimentos de *Data Warehousing* para organização dos dados de modo a apoiar e facilitar a análise. O modelo em questão servirá como base para estudar o desempenho na F1 futuramente, oferecendo uma melhor visão sobre os fatores que influenciam os resultados das corridas.

Introdução

Atualmente, a análise de dados desempenha um papel essencial no desporto, permitindo obter informações valiosas sobre o desempenho de atletas e equipas. No caso da Fórmula 1, onde cada detalhe pode ser decisivo para o sucesso, a utilização de modelos de análise bem estruturados torna-se fundamental para compreender a evolução dos pilotos e das equipas ao longo do tempo.

O objetivo deste projeto é desenvolver um modelo dimensional para um conjunto de dados da Fórmula 1, que possibilite uma análise aprofundada da performance dos pilotos e das equipas/marcas. Através deste modelo, será possível responder a várias questões relevantes, tais como: quantos troféus foram conquistados por um piloto ou por uma equipa/marca ao longo dos anos, o total de pontos acumulados por cada piloto ao longo da sua carreira, a melhor posição alcançada por um piloto em competição assim como a melhor posição atingida por uma equipa/marca em competições.

Para alcançar estes objetivos, será aplicado um modelo dimensional, comum em sistemas de armazenamento de dados (*Data Warehousing*), que permite otimizar a estrutura da informação para facilitar consultas analíticas. O modelo será baseado num esquema em estrela, onde uma tabela de fatos conterá os principais indicadores de desempenho, complementada por tabelas dimensão que fornecerão o contexto necessário para uma análise detalhada.

Metodologia

Temas

- **Modelo Dimensional:** Construção de um modelo dimensional aplicado à análise da performance de pilotos e equipas na Fórmula 1, utilizando conceitos de *Data Warehousing*.
- **Processo de Negócio:** Identificar o fluxo de informação relevante para a análise de desempenho no contexto da Fórmula 1.
- **Dimensões:** Identificar os elementos essenciais para criar as tabelas dimensão.
- **Tabela de Fatos:** Construção da mesma com o grão mais atómico possível dentro do processo de negócio em análise.
- **Esquema em Estrela:** Organização do modelo dimensional com uma tabela de fatos central ligada diretamente às tabelas dimensão

Ferramentas

- **Microsoft Excel:** Utilizado para organizar a informação retirada do talão de venda.
- **Microsoft Word:** Utilizado para redigir o PDF solicitado.

Aplicações e/ou Procedimentos

- **Análise dos Dados:**
 1. Análise dos conjuntos de dados da Fórmula 1.
 2. Identificação dos elementos essenciais presentes nos dados.
 3. Definição do processo de negócio relacionado com a análise do desempenho na modalidade.
- **Definição das Dimensões:**
 4. Identificação dos elementos-chave que fazem parte das dimensões do modelo.
 5. Identificação de atributos relevantes em cada dimensão.
- **Definição do Grão Mais Atómico Possível:**
 6. Definição do grão mais atómico do modelo.
- **Construção da Tabela de Fatos:**
 7. Inclusão das chaves estrangeiras que ligam a tabela de fatos às dimensões.
- **Construção do Esquema em Estrela:**
 8. Estruturação do modelo utilizando o esquema em estrela, onde a tabela de fatos se conecta diretamente a todas as dimensões.

Caso de Estudo

Para melhor entender o esquema de estrela e futura representação da Tabela de Fatos, temos em baixo uma representação de cada uma das Dimensões criadas com as respetivas colunas. Com cada uma está uma pequena explicação indicando o nosso processo de pensamento crítico por detrás da criação das colunas.

Dimensões

Dimensão da Data

Dimensão da Data	A criação da Dimensão Data baseou-se na observação das datas dos Grand Prix de F1, permitindo uma análise detalhada da influência temporal na performance dos pilotos e equipas. Para tal, criámos a Chave da Data como chave primária, garantindo a indexação eficiente da tabela. Para fornecer um contexto mais completo, adicionámos a Data Completa, bem como o Dia da Semana e a Semana do Ano, permitindo identificar padrões de desempenho em diferentes momentos da temporada. Além disso, incluímos colunas para o Mês, Trimestre, Semestre e Ano, facilitando análises comparativas entre diferentes períodos. Considerando a influência do calendário, registámos se a data corresponde a um Feriado e se é um Dia Útil, possibilitando estudar possíveis correlações com a assistência e impacto nos resultados. Para melhor enquadrar a época de cada corrida, criámos a coluna Temporada F1, assim como a Estação do Ano, que pode influenciar fatores como a temperatura e as condições da pista. Incluímos também um indicador para saber se a data corresponde a um Dia de Corrida. Para uma análise mais detalhada, adicionámos informações sobre se a data faz parte de uma Estação Festiva (como Natal ou Ano Novo) ou do período de Férias, fatores que podem influenciar o desempenho do piloto e/ou equipa/marca. Além disso, a Dimensão Data foi estruturada de forma hierárquica, incluindo os níveis Data Completa → Semana do Ano → Mês → Trimestre → Semestre → Ano → Temporada F1, possibilitando análises que vão desde um único dia específico até períodos mais amplos, como temporadas inteiras. Esta estrutura hierárquica permite a realização de comparações entre diferentes épocas, identificando tendências e padrões de desempenho ao longo do tempo.
Chave da Data	
Data Completa	
Dia da Semana	
Semana do Ano	
Mês	
Trimestre	
Semestre	
Ano	
É Feriado	
Dia Útil	
Temporada F1	
Estação do Ano	
É Dia de Corrida	
É Estação Festiva	
É Férias	

Figura 1 - Dimensão da Data

Dimensão da Hora

Dimensão da Hora
Chave da Hora
Hora
Minuto
Segundo

Figura 2 - Dimensão da Hora

A criação da Dimensão Hora permite uma análise detalhada do impacto do tempo exato em que ocorrem os eventos na Fórmula 1. Para isso, definimos a Chave da Hora como chave primária e incluímos a Hora, o Minuto e o Segundo, possibilitando uma análise precisa do tempo em que cada evento ocorre, como o início e fim de uma corrida, tempos de volta, paragens nas boxes, entre outros.

Dimensão do Piloto

Dimensão do Piloto
Chave do Piloto
Nome de Referência do Piloto
Número do Piloto
Código de 3 letras do Piloto
Primeiro Nome do Piloto
Apelido do Piloto
Data de Nascimento
Idade
Nacionalidade
Cidade de Origem
Salário
Se é Ativo
Corrida de Estreia na F1
Ano de Estreia na F1

Figura 3 - Dimensão do Piloto

A criação da Dimensão do Piloto permite uma análise detalhada das características individuais de cada piloto ao longo da história da Fórmula 1. Para isso, definimos a Chave do Piloto como chave primária, assegurando a correta identificação de cada piloto na base de dados e incluímos o Nome de Referência do Piloto, o Número do Piloto e o Código de 3 letras do Piloto, que são frequentemente utilizados para identificação rápida durante as competições. Para uma melhor organização, adicionámos também o Primeiro Nome e o Apelido do Piloto. A dimensão integra ainda dados biográficos, como a Data de Nascimento, a Idade, a Nacionalidade e a Cidade de Origem, permitindo análises demográficas dos pilotos. Além disso, registámos o Salário, um fator relevante para estudos financeiros e de investimento

das equipas. Para indicar a situação competitiva do piloto, adicionámos o campo Se é Ativo, distinguindo os pilotos que ainda competem daqueles que já se retiraram. Por fim, incluímos informações sobre o início da carreira na Fórmula 1, como a Corrida de Estreia na F1 e o Ano de Estreia na F1, permitindo analisar a trajetória e longevidade de cada piloto no desporto.

Dimensão das Equipas/Marcas

Dimensão das Equipas
Chave da Equipa
Nome Completo da Equipa
Nome Abreviado da Equipa
Sede da Equipa
Primeira Entrada da Equipa
Fundador da Equipa
Diretor Atual da Equipa
Nome dos Fornecedores Principais
Networth da Equipa

Figura 4 - Dimensão das Equipas/Marcas

A criação da Dimensão das Equipas permite uma análise aprofundada das equipas que competiram na Fórmula 1 ao longo dos anos. Para tal, definimos a Chave da Equipa como chave primária, assegurando a identificação única de cada equipa na base de dados e incluímos o Nome Completo da Equipa e o Nome Abreviado da Equipa, que são amplamente utilizados em transmissões, classificações e relatórios da modalidade. Para melhor contextualização geográfica,

adicionámos a Sede da Equipa, permitindo análises relacionadas com a origem e a distribuição das equipas. Registámos a Primeira Entrada da Equipa, que indica o ano em que começou a competir na Fórmula 1 e também incluímos o Fundador da Equipa, reconhecendo as figuras responsáveis pela sua criação, e o Diretor Atual da Equipa, fundamental para analisar o impacto da gestão nas performances ao longo do tempo. Além disso, foram acrescentados os Nomes dos Fornecedores Principais, refletindo as parcerias técnicas e comerciais da equipa, e o *Net Worth* da Equipa, permitindo análises financeiras sobre a valorização e investimento das marcas dentro da competição.

Dimensão da Competição

Dimensão da Competição
Chave da Competição
Nome Completo da Competição
Nome Abreviado da Competição
Ano da Competição
Número de Participantes
Número Total de Corridas

Figura 5 - Dimensão da Competição

A Dimensão da Competição permite identificar e organizar os diferentes campeonatos de Fórmula 1 ao longo dos anos. A Chave da Competição é a chave primária, garantindo a identificação única de cada temporada. Incluímos o Nome Completo da Competição e o Nome Abreviado da Competição, utilizados em classificações e análises históricas. O Ano da Competição permite

segmentar dados por temporada, enquanto o Número de Participantes e o Número Total de Corridas ajudam a compreender a escala e evolução da competição ao longo do tempo.

Dimensão dos Circuitos

Dimensão dos Circuitos
Chave do Circuito
Nome de Referência do Circuito
Nome Completo do Circuito
Cidade
País
Comprimento da Pista
Número de Voltas
Ano de Estreia do Circuito

Figura 6 - Dimensão dos Circuitos

A Dimensão dos Circuitos permite organizar e analisar os diferentes circuitos da Fórmula 1. A Chave do Circuito é a chave primária, garantindo a identificação única de cada circuito. Incluímos o Nome de Referência do Circuito e o Nome Completo do Circuito, usados para identificar e distinguir os circuitos em análises. A Cidade e o País proporcionam o contexto geográfico, enquanto o Comprimento da Pista e o Número de Voltas ajudam a caracterizar as corridas em cada circuito. Também registámos o Ano de Estreia do Circuito, permitindo uma visão histórica sobre a introdução de cada circuito no calendário da Fórmula 1.

Dimensão das Corridas

Dimensão das Corridas
Chave da Corrida
Ano da Corrida
Ronda da Corrida
Nome Oficial da Corrida
Número de Voltas Completas
Clima Durante a Corrida
Nome do Circuito
Número de Bilhetes Vendidos

Figura 7 - Dimensão das Corridas

A Dimensão das Corridas organiza as informações relacionadas a cada evento da Fórmula 1. A Chave da Corrida é a chave primária, permitindo a identificação única de cada corrida. Incluímos o Ano da Corrida e a Ronda da Corrida para categorizar a temporada e a posição da corrida no calendário. O Nome Oficial da Corrida e o Nome do Circuito são utilizados para identificar o evento e o local. O Número de Voltas Completas especifica a quantidade de voltas percorridas, e o Clima Durante a Corrida ajuda a analisar as condições climáticas que podem ter influenciado o desempenho. Por fim, o Número de Bilhetes Vendidos oferece uma visão sobre a popularidade e a assistência ao evento.

Dimensão dos Pitstops

Dimensão dos Pitstops
Chave do Pitstop
Tempo Total do Pitstop
Partes do Veículo Trocadas
Número da Volta do Pitstop
Condições no Momento do Pitstop
Nomes dos Membros da Equipa Mecânica

Figura 8 - Dimensão dos Pitstops

A Dimensão dos Pitstops organiza as informações relativas às paragens nas boxes durante as corridas. A Chave do Pitstop é a chave primária, permitindo a identificação única de cada pitstop. Incluímos o Tempo Total do Pitstop, que reflete a eficiência da equipa nas paragens e as Partes do Veículo Trocadas,

detalhando as alterações realizadas no carro. O Número da Volta do Pitstop ajuda a identificar o momento exato da corrida em que a paragem ocorreu. Adicionalmente, registámos as Condições no Momento do Pitstop, que refletem as circunstâncias que motivaram a paragem, como danos no carro ou condições da pista (exemplo: pista molhada, obrigando à troca de pneus). Por fim, incluímos os Nomes dos Membros da Equipa Mecânica, permitindo avaliar o desempenho da equipa responsável pelo pitstop.

Esquema em Estrela

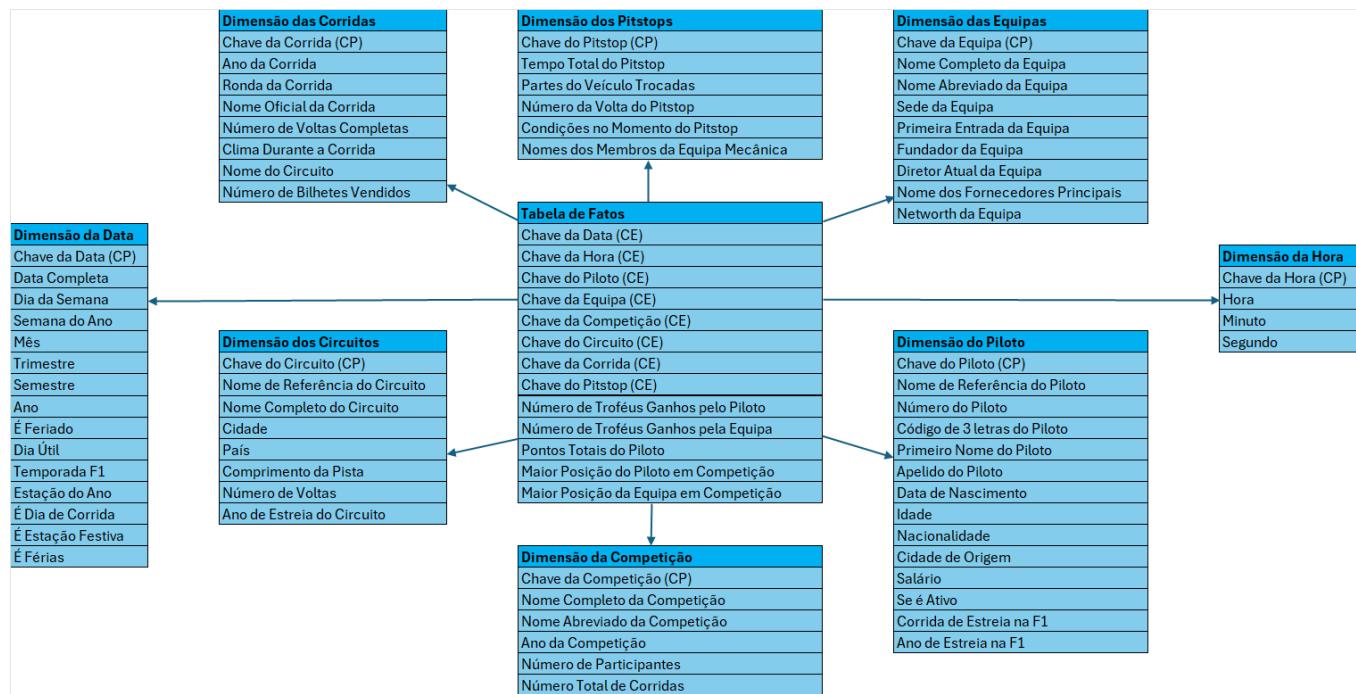


Figura 9 - Esquema em Estrela

Tabela de Fatos

Tabela de Fatos
Chave da Data
Chave da Hora
Chave do Piloto
Chave da Equipa
Chave da Competição
Chave do Circuito
Chave da Corrida
Chave do Pitstop
Número de Troféus Ganhos pelo Piloto
Número de Troféus Ganhos pela Equipa
Pontos Totais do Piloto
Maior Posição do Piloto em Competição
Maior Posição da Equipa em Competição

Figura 10 - Tabela de Fatos

A Tabela de Fatos agrupa os dados quantitativos e os indicadores de desempenho da Fórmula 1, permitindo uma análise detalhada da performance dos pilotos e das equipas ao longo do tempo. Cada registo na tabela é associado a diferentes chaves primárias das dimensões, como Data, Hora, Piloto, Equipa, Competição, Circuito, Corrida e Pitstop, facilitando a análise de diversas perspetivas. A tabela inclui dados importantes, como o Número de Troféus Ganhos pelo Piloto e pela Equipa, os Pontos Totais do Piloto, e as Maior Posição do Piloto e Maior Posição da Equipa em competição, ajudando a avaliar o desempenho nas corridas e nas temporadas. Esta estrutura permite fazer comparações e identificar padrões de desempenho, considerando fatores como o tempo e as condições das corridas.

Grão Mais Atómico Possível

Chave da Data	Chave da Hora	Chave do Piloto	Chave da Equipa	Chave da Competição	Chave do Circuito	Chave da Corrida	Chave do Pitstop
DAT00001	HOR1430	P001	E001	COM2025	CIR001	COR01	PIT001
DAT00004	HOR1430	P001	E001	COM2025	CIR002	COR02	PIT002
DAT00009	HOR1430	P001	E001	COM2025	CIR003	COR03	PIT003
DAT00021	HOR1430	P001	E001	COM2025	CIR004	COR04	PIT004

Figura 11 - Grão Mais Atómico Possível Parte1

Número de Troféus Ganhos pelo Piloto	Número de Troféus Ganhos pela Equipa	Pontos Totais do Piloto	Maior Posição do Piloto em Competição	Maior Posição da Equipa em Competição
1	2	25	1	1
1	1	18	1	2
0	1	12	2	3
2	2	30	1	1

Figura 12 - Grão Mais Atómico Possível Parte2

Conclusões [e Perspetivas Futuras]

As conclusões deste projeto recaem no sucesso em criar um modelo dimensional eficiente para analisar o desempenho de pilotos e equipas/marcas na Fórmula 1, recorrendo a um Esquema em Estrela para organizar e estruturar dados de forma clara e permitindo a análise de fatores como clima, pitstops e características dos envolvidos. A Tabela de Fatos, com indicadores como troféus e pontos, foi essencial para entender as performances.

Os resultados mostram que a aplicação de conceitos de *Data Warehousing* se mostrou eficaz para trabalhar com grandes volumes de dados e proporcionar conclusões sobre o desempenho. Contudo, há espaço para melhorias, como incluir dados sobre o desenvolvimento técnico dos carros, a gestão das equipas e as estratégias executadas nas corridas. A integração de dados em tempo real, como os tempos de volta, também poderiam aumentar a precisão das análises.

Referências

Caldeira, C. P. (2023). *Data Warehousing: Conceitos e Modelos* (3^a ed.). Lisboa: Edições Sílabo.