

FIAP GRADUAÇÃO

TECNOLOGIA EM CIÊNCIA DE DADOS

DATA WAREHOUSING

2º checkpoint: Modelagem Dimensional Star Schema do projeto das Salas de Cinema Five Star

PROF. LEONARDO JOÃO ZAMBONI profleonardo.zamboni@fiap.com.br

PROF. SALVIO PADLIPSKAS salvio@fiap.com.br

MODELAGEM DIMENSIONAL STAR SCHEMA



Nessa aula vamos tratar da seguinte agenda:

- Entendimento do negócio da empresa Five Star (salas de cinema).
- Analisando as perguntas de negócios definidas.
- Criando a modelagem de dados dimensional Star Schema da Five Star.

| ENTENDIMENTO DO NEGÓCIO



Rede de Cinemas





ENTENDIMENTO DO NEGÓCIO

- Depois de se favorecer de um alto crescimento no setor de entretenimento e suas ações aumentarem em mais de 30% em 2024, a Rede de Cinemas Cineplex foi adquirida pela FiveStar, que atualmente é considerada segunda maior empresa de entretenimento da américa latina.
- Os principais fatores que ocasionaram esse crescimento foi o aumento da população jovem e o lançamento incessável de inúmeros filmes de sucesso.
- Com a chegada da Cineplex, a Rede FiveStar obtém agora uma média de 50.000.000 de ingressos vendidos mensais e com uma carteira de clientes de aproximadamente 300.000.000 de pessoas.



ENTENDIMENTO DO NEGÓCIO

- Diante deste cenário, a FiveStar resolveu investir ainda mais e criou um plano de ação radical para se tornar a líder no segmento de entretenimento.
- Como parte estratégica desse crescimento, a empresa resolveu investir em um sistema de informação baseado em tomada de decisão analítica on-line, com base no perfil do cliente que adquire ingressos.
- Sendo assim, a necessidade de um DataMart se faz necessária devido à complexidade dos sistemas legados entre as inúmeras empresas do grupo.
- No slide a seguir temos algumas das necessidades.



NECESSIDADES GERENCIAIS

- 1) É necessário acompanhar a evolução da compra dos ingressos feitos pelos clientes durante um determinado período do tempo. Informações como vendas de ingressos por turno (diurno, vespertino e noturno), dia, dia da semana, se esse dia é um feriado, descrição do clima nesse dia, mês, bimestre, trimestre, semestre e ano são necessárias. Informações como quantidade e valor total de ingressos vendidos devem ser registrados para posterior tomada de decisão. A granularidade necessária é o turno onde ocorreu a venda. Alguns ingressos são brindes dados pelos patrocinadores e devem ser identificados pelo modelo analítico.
- 2) Também é necessário acompanhar os ingressos vendidos por sala de cinema, que tem um código único dentro da empresa, o nome da sala onde o filme será exibido e para ser identificado e temos a localização das salas por País, Estado, Sigla do Estado Cidade, Bairro e Região. A sala de cinema deve ter uma localização específica. Shopping, Loja de Rua, Hotel, Resort, Universidade e Clubes são algumas das localizações que devem fazer parte da localização.
- 3) Os dados do filme também são relevantes para nossa análise: Título do filme, ano e mês de lançamento, Gênero (ação, comédia, drama, terror, etc), duração, classificação indicativa (Live, 12 anos, 16 anos) e País de origem.



NECESSIDADES GERENCIAIS

- 4) Os clientes são essenciais nessa análise pois são aqueles que adquirem os ingressos. Necessitamos obter algumas informações de seu perfil, como: Sexo, Escolaridade, Idade, religião, País, Estado, Sigla do Estado Cidade e Bairro onde ele reside são dados essenciais para oportunidades de novos negócios.
- 5) Um cliente pode adquirir mais de um ingresso e muitas vezes um percentual e valor de desconto é gerado para essas situações. Essas informações são relevantes e necessitam ser armazenadas, ou seja, o valor do desconto e valor total pago pelo ingresso deve estar disponível para análise.
- 6) A forma de pagamento é um assunto importante pois pagamento em Pix ou Dinheiro é a melhor forma de pagamento, seguido de cartão de débito e cartão de crédito.



NECESSIDADES GERENCIAIS

De posse dessas informações pede-se:

- Criar o modelo de dados dimensional do tipo Star Schema contendo: Dimensões, Atributos Descritivos, Fatos e Métricas, todos interconectados.
- A localização deve ser associada ao Cliente e a Sala de Cinema, onde o filme é exibido.

É necessário utilizar as boas práticas aplicadas durante as aulas de laboratório, como: padrão de nomenclatura de colunas e tabelas, tipos de dados adequados ao atributo, obrigatoriedade e comentários para cada coluna do modelo físico (DDL), além dos atributos de rastreabilidade das origens de dados e carga de dados.



HANDS ON: CRIAR MODELO DE DADOS DIMENSIONAL STAR SCHEMA

FIAP

- ❖ REALIZE ESSA ATIVIDADE EM GRUPO E O ENVIO DEVE SER FEITO DENTRO DA ATIVIDADE DO TEAMS DISPONÍVEL E SOMENTE O REPRESENTANTE DO GRUPO FICA RESPONSÁVEL EM FAZER ESSE ENVIO.

Modelo de Dados Dimensional da empresa Five Star



EXERCÍCIO: FERRAMENTA SQL*DATA MODELER



EM 6 PASSOS: DESENVOLVIMENTO DO MODELO DIMENSIONAL PADRÃO STAR SCHEMA

- **PASSO 1: PREENCHIMENTO DA PLANILHA CONTENDO AS DIMENSÕES E A TABELA FATO COM AS ESTRUTURAS DE DADOS DO NEGÓCIO.**
 - **PASSO 2: PREENCHIMENTO DA BUS MATRIX ADAPTADA.**
 - **PASSO 3: DESENVOLVIMENTO DO MODELO DIMENSIONAL LÓGICO ANALÍTICO DA EMPRESA 5 STAR, SEGUINDO TODAS AS BOAS PRÁTICAS UTILIZADAS EM AULA LIVE.**
 - **PASSO 4: CRIAÇÃO DO MODELO DIMENSIONAL FÍSICO ANALÍTICO DA 5 STAR.**
 - **PASSO 5: CRIAÇÃO DO SCRIPT CRIA.SQL E APAGA.SQL.**
 - **PASSO 6: INSTALAÇÃO NO SGBD ORCL DA FIAP.**
 - **CADASTRE ALGUMAS LINHAS NAS TABELAS DE DIMENSÃO E FATO (PEÇA A AJUDA A IA) E DEPOIS DESENVOLVA UMA INSTRUÇÃO SQL QUE EXIBA AS SEGUINTE REGRAS DE NEGÓCIO.**
1. Qual é o valor total de vendas agrupado por ano de uma Sala de Cinema?
 2. Qual é a quantidade total de ingressos vendidos e seu valor total agrupado por Estado, Cidade e Bairro?
 3. Qual é o valor total das vendas agrupada do forma de pagamento, ano e mês de venda do ingresso?

ENTREGAS ESPERADAS

1. ARQUIVO COMPONENTES.TXT CONTENDO O NOME E RM DOS PARTICIPANTES
2. PLANILHA EXCEL CONTENDO AS ESTRUTURAS DE DADOS DE NEGÓCIO
3. PLANILHA BUS MATRIX PREENCHIDA A PARTIR DAS PERGUNTAS DE NEGÓCIO
4. ARQUIVO *.DMD E DIRETÓRIO (ARQUIVO FONTE DO PROJETO)
5. ARQUIVO PROJETO LÓGICO ANALÍTICO (*.PDF) NO PADRÃO BARKER DE MODELAGEM
6. ARQUIVO PROJETO FÍSICO ANALÍTICO (*.PDF) NO PADRÃO ENGENHARIA INFORMAÇÃO
7. ARQUIVO CRIA.SQL E APAGA.SQL
8. EVIDÊNCIA DE INSTALAÇÃO PROJETO ANALÍTICO NO SGBD ORACLE ORCL FIAP
9. COMANDOS INSERTS GERADOS PELA IA PARA POPULAR OS DADOS NAS DIMENSÕES E NA TABELA FATO
10. CODIGO FONTE DAS VIEWS CONTENDO AS INSTRUÇÕES SQL QUE RESOLVAM AS 3 PERGUNTAS EXECUTIVAS INDICADOS NO SLIDE ANTERIOR
11. EVIDÊNCIA DE CADA INSTRUÇÃO SQL COM O RESULTADO DOS DADOS.

Gere um **arquivo *.zip** contendo esses entregáveis.

Se reúna no grupo do Teams e crie o projeto de Modelagem Dimensional Star Schema.
A entrega deve ser feita em grupo apenas pelo representante do grupo.

PONTUAÇÃO DA CORREÇÃO A FEITA PELO PROFESSOR

Entregável Esperado	Desconto em erros identificados
1. Planilha com as Estruturas de Dados e arquivo componentes.txt	Falta de envio: 1 ponto Envio incompleto: de 0,5 a 0,9 pontos
2. Planilha com a Bus Matrix preenchida	Falta de envio: 1 ponto Envio incompleto: de 0,5 a 0,9 pontos
2. Arquivo fonte (*.dmd e diretório)	Falta de envio de um dos arquivos: 4,0 pontos Erros identificado no envios: de 1,0 a 4,9 pontos
3. Arquivo Modelo Lógico e Modelo Físico (*.PDF)	Falta de envio: 1,0 pontos Envio incompleto: 1,0 ponto
4. Arquivo Modelo Físico (*.PDF)	Falta de envio: 1,0 pontos Envio incompleto: 1,0 ponto
5. Arquivo cria.sql e apaga.sql	Falta de envio de um dos arquivos: 1,0 pontos Erros identificado no envios: de 0,5 a 1,0 ponto
6. Evidência instalação ORCL FIAP e instruções SQL fonte e evidência	Falta de envio: 1,0 pontos Envio incompleto: 0,5 ponto



Após concluir a tarefa, faça o upload na área de tarefa do Teams em formato de arquivo *.zip, agrupando todos os arquivos informados acima.

DICAS FINAIS



É necessário utilizar as boas práticas aplicadas durante as aulas de laboratório, como: padrão de nomenclatura de colunas e tabelas, tipos de dados adequados ao atributo, obrigatoriedade e comentários para cada coluna do modelo físico (DDL), além dos atributos de rastreabilidade das origens de dados e carga de dados.

- **VOCÊ TERÁ A AULA FULL DA SEMANA PARA CONCLUIR ESSE PROJETO.**
- **O DESENVOLVIMENTO DO TRABALHO DEVE SER FEITO EM GRUPO E A ENTREGA DEVE SER FEITA NA FERRAMENTA TEAMS SOMENTE PELO REPRESENTANTE DO GRUPO.**
- **CASO TENHA DÚVIDAS, ACIONE O SEU PROFESSOR.**

Copyright © 2025 Prof. Salvio Padlipskas

Todos direitos reservados. Reprodução ou divulgação total ou parcial deste documento é expressamente proibido sem o consentimento formal, por escrito, do Professor (autor).