

Projeto "SAKILA"

**Nome da Empresa Assistida pelo Projeto;**

Sakila DVDs store.

**Equipe Responsável;**

A empresa WODA disponibiliza 3 membros para assistir o projeto “SAKILA”, são os seguintes membros;

**João Mendes: Engenheiro de Dados;**  
 Atribuições: Desenvolver pipelines de ETL, manter bancos de dados, integrar sistemas.

Responsabilidades: Garantir a coleta, transformação e carga confiável dos dados, manter a segurança e o desempenho dos sistemas.

**Maria Prado: Analista de BI;**

Atribuições: Criar relatórios e painéis interativos, desenvolver modelos de dados.

Responsabilidades: Gerar informações relevantes e precisas para a equipe de negócios, colaborar com Pedro para atender às necessidades de análise

.

**Pedro Fonseca: Analista de Dados;**

Atribuições: Coletar, limpar e preparar dados, desenvolver consultas e análises.

Responsabilidades: Garantir dados limpos e acessíveis, criar insights relevantes, colaborar com Maria para atender às necessidades de análise.

**Necessidade da** **Empresa Assistida pelo Projeto;**

A empresa "SAKILA" é uma locadora de DVDs que opera em diversas localidades. Ela atende a uma ampla base de clientes e possui uma extensa coleção de filmes disponíveis para aluguel. A empresa enfrenta desafios relacionados ao gerenciamento de estoque, satisfação do cliente, otimização de receitas e tomada de decisões estratégicas para melhorar suas operações de aluguel de DVDs, de momento ela quer avaliar as vendas, contemplando seus clientes e lojas;

**Desafios;**

Receitas e Rentabilidade: "Sakila" deseja aumentar suas receitas e a lucratividade do negócio. Isso inclui identificar estratégias de preços eficazes e promover filmes de alta demanda.

Tomada de Decisões Estratégicas: A empresa precisa de insights para tomar decisões estratégicas, como abrir ou fechar lojas, expandir sua coleção de filmes e otimizar a logística de entrega de DVDs.

**Análises;**

Análise de Rentabilidade: Calcular a rentabilidade de cada filme, identificar oportunidades de aumento de receita e redução de custos operacionais.

Análise de Preços: Determinar preços competitivos e estratégias de descontos para incentivar a locação de filmes.

Essas análises de dados são essenciais para a "SAKILA" tomar decisões informadas que resultem em um melhor gerenciamento de seu negócio de locação de DVDs, satisfação do cliente e aumento das receitas.

**Solução Proposta;**

Depois da equipe debater com o cliente e entender suas necessidades, foram levantados alguns pontos indispensáveis para solucionar os problemas enfrentados pelo cliente, vale pontuar que o banco de dados usado será o ORACLE, assim como as consultas serão realizadas no SQLDEVELOPER e as integrações de dados serão feitas pelo ODI: Oracle Data Integration.

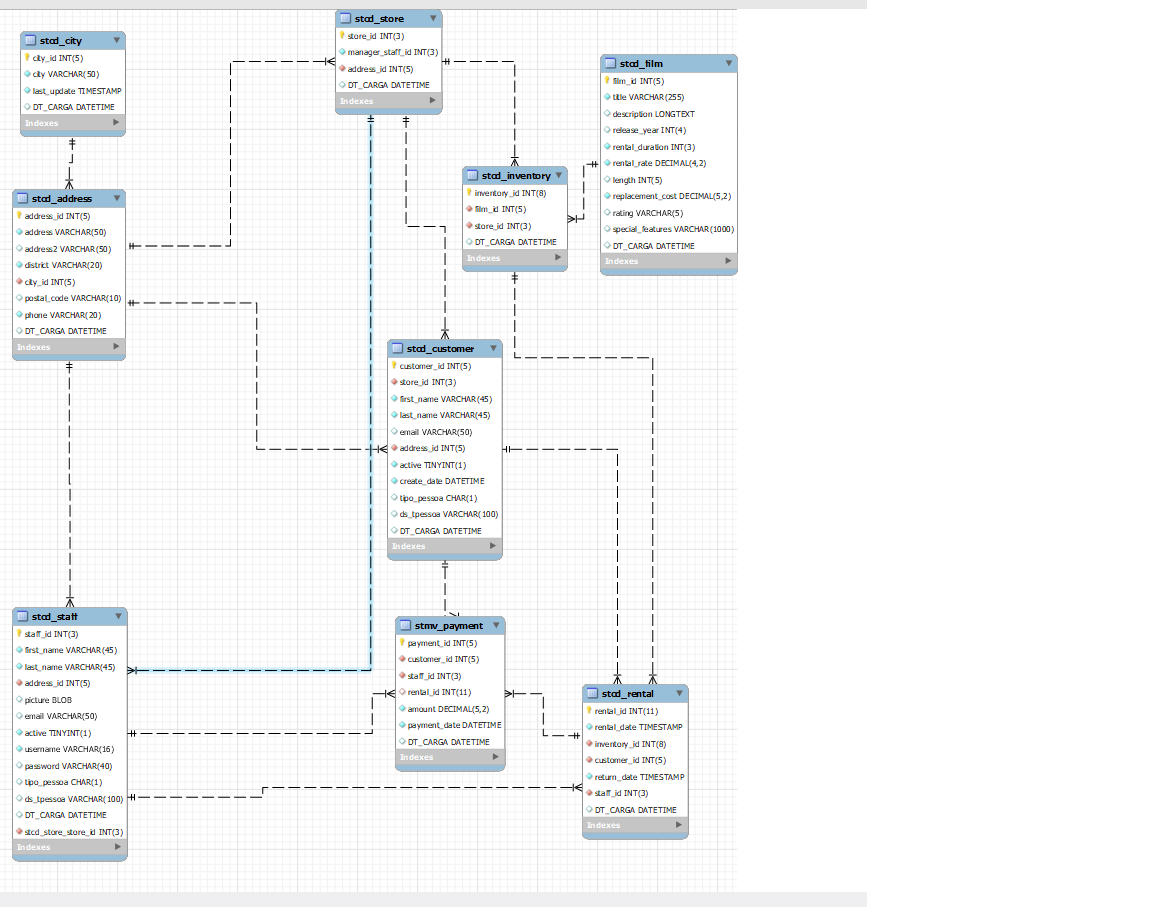
OBS: Modelagens apresentadas em seguidas foram feitas no modelo MYSQL, ou seja, tipo de dados e sintaxe de criação de tabelas são diferentes no banco de dados ORACLE que foram disponibilizadas no GITHUB.

Serão pontuados esses pontos nos tópicos abaixo;

Modelagem de Dados: Nosso Engenheiro de dados planeja implementar três modelagens de dados, modelo STAGE, modelo DW e modelo DM.

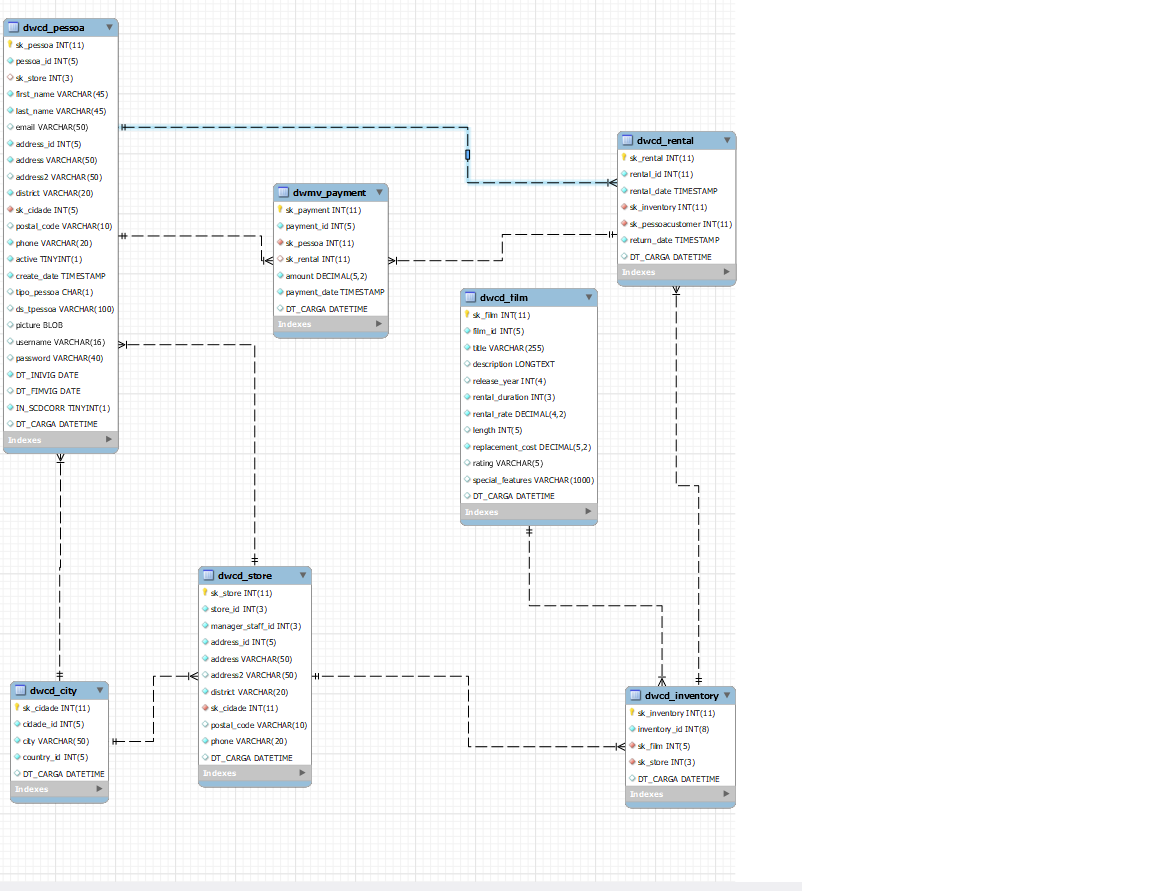
**Modelo STAGE**; responsável por receber os dados providos do sistema OLTP e CSVs de forma rápida e integral, ou seja, os dados no ETL são integrados entre os sistemas o mais rápido possível, evitando transformações nos dados. No ETL essas são as primeiras tabelas a serem carregadas, em síntese, primeiro serão carregadas as STAGE de Cadastro – stcd, e posteriormente as STAGE de movimento – stmv.

Modelagem STAGE



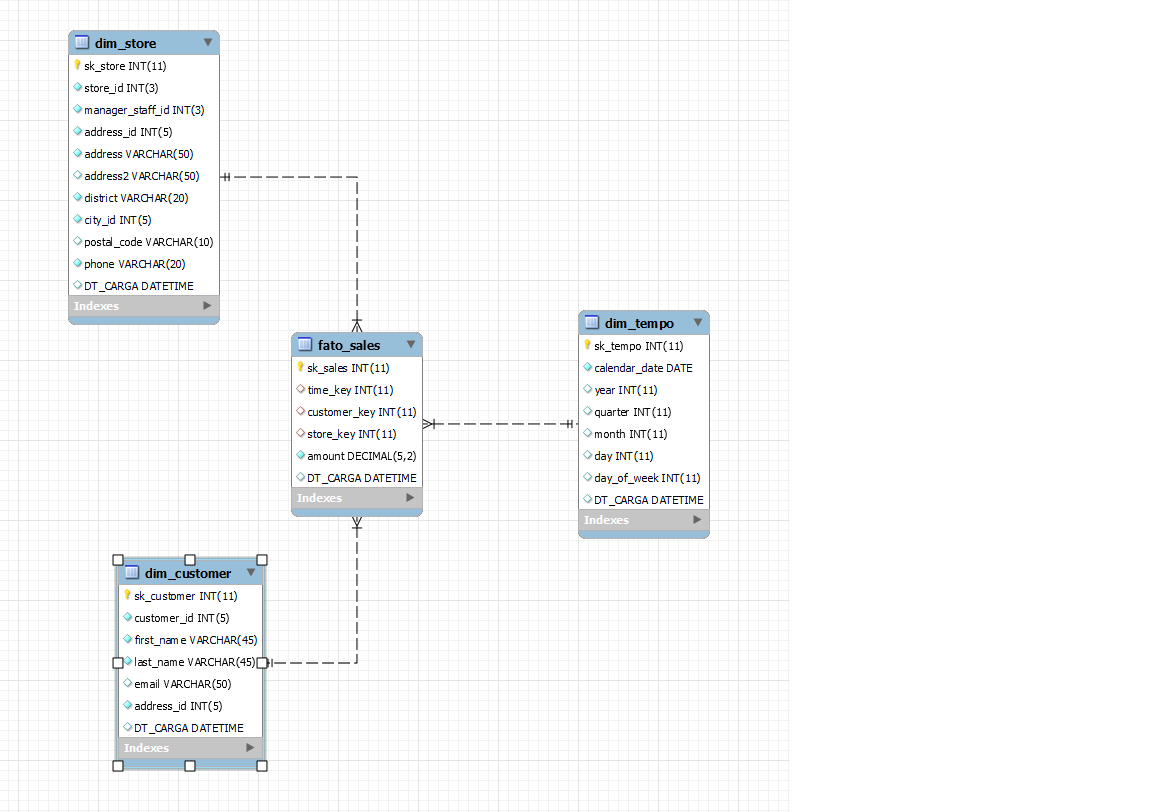
**Modelagem DW;** aqui os dados virão das tabelas do modelo STAGE, e agora começaram os tratamentos de dados para melhorar as análises futuras. Nesse modelo serão implementados conceitos de Surrogate key( SK) e em algumas tabelas de cadastro como de pessoas, implementado o SCD tipo 6. Esse modelo de forma geral, nunca será apagado e manterá dados históricos pertinentes para análises de hoje e futuras. No ETL, as tabelas desses modelos só serão carregadas depois das STAGE, primeiro as DW de cadastro – dwcd e depois as DW de movimento – dwmv.

Modelagem DW



**Modelagem DM;** aqui os dados virão das tabelas do modelo do DW, e aqui os dados já estão totalmente tratados e limpos para a análise, tanto é que a ferramenta de FRONT END se alimentará dos dados do DM. Aqui o Modelo é totalmente dimensional usando como base o STAR SCHEMA desenvolvido com Kimball. No ETL esse é o último modelo a ser carregado, e será carregado primeiro as dimensões e depois a tabela FATO.

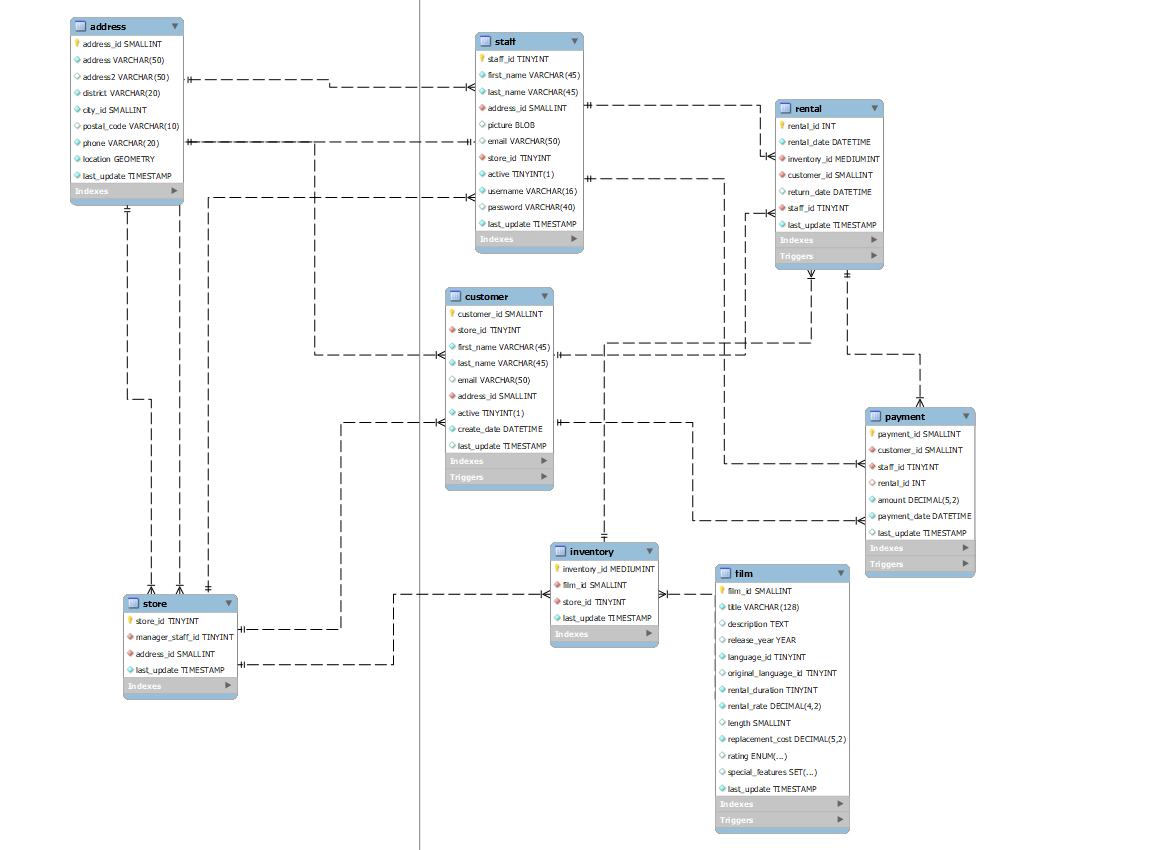
Modelagem DM



**Fontes de Dados:**

Os Modelos apresentados anteriormente foram baseados na modelagem do sistema OLTP do cliente e em um arquivo CSV de cidades atendidas pelo cliente;

Modelagem OLTP



Arquivo CSV GITHUB.

**Automação de ETL (Extração, Transformação e Carga):**

Como citado anteriormente nos Modelos, as cardas de integração de dados ser]ao planejadas pelo Engenheiro de dados;

Teremos dois repositórios principais; DSV e PRD.

DSV; Repositório de desenvolvimento, onde serão feitas as alterações nos mapeamentos e Procedures responsáveis por fazer a integração de dados entre os modelos. Após os mapeamentos serem criados, revisados e homologados, o mapeamento será ESQUEDULADO em um pacote para a definição das prioridades de carga citadas anteriormente (ex. STCD antes das STMV). Depois de ESQUEDULADO este será feito DEPLOY para o repositório PRD.

PRD; Repositório de produção, terá todos os objetos presentes no repositório de desenvolvimento, porém aqui não é possível a manutenção dos mapeamentos e procedures. É desse reportório que a ferramenta de front receberá os dados.

O repositório em DSV rodará nas Quartas e no sábado às 2hr AM e o repositório PRD rodará todos os dias as 2hr AM.

**Visualização de Dados:**

Não realizado.