

Descrição do Modelo Dimensional Estrela do Data Warehouse

Este documento descreve o modelo dimensional em esquema estrela para o Data Warehouse (DW) do grupo de empresas de locação de veículos, detalhando a ligação entre as tabelas do sistema OLTP (Online Transaction Processing) de origem e as tabelas de Fatos e Dimensões do DW, juntamente com a justificativa para cada campo. O modelo visa otimizar as consultas para relatórios gerenciais e análises, conforme os requisitos do projeto.

1. Visão Geral do Modelo Estrela

O modelo estrela é composto por uma ou mais tabelas de fatos no centro, conectadas a múltiplas tabelas de dimensões. As tabelas de dimensões contêm atributos descritivos, enquanto as tabelas de fatos armazenam as medidas numéricas dos eventos de negócio.

Neste DW, identificamos as seguintes tabelas:

- **Tabelas de Dimensão (Atributos Descritivos):**
 - Dim_Tempo
 - Dim_Cliente
 - Dim_Motorista
 - Dim_Patio
 - Dim_Veiculo
- **Tabelas de Fato (Medidas dos Eventos de Negócio):**
 - Fato_Locacao
 - Fato_Reserva
 - Fato_Movimentacao_Patio

2. Ligação entre Fontes de Dados OLTP e Tabelas do DW

As tabelas do DW são populadas a partir das tabelas do esquema OLTP (assumindo o esquema public fornecido na Parte 1, e simulações de outros esquemas através da área de staging). A área de staging atua como uma ponte intermediária para limpeza, padronização e integração de dados de diferentes fontes antes da carga final no DW.

2.1. Fontes para as Dimensões

Dimensão	Tabelas OLTP de Origem (public / staging)	Justificativa da Ligação e Campos
Dim_Tempo	Não diretamente de uma tabela OLTP; gerada programaticamente.	<p>Ligação: População de datas e horas de forma granular.</p> <p>Campos:</p> <p>sk_tempo (Chave Substituta): Identificador único da dimensão.</p> <p>data_completa, ano, trimestre, mes, nome_mes, dia_do_mes, dia_da_semana, nome_dia_da_semana, semana_do_ano, feriado_br, eh_fim_de_semana, hora_completa, hora, minuto:</p> <p>Atributos para permitir análises temporais em diversas granularidades.</p>
Dim_Cliente	public.clientes (e staging.clientes)	<p>Ligação: As informações contratuais e de cobrança dos clientes são extraídas e padronizadas.</p> <p>Campos:</p> <p>sk_cliente (Chave Substituta): Identificador único do cliente no DW.</p> <p>cliente_id_oltp (Chave Natural): clientes.cliente_id da fonte OLTP, para rastreabilidade.</p> <p>nome_completo, cpf_cnpj, tipo_pessoa, email, telefone, cidade_cliente, estado_cliente, data_cadastro: Atributos descritivos do cliente, extraídos diretamente de clientes para análise demográfica e de contato.</p>
Dim_Motorista	public.motoristas (e staging.motoristas)	<p>Ligação: Os dados dos condutores habilitados são extraídos.</p> <p>Campos:</p>

		<p>sk_motorista (Chave Substituta): Identificador único do motorista no DW.</p> <p>motorista_id_oltp (Chave Natural):</p> <p>motoristas.motorista_id da fonte OLTP.</p> <p>sk_cliente_responsavel: FK para Dim_Cliente, permitindo vincular o motorista ao seu cliente responsável (essencial para PJ).</p> <p>nome_completo_motorista, cnh, cnh_categoria, cnh_validade: Atributos descritivos do motorista para análise de habilitação.</p>
Dim_Patio	public.patios (e staging.patios)	<p>Ligação: Informações sobre os locais físicos (pátios) são extraídas.</p> <p>Campos:</p> <p>sk_patio (Chave Substituta): Identificador único do pátio no DW.</p> <p>patio_id_oltp (Chave Natural): patios.patio_id da fonte OLTP.</p> <p>nome_patio, endereco_patio, cidade_patio, estado_patio: Atributos descritivos do pátio; cidade_patio e estado_patio são derivados do endereco_patio para facilitar a análise geográfica.</p>
Dim_Veiculo	public.veiculos, public.grupos_veiculos (e staging.veiculos, staging.grupos_veiculos)	<p>Ligação: Detalhes de cada veículo e seus respectivos grupos são combinados em uma única dimensão conformada.</p> <p>Campos:</p> <p>sk_veiculo (Chave Substituta): Identificador único do veículo no DW.</p> <p>veiculo_id_oltp (Chave Natural): veiculos.veiculo_id da fonte OLTP.</p>

		placa, chassi, marca, modelo, cor, ano_fabricacao, cambio, possui_ar_cond, situacao_atual_oltp: Atributos do veículo. nome_grupo_veiculo, descricao_grupo_veiculo, tarifa_diaria_base_grupo: Atributos do grupo de veículo, extraídos de grupos_veiculos. origem_frota: Atributo derivado (Propria ou Externa, baseado na empresa_id do staging), crucial para o relatório de controle de pátio e para diferenciar a frota.
--	--	---

2.2. Fontes para as Tabelas de Fatos

Fato	Tabelas OLTP de Origem (public / staging)	Justificativa da Ligação e Campos
Fato_Locacao	public.locacoes, public.clientes, public.motoristas, public.veiculos, public.patios (e staging equivalentes)	Ligação: Captura os detalhes de cada aluguel efetivado. Campos: sk_locacao (Chave Substituta): PK da tabela de fatos. sk_tempo_retirada, sk_tempo_devolucao_prevista, sk_tempo_devolucao_real: FKs para Dim_Tempo, permitindo analisar as locações por data de retirada, devolução prevista e real. sk_cliente, sk_motorista, sk_veiculo, sk_patio_retirada, sk_patio_devolucao: FKs para as respectivas dimensões, contextualizando a locação. locacao_id_oltp, reserva_id_oltp (Dimensões Degeneradas): IDs naturais para rastreabilidade de volta ao OLTP; reserva_id_oltp pode

		<p>ser NULL se for "walk-in".</p> <p>valor_previsto, valor_final:</p> <p>Medidas financeiras do aluguel.</p> <p>duracao_em_dias_prevista,</p> <p>duracao_em_dias_real:</p> <p>Medidas calculadas (dias de duração previstos e reais).</p> <p>dias_restantes_para_devolucao:</p> <p>Medida calculada para locações ainda ativas.</p>
Fato_Reserva	<p>public.reservas,</p> <p>public.clientes,</p> <p>public.grupos_veiculos,</p> <p>public.patios (e staging equivalentes)</p>	<p>Ligação: Registra as intenções de locação (reservas).</p> <p>Campos:</p> <p>sk_reserva (Chave Substituta):</p> <p>PK da tabela de fatos.</p> <p>sk_tempo_criacao_reserva,</p> <p>sk_tempo_retirada_prevista,</p> <p>sk_tempo_devolucao_prevista:</p> <p>FKs para Dim_Tempo, para análise temporal das reservas.</p> <p>sk_cliente, sk_veiculo_grupo,</p> <p>sk_patio_retirada: FKs para as dimensões que descrevem quem reservou, qual grupo de veículo e onde.</p> <p>reserva_id_oltp,</p> <p>situacao_reserva_oltp</p> <p>(Dimensões Degeneradas):</p> <p>IDs naturais e status da reserva do OLTP.</p> <p>quantidade_reservas: Medida para contagem de reservas (sempre 1 por linha).</p> <p>duracao_em_dias_prevista:</p> <p>Medida calculada da duração prevista da reserva.</p> <p>dias_ate_retirada_prevista:</p> <p>Medida calculada para analisar a antecedência das reservas.</p>
Fato_Movimentacao_Patio	<p>public.locacoes,</p> <p>public.veiculos, public.patios</p> <p>(e staging equivalentes, focando em devolucao_real)</p>	<p>Ligação: Foco nos movimentos de veículos entre pátios, essencial para a análise de cadeia de Markov.</p>

		<p>Campos:</p> <p>sk_movimentacao_patio (Chave Substituta): PK da tabela de fatos.</p> <p>sk_tempo_devolucao: FK para Dim_Tempo, indicando a data da devolução.</p> <p>sk_veiculo: FK para Dim_Veiculo, identificando o veículo que se moveu.</p> <p>sk_patio_origem, sk_patio_destino: FKs para Dim_Patio, representando o pátio de retirada (origem) e o pátio de devolução (destino).</p> <p>locacao_id_oltp (Dimensão Degenerada): ID da locação do OLTP, para rastreabilidade.</p> <p>quantidade_movimentos: Medida para contagem de movimentos (sempre 1 por linha).</p>
--	--	--

3. Justificativa Detalhada dos Campos

As chaves substitutas (sk_) são inteiros autoincrementáveis, independentes das chaves primárias dos sistemas de origem (chaves naturais _oltp_id). Isso isola o DW de mudanças no esquema OLTP e permite a conformidade de dimensões entre diferentes sistemas de origem.

Dimensões:

- **Dim_Tempo:**
 - sk_tempo (PK): Chave primária substituta.
 - data_completa (DATE): A data em formato YYYY-MM-DD. Única.
 - ano, trimestre, mes, dia_do_mes, dia_da_semana, semana_do_ano (SMALLINT): Componentes numéricos da data para agregações e filtros.
 - nome_mes, nome_dia_da_semana (VARCHAR): Nomes textuais para exibição em relatórios.
 - feriado_br, eh_fim_de_semana (BOOLEAN): Flags para análises específicas de padrões (ex: impacto de feriados/fins de semana nas locações).
 - hora_completa (TIME), hora, minuto (SMALLINT): Componentes de tempo para granularidade horária, se necessária.

- **Dim_Cliente:**
 - sk_cliente (PK): Chave primária substituta.
 - cliente_id_oltp (INTEGER): Chave natural do sistema transacional, para referência.
 - nome_completo (VARCHAR), cpf_cnpj (VARCHAR), tipo_pessoa (CHAR), email (VARCHAR), telefone (VARCHAR): Atributos principais do cliente.
 - cidade_cliente (VARCHAR), estado_cliente (VARCHAR): Dados geográficos para segmentação e análise da origem dos clientes.
 - data_cadastro (TIMESTAMP): Data em que o cliente foi registrado no sistema.
- **Dim_Motorista:**
 - sk_motorista (PK): Chave primária substituta.
 - motorista_id_oltp (INTEGER): Chave natural do sistema transacional.
 - sk_cliente_responsavel (INTEGER): Chave estrangeira para Dim_Cliente, estabelecendo o vínculo.
 - nome_completo_motorista (VARCHAR), cnh (VARCHAR), cnh_categoria (VARCHAR), cnh_validade (DATE): Atributos específicos do motorista.
- **Dim_Patio:**
 - sk_patio (PK): Chave primária substituta.
 - patio_id_oltp (INTEGER): Chave natural do sistema transacional.
 - nome_patio (VARCHAR), endereco_patio (VARCHAR): Dados de identificação e localização.
 - cidade_patio (VARCHAR), estado_patio (VARCHAR): Dados geográficos derivados para análise.
- **Dim_Veiculo:**
 - sk_veiculo (PK): Chave primária substituta.
 - veiculo_id_oltp (INTEGER): Chave natural do sistema transacional.
 - placa (VARCHAR), chassi (VARCHAR), marca (VARCHAR), modelo (VARCHAR), cor (VARCHAR), ano_fabricacao (INTEGER), cambio (VARCHAR), possui_ar_cond (BOOLEAN), situacao_atual_oltp (VARCHAR): Atributos físicos e de status do veículo.
 - nome_grupo_veiculo (VARCHAR), descricao_grupo_veiculo (TEXT), tarifa_diaria_base_grupo (DECIMAL): Atributos do grupo de veículo, promovidos para esta dimensão conformada.
 - origem_frota (VARCHAR): Indica se o veículo é da frota "Propria" (da empresa que é dona do pátio) ou "Externa" (de outra empresa associada), fundamental para o relatório de controle de pátio.

Fatos:

- **Fato_Locacao:**

- sk_locacao (PK): Chave primária da tabela de fatos.
- sk_tempo_retirada, sk_tempo_devolucao_prevista, sk_tempo_devolucao_real (INTEGER): Chaves estrangeiras para Dim_Tempo (granularidade diária), permitindo "fatiar" as locações por diferentes pontos no tempo.
- sk_cliente, sk_motorista, sk_veiculo, sk_patio_retirada, sk_patio_devolucao (INTEGER): Chaves estrangeiras para as dimensões correspondentes, permitindo "fatiar" por cliente, motorista, veículo, e pátios.
- locacao_id_oltp, reserva_id_oltp (INTEGER): Dimensões degeneradas, mantendo o ID original para rastreabilidade sem a necessidade de uma dimensão separada.
- valor_previsto, valor_final (DECIMAL): Medidas aditivas que representam os valores monetários associados à locação.
- duracao_em_dias_prevista, duracao_em_dias_real (DECIMAL): Medidas semianualitivas (podem ser agregadas por soma ou média, mas o "total de dias" é a soma dos dias de cada locação).
- dias_restantes_para_devolucao (DECIMAL): Medida para análise de locações em andamento, indicando a proximidade da devolução.
- **Fato_Reserva:**
 - sk_reserva (PK): Chave primária da tabela de fatos.
 - sk_tempo_criacao_reserva, sk_tempo_retirada_prevista, sk_tempo_devolucao_prevista (INTEGER): Chaves estrangeiras para Dim_Tempo, para análise temporal do ciclo de vida da reserva.
 - sk_cliente, sk_veiculo_grupo, sk_patio_retirada (INTEGER): Chaves estrangeiras para contextualizar a reserva.
 - reserva_id_oltp, situacao_reserva_oltp (VARCHAR): Dimensões degeneradas para rastreabilidade e status da reserva.
 - quantidade_reservas (INTEGER): Medida aditiva, sempre 1 para cada linha de fato.
 - duracao_em_dias_prevista (DECIMAL): Medida semianualitivas da duração prevista da reserva.
 - dias_ate_retirada_prevista (DECIMAL): Medida semianualitivas, indicando a antecedência da reserva.
- **Fato_Movimentacao_Patio:**
 - sk_movimentacao_patio (PK): Chave primária da tabela de fatos.
 - sk_tempo_devolucao (INTEGER): Chave estrangeira para Dim_Tempo, marcando o momento da devolução.
 - sk_veiculo (INTEGER): Chave estrangeira para Dim_Veiculo, identificando qual veículo foi devolvido.
 - sk_patio_origem, sk_patio_destino (INTEGER): Chaves estrangeiras para

Dim_Patio, representando os pátios de retirada e devolução.

- locacao_id_oltp (INTEGER): Dimensão degenerada para rastreabilidade da locação que originou o movimento.
- quantidade_movimentos (INTEGER): Medida aditiva, sempre 1 para cada linha de fato, para contagem de movimentos.

4. Conclusão

O modelo dimensional estrela proposto para o Data Warehouse, com suas tabelas de fatos (Fato_Locacao, Fato_Reserva, Fato_Movimentacao_Patio) e dimensões conformadas (Dim_Tempo, Dim_Cliente, Dim_Motorista, Dim_Patio, Dim_Veiculo), oferece uma arquitetura robusta e flexível para as necessidades de análise do grupo de empresas de locação de veículos.

A separação clara entre dados descritivos (dimensões) e dados de medidas (fatos) otimiza o desempenho das consultas analíticas, permitindo que os relatórios gerenciais globais e as análises preditivas (como a cadeia de Markov) sejam executados de forma eficiente. A inclusão da origem_frota na Dim_Veiculo é um exemplo de como o modelo acomoda requisitos específicos de negócios, transformando dados transacionais em informações estratégicas acionáveis.

Este DW permitirá que a administração tome decisões mais embasadas em dados históricos e tendências unificadas de todas as empresas associadas, superando as limitações dos sistemas OLTP independentes e fornecendo uma visão 360 graus das operações de locação.