

Documentação do Data Warehouse Financeiro

1. Introdução

Este documento descreve a arquitetura, modelagem e dicionário de dados do Data Warehouse desenvolvido para **previsão financeira e detecção de anomalias**.

O Data Warehouse é estruturado em **três camadas lógicas**: - **RAW (Bronze)**: dados brutos, preservados como recebidos das fontes. - **TRUSTED (Silver)**: dados limpos, padronizados e integrados, com integridade referencial. - **REFINED (Gold)**: modelo dimensional (star schema) pronto para análises, dashboards e machine learning.

2. Arquitetura em Camadas

Camada RAW (Bronze)

- Objetivo: armazenar dados brutos para auditoria e reprocessamento.
- Características: nenhum tratamento, valores nulos e inconsistentes permitidos.
- Exemplos: transações financeiras, pagamentos, recebimentos, áreas, fornecedores, categorias, funcionários.

Camada TRUSTED (Silver)

- Objetivo: dados limpos, normalizados e consistentes.
- Características: surrogate keys, chaves estrangeiras, remoção de duplicatas, padronização de datas/moedas/documentos.
- Exemplos: transações financeiras confiáveis, áreas, fornecedores/clientes, categorias contábeis, funcionários.

Camada REFINED (Gold)

- Objetivo: fornecer dados analíticos em modelo dimensional (star schema).
 - Características: dimensões conformadas, tabelas fato, granularidade clara, métricas derivadas.
 - Exemplos: fato financeiro mensal, fato de previsões financeiras, fato de anomalias, dimensões de tempo, áreas, fornecedores, categorias.
-

3. Padrões Globais

- **Naming convention**: raw.<tabela>, trusted.<tabela>, refined.fact_*, refined.dim_*.
- **Surrogate Keys (SK)**: todas as dimensões têm SK (*_sk).
- **Chaves Naturais**: preservadas como atributos (id_*_natural).
- **SCD Tipo 2**: dimensões de negócio preservam histórico com vigencia_inicio, vigencia_fim, is_current.

- **Auditoria/Linhagem:** todas as tabelas têm `etl_batch_id`, `criado_em`, `atualizado_em`, `source_system`.
 - **Particionamento:** tabelas fato particionadas por `ano_mes`.
 - **Padronização:** datas em YYYY-MM-DD, moedas ISO 4217, textos normalizados.
-

4. Dicionário de Dados

4.1 RAW (Bronze)

Objetivo

Armazenar os dados exatamente como vieram da fonte, sem deduplicação ou padronização, mas garantindo rastreabilidade e auditoria.

Todas as tabelas da RAW terão os metadados padrão:

- `ingestion_id` (uuid) → Identificador do lote de ingestão
- `ingestion_ts` (timestamp) → Data/hora da ingestão
- `source_system` (text) → Sistema de origem (ERP, CRM, etc.)
- `source_entity` (text) → Nome do arquivo/tabela/endpoint original
- `row_seq` (bigint) → Número da linha no arquivo/lote
- `raw_row_hash` (text) → Hash da linha original (dedupe/validação)

Tabela: raw.transacoes_financeiras

Descrição:

Registra movimentações financeiras brutas vindas do ERP ou outros sistemas.

Campo	Tipo	Descrição	Observações
id_transacao_raw	int	Identificador único no sistema origem	PK técnica local
data_transacao	text	Data de lançamento (como veio)	Formatos variados
data_competencia	text	Data contábil (como veio)	Pode estar vazia
id_area_raw	int	Código da área/departamento	Não validado
id_fornecedor_raw	int	Código do fornecedor/cliente	Pode não existir
tipo_transacao	text	Receita ou despesa	Pode conter erros
valor_bruto	text	Valor bruto informado	Pode ter separadores errados
valor_liquido	text	Valor líquido informado	Pode estar ausente
moeda	text (até	Código de moeda como veio	Não padronizado

	10)	da fonte	
descricao	text	Texto livre descritivo	Sem restrições
id_categoria_raw	int	Código da categoria contábil	Pode não existir
forma_pagamento	text	Forma de pagamento (ex.: CC, BOL, PIX)	Pode vir abreviado
status_pagamento	text	Pago, pendente, cancelado	Pode estar inconsistente
ingestion_id	uuid	Identificador do lote de ingestão	Auditoria
ingestion_ts	timestamp	Data/hora da ingestão	Auditoria
source_system	text	Sistema de origem	Governança
source_entity	text	Arquivo/API de origem	Linhagem
row_seq	bigint	Número da linha no arquivo	Auditoria
raw_row_hash	text	Hash da linha original	Dedupe

Tabela: raw.areas

Descrição:

Lista bruta de áreas/departamentos extraída dos sistemas de origem.

Campo	Tipo	Descrição	Observações
id_area_raw	int	Identificador único	PK técnica local
codigo_area	text	Código como veio da fonte	Não padronizado
nome_area	text	Nome da área	Pode ter abreviações
gestor_responsavel	text	Nome do gestor responsável	Texto livre
email_gestor	text	Email do gestor	Pode estar vazio
ingestion_id	uuid	Identificador do lote de ingestão	Auditoria
ingestion_ts	timestamp	Data/hora da ingestão	Auditoria
source_system	text	Sistema de origem	Governança
source_entity	text	Arquivo/API de origem	Linhagem
row_seq	bigint	Número da linha no arquivo	Auditoria

raw_row_hash	text	Hash da linha original	Dedupe
---------------------	------	------------------------	--------

Tabela: raw.fornecedores_clientes

Descrição:

Cadastro bruto de fornecedores e clientes.

Campo	Tipo	Descrição	Observações
id_fornecedor_raw	int	Identificador único	PK técnica local
nome_fornecedor	text	Nome como veio	Texto livre
tipo_fornecedor	text	Cliente ou fornecedor	Pode estar inconsistente
cnpj_cpf	text	Documento	Sem validação
contato	text	Telefone/email	Pode ser qualquer formato
endereco	text	Endereço completo	Não normalizado
ingestion_id	uuid	Identificador do lote de ingestão	Auditoria
ingestion_ts	timestamp	Data/hora da ingestão	Auditoria
source_system	text	Sistema de origem	Governança
source_entity	text	Arquivo/API de origem	Linhagem
row_seq	bigint	Número da linha no arquivo	Auditoria
raw_row_hash	text	Hash da linha original	Dedupe

Tabela: raw.categorias_contabeis

Descrição:

Lista bruta de categorias contábeis.

Campo	Tipo	Descrição	Observações
id_categoria_raw	int	Identificador único	PK técnica local
nome_categoria	text	Nome como veio	Texto livre
tipo_categoria	text	Receita ou gasto	Pode estar inconsistente
codigo_contabil	text	Código contábil da fonte	Não padronizado
ingestion_id	uuid	Identificador do lote de ingestão	Auditoria
ingestion_ts	timestamp	Data/hora da ingestão	Auditoria
source_system	text	Sistema de origem	Governança
source_entity	text	Arquivo/API de origem	Linhagem
row_seq	bigint	Número da linha no arquivo	Auditoria
raw_row_hash	text	Hash da linha original	Dedupe

Tabela: raw.pagamentos

Descrição:

Registros brutos de pagamentos.

Campo	Tipo	Descrição	Observações
id_pagamento_raw	int	Identificador único	PK técnica local
id_transacao_raw	int	ID da transação vinculada	Referência bruta
data_pagamento	text	Data do pagamento	Formato livre
valor_pago	text	Valor pago	Sem padronização
metodo_pagamento	text	Método de pagamento	Ex.: CC, Boleto
comprovante	text	Número/Hash do comprovante	Texto livre

ingestion_id	uuid	Identificador do lote de ingestão	Auditoria
ingestion_ts	timestamp	Data/hora da ingestão	Auditoria
source_system	text	Sistema de origem	Governança
source_entity	text	Arquivo/API de origem	Linhagem
row_seq	bigint	Número da linha no arquivo	Auditoria
raw_row_hash	text	Hash da linha original	Dedupe

Tabela: raw.recebimentos

Descrição:

Registros brutos de recebimentos.

Campo	Tipo	Descrição	Observações
id_recebimento_raw	int	Identificador único	PK técnica local
id_transacao_raw	int	ID da transação vinculada	Referência bruta
data_recebimento	text	Data do recebimento	Formato livre
valor_recebido	text	Valor recebido	Sem padronização
metodo_recebimento	text	Método de recebimento	Ex.: PIX, TED
comprovante	text	Número/Hash do comprovante	Texto livre
ingestion_id	uuid	Identificador do lote de ingestão	Auditoria
ingestion_ts	timestamp	Data/hora da ingestão	Auditoria
source_system	text	Sistema de origem	Governança
source_entity	text	Arquivo/API de origem	Linhagem
row_seq	bigint	Número da linha no arquivo	Auditoria

raw_row_hash	text	Hash da linha original	Dedupe
---------------------	------	------------------------	--------

Tabela: raw.funcionarios

Descrição:

Cadastro bruto de funcionários.

Campo	Tipo	Descrição	Observações
id_funcionario_raw	int	Identificador único	PK técnica local
nome	text	Nome completo	Texto livre
cpf	text	Documento	Não validado
cargo	text	Cargo atual	Texto livre
departamento	text	Departamento	Texto livre
data_admissao	text	Data de entrada	Formato livre
salario	text	Valor do salário	Pode vir em diferentes formatos
tipo_contrato	text	CLT, PJ, estagiário	Texto livre
ingestion_id	uuid	Identificador do lote de ingestão	Auditoria
ingestion_ts	timestamp	Data/hora da ingestão	Auditoria
source_system	text	Sistema de origem	Governança
source_entity	text	Arquivo/API de origem	Linhagem
row_seq	bigint	Número da linha no arquivo	Auditoria
raw_row_hash	text	Hash da linha original	Dedupe

4.2 TRUSTED (Silver)

Objetivo

Armazenar os dados **limpos, padronizados e integrados**, prontos para consumo corporativo.

Nesta camada já foram aplicadas as principais **regras de qualidade e governança**:

- Padronização de formatos (datas, moedas, valores numéricos).
- Deduplicação e consolidação de registros.
- Enriquecimento com atributos de referência.
- Criação de **Surrogate Keys (SK)** para unificação.
- Preservação das **chaves naturais** para rastreabilidade (id_*_natural).
- Implementação de **SCD Tipo 2** em dimensões de negócio (histórico).

A camada **TRUSTED** é a “verdade única corporativa” (Single Source of Truth) dos dados transacionais e cadastrais, servindo como base para a modelagem **Refined (dimensional)**.

Metadados padrão

Todas as tabelas da TRUSTED terão os seguintes campos de auditoria e linhagem:

- **etl_batch_id (uuid)** → Identificador do lote de processamento ETL.
- **criado_em (timestamp)** → Data/hora da inserção do registro na tabela.
- **atualizado_em (timestamp)** → Data/hora da última atualização.
- **source_system (text)** → Sistema de origem consolidado (ERP, CRM, RH etc.).

Além disso, em dimensões de negócio, utilizamos:

- **vigencia_inicio (date)** → Data de início da validade do registro (SCD2).
- **vigencia_fim (date)** → Data de fim da validade do registro (SCD2).
- **is_current (boolean)** → Flag indicando se o registro é a versão atual (SCD2)

Tabela: trusted.areas

Descrição: Lista oficial de áreas/departamentos da empresa.

Campo	Tipo	Descrição	Observações
area_sk	bigint (PK)	Surrogate Key (chave substituta)	PK
id_area_natural	text	Código/número da área no ERP	Chave natural

codigo_area	text	Código padronizado da área	Único
nome_area	text	Nome oficial da área	Normalizado
gestor_responsavel	text	Nome completo do gestor	Texto limpo
email_gestor	text	Email válido do gestor	Validado
vigencia_inicio	date	Data de início da validade (SCD2)	Histórico
vigencia_fim	date	Data de fim da validade (SCD2)	Histórico
is_current	boolean	Indica se é a versão atual	SCD2
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora da criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: *trusted.fornecedores_clientes*

Descrição: Cadastro único de fornecedores e clientes, padronizado e validado.

Campo	Tipo	Descrição	Observações
parceiro_sk	bigint (PK)	Surrogate Key	PK

id_parceiro_natural	text	Código/ID no sistema origem	Chave natural
nome_fornecedor	text	Nome oficial	Normalizado
tipo_fornecedor	text	CLIENTE ou FORNECEDOR	Padronizado
documento	text	CPF/CNPJ validado	Único
contato_email	text	Email validado	Padronizado
contato_telefone	text	Telefone padronizado	Normalizado
endereco	text	Endereço padronizado	Texto limpo
pais	text	País	Padronizado
estado	text	Estado	Padronizado
cidade	text	Cidade	Padronizado
setor_atuacao	text	Setor econômico	Opcional
rating_credito	text	Avaliação de crédito	Opcional
prazo_medio	int	Prazo médio (dias)	Opcional
vigencia_inicio	date	Data de início (SCD2)	Histórico
vigencia_fim	date	Data de fim (SCD2)	Histórico
is_current	boolean	Registro atual?	SCD2
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora de criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: trusted.categorias_contabeis

Descrição: Categorias contábeis unificadas e validadas.

Campo	Tipo	Descrição	Observações
categoria_sk	bigint (PK)	Surrogate Key	PK
id_categoria_natural	text	Código da categoria no ERP	Chave natural
nome_categoria	text	Nome oficial	Normalizado
tipo_categoria	text	RECEITA ou DESPESA	Padronizado
codigo_contabil	text	Código contábil oficial	Padronizado
vigencia_inicio	date	Data de início (SCD2)	Histórico
vigencia_fim	date	Data de fim (SCD2)	Histórico
is_current	boolean	Registro atual?	SCD2
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora de criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: trusted.funcionarios

Descrição: Cadastro oficial de funcionários.

Campo	Tipo	Descrição	Observações
funcionario_sk	bigint (PK)	Surrogate Key	PK
id_funcionario_natural	text	Código/matrícula do funcionário no RH	Chave natural
nome	text	Nome completo	Normalizado

cpf	text	CPF validado	Único
cargo	text	Cargo oficial	Padronizado
area_sk	bigint (FK)	FK para trusted.areas	Relacionamento
data_admissao	date	Data de admissão	Validado
data_demissao	date	Data de desligamento	Pode ser NULL
status_funcionario	text	ATIVO ou INATIVO	Padronizado
tipo_contrato	text	CLT, PJ, ESTÁGIO	Padronizado
vigencia_inicio	date	Data de início (SCD2)	Histórico
vigencia_fim	date	Data de fim (SCD2)	Histórico
is_current	boolean	Registro atual?	SCD2
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora de criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: trusted.transacoes_financeiras

Descrição: Movimentações financeiras já limpas e padronizadas.

Campo	Tipo	Descrição	Observações
transacao_sk	bigint (PK)	Surrogate Key	PK
id_transacao_natural	text	Código da transação no ERP	Chave natural
area_sk	bigint (FK)	Área da empresa	Relacionamento
parceiro_sk	bigint (FK)	Cliente/Fornecedor	Relacionamento

categoria_sk	bigint (FK)	Categoria contábil	Relacionamento
funcionario_sk	bigint (FK)	Funcionário responsável	Relacionamento
moeda_sk	bigint (FK)	Moeda da transação (ISO 4217)	Relacionamento
data_transacao	date	Data da transação	Padronizada
data_competencia	date	Data contábil	Padronizada
valor_bruto	numeric(18,2)	Valor bruto	Padronizado
valor_liquido	numeric(18,2)	Valor líquido	Padronizado
taxa_cambio	numeric(18,6)	Taxa de câmbio	Opcional
tipo_transacao	text	RECEITA ou DESPESA	Padronizado
forma_pagamento	text	Método de pagamento padronizado	Validado
status_pagamento	text	Pago, Pendente, Cancelado	Padronizado
status_contabil	text	Lançado, Apurado, Fechado	Padronizado
numero_documento	text	Nota/fatura	Degenerate dimension
descricao_limpa	text	Descrição da transação	Texto limpo
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora de criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: trusted.pagamentos

Descrição: Pagamentos validados e vinculados a transações.

Campo	Tipo	Descrição	Observações
pagamento_sk	bigint (PK)	Surrogate Key	PK
id_pagamento_natural	text	Código/ID no sistema origem	Chave natural
transacao_sk	bigint (FK)	FK para transações financeiras	Relacionamento
data_pagamento	date	Data efetiva do pagamento	Padronizada
valor_pago	numeric(18,2)	Valor pago	Padronizado
metodo_pagamento	text	Método padronizado	Validado
moeda_sk	bigint (FK)	Moeda utilizada	Relacionamento
comprovante	text	Identificador do comprovante	Degenerate dimension
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora de criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: trusted.recebimentos

Descrição: Recebimentos validados e vinculados a transações.

Campo	Tipo	Descrição	Observações
recebimento_sk	bigint (PK)	Surrogate Key	PK
id_recebimento_natural	text	Código/ID no sistema origem	Chave natural
transacao_sk	bigint (FK)	FK para transações financeiras	Relacionamento
data_recebimento	date	Data efetiva do recebimento	Padronizada

valor_recebido	numeric(18,2)	Valor recebido	Padronizado
metodo_recebimento	text	Método padronizado	Validado
moeda_sk	bigint (FK)	Moeda utilizada	Relacionamento
comprovante	text	Identificador do comprovante	Degenerate dimension
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora de criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

4.3 REFINED (Gold)

Objetivo

Armazenar os dados em formato **dimensional (Star Schema)**, otimizados para análise, relatórios, previsão financeira e detecção de anomalias.

Nesta camada:

- **Tabelas fato** concentram **medidas numéricas** (valores financeiros, quantidades, saldos).
- **Tabelas dimensão** armazenam **atributos descritivos** (datas, áreas, fornecedores, moedas, funcionários etc.).
- O modelo é projetado para **consultas rápidas, consistentes e escaláveis**, com foco em **BI, Analytics e Data Science**.

A camada REFINED é a que alimenta dashboards (Power BI, Tableau, Superset), APIs analíticas e modelos de machine learning.

Padrões de nomenclatura

- **refined.fact_*** → Tabelas fato.
- **refined.dim_*** → Tabelas dimensão.

Metadados padrão

Todas as tabelas da REFINED (fato e dimensão) possuem colunas de auditoria e linhagem:

- **etl_batch_id (uuid)** → Identificador do lote de processamento ETL.
- **criado_em (timestamp)** → Data/hora da inserção do registro.
- **atualizado_em (timestamp)** → Data/hora da última atualização.
- **source_system (text)** → Sistema(s) de origem do dado.

Adicional em dimensões (SCD Tipo 2):

- **vigencia_inicio (date)** → Data de início da validade do registro.
- **vigencia_fim (date)** → Data de fim da validade do registro.
- **is_current (boolean)** → Indica se o registro é a versão atual.

Adicional em fatos:

- **Grão (granularidade)** deve estar claramente definido (ex.: transação, pagamento, recebimento, saldo diário, etc.).
- Sempre usam **chaves substitutas (SK)** das dimensões, nunca chaves naturais.
- Devem conter **medidas numéricas aditivas ou semi-aditivas** (valores, quantidades).

Boas práticas na REFINED

1. **Surrogate Keys (SKs)** → Todas as dimensões têm SK (chave técnica).
 2. **Histórico preservado (SCD2)** → Para permitir análises temporais corretas.
 3. **Degenerate Dimensions (DD)** → IDs de documentos (ex.: número de fatura, nota fiscal) podem ficar na fato sem gerar uma dimensão própria.
 4. **Granularidade clara** → Cada fato deve ter um grão bem definido (ex.: fact_transacao no nível de cada lançamento financeiro).
 5. **Conversão de moeda** → Dimensão de moedas + tabela de taxas de câmbio permitem consolidação em moeda corporativa.
 6. **Performance** → Particionamento por data, compressão e índices em SKs.
-

REFINED - Dimensões

Tabela: refined.dim_tempo

Descrição: Dimensão de tempo padronizada, utilizada em todas as análises.

Campo	Tipo	Descrição	Observações
tempo_sk	bigint (PK)	Surrogate Key	PK
data	date	Data completa (YYYY-MM-DD)	Única
ano	int	Ano	Derivado
semestre	int	Semestre (1 ou 2)	Derivado
trimestre	int	Trimestre (1 a 4)	Derivado
mes	int	Número do mês (1 a 12)	Derivado
nome_mes	text	Nome do mês (Janeiro, Fevereiro, ...)	Padronizado
dia	int	Dia do mês	Derivado
dia_semana	int	Dia da semana (1=Segunda, 7=Domingo)	Padronizado
nome_dia_semana	text	Nome do dia da semana	Padronizado
semana_ano	int	Semana do ano (1 a 53)	Derivado
eh_feriado	boolean	Indicador de feriado nacional/local	Opcional
etl_batch_id	uuid	Identificador do lote ETL	Auditoria

criado_em	timestamp	Data/hora de criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: refined.dim_tempo

Descrição: Dimensão de tempo padronizada, utilizada em todas as análises.

Campo	Tipo	Descrição	Observações
tempo_sk	bigint (PK)	Surrogate Key	PK
data	date	Data completa (YYYY-MM-DD)	Única
ano	int	Ano	Derivado
semestre	int	Semestre (1 ou 2)	Derivado
trimestre	int	Trimestre (1 a 4)	Derivado
mes	int	Número do mês (1 a 12)	Derivado
nome_mes	text	Nome do mês (Janeiro, Fevereiro, ...)	Padronizado
dia	int	Dia do mês	Derivado
dia_semana	int	Dia da semana (1=Segunda, 7=Domingo)	Padronizado
nome_dia_semana	text	Nome do dia da semana	Padronizado
semana_ano	int	Semana do ano (1 a 53)	Derivado

eh_feriado	boolean	Indicador de feriado nacional/local	Opcional
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora de criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: refined.dim_area

Descrição: Dimensão de áreas/departamentos.

Campo	Tipo	Descrição	Observações
area_sk	bigint (PK)	Surrogate Key	PK
id_area_natural	text	Código da área no sistema origem	Chave natural
codigo_area	text	Código padronizado da área	Único
nome_area	text	Nome oficial da área	Normalizado
gestor_responsavel	text	Nome do gestor	Texto limpo
email_gestor	text	Email do gestor	Validado
vigencia_inicio	date	Data início validade (SCD2)	Histórico
vigencia_fim	date	Data fim validade (SCD2)	Histórico
is_current	boolean	Registro atual?	SCD2
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria

source_system	text	Sistema de origem	Governança
----------------------	------	-------------------	------------

Tabela: refined.dim_fornecedor_cliente

Descrição: Dimensão de fornecedores e clientes.

Campo	Tipo	Descrição	Observações
parceiro_sk	bigint (PK)	Surrogate Key	PK
id_parceiro_natural	text	Código no sistema origem	Chave natural
nome_parceiro	text	Nome oficial	Normalizado
tipo_parceiro	text	CLIENTE ou FORNECEDOR	Padronizado
documento	text	CPF/CNPJ validado	Único
contato_email	text	Email validado	Padronizado
contato_telefone	text	Telefone padronizado	Normalizado
endereco	text	Endereço	Padronizado
pais	text	País	Padronizado
estado	text	Estado	Padronizado
cidade	text	Cidade	Padronizado
setor_atuacao	text	Setor econômico	Opcional
rating_credito	text	Avaliação de crédito	Opcional
prazo_medio	int	Prazo médio em dias	Opcional
vigencia_inicio	date	Data início validade (SCD2)	Histórico
vigencia_fim	date	Data fim validade (SCD2)	Histórico
is_current	boolean	Registro atual?	SCD2

etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: refined.dim_categoria_contabil

Descrição: Dimensão de categorias contábeis.

Campo	Tipo	Descrição	Observações
categoria_sk	bigint (PK)	Surrogate Key	PK
id_categoria_natural	text	Código no ERP	Chave natural
nome_categoria	text	Nome oficial	Normalizado
tipo_categoria	text	RECEITA ou DESPESA	Padronizado
codigo_contabil	text	Código contábil	Padronizado
vigencia_inicio	date	Data início validade (SCD2)	Histórico
vigencia_fim	date	Data fim validade (SCD2)	Histórico
is_current	boolean	Registro atual?	SCD2
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: refined.dim_funcionario

Descrição: Dimensão de funcionários.

Campo	Tipo	Descrição	Observações
funcionario_sk	bigint (PK)	Surrogate Key	PK
id_funcionario_natural	text	Matrícula/ID no sistema origem	Chave natural
nome	text	Nome completo	Normalizado
cpf	text	CPF validado	Único
cargo	text	Cargo oficial	Padronizado
area_sk	bigint (FK)	FK para dim_area	Relacionamento
data_admissao	date	Data de admissão	Validado
data_demissao	date	Data de desligamento	Pode ser nulo
status_funcionario	text	ATIVO ou INATIVO	Padronizado
tipo_contrato	text	CLT, PJ, Estágio	Padronizado
vigencia_inicio	date	Data início validade (SCD2)	Histórico
vigencia_fim	date	Data fim validade (SCD2)	Histórico
is_current	boolean	Registro atual?	SCD2
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: refined.dim_moeda

Descrição: Dimensão de moedas conforme padrão ISO 4217.

Campo	Tipo	Descrição	Observações
moeda_sk	bigint (PK)	Surrogate Key	PK
codigo_iso	char(3)	Código ISO 4217 (ex.: BRL, USD)	Único
nome_moeda	text	Nome oficial (Real, Dólar, etc.)	Padronizado
simbolo	text	Símbolo monetário (R\$, \$, €)	Opcional
pais_referencia	text	País de referência	Opcional
vigencia_inicio	date	Data início validade (SCD2)	Histórico
vigencia_fim	date	Data fim validade (SCD2)	Histórico
is_current	boolean	Registro atual?	SCD2
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

REFINED – Fatos

Tabela: refined.fact_transacoes_financeiras

Descrição: Fato principal com granularidade em cada transação financeira.

Campo	Tipo	Descrição	Observações
transacao_sk	bigint (PK)	Surrogate Key	PK
tempo_sk	bigint (FK)	Data da transação (dim_tempo)	Relacionamento

area_sk	bigint (FK)	Área da empresa (dim_area)	Relacionamento
parceiro_sk	bigint (FK)	Cliente/Fornecedor (dim_fornecedor)	Relacionamento
categoria_sk	bigint (FK)	Categoria contábil (dim_categoria)	Relacionamento
funcionario_sk	bigint (FK)	Funcionário responsável	Relacionamento
moeda_sk	bigint (FK)	Moeda (dim_moeda)	Relacionamento
valor_bruto	numeric(18,2)	Valor bruto	Medida
valor_liquido	numeric(18,2)	Valor líquido	Medida
taxa_cambio	numeric(18,6)	Taxa de câmbio	Opcional
tipo_transacao	text	RECEITA ou DESPESA	Degenerate dimension
forma_pagamento	text	Método de pagamento	Degenerate dimension
status_pagamento	text	Pago, Pendente, Cancelado	Degenerate dimension
numero_documento	text	Número da nota/fatura	Degenerate dimension
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: refined.fact_pagamentos

Descrição: Fato de pagamentos, granularidade em cada evento de pagamento.

Campo	Tipo	Descrição	Observações
pagamento_sk	bigint (PK)	Surrogate Key	PK
tempo_sk	bigint (FK)	Data do pagamento (dim_tempo)	Relacionamento
transacao_sk	bigint (FK)	Transação vinculada	Relacionamento
parceiro_sk	bigint (FK)	Cliente/Fornecedor (dim_fornecedor)	Relacionamento
funcionario_sk	bigint (FK)	Funcionário responsável	Relacionamento
moeda_sk	bigint (FK)	Moeda utilizada	Relacionamento
valor_pago	numeric(18,2)	Valor pago	Medida
metodo_pagamento	text	Método de pagamento	Degenerate dimension
comprovante	text	Comprovante do pagamento	Degenerate dimension
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: refined.fact_recebimentos

Descrição: Fato de recebimentos, granularidade em cada evento de recebimento.

Campo	Tipo	Descrição	Observações
recebimento_sk	bigint (PK)	Surrogate Key	PK
tempo_sk	bigint (FK)	Data do recebimento (dim_tempo)	Relacionamento

transacao_sk	bigint (FK)	Transação vinculada	Relacionamento
parceiro_sk	bigint (FK)	Cliente/Fornecedor	Relacionamento
funcionario_sk	bigint (FK)	Funcionário responsável	Relacionamento
moeda_sk	bigint (FK)	Moeda utilizada	Relacionamento
valor_recebido	numeric(18,2)	Valor recebido	Medida
metodo_recebimento	text	Método de recebimento	Degenerate dimension
comprovante	text	Comprovante do recebimento	Degenerate dimension
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

Tabela: refined.fact_saldos_diarios

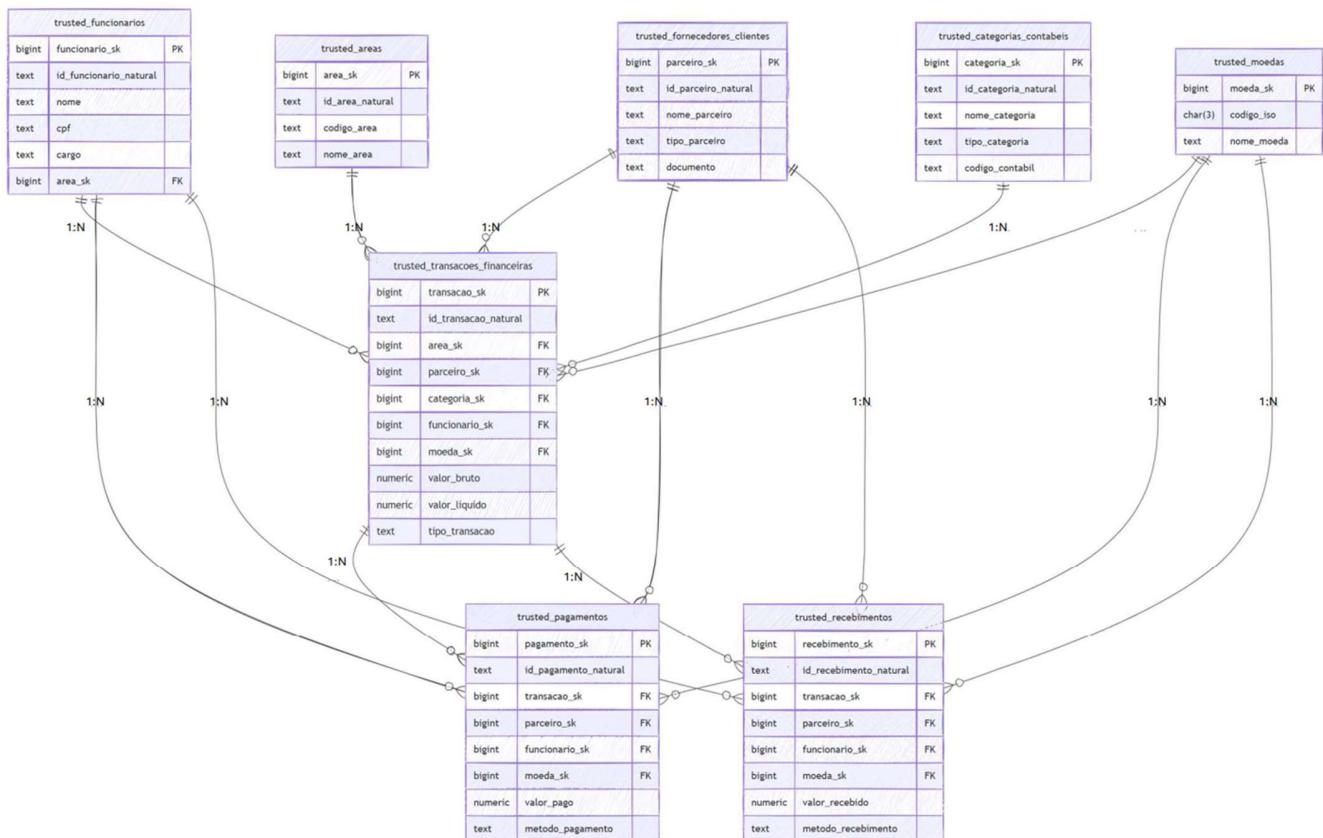
Descrição: Snapshot de saldos consolidados por dia e moeda.

Campo	Tipo	Descrição	Observações
saldo_sk	bigint (PK)	Surrogate Key	PK
tempo_sk	bigint (FK)	Data (dim_tempo)	Relacionamento
area_sk	bigint (FK)	Área/Unidade de negócio	Relacionamento
moeda_sk	bigint (FK)	Moeda	Relacionamento
saldo_inicial	numeric(18,2)	Saldo inicial do dia	Medida
saldo_final	numeric(18,2)	Saldo final do dia	Medida
entradas	numeric(18,2)	Total recebido no dia	Medida

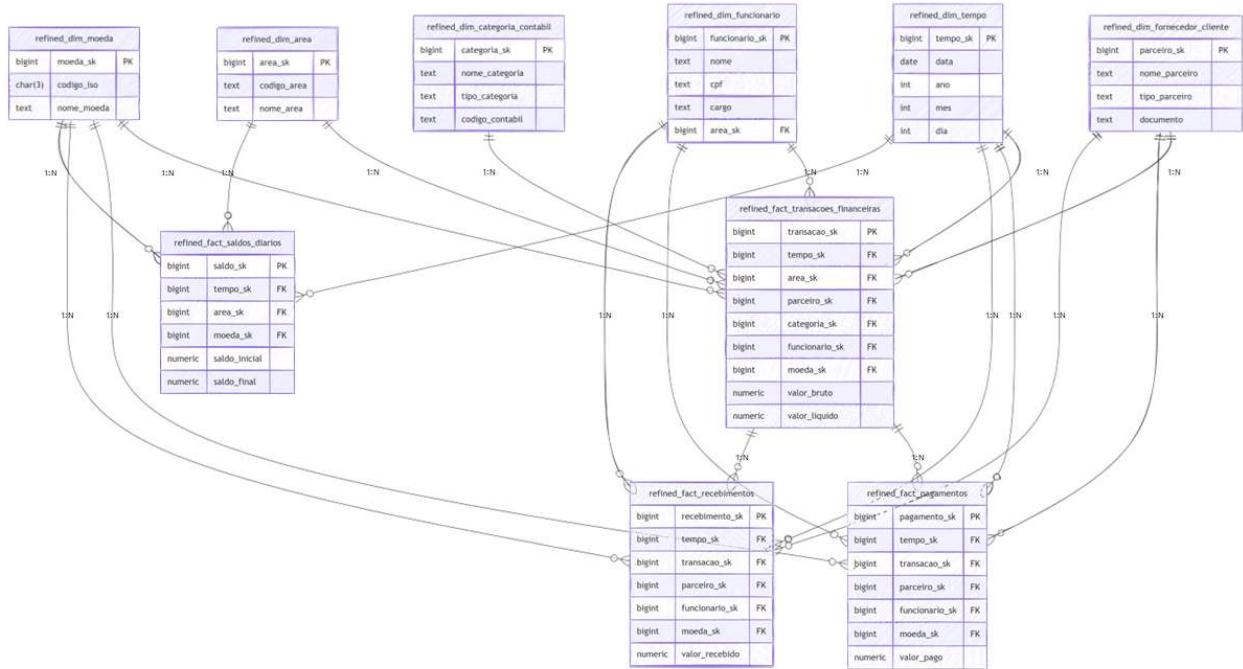
saidas	numeric(18,2)	Total pago no dia	Medida
etl_batch_id	uuid	Identificador do lote ETL	Auditoria
criado_em	timestamp	Data/hora criação no DW	Auditoria
atualizado_em	timestamp	Última atualização	Auditoria
source_system	text	Sistema de origem	Governança

5. DERs e Modelos

- Trusted



- **Refined**



6. Processos e Governança

- **Transformação RAW → TRUSTED**: limpeza, padronização, deduplicação, criação de SK e FKs.
- **Transformação TRUSTED → REFINED**: agregações, cálculos de métricas, inserção de previsões e anomalias.
- **Qualidade de dados**: unicidade, preenchimento obrigatório, validação de chaves.
- **Histórico (SCD2)**: preservação de alterações em dimensões.
- **Auditoria**: ETL batch, timestamps, fonte de origem.
- **Performance**: particionamento de fatos, índices em FKs, materialized views.