

**amcom**

inspired by tech, powered by people

## Sumário

---

<b>1.</b>	<b>Visão Geral.....</b>	<b>3</b>
<b>2.</b>	<b>Origem.....</b>	<b>3</b>
2.1	Tabela.....	3
2.2	Qualidade origem.....	5
<b>3.</b>	<b>Raw.....</b>	<b>6</b>
3.1	Tabela.....	6
<b>4.</b>	<b>Trusted .....</b>	<b>7</b>
4.1	Tabela.....	8
<b>5.</b>	<b>Refined .....</b>	<b>8</b>
5.1	Tabela.....	8
<b>6.</b>	<b>QUALIDADE.....</b>	<b>8</b>
<b>7.</b>	<b>DETALHAMENTO .....</b>	<b>9</b>

## 1. Visão Geral

---

Este documento apresenta a arquitetura de dados da tabela **Clientes**, especificando as fontes de origem e as estruturas existentes em cada camada do pipeline de tratamento e processamento de dados.

### Clientes:

*“Cadastro de Clientes. Tabela apenas para identificação e análise de particularidades de clientes caso seja necessário”.*

## 2. Origem

---

**Banco:** SQL Server

**Base de dados:** DBO

**Tabela:** Clientes

**Select simples:** `select * from dbo.clientes`

### 2.1 Tabela

A tabela abaixo contempla a estrutura atual da fonte Clientes com a definição de tipo de dado correspondente, onde:

- **Coluna:**  
Representa o nome de cada coluna disponível na tabela original.
- **Tipo:**  
Apresenta qual o tipo de dado da coluna relacionada, exemplo: int, varchar, etc.
- **Necessária:**  
Campo preenchido pela ADAMI, onde informa se a coluna relacionada será necessária para execução do projeto.
- **Ação Sugerida:**  
Sugestão preenchida pela AmCom, onde informa se a coluna relacionada deveria ser mantida ou excluída. Para isto foram considerados os seguintes critérios de exclusão:  
  
Muitos Nulos: > 90% de valores nulos  
Valor Único: Coluna possui apenas 1 valor único  
Muitos zeros: > 80% de valores zero (para colunas numéricas)  
Strings vazias: > 80% de strings vazias (para colunas de texto)  
  
Colunas que não atendem a esses critérios são mantidas.
- **Observação:**  
Observações relacionadas aos dados e incluídas pela AmCom e/ou ADAMI.

Coluna	Tipo	Necessária?	Ação Sugerida (AmCom)	Observação
IDCliente	int	Sim	Manter	
Cliente	varchar	Sim	Manter	
CodCliente	varchar	Sim	Manter	
CodRepresentante	int	Sim	Manter	
TolMais	tinyint	Não	Manter	
TolMenos	tinyint	Não	Manter	
TipoObs1	varchar	Não	Excluir	
Obs1	varchar	Não	Excluir	
TipoObs2	varchar	Não	Excluir	
Obs2	varchar	Não	Excluir	
TipoObs3	varchar	Não	Excluir	
Obs3	varchar	Não	Excluir	
TipoObs4	varchar	Não	Excluir	
Obs4	varchar	Não	Excluir	
TipoABC	varchar	Não	Excluir	
CodSegmento	varchar	Não	Excluir	
PreferencialSN	bit	Não	Excluir	
ClienteListaSN	bit	Não	Excluir	
Agrupamento	varchar	Não	Excluir	
Cliente_Fornecedor	tinyint	Não	Excluir	
DescrCliente_Fornecedor	varchar	Não	Excluir	
RazaoSocial	varchar	Não	Manter	
CodigoERP	varchar	Não	Excluir	
Etiqueta	varchar	Não	Excluir	
ExigeLaudo	bit	Não	Manter	
Laudo	varchar	Não	Excluir	
EmailLaudo	varchar	Não	Excluir	
CartaOrcamento	varchar	Não	Excluir	
NrEtiquetas	tinyint	Não	Excluir	
IDUnicoRegistro	uniqueidentifier	Não	Manter	
AceitaPaqueteIncompleto	tinyint	Não	Manter	
DescAceitaPaqueteIncompleto	varchar	Não	Manter	
idAtendenteVenda	int	Não	Excluir	
ContatoFaturamento	varchar	Não	Excluir	
IDClassFiscal	int	Não	Excluir	
DataCriacao	datetime	Não	Manter	
UsuarioCriacao	varchar	Não	Excluir	
DataUltModif	datetime	Não	Manter	
UsuarioUltModif	varchar	Não	Excluir	
RazaoTravaComposicao	varchar	Não	Excluir	
CustoFinanceiroAdicional	float	Não	Excluir	
CustoAdicTransportePorPaquete	float	Não	Excluir	

CustoAdicTransportePorKg_Paletizado	float	Não	Excluir	
CustoAdicTransportePorKg_NaoPaletizado	float	Não	Excluir	
CustoAdicTransportePorM2_Paletizado	float	Não	Excluir	
CustoAdicTransportePorM2_NaoPaletizado	float	Não	Excluir	
CustoAdicTransportePorPeca_Paletizado	float	Não	Excluir	
CustoAdicTransportePorPeca_NaoPaletizado	float	Não	Excluir	
UsoGeralStr1	varchar	Não	Excluir	
UsoGeralStr2	varchar	Não	Excluir	
UsoGeralNum1	float	Não	Excluir	
UsoGeralNum2	float	Não	Excluir	
ExigeChapaUnitizada	bit	Não	Excluir	
ExigeAgendamentoEntregas	bit	Não	Excluir	
ClienteProspecto	bit	Não	Excluir	
ImprimeLogotipo	bit	Não	Excluir	
Paletizado	smallint	Não	Excluir	
FacesOpostas	smallint	Não	Excluir	
Espelho	smallint	Não	Excluir	
Cantoneiras	smallint	Não	Excluir	
Filme	smallint	Não	Excluir	
OrelhasInvertidas	smallint	Não	Excluir	
PacotesLargura	int	Não	Excluir	
PacotesComprimento	int	Não	Excluir	
PacotesAltura	int	Não	Excluir	
IDPaleta	int	Não	Excluir	
EmpilhamentoMaximo	int	Não	Excluir	
PressaoDeArqueamento	int	Não	Excluir	
FitasLargPaleta	int	Não	Excluir	
FitasCompPaleta	int	Não	Excluir	

## 2.2 Qualidade origem

Com o intuito de compreender e gerenciar a qualidade dos dados disponíveis na origem, foi conduzido um trabalho de validação detalhado. Cada coluna foi avaliada individualmente quanto à completude, consistência, padronização e presença de valores inválidos ou ausentes. Esse processo permitiu identificar inconsistências e definir as regras de limpeza e transformação necessárias para garantir a confiabilidade dos dados nas etapas subsequentes do pipeline.

O artefato correspondente a esta validação está disponível em:

[https://github.com/Adami-S-A/Projeto\\_IA\\_AMCOM/tree/main/project\\_data\\_science/docs/data\\_quality](https://github.com/Adami-S-A/Projeto_IA_AMCOM/tree/main/project_data_science/docs/data_quality)

### 3. Raw

A camada *Raw* é a primeira etapa do pipeline de dados, onde as informações são armazenadas exatamente como são recebidas da fonte, sem qualquer tipo de transformação ou limpeza. Essa camada preserva o dado em seu formato original, incluindo possíveis inconsistências, duplicidades ou valores faltantes. Seu principal objetivo é servir como backup confiável do dado bruto, garantindo rastreabilidade e permitindo auditoria e reconstrução de dados caso seja necessário.

#### 3.1 Tabela

A tabela abaixo contempla a estrutura disponibilizada na camada *raw* para a tabela Clientes.

Nome	Tipo
IDCLIENTE	NUMBER
CLIENTE	VARCHAR2(50)
CODCLIENTE	VARCHAR2(30)
CODREPRESENTANTE	NUMBER
TOLMAIS	NUMBER(3)
TOLMENOS	NUMBER(3)
TIPOOBS1	VARCHAR2(15)
OBS1	VARCHAR2(255)
TIPOOBS2	VARCHAR2(15)
OBS2	VARCHAR2(255)
TIPOOBS3	VARCHAR2(15)
OBS3	VARCHAR2(255)
TIPOOBS4	VARCHAR2(15)
OBS4	VARCHAR2(255)
TIPOABC	VARCHAR2(5)
CODSEGMENTO	VARCHAR2(20)
PREFERENCIALSN	CHAR(1)
CLIENTELISTASN	CHAR(1)
AGRUPAMENTO	VARCHAR2(20)
CLIENTE_FORNECEDOR	NUMBER(3)
DESCRCLIENTE_FORNECEDOR	VARCHAR2(30)
RAZAOSOCIAL	VARCHAR2(99)
CODIGOERP	VARCHAR2(60)
ETIQUETA	VARCHAR2(50)
EXIGELAUDO	CHAR(1)
LAUDO	VARCHAR2(50)
EMAILLAUDO	VARCHAR2(1000)
CARTAORCAMENTO	VARCHAR2(50)
NRETIQUETAS	NUMBER(3)
IDUNICOREGISTRO	RAW(16 BYTE)
ACEITAPALETEINCOMPLETO	NUMBER(3)
DESCACEITAPALETEINCOMPLETO	VARCHAR2(100)
IDATENDENTEVENDA	NUMBER

CONTATOFATURAMENTO	VARCHAR2(100)
IDCLASSFISCAL	NUMBER
DATACRIACAO	DATE
USUARIOCRIACAO	VARCHAR2(102)
DATAULTMODIF	DATE
USUARIOULTMODIF	VARCHAR2(102)
RAZAOTRAVACOMPOSICAO	VARCHAR2(50)
CUSTOFINANCEIROADICIONAL	BINARY_DOUBLE
CUSTOADICTRANSPORTEPORPALETE	BINARY_DOUBLE
CUSTOADICTRANSPORKG_PALETIZ	BINARY_DOUBLE
CUSTOADICTRANSPORKG_NPALETIZ	BINARY_DOUBLE
CUSTOADICTRANSPORM2_PALETIZ	BINARY_DOUBLE
CUSTOADICTRANSPORM2_NPALETIZ	BINARY_DOUBLE
CUSTOADICTRANSPORPECA_PALETIZ	BINARY_DOUBLE
CUSTOADICTRANSPORPECANPALETIZ	BINARY_DOUBLE
USOGERALSTR1	VARCHAR2(255)
USOGERALSTR2	VARCHAR2(255)
USOGERALNUM1	BINARY_DOUBLE
USOGERALNUM2	BINARY_DOUBLE
EXIGECHAPAUNITIZADA	CHAR(1)
EXIGEAGENDAMENTOENTREGAS	CHAR(1)
CLIENTEPROSPECTO	CHAR(1)
IMPRIMELOGOTIPO	CHAR(1)
PALETIZADO	NUMBER(5)
FACESOPPOSTAS	NUMBER(5)
ESPELHO	NUMBER(5)
CANTONEIRAS	NUMBER(5)
FILME	NUMBER(5)
ORELHASINVERTIDAS	NUMBER(5)
PACOTESLARGURA	NUMBER
PACOTOSCOMPRIMENTO	NUMBER
PACOTESALTURA	NUMBER
IDPALETE	NUMBER
EMPILHAMENTOMAXIMO	NUMBER
PRESSAODEARQUEAMENTO	NUMBER
FITASLARGPALETE	NUMBER
FITASCOMMPALETE	NUMBER

## 4. Trusted

A camada Trusted é composta por dados que passaram por processos de validação e limpeza, tornando-os consistentes e confiáveis para análises. Nessa camada, são aplicadas regras de qualidade, padronização de formatos, remoção de duplicidades e tratamento de valores nulos críticos. Todas as colunas sinalizadas com a opção de exclusão durante a análise de entendimento foram desconsideradas.

## 4.1 Tabela

A tabela abaixo contempla a estrutura disponibilizada na camada *trusted* para a tabela Clientes.

Name	Type
ID_CLIENTE	NUMBER
TX_CLIENTE	VARCHAR2(50)
CD_CLIENTE	VARCHAR2(30)
CD_REPRESENTANTE	NUMBER
TOLMAIS	NUMBER(3)
TOLMENOS	NUMBER(3)
TX_RAZAOSOCIAL	VARCHAR2(99)
FL_EXIGELAUDO	CHAR(1)
ID_UNICOREGISTRO	RAW(16 BYTE)
VL_ACEITAPALETEINCOMPLETO	NUMBER(3)
TX_DESCACEITAPALETEINCOMPLETO	VARCHAR2(100)
DT_CRIACAO	TIMESTAMP(6)
DT_ULTMODIF	TIMESTAMP(6)

## 5. Refined

A camada Refined é a etapa final do pipeline, onde os dados são transformados e otimizados para consumo analítico avançado. Inclui cálculos, agregações, junções entre múltiplas fontes e criação de indicadores ou métricas derivadas. Nesta camada, os dados estão modelados para suportar dashboards, relatórios estratégicos, machine learning e tomada de decisão de negócio, proporcionando insights de alto valor a partir de informações confiáveis e enriquecidas.

### 5.1 Tabela

A tabela abaixo contempla a estrutura disponibilizada na camada *refined* para a tabela Clientes.

Name	Type
ID_CLIENTE	VARCHAR2(20)
TX_CLIENTE	VARCHAR2(255)
CD_CLIENTE	VARCHAR2(50)
CD_REPRESENTANTE	VARCHAR2(50)

## 6. QUALIDADE

Foi realizada a análise da qualidade dos dados disponíveis na camada *refined*.

Name	Qualidade verificada
ID_CLIENTE	Valores NULL: 0 (0%) Tamanho: 5 Valores textuais nulos: 0 Valores NaN em texto: 0



TX_CLIENTE	Valores NULL: 0 (0%) Maior tamanho de texto: 50 Valores textuais nulos: 0 Valores NaN em texto: 0
CD_CLIENTE	Valores NULL: 0 (0%) Maior tamanho de texto: 6 Valores textuais nulos: 0 Valores NaN em texto: 0 Duplicados: 0
CD_REPRESENTANTE	Valores NULL: 0 (0%) Maior tamanho de texto: 3 Valores textuais nulos: 0 Valores NaN em texto: 0

## 7. DETALHAMENTO

Para mais detalhes sobre esta tabela, consulte o documento:  
[project\\_data\\_science/docs/Analise/dicionarioDados.html](project_data_science/docs/Analise/dicionarioDados.html)