

**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.....	4
<b>3.</b>	<b>Raw.....</b>	<b>5</b>
3.1	Tabela.....	5
<b>4.</b>	<b>Trusted .....</b>	<b>5</b>
4.1	Tabela.....	5
<b>5.</b>	<b>Refined .....</b>	<b>6</b>
5.1	Tabela.....	6
<b>6.</b>	<b>QUALIDADE.....</b>	<b>6</b>
<b>7.</b>	<b>DETALHAMENTO .....</b>	<b>7</b>

## 1. Visão Geral

---

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

### Paradas:

*“Cadastros dos eventos de paradas. Descrição dos códigos relacionados aos eventos de parada apontados durante as produções dos itens”.*

## 2. Origem

---

**Banco:** SQL Server

**Base de dados:** DBO

**Tabela:** Paradas

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

### 2.1 Tabela

A tabela abaixo contempla a estrutura atual da fonte Paradas 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
idParada	int	Não	Manter	
Parada	int	Sim	Manter	
Descricao	varchar	Sim	Manter	
Setor	varchar	Não	Manter	
Tipo	int	Não	Manter	
CodigoERP	varchar	Não	Excluir	
ModoValidacao	int	Não	Excluir	
TextoValidacao	varchar	Não	Excluir	
UsadaOnduladeira	bit	Não	Manter	
UsadaConversao	bit	Sim	Excluir	"=1"
Desativada	bit	Sim	Excluir	"=0"
flagAjuste	bit	Não	Excluir	Identifica quando Ajuste (Início nova Ordem Prd.)
flagExterna	bit	Sim	Excluir	
FlagContinuacaoAjuste	bit	Não	Excluir	
FlagAssociadoAoProduto	bit	Não	Excluir	
ForcaTrocaConversaoOnline	bit	Não	Excluir	
ExcluiOEE	bit	Não	Excluir	
ExigeInformacaoSecaoMaqCon	bit	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 Paradas.

Nome	Tipo
IDPARADA	NUMBER
PARADA	NUMBER
DESCRICAO	VARCHAR2(30)
SETOR	VARCHAR2(30)
TIPO	NUMBER
CODIGOERP	VARCHAR2(30)
MODOVALIDACAO	NUMBER
TEXTOVALIDACAO	VARCHAR2(255)
USADAONDULADEIRA	CHAR(1)
USADACONVERSAO	CHAR(1)
DESATIVADA	CHAR(1)
FLAGAJUSTE	CHAR(1)
FLAGEXTERNA	CHAR(1)
FLAGCONTINUACAOAJUSTE	CHAR(1)
FLAGASSOCIADOAOPRODUTO	CHAR(1)
FORCATROCACONVERSAOONLINE	CHAR(1)
EXCLUIOEE	CHAR(1)
EXIGEINFORMACAOSECAOMAQCONV	CHAR(1)

### 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 Paradas.

Name	Type
IDPARADA	NUMBER

PARADA	NUMBER
DESCRICAO	VARCHAR2(30)
SETOR	VARCHAR2(30)
TIPO	NUMBER
USADAONDULADEIRA	CHAR(1)
USADACONVERSAO	CHAR(1)
DESATIVADA	CHAR(1)
FLAGAJUSTE	CHAR(1)
FLAGEXTERNA	CHAR(1)

## 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 Paradas.

Name	Type
ID_PARADA	VARCHAR2(20)
TX_DESCRICAO	VARCHAR2(125)
FL_USADACONVERSAO	VARCHAR2(1)
FL_DESATIVADA	VARCHAR2(1)
FL_EXTERNA	VARCHAR2(1)

## 6. QUALIDADE

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

Name	Qualidade verificada
ID_PARADA	Valores nulos: 0 Comprimento máximo encontrado: 6 (limite: 20)
TX_DESCRICAO	Valores nulos: 1 Comprimento máximo encontrado: 30 (limite: 125)
FL_USADACONVERSAO	Valores nulos: 0 Valores inválidos (diferentes de 0/1/S/N): 0

FL_DESATIVADA	Valores nulos: 0 Valores inválidos (diferentes de 0/1/S/N): 0
FL_EXTERNA	Valores nulos: 0 Valores inválidos (diferentes de 0/1/S/N): 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)