Documento de Dados do Projeto

Avaliações de vinhos

Beatriz Guimarães Tricoli

João Pedro Vendramini

Weslley de Souza Gomes Luiz

Sumário

[1. Descrição do Sistema 3](#_Toc63717609)

[2. Modelagem Lógica do Banco de Dados Transacional 4](#_Toc63717610)

[3. Modelagem Relacional do Banco de Dados Transacional 5](#_Toc63717611)

[4. Dicionário de Dados 6](#_Toc63717612)

[5. Sequences e Triggers 8](#_Toc63717613)

[6. Tabelas Temporais 9](#_Toc63717614)

[6.1. Modelo Relacional para Tabela de Historiamento 9](#_Toc63717615)

[6.2. Triggers, Stored Procedures e Functions de Historimanto 9](#_Toc63717616)

[7. Tabelas de Auditoria 10](#_Toc63717617)

[7.1. Modelo Relacional para Tabela de Auditoria 10](#_Toc63717618)

[7.2. Triggers, Stored Procedures e Functions de Auditoria 10](#_Toc63717619)

[8. Tabela Stage 11](#_Toc63717620)

[8.1. Tabela Stage 11](#_Toc63717621)

[8.2. Rotina para Armazenamento de Dados na Stage 11](#_Toc63717622)

[9. Queries de Informações Estratégicas 12](#_Toc63717623)

[9.1. Descrição da consulta estratégica 1 12](#_Toc63717624)

[9.2. Descrição da consulta estratégica 2 12](#_Toc63717625)

[9.3. Descrição da consulta estratégica 3 12](#_Toc63717626)

[9.4. Descrição da consulta estratégica 4 12](#_Toc63717627)

[9.5. Descrição da consulta estratégica 5 12](#_Toc63717628)

[9.6. Descrição da consulta estratégica 6 12](#_Toc63717629)

[9.7. Descrição da consulta estratégica 7 13](#_Toc63717630)

[9.8. Descrição da consulta estratégica 8 13](#_Toc63717631)

[10. Queries para Gráficos de Análise 14](#_Toc63717632)

[10.1. Descrição da consulta de análise 1 14](#_Toc63717633)

[10.2. Descrição da consulta de análise 2 14](#_Toc63717634)

[10.3. Descrição da consulta de análise 3 14](#_Toc63717635)

[10.4. Descrição da consulta de análise 4 14](#_Toc63717636)

[10.5. Descrição da consulta de análise 5 14](#_Toc63717637)

[10.6. Descrição da consulta de análise 6 15](#_Toc63717638)

[10.7. Descrição da consulta de análise 7 15](#_Toc63717639)

[10.8. Descrição da consulta de análise 8 15](#_Toc63717640)

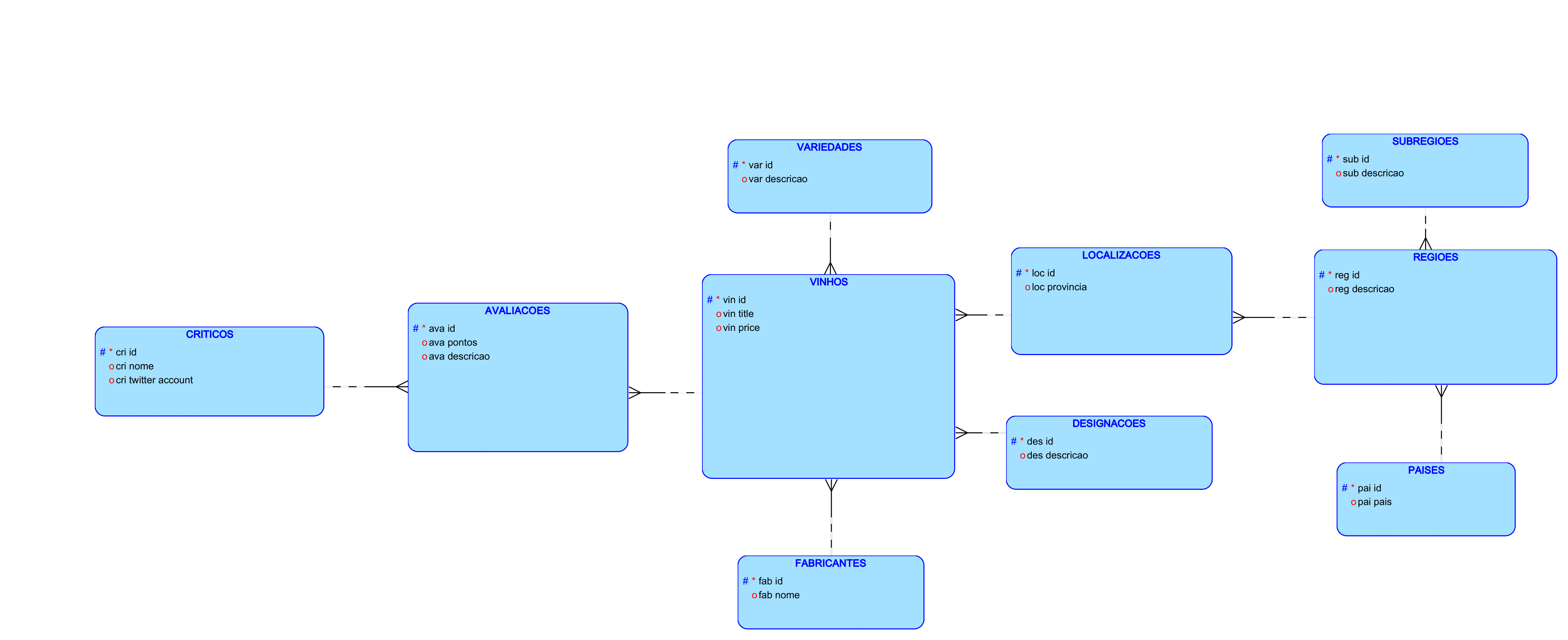
# Descrição do Sistema

O sistema é modelado tendo como base 130 mil avaliações de vinhos e suas origens realizadas por pessoas ao redor do mundo. A avaliação é realizada por usuários e possui como métricas a região e sub-região dos vinhedos em que foi realizado a fermentação do vinho, tal como a nota dos usuários.

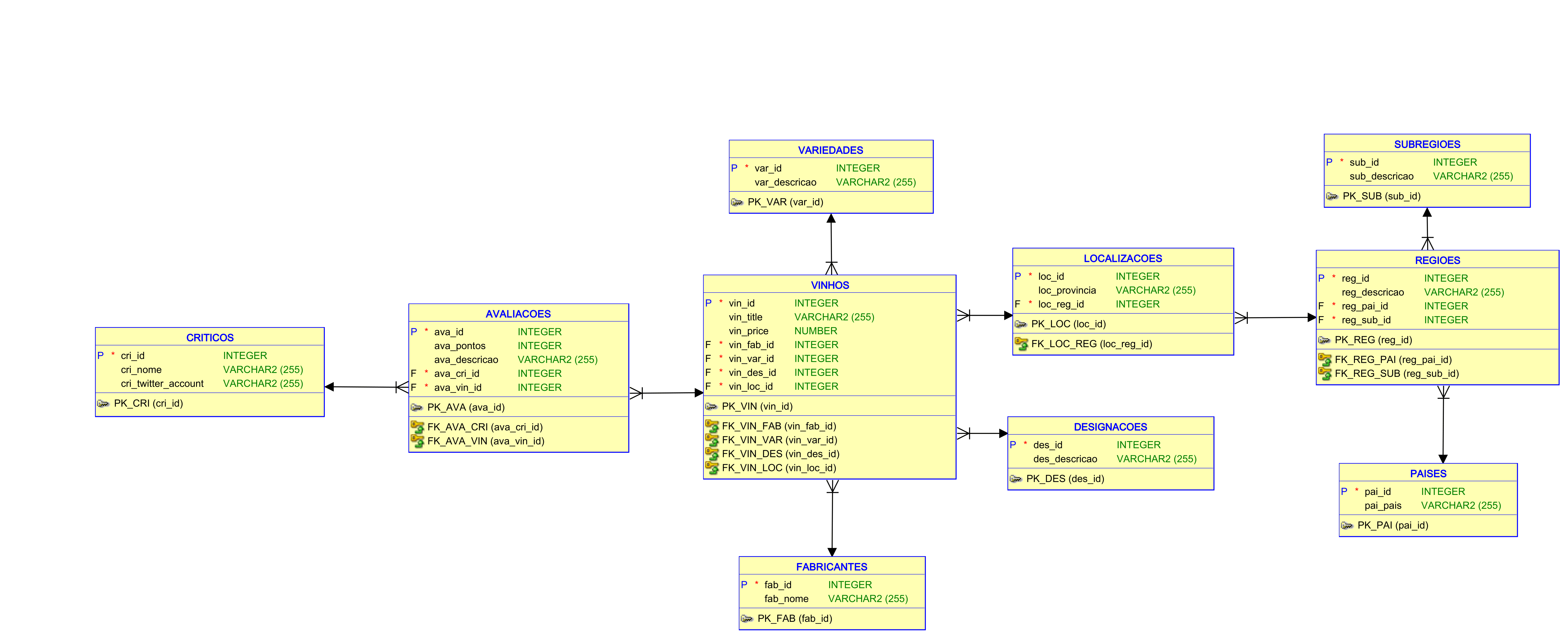
A respeito das informações dos críticos, o Dataset contém apenas o nome do avaliador e o nome de usuário no Twitter. Em contrapartida, os dados de cada vinho possuem informações mais detalhadas como o vinhedo, a cidade (sub-região), estado (região), país, fabricante, descrição do produto provida pelos avaliadores e uma nota também atribuída pelos usuários.

Esse conjunto de dados oferece ótimas oportunidades para análise de sentimentos e outros modelos preditivos relacionados a texto. O objetivo geral é criar um modelo que possa identificar a variedade, adega e localização de um vinho com base em uma descrição.

# Modelagem Lógica do Banco de Dados Transacional



# Modelagem Relacional do Banco de Dados Transacional



# Dicionário de Dados

**Tabela avaliacoes:**

Esta tabela armazena dados referentes a avaliação de cada vinho por cada crítico.

|  |  |  |  |
| --- | --- | --- | --- |
| **Coluna** | **Tipo (tamanho)** | **Constraints** | **Descrição** |
| ava\_id | INTEGER | PK | Esta coluna é a pk da relação. |
| ava\_pontos | INTEGER | - | Esta coluna armazena o dado de pontuação da avaliação pelo crítico. |
| ava\_descricao | VARCHAR2(255) | - | Esta coluna armazena o dado de descrição da avaliação do crítico. |
| ava\_cri\_id | INTEGER | FK referenciando tabela de críticos,  NOT NULL | Esta coluna armazena a chave estrangeira para a tabela de críticos. |
| ava\_vin\_id | INTEGER | FK referenciando tabela de vinícolas,  NOT NULL | Esta coluna armazena a chave estrangeira para a tabela de vinícolas. |

**Tabela criticos:**

Esta tabela armazena dados referentes aos avaliadores dos vinhos.

|  |  |  |  |
| --- | --- | --- | --- |
| **Coluna** | **Tipo (tamanho)** | **Constraints** | **Descrição** |
| cri\_id | INTEGER | PK | Esta coluna é a pk da relação. |
| cri\_nome | VARCHAR2(255) | - | Esta coluna armazena o dado nome do crítico. |
| cri\_twitter\_account | VARCHAR2(255) | - | Esta coluna armazena o dado da conta do twitter do avaliador. |

**Tabela designacoes:**

Esta tabela armazena dados referentes aos vinhedos dos fabricantes.

|  |  |  |  |
| --- | --- | --- | --- |
| **Coluna** | **Tipo (tamanho)** | **Constraints** | **Descrição** |
| des\_id | INTEGER | PK | Esta coluna é a pk da relação. |
| des\_descricao | VARCHAR2(255) | - | Esta coluna armazena o dado de descrição da designação. |

**Tabela fabricantes**

Esta tabela armazena dados referentes aos fabricantes do vinho.

|  |  |  |  |
| --- | --- | --- | --- |
| **Coluna** | **Tipo (tamanho)** | **Constraints** | **Descrição** |
| fab\_id | INTEGER | PK | Esta coluna é a pk da relação. |
| fab\_nome | VARCHAR2(255) | - | Esta coluna armazena o dado nome do fabricante. |

**Tabela localizacoes:**

Esta tabela armazena dados referentes à localização do fabricante.

|  |  |  |  |
| --- | --- | --- | --- |
| **Coluna** | **Tipo (tamanho)** | **Constraints** | **Descrição** |
| loc\_id | INTEGER | PK, FK referenciando Tabn(coln), etc | Esta coluna é a pk da relação. |
| loc\_provincia | VARCHAR2(255) | - | Esta coluna armazena o dado da província. |
| loc\_reg\_id | INTEGER | FK referenciando tabela de regiões,  NOT NULL | Esta coluna armazena o dado fk da tabela de regiões. |

**Tabela paises:**

Esta tabela armazena dados referentes aos países.

|  |  |  |  |
| --- | --- | --- | --- |
| **Coluna** | **Tipo (tamanho)** | **Constraints** | **Descrição** |
| pai\_id | INTEGER | PK | Esta coluna é a pk da relação. |
| pai\_pais | VARCHAR2(255) | - | Esta coluna armazena o dado do nome do país. |

**Tabela regioes:**

Esta tabela armazena dados referentes à região.

|  |  |  |  |
| --- | --- | --- | --- |
| **Coluna** | **Tipo (tamanho)** | **Constraints** | **Descrição** |
| reg\_id | INTEGER | PK | Esta coluna é a pk da relação. |
| reg\_descricao | VARCHAR2(255) | - | Esta coluna armazena o dado de descrição da região. |
| reg\_pai\_id | INTEGER | FK referenciando tabela de países.,  NOT NULL | Esta coluna armazena o dado id do país na tabela de países. |

**Tabela subregiões:**

Esta tabela armazena dados referentes à subregiões das regiões.

|  |  |  |  |
| --- | --- | --- | --- |
| **Coluna** | **Tipo (tamanho)** | **Constraints** | **Descrição** |
| sub\_id | INTEGER | PK | Esta coluna é a pk da relação. |
| sub\_descricao | VARCHAR2(255) | - | Esta coluna armazena o dado de descrição da sub-região. |

**Tabela variedades:**

Esta tabela armazena dados referentes às variedades dos vinhos.

|  |  |  |  |
| --- | --- | --- | --- |
| **Coluna** | **Tipo (tamanho)** | **Constraints** | **Descrição** |
| var\_id | INTEGER | PK | Esta coluna é a pk da relação |
| var\_descricao | VARCHAR2(255) | - | Esta coluna armazena o dado de descrição da variedade |

**Tabela vinhos:**

Esta tabela armazena dados referentes aos vinhos.

|  |  |  |  |
| --- | --- | --- | --- |
| **Coluna** | **Tipo (tamanho)** | **Constraints** | **Descrição** |
| vin\_id | INTEGER | PK | Esta coluna é a pk da relação |
| vin\_title | VARCHAR2(255) | - | Esta coluna armazena o dado de título do vinho |
| vin\_price | NUMBER |  | Esta coluna armazena o dado de preço do vinho |
| vin\_fab\_id | INTEGER | FK referenciando tabela de fabricantes,  NOT NULL | Esta coluna armazena o dado de id do fabricante |
| vin\_var\_id | INTEGER | FK referenciando tabela de variedades,  NOT NULL | Esta coluna armazena o dado de id da variedade |
| vin\_des\_id | INTEGER | FK referenciando tabela de designações,  NOT NULL | Esta coluna armazena o dado de id da designação |
| vin\_loc\_id | INTEGER | FK referenciando tabela de localizações,  NOT NULL | Esta coluna armazena o dado de id da localização |

# Sequences e Triggers

<<Nesta seção, inserir a relação de todas as sequências e o nome de todos os gatilhos criados para resgatar o valor da sequência na inserção de dados.>>

|  |  |  |
| --- | --- | --- |
| **Nome da Tabela** | **Nome da Sequência** | **Nome do Gatilho** |
| Tabela A | Seq\_tabA | Tg\_seq\_tabA |
| Tabela B | Seq\_tabB | Tg\_seq\_tabB |
| Tabela C | Seq\_tabC | Tg\_seq\_tabC |
| Tabela D | Seq\_tabD | Tg\_seq\_tabD |
| Tabela E | Seq\_tabE | Tg\_seq\_tabE |
| Tabela F | Seq\_tabF | Tg\_seq\_tabF |
| Tabela G | Seq\_tabG | Tg\_seq\_tabG |

# Tabelas Temporais

# Modelo Relacional para Tabela de Historiamento

<<Nesta seção, inserir o modelo das tabelas de histórico e seu relacionamento com as tabelas transacionais.>>

# Triggers, Stored Procedures e Functions de Historimanto

<<Nesta seção, inserir a relação de todos os gatilhos, procedimentos armazenados e funções criados para o mecanismo de armazenamento dos dados históricos no banco de dados.>>

|  |  |  |
| --- | --- | --- |
| Nome do trigger | Tabela de negócio | Tabela de histórico |
|  |  |  |
|  |  |  |
|  |  |  |

# Tabelas de Auditoria

# Modelo Relacional para Tabela de Auditoria

<<Nesta seção, inserir o modelo da tabela de auditoria do sistema.>>

# Triggers, Stored Procedures e Functions de Auditoria

<<Nesta seção, inserir a relação de todos os gatilhos, procedimentos armazenados e funções criados para o mecanismo de auditoria do banco de dados.>>

# Tabela Stage

# Tabela Stage

<<Nesta seção, inserir a especificação da tabela stage de seu projeto. A stage é uma tabela intermediária entre o mundo transacional e o banco de data warehouse. Este tópico deve conter uma tabela semelhante ao dicionário de dados, com a especificação de cada campo da tabela de stage e, quando aplicável, deve ser especificada a transformação que foi feita nos dados de origem antes de persistir o dado na stage.>>

**Tabela Stage:**

Esta tabela armazena dados referentes a lalala lalala lalala.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Coluna** | **Tipo (tamanho)** | **Constraints** | **Descrição** | **Transformação** | **Coluna(s) de origem** |
| CA1 | Tipo (tam) | PK, FK referenciando Tabn(coln), etc | Esta coluna é a pk da relação |  |  |
| CA2 | Tipo (tam) | PK, FK referenciando Tabn(coln), etc | Esta coluna armazena o dado x de lalala |  |  |
| CA3 | Tipo (tam) | PK, FK referenciando Tabn(coln), etc | Esta coluna armazena o dado y de lalala |  |  |

# Rotina para Armazenamento de Dados na Stage

<<Nesta seção, deve ser colocado o código SQL da rotina criada para alimentar a tabela de stage do sistema.>>

# Queries de Informações Estratégicas

# Descrição da consulta estratégica 1

Autor:

SQL:

Resultado:

# Descrição da consulta estratégica 2

Autor:

SQL:

Resultado:

# Descrição da consulta estratégica 3

Autor:

SQL:

Resultado:

# Descrição da consulta estratégica 4

Autor:

SQL:

Resultado:

# Descrição da consulta estratégica 5

Autor:

SQL:

Resultado:

# Descrição da consulta estratégica 6

Autor:

SQL:

Resultado:

# Descrição da consulta estratégica 7

Autor:

SQL:

Resultado:

# Descrição da consulta estratégica 8

Autor:

SQL:

Resultado:

# Queries para Gráficos de Análise

# Descrição da consulta de análise 1

Autor:

SQL:

Resultado:

Gráfico:

# Descrição da consulta de análise 2

Autor:

SQL:

Resultado:

Gráfico:

# Descrição da consulta de análise 3

Autor:

SQL:

Resultado:

Gráfico:

# Descrição da consulta de análise 4

Autor:

SQL:

Resultado:

Gráfico:

# Descrição da consulta de análise 5

Autor:

SQL:

Resultado:

Gráfico:

# Descrição da consulta de análise 6

Autor:

SQL:

Resultado:

Gráfico:

# Descrição da consulta de análise 7

Autor:

SQL:

Resultado:

Gráfico:

# Descrição da consulta de análise 8

Autor:

SQL:

Resultado:

Gráfico: