

**Autores:** Andressa Santos Cucharro , Aline Luisa Mansur linebio9@gmail.com  
gutembergalmteidapereira@gmail.com luischurros@gmail.com  
gustavo.souz.camargo@gmail.com

**Revisores:** Professores Allan e Edgar

**Última Atualização:** 23 de out. de 2025

# FIAP 11DTAT - Pós Tech

## Tech Challenge

### Análise - O Mercado Mundial de Vinhos

Este documento detalha a criação do projeto de análise de dados do Mercado Mundial de Vinhos utilizando a metodologia Knowledge Discovery in Databases (KDD) e que será entregue ao público final (acionistas e investidores) através de uma apresentação executiva. O objetivo é fornecer uma estrutura clara para a execução do projeto, desde a compreensão dos dados até a implantação do conhecimento descoberto.

#### 1. Entendimento dos Dados (Data Understanding)

Nesta fase, o foco é compreender os dados disponíveis e os objetivos do negócio.

##### 1.1 Definição dos Objetivos do Negócio

- **Objetivo Principal:** Gerar uma apresentação com os relatórios iniciais para investidores e acionistas onde serão demonstradas as quantidades de vinhos exportados nos últimos 15 anos por país, trazendo as prospecções futuras e possíveis ações para uma melhoria nas exportações.
- **Perguntas de Negócio:**
  - Quais são nossos maiores clientes?
  - Quais são os nossos clientes mais rentáveis?
  - Qual é o percentual da produção que é exportado? Aqui precisamos entender capacidade de ampliação da exportação
  - Qual é o preço médio por litro de vinho nos últimos 23 anos?

- Como foi a evolução em volume e valor de vendas nos últimos anos desses nossos maiores clientes?
- Quais fatores externos - climáticos, econômicos e demográficos - influenciam nossas exportações? E como influenciam?
- O desempenho das exportações em volume e valor acompanham as avaliações de qualidade dos vinhos?
- Relacionando as tendências históricas e projeções econômicas quais os mercados que apresentam melhor potencial?
- Quais as categorias de vinho são as mais rentáveis e qual a relação destas com a aceitação internacional?
- Quais os principais riscos e oportunidades futuras que podem impactar o mercado de exportação de vinhos?
- Qual tipo de vinho teve maior crescimento (produção, importação, exportação)?
- Há anos com queda geral na produção?
- Anos de pandemia afetaram os padrões de produção?

## 1.2 Coleta de Dados

- Fontes de Dados:
  - [https://vitibrasil.cnpuv.embrapa.br/index.php?opcao=opt\\_01](https://vitibrasil.cnpuv.embrapa.br/index.php?opcao=opt_01)
  - [Comex Stat](#)

## 1.3 Exploração Inicial dos Dados - Análise Descritiva e Diagnóstica

- Análise Descritiva: [Resumo das principais características dos dados, como média, mediana, desvio padrão, etc.]
- Seria interessante ter os dados de exportação abertos por mês para entender com maiores detalhes a frequência de compras e as sazonalidades por país para ações mais direcionadas e segmentação mais detalhada dos nossos clientes.
- Visualização de Dados da Análise Descritiva: [Mencione os tipos de gráficos e visualizações utilizados, por exemplo, "histogramas, gráficos de dispersão"].
- Análise Diagnóstica: [Passado - porque algo identificado na análise descritiva aconteceu?, correlações, análise de dispersão]

## Rússia x Anexação da Crimeia

- **Anexação da Crimeia (2014):** Com a anexação da Crimeia, a Rússia incorporou uma região com uma riquíssima e histórica tradição vinícola. As vinícolas da Crimeia, algumas muito famosas como Massandra e Novy Svet, passaram a fazer parte do mercado russo.

- **Sanções Internacionais e Contrassanções:** Em resposta à anexação, países ocidentais impuseram sanções econômicas à Rússia. O governo russo reagiu com contrassanções, proibindo a importação de uma vasta gama de produtos alimentícios e agrícolas da União Europeia e de outros países. Isso **criou um vácuo no mercado interno**.
- **Política de "Substituição de Importações":** Com menos vinhos europeus disponíveis nas prateleiras, o governo russo lançou um forte programa de incentivo à produção local. Eles começaram a oferecer subsídios generosos, empréstimos com juros baixos e apoio estatal para quem quisesse plantar novos vinhedos, modernizar vinícolas e investir em tecnologia.
- **A Lei do Vinho de 2020:** Para solidificar essa tendência, em junho de 2020 entrou em vigor uma lei federal sobre viticultura e vinificação. Essa lei foi um marco, pois definiu oficialmente o "vinho russo" como sendo um produto feito **exclusivamente com uvas cultivadas na Rússia**. Isso protegeu ainda mais os produtores locais da concorrência com vinhos importados a granel e fortaleceu a identidade do vinho nacional.

Portanto, embora a Rússia sempre tenha tido uma história de produção de vinho, o período **pós-2014** foi a verdadeira "virada de chave" que iniciou uma nova era de investimento, modernização e aumento expressivo tanto na quantidade quanto, mais importante, na qualidade dos vinhos russos.

- 
- Visualização de Dados da Análise Diagnóstica: [\[Mencione os tipos de gráficos e visualizações utilizados, por exemplo, "histogramas, gráficos de dispersão"\]](#).

## 2. Pré-processamento dos Dados (Data Preprocessing)

Esta etapa envolve a limpeza, transformação e redução dos dados para prepará-los para a mineração.

### 2.1 Limpeza de Dados

- Tratamento de Valores Ausentes: [\[Método a ser utilizado, por exemplo, "remoção de linhas, imputação de média/mediana"\]](#).
- Tratamento de Ruídos: [\[Método a ser utilizado, por exemplo, "filtragem, suavização"\]](#).
- Identificação e Correção de Inconsistências: [\[Descreva as inconsistências encontradas e as ações para corrigi-las\]](#).

### 2.2 Transformação de Dados

- Normalização/Padronização: [\[Método a ser utilizado, por exemplo, "escalonamento min-max, padronização Z-score"\]](#).

- Discretização: [Se aplicável, descreva como as variáveis contínuas serão discretizadas].
- Criação de Novas Features: [Descreva as novas variáveis que serão criadas a partir das existentes].

## 2.3 Redução de Dados

- Seleção de Atributos: [Método a ser utilizado para selecionar os atributos mais relevantes].
- Redução de Dimensionalidade: [Técnicas como PCA, se aplicável].

## 3. Mineração de Dados (Data Mining)

Nesta fase, os algoritmos de mineração de dados são aplicados para descobrir padrões nos dados pré-processados.

### 3.1 Seleção da Tarefa de Mineração

- Classificação: [Se aplicável, descreva a tarefa, por exemplo, "prever churn de clientes"].
- Regressão: Prever demanda futura de exportação.
- Agrupamento (Clustering): [Se aplicável, descreva a tarefa, por exemplo, "segmentar clientes em grupos com base em seu comportamento"].
- Associação: [Se aplicável, descreva a tarefa, por exemplo, "identificar produtos frequentemente comprados juntos"].

### 3.2 Seleção do Algoritmo de Mineração

- Algoritmos Propostos:
  - Regressão linear múltipla: Simples de implementar e interpretar. Excelente para começar a identificar correlações.
  - ARIMAX: Combina a inércia do histórico de vendas com o efeito de fatores externos.
  - Suavização Exponencial (Holt): Bom para prever apenas a continuação da tendência observada (crescimento ou declínio).
- Ferramentas de Mineração: [Nome das ferramentas, por exemplo, "Python com scikit-learn, R com caret"].

### 3.3 Execução e Avaliação dos Modelos

- Métricas de Avaliação: [Métricas relevantes para a tarefa, por exemplo, "acurácia, precisão, recall, F1-score, silhueta"].
- Resultados Iniciais: [Descreva os primeiros resultados obtidos e as observações].

## 4. Avaliação dos Padrões (Pattern Evaluation)

Os padrões descobertos são avaliados para determinar sua significância e validade.

### 4.1 Interpretação dos Padrões

- Significância Estatística: [\[Como os padrões serão avaliados em termos estatísticos\]](#).
- Relevância para o Negócio: [\[Como os padrões se relacionam com os objetivos de negócio definidos\]](#).

### 4.2 Validação dos Padrões

- Validação Cruzada: [\[Método de validação, por exemplo, "k-fold cross-validation"\]](#).
- Comparação com Especialistas de Domínio: [\[Como os padrões serão validados com o conhecimento de especialistas\]](#).

## 5. Cronograma do Projeto

Fase	Início	Término	Responsável
Análise descritiva e diagnóstica	13 de out. de 2025	17 de out. de 2025	Todos
Pré-processamento de dados e Mineração	20 de out. de 2025	24 de out. de 2025	Todos
Avaliação dos padrões e geração dos gráficos	27 de out. de 2025	31 de out. de 2025	Todos
Revisão Final para entrega	3 de nov. de 2025	3 de nov. de 2025	Todos

## 6. Entregáveis

- [Repositório no GITHUB](#) com todos os notebooks/código utilizados
- DashBoard no Looker ou Gráficos gerados no Google Sheets (já que se trata de uma análise AdHoc)