

Relatório Técnico - Projeto Aplicado II

1. Introdução

Nesse relatório técnico vamos documentar todas as etapas realizadas no desenvolvimento do Projeto Aplicado II, cujo objetivo foi a análise e desenvolvimento de insights baseados na base de dados do YouTube. Este relatório sintetiza as definições iniciais, metodologias empregadas, análise exploratória, preparação dos dados e apresentação de resultados obtidos ao longo das entregas.

2. Grupo de Trabalho

- **Integrantes:** Mariana Simões Rubio;
Patrícia Corrêa França;
Valdiney Atílio Pedro.
- **Funções:**
 - Coleta e organização de dados;
 - Análise exploratória;
 - Modelagem e validação;
 - Storytelling e relatórios.

3. Premissas do Projeto

Organização Escolhida

- **Nome:** Plataformas de Canais do YouTube.
- **Área de atuação:** Entretenimento digital e redes sociais.
- **Descrição:** Os dados analisados foram extraídos de canais do YouTube, abrangendo categorias como música, canais, lifestyle, entre outras.
- **Dados utilizados:** Colunas como Subscribers, Video Views, Country, Category, Earnings, e Creation Year oferecem informações sobre o desempenho e contexto dos canais. Dados complementares sobre características demográficas e econômicas, como população e urbanização dos países, foram utilizados para contextualizar tendências globais.

Objetivos e Metas

- 1. Analisar o desempenho dos canais em diferentes categorias.
- 2. Explorar as relações entre variáveis, como inscritos, visualizações e ganhos.
- 3. Desenvolver um modelo preditivo para estimar ganhos anuais dos canais com base nas características fornecidas.
- 4. Apresentar insights de negócio para monetização e otimização de estratégias de canais do YouTube.

Cronograma de Atividades

Etapa	Descrição	Prazo
Definição de premissas	Escolha de dados e objetivos	A1
Análise exploratória	Entendimento inicial dos dados	A2
Modelagem e validação	Desenvolvimento de algoritmos	A3
Storytelling e relatório	Apresentação de resultados	A4

4. Metodologia

Definição da Linguagem e Ferramentas

- **Linguagem de programação:** Python
- **Ferramentas utilizadas:**
 - Bibliotecas: Pandas, Matplotlib, Scikit-Learn, Seaborn.

Análise Exploratória

- Identificação de variáveis com maior impacto, como Subscribers e Video Views.
- Visualizações iniciais mostraram que categorias como Music e Gaming possuem os maiores ganhos potenciais.

Tratamento e Preparação dos Dados

- 1. **Limpeza:**
 - a. Remoção de dados ausentes ou inconsistentes.

2. Transformação:

- a. Escalonamento de variáveis como visualizações e ganhos usando normalização Min-Max.

Métodos e Cálculo da Acurácia

- **Modelos testados:**
 - Regressão Linear: Estimativa de ganhos baseados em Subscribers e Video Views.
 - Árvore de Decisão: Para explorar não-linearidades.
- **Medida de acurácia:** R^2 (coeficiente de determinação) e RMSE (Root Mean Squared Error).

5. Aplicação dos Métodos e Resultados

Análise e Resultados Preliminares

1. **Relação entre Inscritos e Ganhos:** Identificamos uma correlação positiva ($r = 0.85$), sugerindo que canais com mais inscritos tendem a gerar maior receita.
2. **Categorias mais lucrativas:** Music e Gaming apresentam maior estimativa de ganhos médios anuais.
3. **Análise por país:** Canais em países com maior urbanização apresentam maior média de inscritos e ganhos.

Modelo de Negócios

Proposta de segmentação de canais para estratégias específicas:

- **Pequenos Criadores:** Foco em crescimento orgânico e parcerias.
- **Médios Criadores:** Incentivar monetização por assinaturas e produtos.
- **Grandes Criadores:** Expansão para conteúdo exclusivo e internacionalização.

Medidas de Acurácia

- Modelo Regressão Linear:
 - $R^2 = 0.78$
 - RMSE = 15,000 USD
- Modelo Árvore de Decisão:
 - $R^2 = 0.82$

- RMSE = 12,500 USD

6. Storytelling e Visualização

- **Resumo gráfico:**
 - Gráficos de coluna relacionando a ganhos e inscritos por categoria, país e crescimento.
- **Narrativa:** "Canais do YouTube não são apenas entretenimento; eles refletem tendências globais de consumo e podem ser otimizados para aumentar a rentabilidade."

7. Conclusão

Este projeto demonstrou como a análise de dados pode oferecer insights valiosos para o mercado digital. Os resultados reforçam a importância de dados estruturados e modelos preditivos no suporte à tomada de decisão.

8. Repositório do Projeto

- Link para o GitHub: <https://github.com/valdineyatilio/ProjetoAplicado-II>