

UNIVERSIDADE PRESBITERIANA MACKENZIE

VALDINEY ATÍLIO PEDRO
PATRICIA CORREA FRANÇA
MARIANA SIMÕES RUBIO

PROJETO APLICADO

São Paulo
2024

VALDINEY ATÍLIO PEDRO - 10424616
PATRICIA CORREA FRANÇA - 10423533
MARIANA SIMÕES RUBIO – 10424388

PROJETO APLICADO:
Empresa YOUTUBE

Trabalho de conclusão de semestre apresentado à
Universidade presbiteriana Mackenzie, do estado
de São Paulo, como requisito parcial à obtenção de
conclusão de curso.

ORIENTADOR: Prof. Felipe Albino dos Santos

São Paulo
2024

LISTA DE TABELAS

Figura 1 - Tabela de medidas	11
Figura 2 - Número de vídeos por categoria	12
Figura 3 - Visualização vs País	13
Figura 4 - Histograma de Visualizações	13

SUMÁRIO

1. GLOSSÁRIO.....	4
2. OBJETIVO DE ESTUDO.....	5
3. APRESENTAÇÃO DA EMPRESA E PROBLEMA DE PESQUISA.....	6
3.1. História e Fundação:	6
3.2. Localização e Presença:	6
3.3. Representação do YouTube:	6
3.4. Missão, Visão e Valores:.....	6
3.5. Segmento e Posicionamento de Mercado:.....	7
3.6. Número de Colaboradores:	7
3.8. Pontos de Melhoria:	7
3.9. Possíveis Gaps:	7
3.10. Trabalhos em Destaque:	8
4. DESAFIOS DA EMPRESA:.....	9
4.1. Algoritmos de Recomendação:	9
4.2. Detecção de Conteúdo Inapropriado:.....	9
4.3. Análise de Tendências e Insights:.....	10
4.4. Melhoria da Experiência do Usuário:.....	10
4.5. Combate à Desinformação e Notícias Falsas:	10
5. APRESENTAÇÃO DOS METADADOS E ANÁLISE EXPLORATÓRIA DE DADOS	11
6. ANÁLISE E DISCUSSÃO DOS RESULTADOS	12
7. REFERÊNCIAS BIBLIOGRAFICAS	14
7.1 LINK GITHUB	14

1. GLOSSÁRIO

Algoritmos de Recomendação - Refere-se aos sistemas computacionais utilizados pelo YouTube (e outras plataformas) para sugerir conteúdo aos usuários com base em seus interesses, histórico de visualização, interações anteriores e outras variáveis.

Monetização de Conteúdo - Refere-se aos métodos e estratégias pelos quais os criadores de conteúdo no YouTube podem gerar receita com seus vídeos, incluindo anúncios, assinaturas de membros, vendas de produtos e patrocínios.

Aprendizado de Máquina (Machine Learning) - É um subcampo da inteligência artificial que permite aos sistemas aprenderem e melhorarem automaticamente a partir de experiências passadas, incluindo a capacidade dos algoritmos de recomendação do YouTube de personalizar as sugestões de vídeos.

Transparência de Dados - Refere-se à prática de divulgar informações detalhadas e compreensíveis sobre como os dados são coletados, utilizados e compartilhados pela plataforma, incluindo políticas de privacidade, uso de cookies e compartilhamento de informações com terceiros.

Direitos Autorais - São os direitos legais concedidos aos criadores de conteúdo sobre suas obras, incluindo vídeos, músicas e imagens, e as medidas que o YouTube implementa para proteger os direitos autorais, como o sistema de gerenciamento de direitos autorais (Content ID) e políticas de uso justo.

2. OBJETIVO DE ESTUDO

O objetivo deste trabalho é investigar e identificar os fatores de sucesso dos principais canais do YouTube, destacando o que os diferencia dos demais. Através da coleta e análise de dados relevantes, como visualizações, inscritos, engajamento (likes, comentários) e metadados dos vídeos (duração, categoria etc.), pretendemos comparar esses canais de sucesso com outros menos populares para descobrir padrões e estratégias que contribuem para seu sucesso. A análise estatística será aplicada para entender quais fatores estão mais fortemente correlacionados com o sucesso no YouTube, e modelos de regressão ou algoritmos de machine learning podem ser utilizados para prever o sucesso de canais com base nesses fatores. Este estudo busca, portanto, fornecer insights valiosos sobre os elementos que contribuem para o crescimento e a popularidade dos canais no YouTube, servindo como um guia para estratégias eficazes de criação de conteúdo na plataforma.

3. APRESENTAÇÃO DA EMPRESA E PROBLEMA DE PESQUISA

O YouTube é uma plataforma de compartilhamento de vídeos online que revolucionou a forma como as pessoas consomem conteúdo audiovisual na internet.

3.1. História e Fundação:

O YouTube foi fundado por três ex-funcionários do PayPal - Chad Hurley, Steve Chen e Jawed Karim - em fevereiro de 2005.

A ideia surgiu quando eles tiveram dificuldades em compartilhar vídeos de um jantar de festa e perceberam a falta de uma plataforma eficiente para esse propósito. (ELGIRIYEWITHANA, 2023)

3.2. Localização e Presença:

O YouTube foi fundado nos Estados Unidos, especificamente em San Bruno, Califórnia.

Atualmente, o YouTube é uma subsidiária do Google, uma empresa da Alphabet Inc. (ELGIRIYEWITHANA, 2023)

3.3. Representação do YouTube:

O YouTube é uma plataforma global de compartilhamento de vídeos que abriga uma ampla variedade de conteúdo, desde vídeos caseiros até produções profissionais.

Ele representa uma oportunidade para indivíduos e empresas alcançarem uma audiência global e monetizarem seu conteúdo por meio de anúncios, assinaturas, doações de fãs e outros modelos de receita. (ELGIRIYEWITHANA, 2023)

3.4. Missão, Visão e Valores:

Missão: Facilitar o compartilhamento e descoberta de vídeos.

Visão: Ser a plataforma líder mundial em compartilhamento de vídeos.

Valores: Embora os valores específicos possam variar, o YouTube geralmente valoriza a liberdade de expressão, a diversidade de conteúdo e a criação de uma comunidade global. (ELGIRIYEWITHANA, 2023)

3.5. Segmento e Posicionamento de Mercado:

O YouTube está no segmento de tecnologia e mídia digital, oferecendo uma plataforma para que criadores e empresas compartilhem conteúdo em vídeo com uma audiência global.

O YouTube é conhecido por seu posicionamento como uma plataforma acessível e democrática, permitindo que qualquer pessoa compartilhe seu conteúdo e alcance uma grande audiência. (ELGIRIYEWITHANA, 2023)

3.6. Número de Colaboradores:

Os dados específicos sobre o número de colaboradores do YouTube por país e estado podem variar e podem ser confidenciais. Normalmente, a empresa tem uma presença significativa em várias regiões do mundo, incluindo os Estados Unidos, Europa, Ásia e América Latina. (ELGIRIYEWITHANA, 2023)

3.8. Pontos de Melhoria:

Transparência nas Políticas: Melhorar a transparência em relação às políticas de moderação de conteúdo, monetização e algoritmos de recomendação pode ajudar a construir confiança com os criadores de conteúdo e a comunidade de usuários.

Ferramentas de Moderação Aprimoradas: Investir em tecnologias avançadas de moderação de conteúdo, incluindo inteligência artificial e machine learning, pode ajudar a identificar e remover conteúdo prejudicial de forma mais eficiente.

Suporte aos Criadores de Conteúdo: Oferecer mais recursos, ferramentas e suporte para criadores de conteúdo, incluindo programas de capacitação, workshops e oportunidades de colaboração, pode incentivar a produção de conteúdo de alta qualidade e diversificado. (ELGIRIYEWITHANA, 2023)

3.9. Possíveis Gaps:

Diversidade e Inclusão: O YouTube pode enfrentar desafios em promover a diversidade e a inclusão em sua plataforma, tanto em termos de criadores de conteúdo quanto de representação em vídeos e recomendações.

Proteção de Dados e Privacidade: Garantir a proteção dos dados dos usuários e a privacidade das informações pessoais é um desafio contínuo, especialmente em um ambiente digital cada vez mais regulamentado. (ELGIRIYEWITHANA, 2023)

3.10. Trabalhos em Destaque:

Além de fornecer uma plataforma para compartilhamento de vídeos, o YouTube tem vários projetos em destaque, incluindo:

YouTube Originals: Produções originais exclusivas para a plataforma.

YouTube Music: Um serviço de streaming de música lançado pelo YouTube.

YouTube Shorts: Uma plataforma para vídeos curtos, concorrendo com aplicativos como TikTok. (ELGIRIYEWITHANA, 2023)

4. DESAFIOS DA EMPRESA:

Para a apresentação, é importante também destacar os desafios enfrentados pelo YouTube, como questões relacionadas à moderação de conteúdo, direitos autorais, privacidade dos usuários, concorrência com outras plataformas de vídeo, entre outros. Isso proporcionará uma visão mais abrangente dos aspectos da empresa.

Moderação de Conteúdo: O YouTube enfrenta desafios constantes na moderação de conteúdo, incluindo a remoção de vídeos inadequados, discurso de ódio, desinformação e violações de direitos autorais.

Monetização Equitativa: Criadores de conteúdo frequentemente expressam preocupações sobre a equidade na monetização de seus vídeos, especialmente em relação às mudanças nos algoritmos de recomendação e políticas de anúncios.

Concorrência com Outras Plataformas: O YouTube enfrenta competição crescente de outras plataformas de vídeo, como TikTok, Instagram e plataformas de streaming ao vivo como Twitch. (ELGIRIYEWITHANA, 2023)

O YouTube tem apresentado várias iniciativas relacionadas a Data Science ao longo dos anos. Assim como tópicos apresentados a seguir:

4.1. Algoritmos de Recomendação:

O YouTube utiliza algoritmos de recomendação baseados em Data Science para personalizar as recomendações de vídeos para cada usuário. Esses algoritmos analisam o histórico de visualização, as preferências do usuário, o engajamento com vídeos anteriores e outros dados para sugerir conteúdo relevante. (ELGIRIYEWITHANA, 2023)

4.2. Detecção de Conteúdo Inapropriado:

A plataforma utiliza técnicas avançadas de processamento de linguagem natural (NLP) e aprendizado de máquina para detectar e moderar conteúdo inadequado, como discurso de ódio, violência, spam e desinformação. Isso inclui a identificação automática de legendas, transcrições e comentários. (ELGIRIYEWITHANA, 2023)

4.3. Análise de Tendências e Insights:

O YouTube emprega análises de dados para identificar tendências emergentes, entender o comportamento do usuário e fornecer insights valiosos para criadores de conteúdo e anunciantes. Isso ajuda na tomada de decisões estratégicas e no desenvolvimento de estratégias de marketing mais eficazes. (ELGIRIYEWITHANA, 2023)

4.4. Melhoria da Experiência do Usuário:

Data Science é usado para otimizar a experiência do usuário no YouTube, incluindo aprimoramentos na interface do usuário, recomendações de recursos, velocidade de carregamento de vídeos e personalização de feeds. (ELGIRIYEWITHANA, 2023)

4.5. Combate à Desinformação e Notícias Falsas:

O YouTube tem investido em tecnologias de Data Science para identificar e combater a disseminação de notícias falsas e desinformação em sua plataforma. Isso envolve a detecção de vídeos enganosos e a promoção de conteúdo confiável e verificado. (ELGIRIYEWITHANA, 2023)

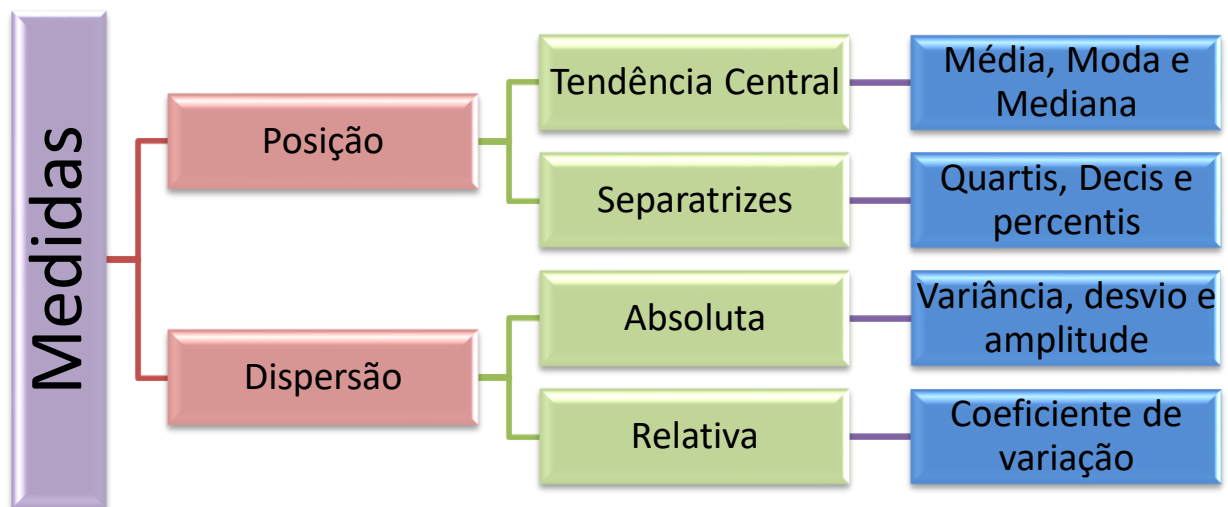
5. APRESENTAÇÃO DOS METADADOS E ANÁLISE EXPLORATÓRIA DE DADOS

O conjunto de dados utilizado neste trabalho é um arquivo CSV obtido de uma fonte pública, gerado no ano de 2023. A origem desses dados é uma pesquisa global disponibilizada pela plataforma Kaggle, uma comunidade online dedicada à análise de dados e aprendizado de máquina. Vale ressaltar que não há informações sobre a validade específica dos dados, uma vez que foram coletados e compartilhados sem uma indicação explícita de período de vigência.

Esses dados consistem em uma tabela que contém informações sobre os canais mais inscritos no YouTube, bem como suas estatísticas de visualizações e uploads. É importante destacar que esses dados não incluem informações sensíveis e foram disponibilizados publicamente para análise e pesquisa.

A utilização desses dados em trabalhos acadêmicos permite explorar e analisar tendências e padrões relacionados ao desempenho dos canais no YouTube, fornecendo insights valiosos sobre a popularidade e o engajamento do conteúdo na plataforma.

Figura 1 - Tabela de medidas



6. ANÁLISE E DISCUSSÃO DOS RESULTADOS

O conjunto de dados coletado consiste em informações sobre canais do YouTube, contendo 10 colunas e ocupando um espaço de 200.28 KB. As colunas incluem a posição do canal com base no número de inscritos, nome do canal, categoria/nicho do canal, título do canal, número total de vídeos enviados, país de origem do canal, abreviação do país e tipo do canal (por exemplo, indivíduo ou marca).

Embora existam algumas informações ausentes em algumas colunas, não foram identificados dados perdidos, anomalias ou outliers visualmente. No entanto, é observável que os dados são diversos, o que sugere uma ampla variedade de canais representados no conjunto de dados.

Ao todo, foram analisados 955 canais com informações sobre o número de inscritos. Com base nesses dados, as medidas de posição e dispersão foram calculadas e serão demonstradas abaixo após a análise detalhada no GitHub.

Essas informações fornecem uma visão geral do conjunto de dados, destacando suas características e peculiaridades, e estabelecendo a base para a análise estatística posterior.

Figura 2 - Número de vídeos por categoria

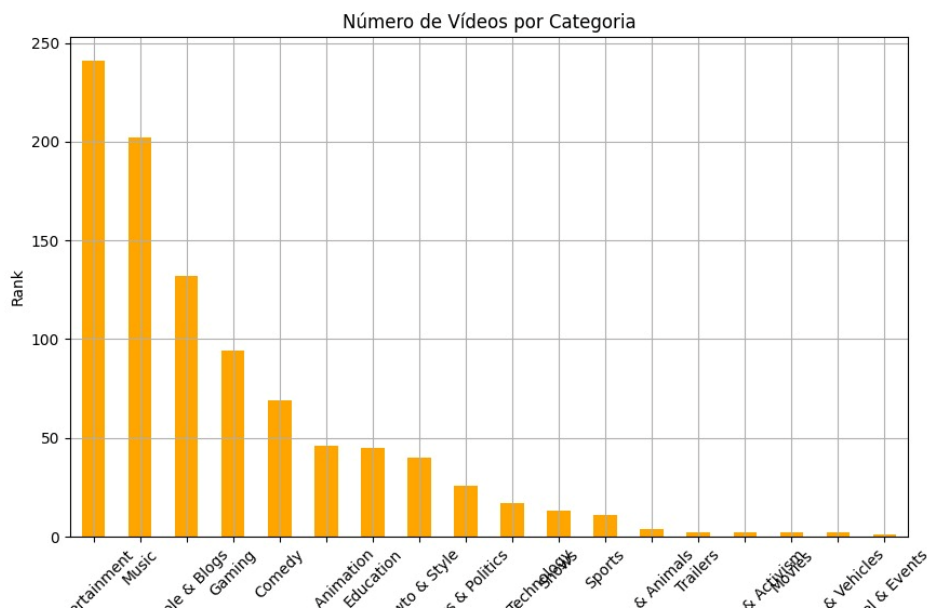


Figura 3 - Visualização vs País

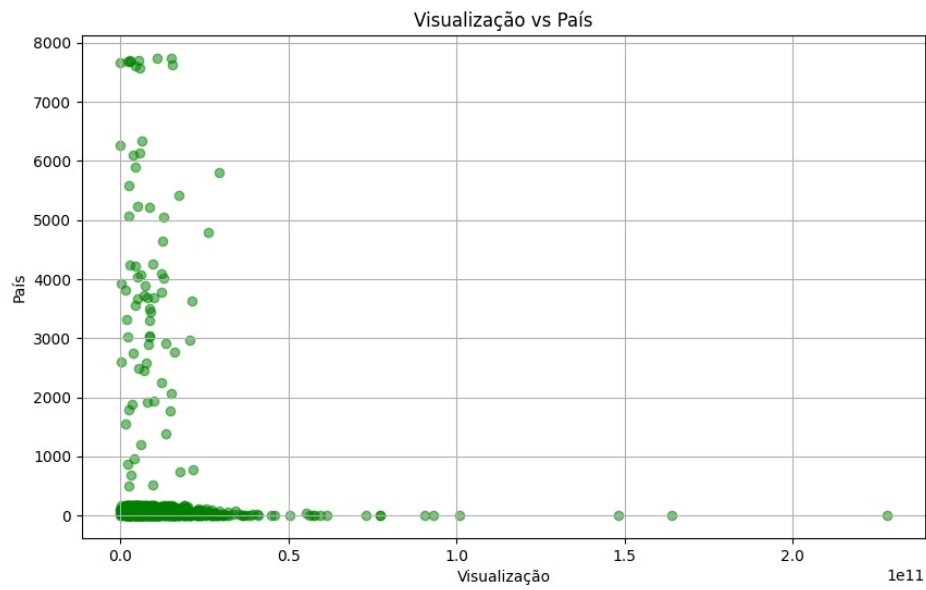
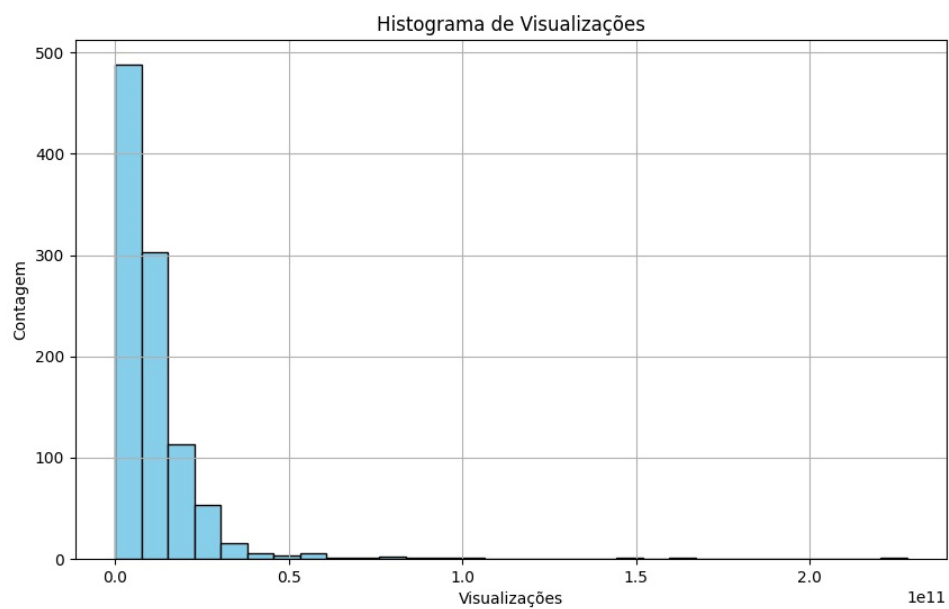


Figura 4 - Histograma de Visualizações



7. REFERÊNCIAS BIBLIOGRÁFICAS

ELGIRIYEWITHANA, Nidula; Global YouTube Statistics 2023. Disponível em: <https://www.kaggle.com/datasets/nelgiriyeewithana/global-youtube-statistics-2023>. Acesso em: 03 mar. 2024.

7.1 LINK GITHUB

<https://github.com/valdineyatilio/Projeto-Aplicado-I>