



PROJETO I - ANÁLISE EXPLORATÓRIA

INSIGHTS E DATASETS SOBRE O COMPORTAMENTO DE ASSINANTES DA NETFLIX

São Paulo 2024

PROJETO I - ANÁLISE EXPLORATÓRIA

INSIGHTS E DATASETS SOBRE O COMPORTAMENTO DE ASSINANTES DA NETFLIX

Curso:

- **Ciências de Dados**

Componente Curricular:

- **Projeto Aplicado**

Professor:

- **Thiago Graziani Traue**

Autores:

- **Gustavo Goes: 10442973.**
- **Tatiane Okochi: 10444313.**
- **Julia Raissa Silva: 10444671.**

São Paulo 2024

Sumário

Glossário.....	4
Objetivo	5
Descrição	5
História	5
Área de Atuação e Serviços.....	5
Serviços/Produtos	5
Fonte e Dados Apresentação dos Dados (Metadados)	6
Tipos de Dados	6
Formato dos Dados	6
Link para o GitHub do Projeto	6
Apresentação Da Empresa E Problemas De Pesquisa	7
Abordagem do Pensamento Computacional	7
Objetivo	8
Sensibilidade e Validade dos Dados.....	8
Proprietário dos dados e restrições de uso	8
Descrição dos Atributos	8

Glossário

- **Netflix:** Plataforma global de streaming que oferece filmes, séries, documentários e conteúdo original, com milhões de assinantes em mais de 190 países.
- **Dataset:** Conjunto de dados organizados que podem ser usados para análises e insights. No projeto, inclui informações sobre assinantes, como tipo de assinatura, demografia e dispositivos utilizados.
- **Análise de Dados:** Processo de inspecionar, limpar e modelar dados com o objetivo de descobrir informações relevantes e orientar decisões estratégicas. No caso, foca em retenção e comportamento de usuários.
- **Churn:** Taxa de cancelamento de assinantes, um indicador importante para avaliar a retenção e identificar padrões de abandono.
- **Padrões de Uso:** Tendências ou comportamentos recorrentes observados nos dados dos assinantes, como frequência de visualização ou dispositivos preferidos.
- **Personalização:** Processo de adaptar a oferta de conteúdo e campanhas com base no comportamento e perfil dos usuários, melhorando a experiência do cliente.
- **Validade dos Dados:** Avaliação sobre se os dados estão atualizados e ainda aplicáveis para a análise, essencial para garantir resultados precisos.
- **Sensibilidade dos Dados:** Grau de confidencialidade dos dados, especialmente os que envolvem informações pessoais, exigindo conformidade com a LGPD.
- **Legislação LGPD:** Lei Geral de Proteção de Dados, que estabelece regras para o tratamento de dados pessoais e protege a privacidade dos usuários no Brasil.
- **Dispositivo Utilizado:** Aparelho (smartphone, TV, tablet, etc.) usado pelo assinante para acessar a plataforma, importante para entender preferências de consumo.
- **Proposta Analítica:** Sugestão de métodos e estratégias para analisar dados e gerar insights que orientem ações, como campanhas de retenção e personalização de marketing.

Objetivo

Este projeto tem como foco a análise exploratória dos dados de assinantes da Netflix, utilizando datasets disponíveis no Kaggle. O objetivo é identificar padrões de comportamento e consumo, contribuindo para a retenção de clientes e a personalização de campanhas de marketing. Os resultados serão apresentados por meio de dashboards interativos e relatórios automatizados para oferecer suporte à tomada de decisão estratégica.

Sobre a Netflix

Descrição

A Netflix é uma plataforma global de streaming que oferece uma ampla variedade de conteúdo, incluindo filmes, séries, documentários e produções originais. A empresa foi fundada em 1997, iniciando suas operações como serviço de aluguel de DVDs. Em 2007, migrou para o modelo de streaming, consolidando-se como uma das maiores plataformas de entretenimento digital do mundo. Atualmente, opera em mais de 190 países, atendendo milhões de assinantes com diferentes perfis.

História

A Netflix passou por diversas transformações ao longo dos anos. A expansão para o streaming digital e o investimento em produções originais consolidaram seu modelo de negócios inovador. A empresa continua evoluindo para acompanhar as mudanças nas preferências de consumo de entretenimento em um mercado altamente competitivo.

Área de Atuação e Serviços

Serviços/Produtos

A Netflix opera no **setor de entretenimento digital**, oferecendo conteúdo por meio de diferentes tipos de assinatura:

- **Basic:** Acesso em uma tela por vez, sem resolução HD.
- **Standard:** Acesso simultâneo em duas telas, com resolução Full HD (1080p).
- **Premium:** Acesso simultâneo em até quatro telas, com conteúdo em 4K HDR.

Mercado-Alvo: Assinantes globais de diferentes faixas etárias, perfis demográficos e interesses culturais.

Fonte e Dados

Apresentação dos Dados (Metadados)

Os dados analisados neste projeto foram obtidos do Kaggle. Eles representam informações demográficas e comportamentais dos assinantes da Netflix, possibilitando uma análise detalhada do consumo por diferentes perfis.

Tipos de Dados

- ID do Assinante
- Tipo de Assinatura: Basic, Standard, Premium
- Valor da Assinatura/mês
- Data de Adesão e Última Data de Pagamento
- Meses de Utilização
- País, Idade e Gênero
- Dispositivo Utilizado: Smartphone, TV, tablet, entre outros.

Formato dos Dados

- CSV
- JSON
- CSV: Formato principal para análise em Python e Power BI.

Qualidade dos Dados

Os dados foram avaliados quanto à integridade e consistência, garantindo que apenas informações relevantes e válidas sejam utilizadas na análise. Foram aplicadas técnicas de limpeza e normalização para remover valores inconsistentes.

Link para o GitHub do Projeto

Repositório do GitHub: [goessgg/PROJETO-APLICADO---NETFLIX](https://github.com/goessgg/PROJETO-APLICADO---NETFLIX): Este projeto visa desenvolver uma proposta de solução analítica utilizando serviços da Amazon Web Services (AWS). Através da implementação de tecnologias de computação em nuvem, este projeto buscará resolver um problema específico ou otimizar um processo em uma organização.

Apresentação Da Empresa E Problemas De Pesquisa

Nome da Empresa: Netflix

- Missão: Levar entretenimento de alta qualidade a todos, em qualquer lugar e a qualquer momento.
- Visão: Ser a plataforma de entretenimento mais amada e acessível do mundo.

Problemas de Pesquisa

- Como a Netflix pode melhorar a retenção de assinantes?
- Quais dispositivos são mais utilizados pelos assinantes?
- Como o perfil demográfico afeta a escolha do tipo de assinatura?
- Quais padrões podem ser identificados nos cancelamentos e renovações?
-

Abordagem do Pensamento Computacional

1. Divisão de Problemas:

- Analisar comportamento por **tipo de assinatura** (Basic, Standard, Premium).
- Estudar padrões por **país e dispositivo** utilizado.
- Identificar sazonalidade em **cancelamentos e pagamentos**.

2. Observação de Padrões:

- Identificar **picos de adesão** e **motivos de cancelamento**.
- Detectar padrões demográficos que influenciam o consumo.

3. Abstração:

- Simplificar a análise focando em **atributos comuns**, como tipo de assinatura e perfil de consumo.

4. Algoritmos:

- Desenvolver scripts em **Python** para análises automatizadas.
- Usar **Power BI** para criar dashboards interativos com insights em tempo real.

Objetivo

O projeto visa fornecer insights para:

- Melhorar a retenção de assinantes e reduzir o churn.
- Otimizar campanhas de marketing com foco em personalização.
- Aprimorar a oferta de conteúdo com base no comportamento dos usuários.

Tipo de Arquivo:

- Formato: CSV.
- Origem: Dataset público do Kaggle.

Sensibilidade e Validade dos Dados

- Sensibilidade: O tratamento dos dados segue as normas da LGPD, garantindo a privacidade dos assinantes.
- Validade: Os dados foram validados e atualizados para garantir que as análises reflitam a realidade atual.

Proprietário dos dados e restrições de uso

- Proprietário: Netflix
- Restrições: O uso do dataset é restrito a fins educacionais, não podendo ser distribuído sem autorização.

Descrição dos Atributos

- ID do Assinante: Identificador único do cliente.
- Tipo de Assinatura: Basic, Standard ou Premium.
- Valor da Assinatura/mês: Custo mensal do plano escolhido.
- Data de Adesão: Data inicial da assinatura.
- Última Data de Pagamento: Data do pagamento mais recente.
- Meses de Utilização: Tempo total de permanência como assinante.
- País, Idade e Gênero: Dados demográficos do cliente.
- Dispositivo Utilizado: Equipamento mais utilizado para acessar a Netflix.