Logotipo

Descrição gerada automaticamente

***Documentação do Código: Integração de APIs e Geração de Banco de Dados***

Kevelin Suellen de Souza Oliveira

Turma: 58875

Data de criação: 22/07/2024

Última modificação: 17/08/2024

Sumário

[Objetivo 3](#_Toc175040630)

[Público alvo 3](#_Toc175040631)

[Saída 3](#_Toc175040632)

[Nível de privacidade 3](#_Toc175040633)

[Pré requisitos 3](#_Toc175040634)

[Ambientes 3](#_Toc175040635)

[Bibliotecas 3](#_Toc175040636)

[Arquivos 4](#_Toc175040637)

[API’s utilizadas 4](#_Toc175040638)

[Funções criadas 4](#_Toc175040639)

[Tratamentos aplicados 4](#_Toc175040640)

[ **Tratamento de datas** 4](#_Toc175040641)

[ **Tratamento de Valores Ausentes** 4](#_Toc175040642)

[ **Conversão de Tipos** 5](#_Toc175040643)

[ **Agregação** 5](#_Toc175040644)

[Método de saída 5](#_Toc175040645)

[Exemplo de consulta 6](#_Toc175040646)

[Versionamento 6](#_Toc175040647)

[Referências 6](#_Toc175040648)

# Objetivo

Este projeto tem por finalidade extrair dados da API do YouTube para analisar e comparar canais na plataforma. O objetivo é reunir informações relevantes sobre o desempenho dos canais e dos vídeos, possibilitando uma análise detalhada e comparativa. Foi necessário a instalação das bibliotecas Pandas, google-api-python-client.

## Público alvo

Este código é voltado para analistas de dados e profissionais de marketing digital que desejam fornecer insights valiosos para criadores de conteúdo. Ele permite coletar métricas importantes para acompanhar o crescimento de canais e identificar tópicos que possam aumentar o engajamento do público.

## Saída

O código gerado nos fornece um arquivo chamado 'Análise\_Youtube.xlsx' e 4 abas sendo elas:

'Estatísticas dos Canais' – com as colunas channel\_id, channel\_name, published\_date, published\_tim, country, subscribers, view, total\_videos e playlist\_id;

'Detalhes dos Vídeos' - channel\_name, vídeo\_id, title, published\_date, published\_time, duration, category\_id, views, likes e comments;

'Uploads Mensais' - channel\_name, month e uploads;

'Média de Engajamento' – metric, avarege e channel\_name.

## Nível de privacidade

Os dados coletados são públicos, pois podem ser acessados por qualquer usuário que navegue pelo Youtube.

Entretando, para conseguir acesso a API, é necessário solicitar uma chave. O acesso a chave é feito através do google clound.

# Pré requisitos

## Ambientes

O código foi executado em Python 3.7 utilizando o Visual Code como editor.

## Bibliotecas

As bibliotecas utilizadas são:

* googleapiclient.discovery: v Interação com APIs do Google
* Pandas: Manipulação e análise de dados.
* Requests: Realização de requisições HTTP.
* Isodate: Manipulação de datas e horários em formato ISO 8601.
* Sqlite3: Interação com bancos de dados SQLite.

## Arquivos

**Chave da API:** O código requer que a chave da API do YouTube seja inserida diretamente no código, substituindo 'YOUR\_API\_KEY' pela chave válida que você obtém no Google Cloud Console.

## API’s utilizadas

A API utilizada foi a YouTube Data API v3.Ela permite o acesso e interação com os dados do YouTube, como informações sobre vídeos, canais, playlists, comentários, etc.

Para acessar a API, é necessário criar um projeto no Google Cloud Console, habilitar a YouTube Data API v3 e gerar credenciais de API.

## Funções criadas

* **get\_channel\_id(youtube, custom\_url):**

Busca o ID do canal usando a API e retorna o ID encontrado.

* **get\_channel\_stats(youtube, channel\_ids):**

Obtém estatísticas dos canais, como número de inscritos, visualizações e vídeos publicados.

* **get\_video\_ids(youtube, playlist\_id):**

Retorna uma lista contendo os IDs dos vídeos encontrados.

* **get\_video\_details(youtube, video\_ids, channel\_name):**

Obéem detalhes dos vídeos, como título, data de publicação, visualizações, curtidas e comentários.

* **get\_channel\_activity(youtube, playlist\_id):**

Calcula a frequência mensal de uploads e o engajamento médio (visualizações, curtidas e comentários) dos vídeos do canal.

## Tratamentos aplicados

* **Tratamento de datas**

Formatação**:** A parte da A data é convertida para um objeto datetime do Pandas e formatada para o padrão dia/mês/ano. E como as datas são separadas usando o separador 'T' elas são tratadas para que venham separadas por barras (/).

* **Tratamento de Valores Ausentes**

Uso de get: O método get é usado para obter um valor de um dicionário, se a chave 'publishedAt' não existir no dicionário, o valor 'N/A' será atribuído.

* **Conversão de Tipos**

O número de inscritos, que é uma string no JSON, é convertido para um número inteiro usando a função int(). Se a chave 'subscriberCount' não existir, o valor padrão será 0.

* **Agregação**

Os dados são agrupados por canal e mês usando o método groupby.

* **Erro de conexão**

Verificação de falhas na comunicação com a API.

## Método de saída

Os dados extraídos da API, irão gerar um arquivo .XLSX com 4 abas.

Primeira aba - Estatísticas dos Canais

channel\_id - object

channel\_name - object

published\_date - object

published\_time - object

country - object

subscribers - int64

views - int64

total\_videos - int64

playlist\_id – object

Segunda aba – Destalhes dos vídeos

channel\_name - object

video\_id - object

title - object

published\_date - object

published\_time - object

duration - object

category\_id - object

views - int64

likes - int64

comments - int64

Terceira aba – Uploads mensais

channel\_name - object

month - object

uploads - int64

Quarta aba – Média de enganjamento

Metric - object

average - float64

channel\_name – object

## Exemplo de consulta

Para executar uma consulta as tabelas, foi usado o método head() para visualizar as primeiras linhas do código a fim de verificar como estão sendo extraídos os dados.

Exemplo: avg\_engagement\_df.head(3) – Nos traz as primeiras 3 linhas da tabela Média de enganjamento.

## Versionamento

Um arquivo requiriment.txt foi gerado e será fornecido juntamente com a documentação deste projeto.

## Referências

Foi utilizado como referência o tutorial do canal Programação dinâmica do Youtube

https://www.youtube.com/watch?v=olDCJ1w3FLM&t=538s