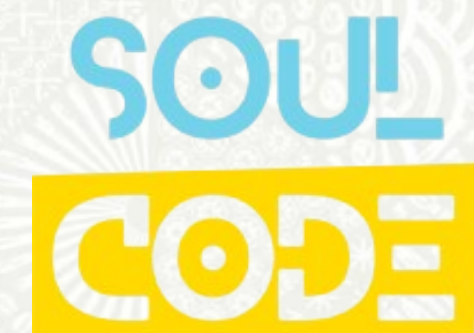


Bem-vindos

PROJETO FINAL

O CURSO



Enquanto membros do Curso de Engenharia de Dados, é um dos critérios da **Soul Code Academy**, capacitar e apoiar a **empregabilidade** na área de Tecnologia.

Algumas qualidades que adquirimos

Comunicação / Iniciativa
Adaptabilidade / Autonomia
Flexibilidade / Persistência

Desenvolvendo Qualidades



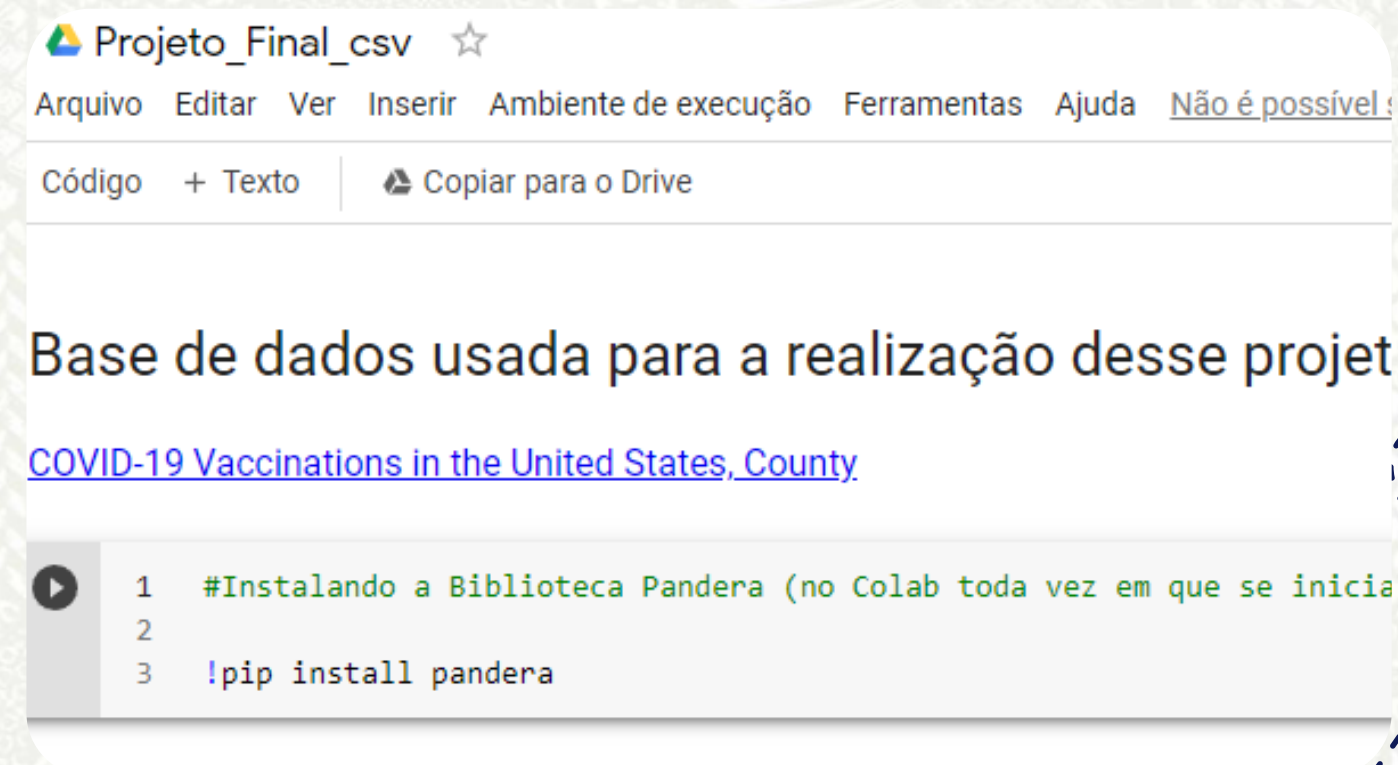
O DESAFIO



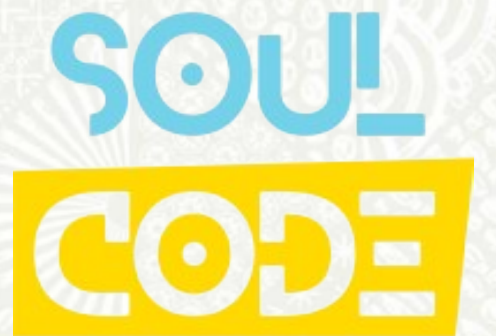
Utilizar os conceitos e ferramentas vistos durante o curso para tratar, **organizar e modelar os dados** de dois datasets.

Clicando na imagem e/ou copiando o link, terá acesso ao **notebook** deste desafio

<https://bit.ly/NotebooksColab>



DATASETS



Tema: Saúde

Escolhemos dois datasets relacionados a **hospitalização**, **óbitos** e **vacinação** de Covid-19 nos EUA em Connecticut.

Plataforma escolhida para extração dos datasets

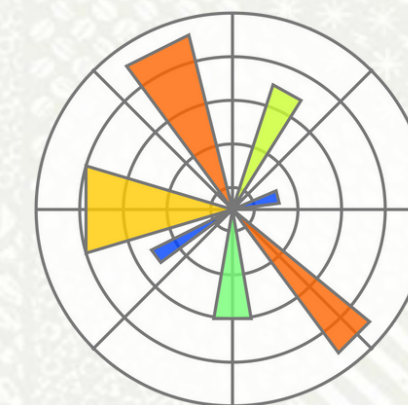


Extração de Dados

TIMELINE



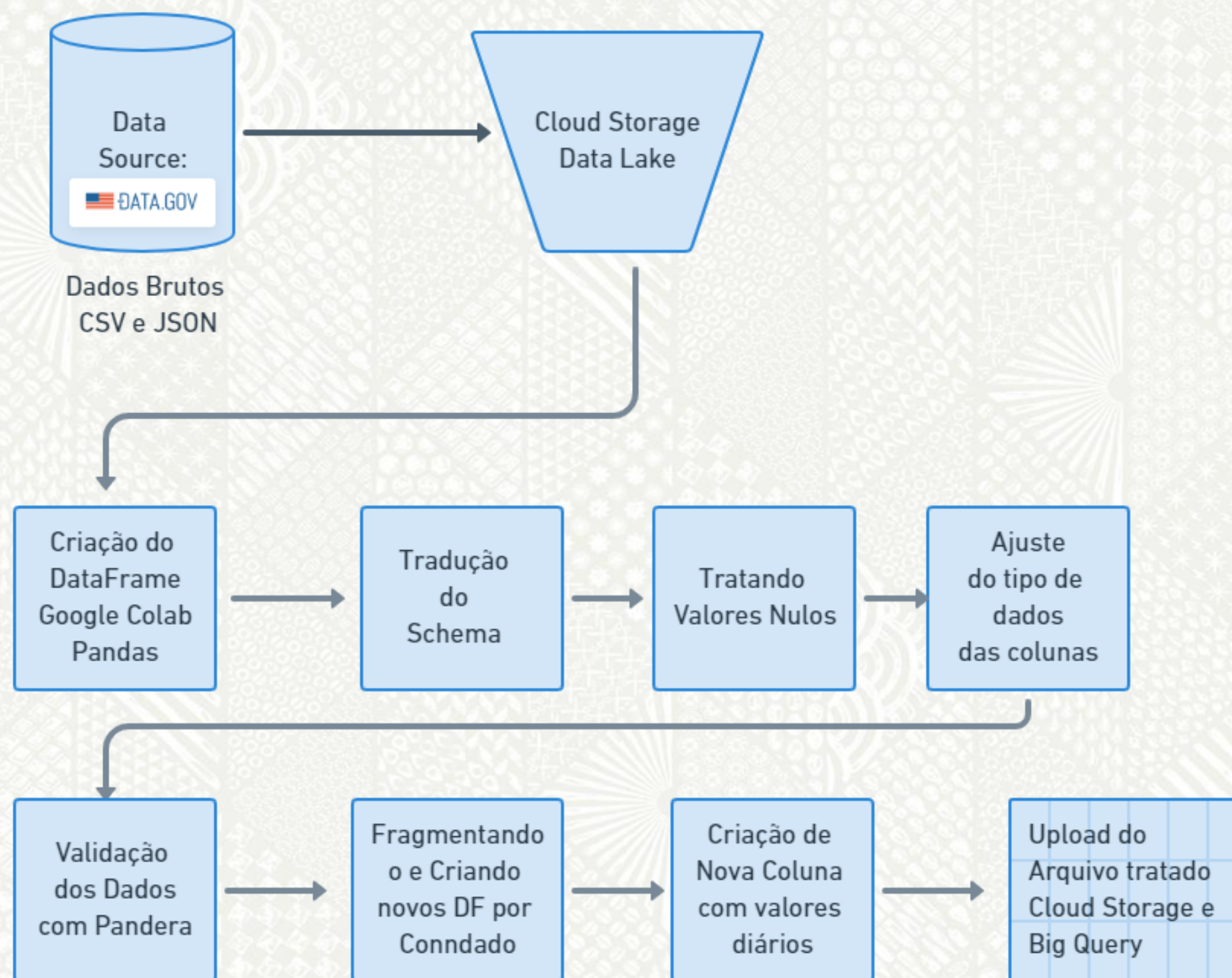
FERRAMENTAS



SQL



WORKFLOW - ETL



ETL



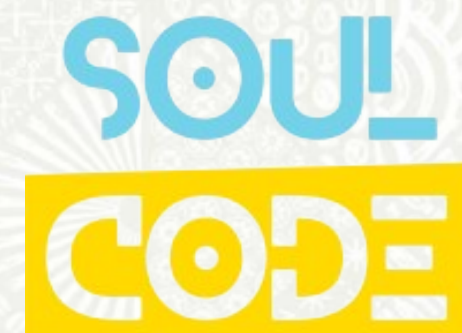
Importando Bibliotecas para a utilização do bucket do GCP

```
[ ] 1  # Importando Bibliotecas
      2
      3  import pandas as pd
      4  import pandera as pa
      5  import json
      6  from google.cloud import storage
      7
```

Vamos ao Colab?

Tratamento dos Dados

ETL

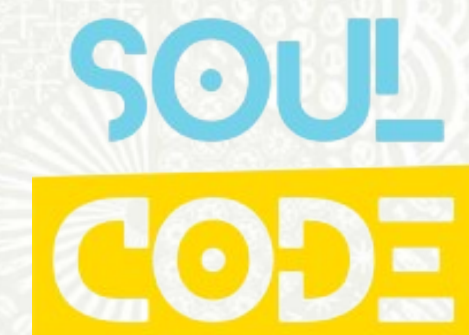


Vamos ao **BigQuery**?

```
4  /* Comando utilizado para filtrar apenas Condados c
5
6  SELECT
7  Data_Atualizacao,
8  Condado, Casos_Confirmados
9  FROM `projeto-finalsc01.hospitalizacao.Hospitalizac
10 WHERE Casos_Confirmados > 5000;
11
12 /* Comando utilizado para Somar os Casos Confirmado
13
14 SELECT Condado,
15 SUM(Casos_Confirmados)
16 AS Soma_Casos_Confirmados
17 FROM `projeto-finalsc01.hospitalizacao.Hospitalizac
18 GROUP BY Condado
```

Análise dos Dados

DASHBOARD

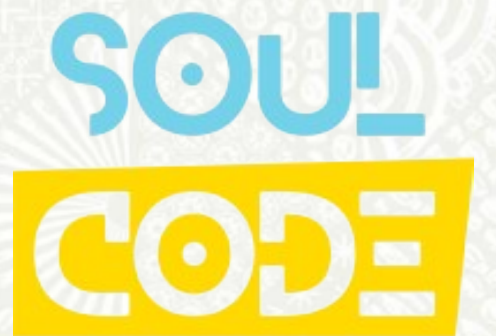


Vamos ao **DataStudio**?

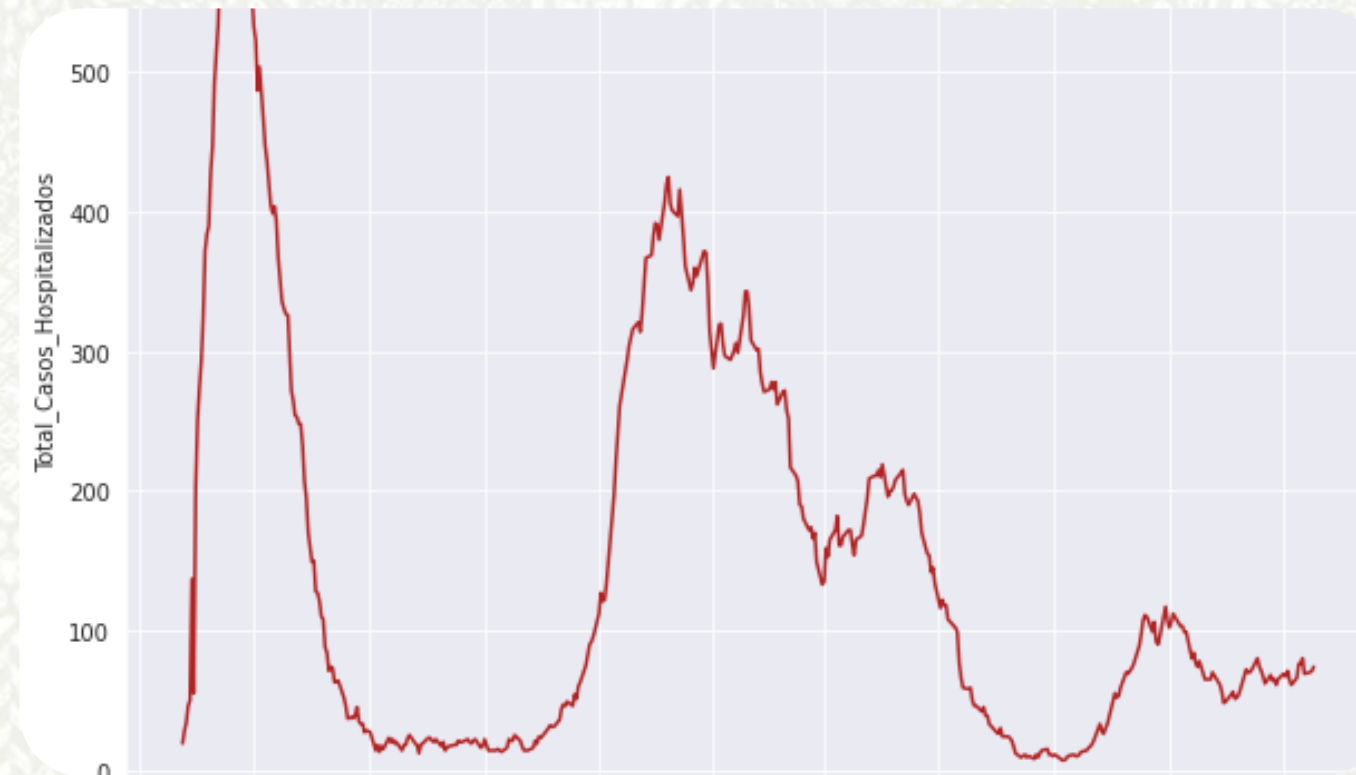
Vacinação	
Vacinados	Idosos
1,9 mi	426.416
Idosos (%)	Média por dia
93,6	15,7 mil

Análise dos Dados

DASHBOARD

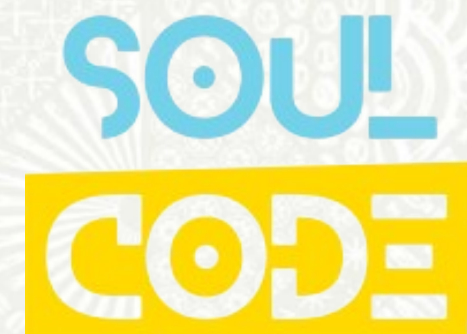


Vamos ao **seaborn**?



Análise dos Dados

AS CONSIDERAÇÕES



O projeto teve como objetivo flexibilizar a **capacitação** em projetar, criar, operacionalizar, proteger e monitorar sistemas de **processamento de dados** com ênfase particular em conformidade, escalabilidade, eficiência, confiabilidade e portabilidade. Utilizando todas as ferramentas e recursos lecionadas no curso.

DÚVIDAS?



Contatos

Humberto Oliveira | humberto-barbosa-oliveira



Phylyp Cavalcante | phylyp-cavalcante



Tyciano Junior | tycianojunior



Vitor Siqueira | vitorsiqueirads

