DAC Phase 3   
Problem Statement: COVID\_19 CASES ANALYSIS

**Loading and Pre-processing of data:**

from google.colab import drive

drive.mount('/content/drive')

**Loading data**

import pandas as pd

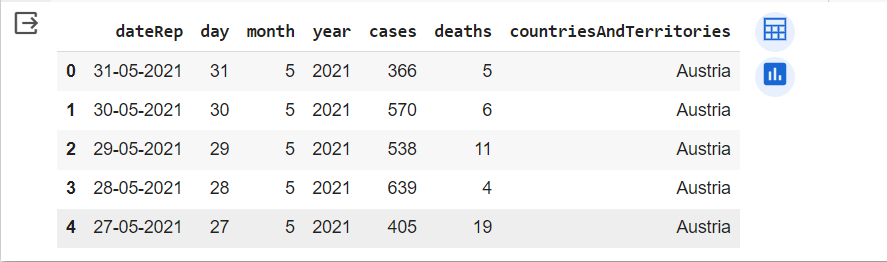
import numpy as np

import matplotlib.pyplot as plt

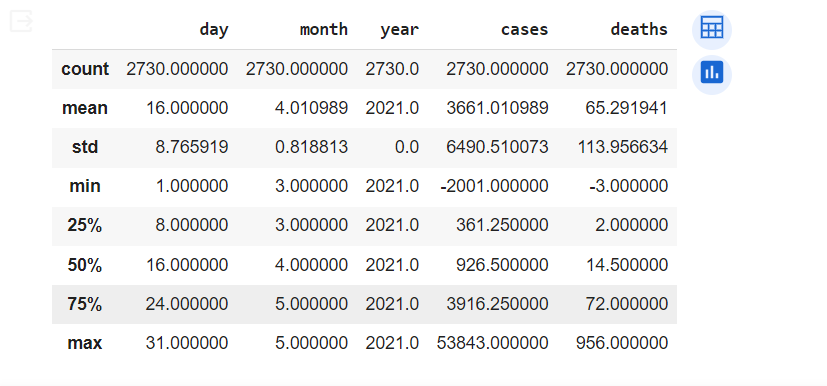
data = pd.read\_csv('/content/drive/MyDrive/NAAN MUDHALVAN/Covid\_19\_cases4.csv')

data.head(5)

**OUTPUT**

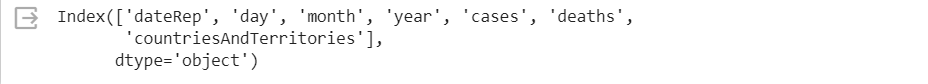
****

data.describe()

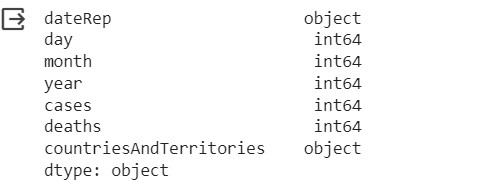
****

This command is used to view the brief summary of the dataset. We can see the mathematical parameters such as percentiles, standard deviation , mean, minimum and maximum values and count of each column.

data.columns

****

data.dtypes

****

Info command is used check the datatype of every column and the count of each column. The difference between the describe() and info() is that describe command will give the mathematical parameters but info command will not give the mathematical parameters such as mean and standard deviation

data.info()

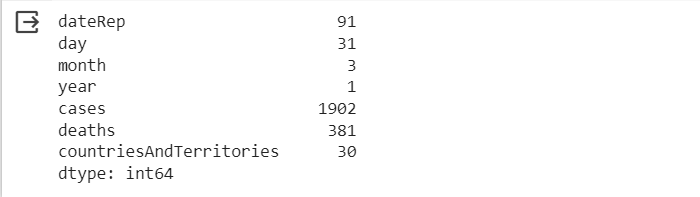
****

data.isna().sum()

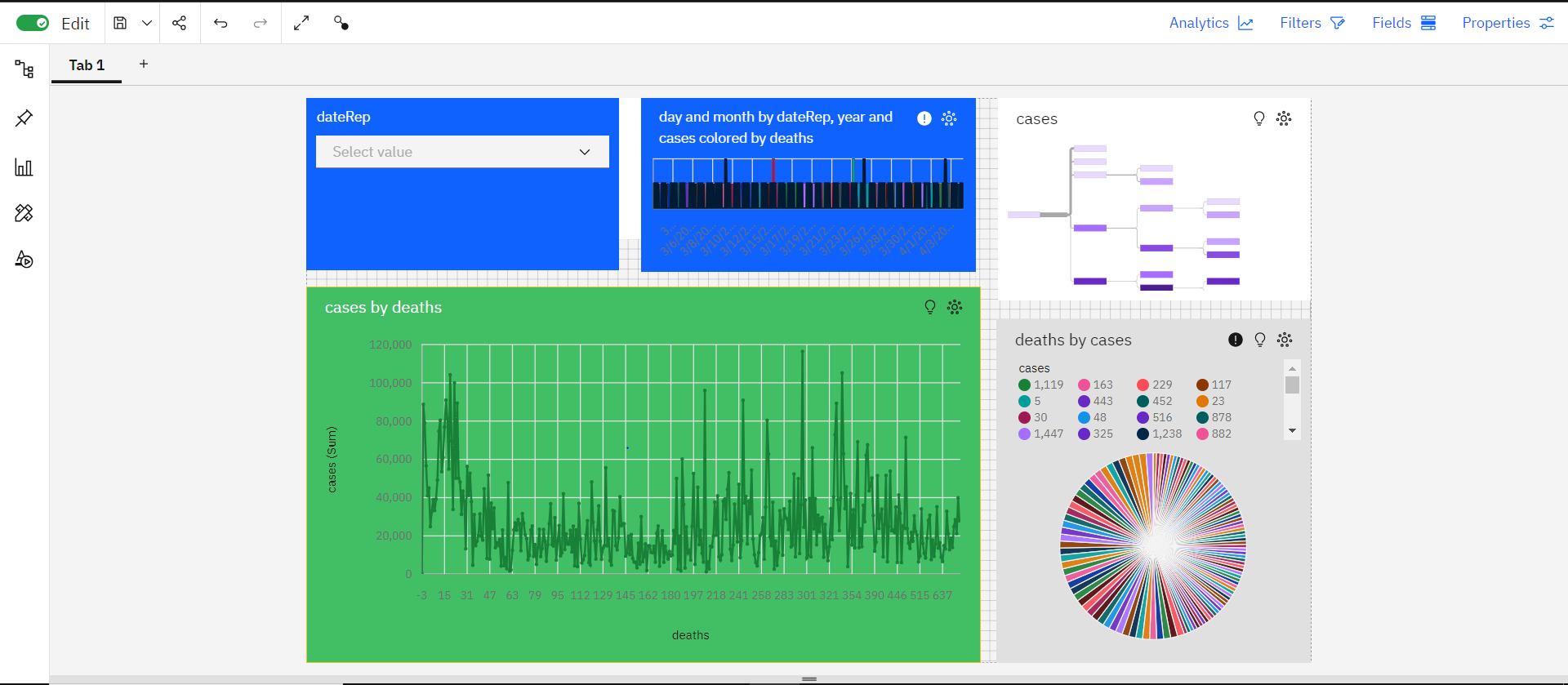
****

The above command is used to check for null values in each column. We can see that there are null values in the columns such as day, month, year. It is very necessary to take action to clear the null values in the data set

data.nunique()

****

The data.nunique() command is used in pandas to count the number of unique values in each column of a DataFrame. When you execute data.nunique(), it returns a Series where the index is the column names, and the values represent the count of unique values in each respective column.

****

1. Start by integrating COVID-19 data from various sources (e.g., John Hopkins University, government agencies) into the IBM Cognos platform.
2. Use IBM Cognos to create interactive dashboards that visualize COVID-19 data
3. Leverage geospatial capabilities in IBM Cognos to analyze the geographic spread of COVID-19.
4. Analyze the impact of COVID-19 on different demographics, such as age groups, gender, and pre-existing health conditions. Identify vulnerable populations and tailor interventions accordingly.
5. IBM Cognos allows for easy data sharing and collaboration. Ensure that your analysis is accessible to decision-makers, public health authorities, and the general public, as needed.