COVID-19 CASE ANALYSIS USING COGNOS

INTRODUCTION:

Analyzing COVID-19 cases is essential in the ongoing global pandemic. It involves collecting and interpreting data to understand the virus's spread, make informed decisions, and allocate resources effectively. This data-driven approach is vital for policymakers, healthcare professionals, and the public to combat the crisis. In a world connected by data and technology, we rely on tools and insights to effectively track and respond to COVID-19.

DATA COLLECTION:

To analyze COVID-19 cases comprehensively, a variety of data types are collected. These data points are critical for understanding the virus's spread and impact on public health and making informed decisions. The primary data collected for COVID-19 case analysis include:

Case Counts: The number of confirmed, probable, and suspected COVID-19 cases, which includes both active and resolved cases.

Demographic Information: Data on age, gender, race, ethnicity, and other demographic factors to identify patterns and disparities in infection rates.

Geographic Data: Information about the location of cases, including city, county, or region. This helps identify hotspots and track the virus's spread.

Testing Data: The number of tests conducted, testing methods, and results (positive, negative, or inconclusive).

Hospitalization Data: Information on the number of COVID-19 patients hospitalized, ICU admissions, and ventilator usage.

Mortality Data: The number of COVID-19-related deaths, including age, gender, and comorbidities of the deceased.

Vaccination Data: Data related to the distribution and administration of COVID-19 vaccines, including the number of people vaccinated, vaccine types, and vaccination rates.

Contact Tracing Information: Details about individuals who have been in contact with confirmed cases, aiding in identifying potential sources of transmission.

Feature Selection:

Open a new Syntax Editor session in SPSS Statistics by selecting File > New > Syntax.

Copy the following syntax into the Syntax Editor dialog box.

```
import pandas as pd

import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt

print('Modules are imported.')

Modules are imported.
```

importing "Covid19_Confirmed_dataset.csv" from "./Dataset" folder.

```
In[2]

df=pd.read_csv("../input/covid19/covid19_Confirmed_datas

et.csv")

df.head()
```

Out[2]:

Province/State□Country/Region□Lat□Long□1/22/20□1/23/20□1/24/20□1/25/20

1/26/20□1/27/20□...□4/21/20□4/22/20□4/23/20□4/24/20□4/25/20□4/26/20

4/27/20□4/28/20□4/29/20□4/30/20

0□NaN□Afghanistan□33.0000□65.0000□0□0□0□0□0□0□...□1092□1176□1279□1351

1463□1531□1703□1828□1939□2171

1□NaN□Albania□41.1533□20.1683□0□0□0□0□0□0□...□609□634□663□678□712□726

736□750□766□773

2□NaN□Algeria□28.0339□1.6596□0□0□0□0□0□0□...□2811□2910□3007□3127□3256

3382□3517□3649□3848□4006

3□NaN□Andorra□42.5063□1.5218□0□0□0□0□0□0□...□717□723□723□731□738□738

743□743□743□745

4□NaN□Angola□-11.2027□17.8739□0□0□0□0□0□0□...□24□25□25□25□25□26□27□27

27□27

Let's check the shape of the dataframe

In[3]:

df.shape

Out[3]:

(266, 104)

Model training:

Delete the useless columns
In[4]:
df.drop(["Lat","Long"],axis=1,inplace=True)
In[5]:
df.head()

Out[5]:

Province/State□Country/Region□1/22/20□1/23/20□1/24/20□1/25/20□1/26/20□1/27/20

1/28/20□1/29/20□...□4/21/20□4/22/20□4/23/20□4/24/20□4/25/20□4/26/20□4/27/20

4/28/20□4/29/20□4/30/20

0□NaN□Afghanistan□0□0□0□0□0□0□0□0□...□1092□1176□1279□1351□1463□1531□1703

1828□1939□2171

1□NaN□Albania□0□0□0□0□0□0□0□0□...□609□634□663□678□712□726□736□750□766□773

2□NaN□Algeria□0□0□0□0□0□0□0□0□...□2811□2910□3007□3127□3256□3382□3517□3649

3848□4006

3□NaN□Andorra□0□0□0□0□0□0□0□0□...□717□723□723□731□738□738□743□743□743□745

4□NaN□Angola□0□0□0□0□0□0□0□0□...□24□25□25□25□25□26

Aggregating the rows by the country

In[6]:

aggregating=df.groupby("Country/Region").sum()

In[7]:

aggregating.head()

Out[7]:

1/22/20□1/23/20□1/24/20□1/25/20□1/26/20□1/27/20□1/28/20□1/29/20

1/30/20□1/31/20□...□4/21/20□4/22/20□4/23/20□4/24/20□4/25/20□4/26/20

4/27/20□4/28/20□4/29/20□4/30/20

Country/Region

Afghanistan□0□0□0□0□0□0□0□0□0□0□...□1092□1176□1279□1351□1463□1531

1703□1828□1939□2171

Albania□0□0□0□0□0□0□0□0□0□0□...□609□634□663□678□712□726□736□750

766□773

Algeria□0□0□0□0□0□0□0□0□0□0□...□2811□2910□3007□3127□3256□3382

3517□3649□3848□4006

Andorra□0□0□0□0□0□0□0□0□0□0□...□717□723□723□731□738□738□743□743

743□745

Angola□0□0□0□0□0□0□0□0□0□0□...□24□25□25□25□25□26□27□27□27□27

In[8]:

aggregating.shape

Out[8]:

(187, 100)

Visualizing data related to a country for example China

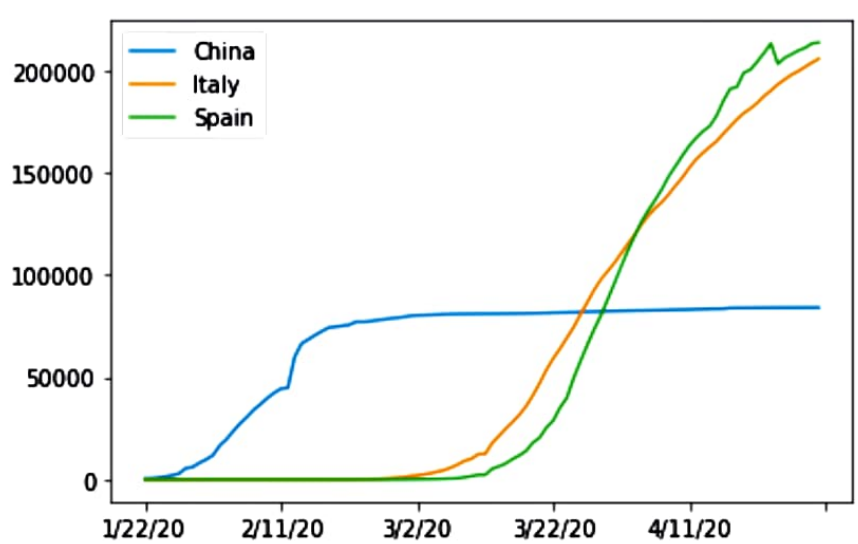
visualization always helps for better understanding of our data.

In[9]:

```
aggregating.loc["China"].plot()
aggregating.loc["Italy"].plot()
aggregating.loc["Spain"].plot()
plt.legend()
```

Out[9]:

<matplotlib.legend.Legend at 0x7f482e1e3990>



Calculating a good measure

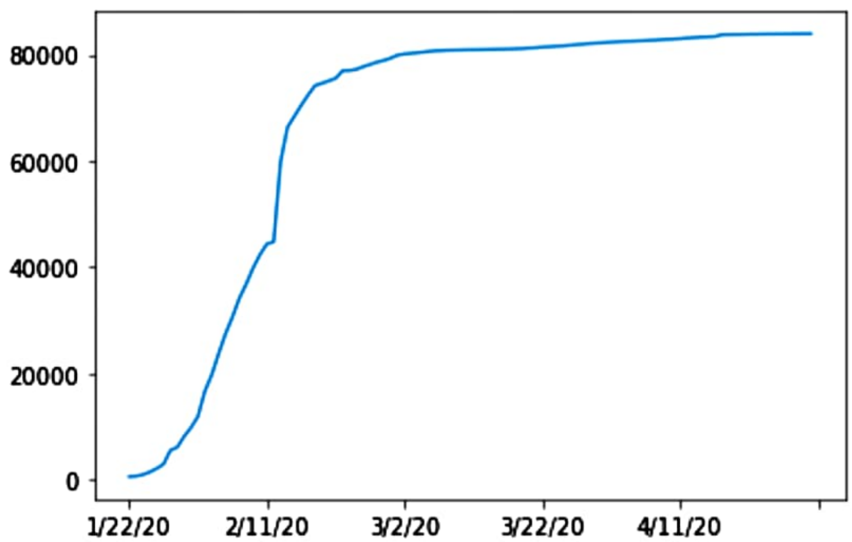
we need to find a good measure reperestend as a number, describing the spread of the virus in a country.

In[10]:

aggregating.loc['China'].plot()

Out[10]:

<matplotlib.axes._subplots.AxesSubplot at 0x7f482df94d90>

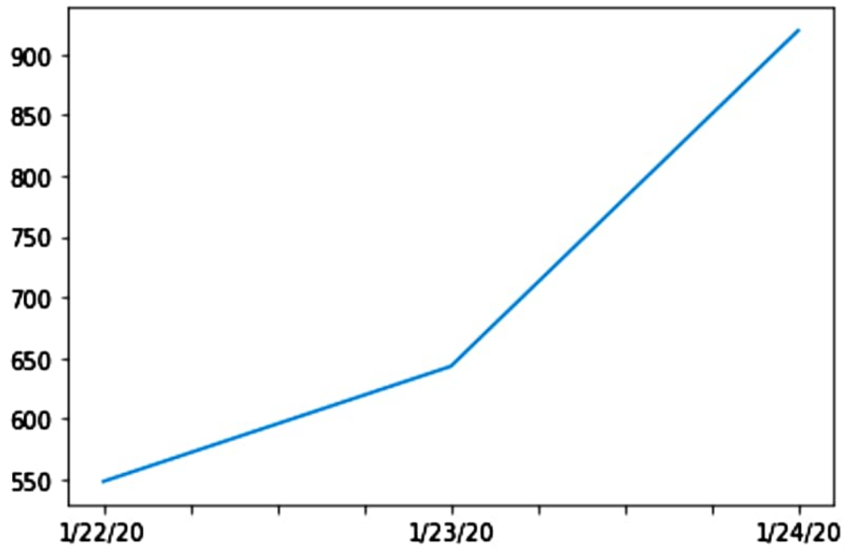


In[11]:

aggregating.loc['China'][:3].plot()

Out[11]:

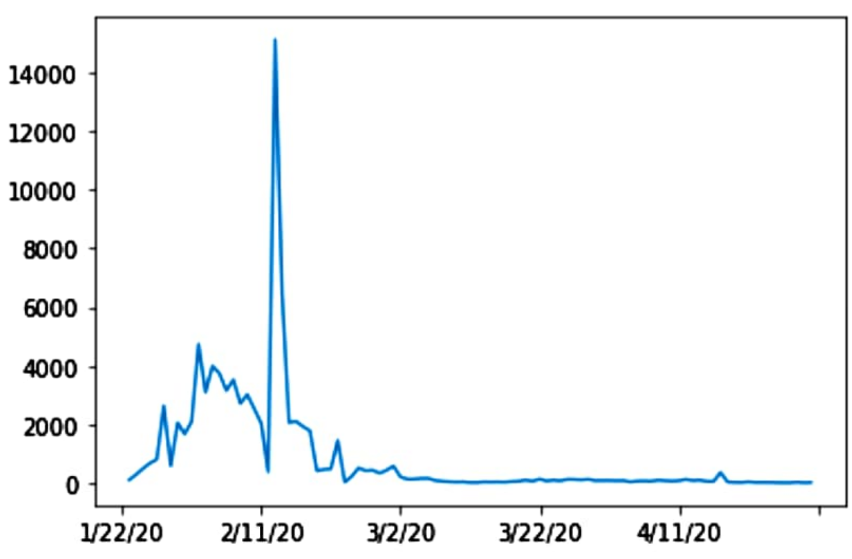
<matplotlib.axes._subplots.AxesSubplot at 0x7f482df83990>



caculating the first derivative of the curve

In[12]:
```
aggregating.loc['China'].diff().plot()
```

Out[12]:
<matplotlib.axes._subplots.AxesSubplot at 0x7f482df09290>



Genomic Sequencing: Genetic data from the virus to track mutations and variants can influence the virus's behavior and vaccine efficacy.

Symptom Data: Information on symptoms reported by COVID-19 patients, which can provide insights into the disease's presentation and severity.

Travel and Exposure History: Data on recent travel and potential exposure to COVID-19, helps to trace and prevent the virus's spread.

Healthcare System Capacity: Information on hospital bed availability, ventilators, and other critical resources to manage patient influx.

Public Health Measures: Data on the implementation and compliance with preventive measures, such as mask mandates and lockdowns.

Socioeconomic Data: Information on income, employment, education, and housing conditions to assess how socioeconomic factors influence infection rates.

Weather and Environmental Data: Climate and air quality data to explore potential environmental influences on the virus's spread.

Surveillance Data: Data from public health surveillance systems for monitoring trends and outbreaks.

Collecting and analyzing these various data types is crucial for assessing the pandemic's impact, identifying trends, and guiding public health responses.

DATA PROCESSING:

Data Collection: Gathering COVID-19 data from various sources, such as health departments, testing centers, and research organizations.

Data Cleaning: Removing duplicates, correcting errors, and handling missing or inconsistent data to ensure the dataset's quality.

Data Integration: Combining data from multiple sources into a single dataset for analysis.

Data Transformation: Converting data into a suitable format, such as date formatting, geospatial data integration, and creating new variables if needed.

Exploratory Data Analysis (EDA): Analyzing the data to understand its characteristics, such as trends, distributions, and correlations.

Data Visualization: Creating charts, graphs, and maps to visually represent the data and convey insights effectively.

Reporting and Interpretation: Summarizing findings and providing actionable insights for public health decision-making.

CONCLUSION:

In conclusion, COVID-19 case analysis is a multifaceted process with far-reaching implications for public health and societal well-being. Through real-time tracking, trend analysis, resource allocation, vaccination distribution, predictive modeling, contact tracing, mutational analysis, policy development, public awareness, and scientific research, data processing plays a central role in managing and controlling the pandemic. It empowers governments, healthcare systems, and the public to make informed decisions, adapt strategies, and ultimately work together to combat this global health crisis. As the pandemic evolves, data analysis remains an indispensable tool in our ongoing efforts to safeguard public health and overcome the challenges posed by COVID-19.