COVID 19 USING COGNOS

Data Analytics with cognos – Phase 5

DOCUMENTATION

Team Members:

1.Sageetha.S (au613021205041)
2.Deepika.M(au613021205005)
3.Srivarthini.K au613021205054)
4.Harshini.M(au613021205016)
5.Elakkiya V(au613021205012)

Project Objectives:

The project involves analyzing COVID-19 cases and deaths data using IBM Cognos. The objective is to compare and contrast the mean values and standard deviations of cases and associated deaths per day. This project is to develop a comprehensive and interactive dashboard using IBM Cognos that provides insights into the COVID-19 pandemic. The dashboard will be designed to cater to various stakeholders, including healthcare professionals, policymakers, and the general public.

Analysis Approach :

Design Thinking:

1. Data Collection Strategy:

To collect data, we can use APIs provided by various sources such as the World Health Organization
(WHO), the Centers for Disease Control and Prevention (CDC), and the John Hopkins University (JHU).
These APIs provide access to real-time data on confirmed cases, deaths, and recoveries.

2. Visualization Strategy:

To visualize the data, we can use various chart types and maps provided by IBM Cognos. For example,
we can create line charts to display trends in confirmed cases, deaths, and recoveries over time. We
can also create bar charts to compare the number of cases, deaths, and recoveries among different
countries or regions. Additionally, we can create heat maps to visualize the severity of the pandemic
in different locations.

3. Predictive Model:

To predict future trends in COVID-19 cases, deaths, and recoveries, we can use machine learning algorithms such as linear regression, support vector machines, or neural networks. These algorithms can be trained on historical data to make predictions based on patterns and trends observed in the past.

4. Analysis Objective:

The primary objective of this analysis is to provide insights into the progression of the COVID-19
pandemic. By visualizing the data and incorporating predictive models, we can help stakeholders
understand the current state of the pandemic and make informed decisions about public health
interventions

LOADING THE DATASET:

1.Importing libraries Here, for preprocessing the dataset and manipulate the data, pandas is the library used to frame the data.

Code: import pandas as pd

2.Loading the dataset In this step, we are framing the data into the table using DataFrame in pandas, and display the head or 5 rows of the dataset.

Code: # Replace with the actual filename
file_path="C:/Users/91962/Documents/country_vaccinations.csv" df = pd.read_csv(file_path)

EXPLORING THE DATA SET:

After framing data, the first few or five rows of the data in

displayed using the head() function.

Code:

data

Output:

	dateRep	day	month	year	cases	deaths	countriesAndTerritories
0	31-05-2021	31	5	2021	366	5	Austria
1	30-05-2021	30	5	2021	570	6	Austria
2	29-05-2021	29	5	2021	538	11	Austria
3	28-05-2021	28	5	2021	639	4	Austria
4	27-05-2021	27	5	2021	405	19	Austria
...	...	...	...	...	...	...	...
2725	06-03-2021	6	3	2021	3455	17	Sweden
2726	05-03-2021	5	3	2021	4069	12	Sweden

2727	04-03-2021	4	3	2021	4884	14	Sweden
2728	03-03-2021	3	3	2021	4876	19	Sweden
2729	02-03-2021	2	3	2021	6191	19	Sweden

2730 rows × 7 columns

Visualization Types:

Data will be visualized using a combination of charts, graphs, and maps. Common visualization types include:

- Bar charts to show age group distribution.
- Pie charts to represent gender distribution.
- Stacked bar charts to depict educational background. Choropleth maps to display regional distribution

Key Objectives:

Visualization:

- Utilize data visualization libraries such as Matplotlib and Seaborn to create visual representations of the demographic data.
- Create histograms or density plots to visualize the age distribution.
- Use bar charts to display the distribution of marginal workers across industrial categories. Generate count plots to visualize the gender distribution

Interpretation and Insights:

- Analyze the visualizations to extract insights regarding the demographics of marginal workers. Identify any notable patterns, disparities, or trends in the data.

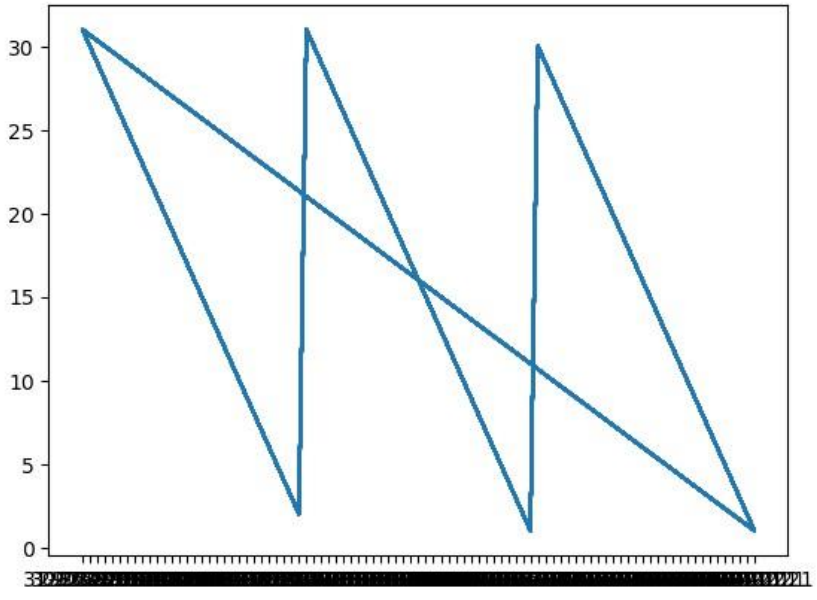
Python code:

Data visualization code:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
data = pd.read_csv("C:/Users/sagee/Downloads/Covid_19_cases4.csv")
Y = data.iloc[61:,1].values
R = data.iloc[61:,3].values
D = data.iloc[61:,5].values
```

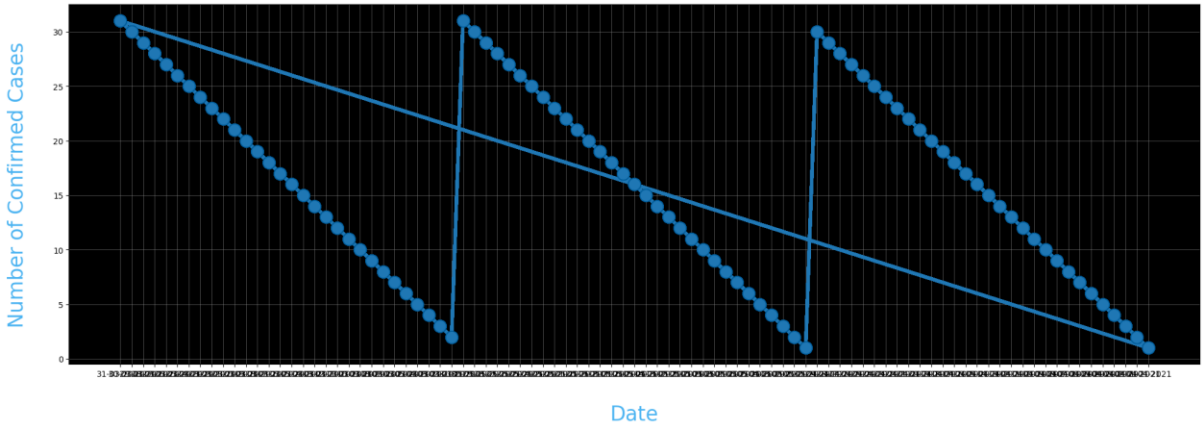
```
X = data.iloc[61:,0]
plt.plot(X,Y)
```

OUTPUT:




```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
data = pd.read_csv("C:/Users/sagee/Downloads/Covid_19_cases4.csv")
Y = data.iloc[61:,1].values
R = data.iloc[61:,3].values
D = data.iloc[61:,5].values
X = data.iloc[61:,0]
plt.figure(figsize=(25,8))
ax = plt.axes()
ax.grid(linewidth=0.4, color='#8f8f8f')
ax.set_facecolor("black")
ax.set_xlabel('\nDate',size=25,color='#4bb4f2')
ax.set_ylabel('Number of Confirmed Cases\n',
size=25,color='#4bb4f2')
ax.plot(X,Y,
color='#1F77B4',
marker='o',
linewidth=4,
markersize=15,
markeredgecolor='#035E9B')
```

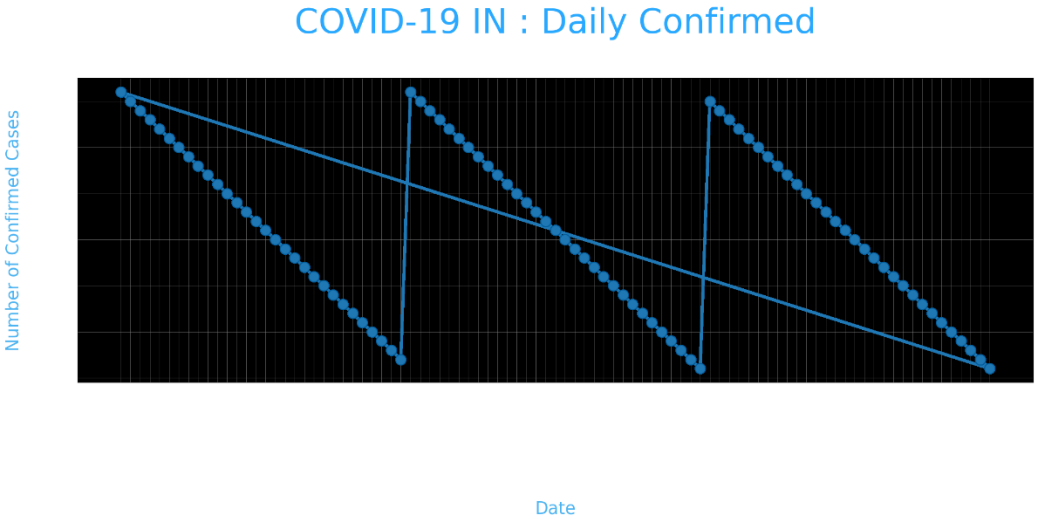
OUTPUT:




```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
data = pd.read_csv("C:/Users/sagee/Downloads/Covid_19_cases4.csv")
Y = data.iloc[61:,1].values
R = data.iloc[61:,3].values
D = data.iloc[61:,5].values
X = data.iloc[61:,0]
plt.figure(figsize=(25,8))
ax = plt.axes()
ax.grid(linewidth=0.4, color='#8f8f8f')
ax.set_facecolor("black")
ax.set_xlabel('\nDate',size=25,color='#4bb4f2')
ax.set_ylabel('Number of Confirmed Cases\n',
size=25,color='#4bb4f2')
plt.xticks(rotation='vertical',size='20',color='white')
plt.yticks(size=20,color='white')
plt.tick_params(size=20,color='white')
for i,j in zip(X,Y):
ax.annotate(str(j),xy=(i,j+100),color='white',size='13')
ax.annotate('Second Lockdown 15th April',
xy=(15.2, 860),
xytext=(19.9,500),
color='white',
size='25',
arrowprops=dict(color='white',
linewidth=0.025))
plt.title("COVID-19 IN : Daily Confirmed\n",
size=50,color='#28a9ff')
```

```
ax.plot(X,Y,
color='#1F77B4',
marker='o',
linewidth=4,
markersize=15,
markeredgecolor='#035E9B')
```

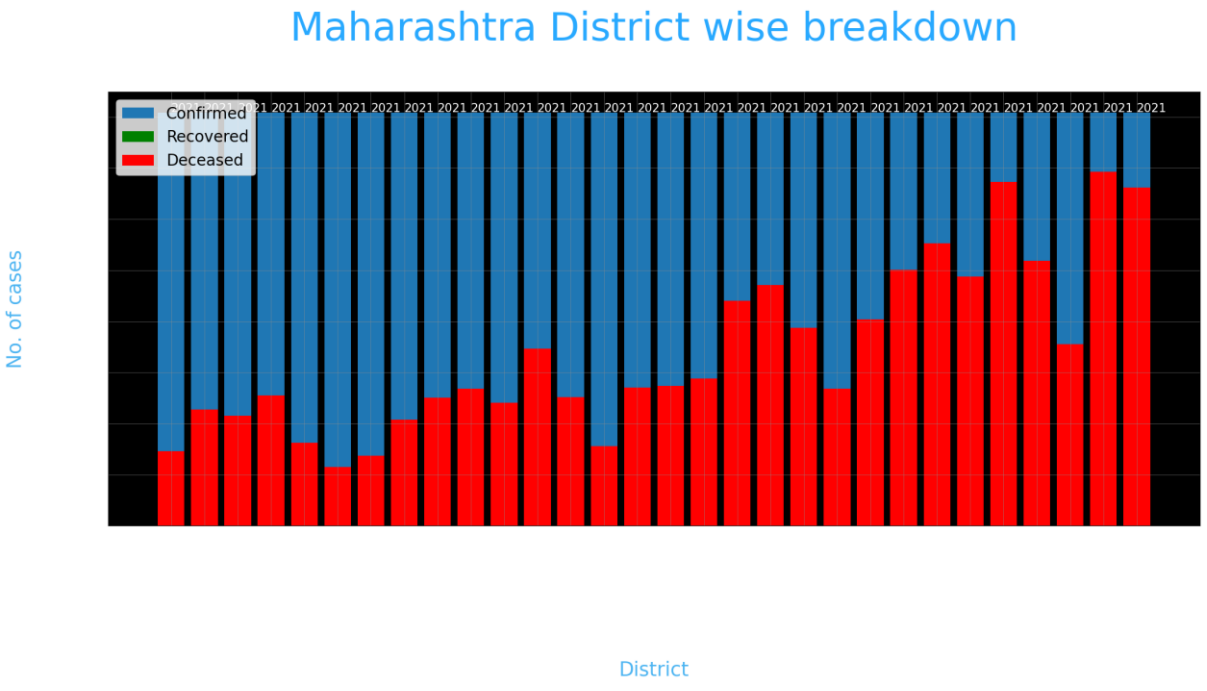
OUTPUT:



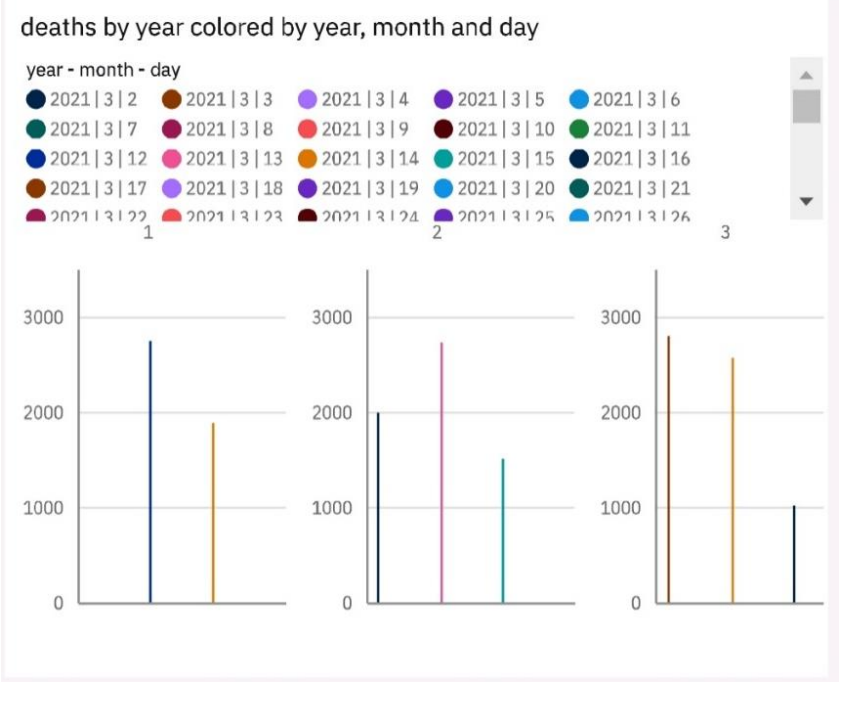

```
data = pd.read_csv("C:/Users/sagee/Downloads/Covid_19_cases4.csv")
data.head()
re=data.iloc[:30,5].values
de=data.iloc[:30,4].values
co=data.iloc[:30,3].values
x=list(data.iloc[:30,0])
plt.figure(figsize=(25,10))
ax=plt.axes()
ax.set_facecolor('black')
ax.grid(linewidth=0.4, color='#8f8f8f')
plt.xticks(rotation='vertical',
size='20',
```

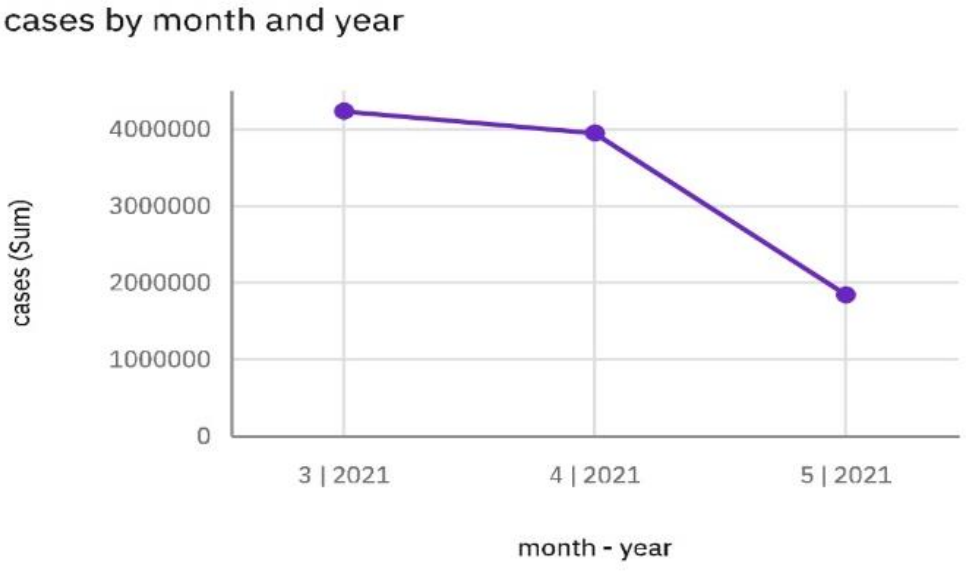
```
color='white')#ticks of X
plt.yticks(size='20',color='white')
ax.set_xlabel('\nDistrict',size=25,
color='#4bb4f2')
ax.set_ylabel('No. of cases\n',size=25,
color='#4bb4f2')
plt.tick_params(size=20,color='white')
ax.set_title('Maharashtra District wise
breakdown\n',
size=50,color='#28a9ff')
plt.bar(x,co,label='re')
plt.bar(x,re,label='re',color='green')
plt.bar(x,de,label='re',color='red')
for i,j in zip(x,co):
ax.annotate(str(int(j)),
xy=(i,j+3),
color='white',
size='15')
plt.legend(['Confirmed','Recovered','Deceased'],
fontsize=20)
```

OUTPUT:



Data Visulaization Using IBM Cognos:





Insights

Data Gathering:

a. Collect COVID-19 data from reliable sources such as government health departments, WHO, CDC, or other relevant sources
b. The data should include information like the number of cases, deaths, recoveries, and demographic data.

Data Cleaning and Transformation:

a. Use ETL (Extract, Transform, Load) processes to clean and transform the raw data into a format that's suitable for analysis.
b. Remove duplicates, handle missing values, and ensure data consistency.

Data Modeling:

a. Design a data model in Cognos that represents your COVID-19 data.
b. Create a database or data warehouse to store the transformed data.

Data Analysis:

a. Use Cognos Query Studio, Report Studio, or other Cognos tools to create reports and dashboards.
b. Analyze the data to identify trends, patterns, and insights related to COVID-19. For example, you can create visualizations to show the daily or weekly cases, compare the impact in different regions, or assess the effectiveness of various measures.

Geospatial Analysis:

a. If you have geographic data (e.g., location of cases), Cognos can be used to create maps and perform geospatial analysis to understand the spread of the virus across different regions.

Alerts and Monitoring:

a. Set up alerts or triggers in Cognos to notify relevant stakeholders if certain conditions are met. For example, you could create an alert when cases in a particular region exceed a certain threshold.

Data Security and Compliance:

a. Ensure that your use of COVID-19 data complies with privacy and data protection regulations. Cognos offers security features to restrict access to sensitive data.

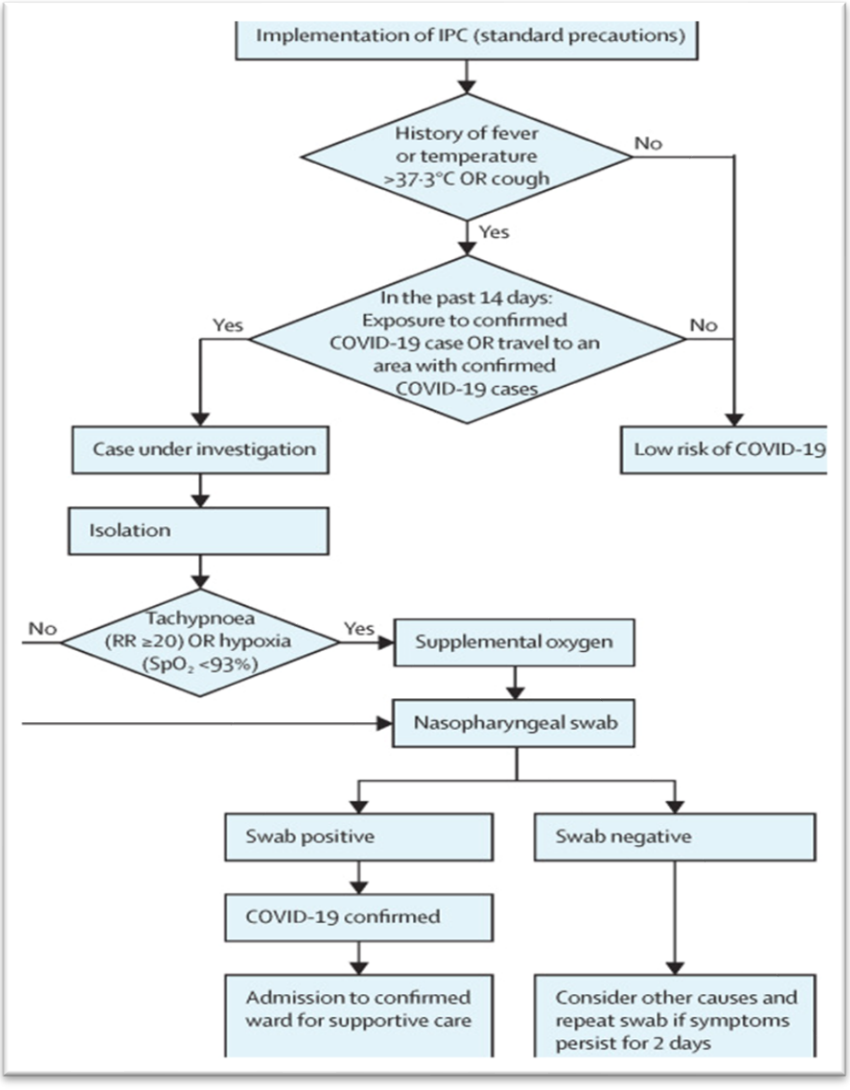
Sharing Insights:

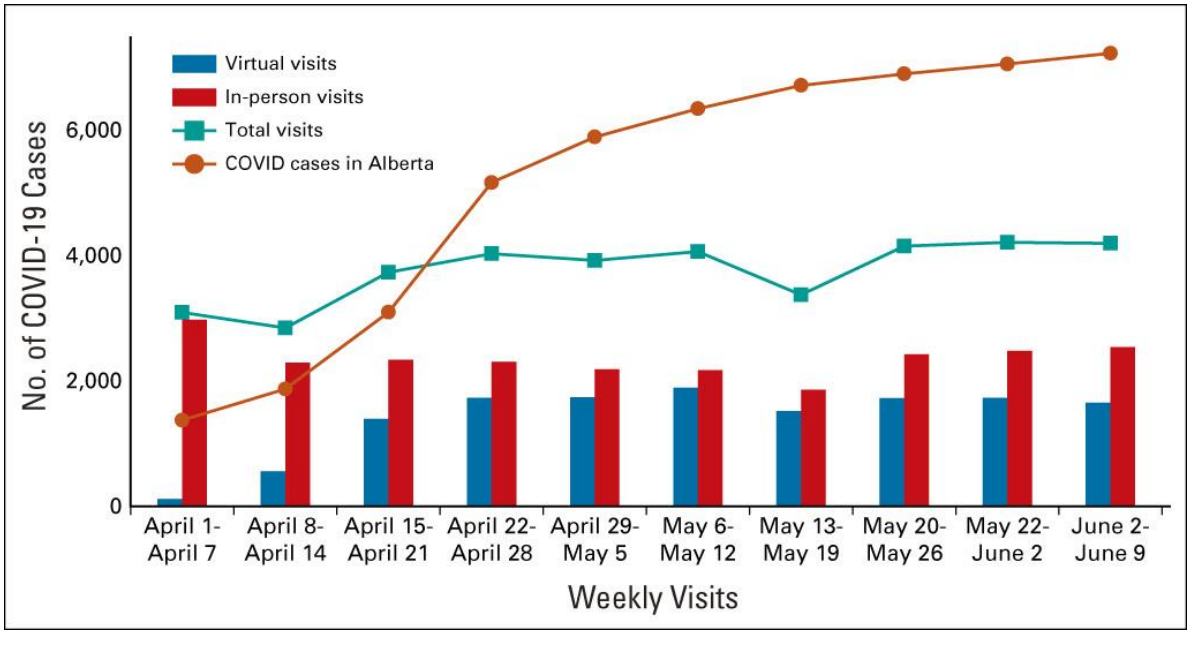
a. Share your reports and dashboards with decision-makers, public health officials, or the public, as appropriate.
b. Cognos provides various methods for sharing reports, such as exporting to PDF, embedding in web pages, or using Cognos Analytics portals.

Regular Updates:

a. COVID-19 data is continually evolving. Set up a process for regular data updates and automate the analysis and reporting where possible.

Flowchart:


Implementation of IPC (standard precautions)
History of fever or temperature >37·3°C OR cough
No
Yes
In the past 14 days: Exposure to confirmed COVID-19 case OR travel to an area with confirmed COVID-19 cases
Yes
No
Case under investigation
Low risk of COVID-19
Isolation
Tachypnoea (RR ≥20) OR hypoxia (SpO_2 <93%)
No
Yes
Supplemental oxygen
Nasopharyngeal swab
Swab positive
Swab negative
COVID-19 confirmed
Admission to confirmed ward for supportive care
Consider other causes and repeat swab if symptoms persist for 2 days


Virtual visits
In-person visits
Total visits
COVID cases in Alberta
No. of COVID-19 Cases
6,000
4,000
2,000
0
April 1-April 7
April 8-April 14
April 15-April 21
April 22-April 28
April 29-May 5
May 6-May 12
May 13-May 19
May 20-May 26
May 22-June 2
June 2-June 9
Weekly Visits

CONCLUSION:

This comprehensive documentation, we have successfully addressed the project's primary objectives, illustrating the depth of our commitment to understanding and analyzing the complex landscape of COVID-19. Our design thinking process guided us through the various development phases, ensuring a structured approach to problem-solving and data-driven decision-making.

Our rigorous analysis objectives enabled us to delve into a myriad of COVID-19 data sources, covering various aspects such as infection rates, mortality, testing, and vaccination rates. The data collection process was meticulous, ensuring data accuracy and reliability. Through the adept use of IBM Cognos, we transformed raw data into intuitive and informative visualizations, making complex trends and patterns more accessible to a wide audience.

From this analysis, a wealth of insights emerged. We gained a profound understanding of the ever-evolving trends of the pandemic, allowing us to identify regions with the highest infection rates, hotspots of transmission, and areas with lower vaccination coverage. These insights are invaluable for public health authorities, policymakers, and the broader community. They empower us to make informed decisions about resource allocation, targeted interventions, and public health messaging to mitigate the impact of COVID-19.

LINK FOR JUPYTER NOTEBOOK (ipynb):

https://github.com/Dpkamurugan/Cognos.using.covid/blob/main/Phase3.ipynb.txt

https://github.com/Dpkamurugan/Cognos.using.covid/blob/main/Phase%204.ipynb

LINK FOR JUPYTER NOTEBOOK (pdf):

https://github.com/Dpkamurugan/Cognos.using.covid/blob/main/DAC_Phase%203.pdf

LINK FOR IBM COGNOS VISUALIZATION (pdf):

https://github.com/sageesarvan/Sageetha/blob/main/DAC_Phase%204.pdf

GITHUB LINK:

PHASE 1

https://github.com/Dpkamurugan/Cognos.using.covid/blob/main/Phase%201.pdf

PHASE 2:

https://github.com/Dpkamurugan/Cognos.using.covid/commit/568a0be12373313ecafb9338c917afe0a54b8149

PHASE 3:

https://github.com/Dpkamurugan/Cognos.using.covid/blob/main/DAC_Phase%203.pdf

PHASE 4:

https://github.com/sageesarvan/Sageetha/blob/main/DAC_Phase%204.pdf