DAV Test

Date. __/__/____

Covid - 19

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
% matplotlib inline

# Load dataset
dataset_url = "./covid_19_india.csv"
df = pd.read_csv(dataset_url)
df
```

1.

```
three_months_df = df[(df['Date'] >= '2020-05-01')
                     & (df['Date'] < '2020-08-01')]

cured = three_months_df.groupby(['State/UnionTerritory'
    ])['Date', 'Cured'].max()
cured

deaths = three_months_df.groupby(['State/UnionTerritory'],
    ['Date', 'Deaths'].max()
deaths

confirmed = three_months_df.groupby(['State/UnionTerritory'])
    ['Date', 'Confirmed'].max()
confirmed
```

Yoyo

Date. __/__/____

2. ## Extracting year, month, and day

```
extracted_df = df.copy()

year = pd.to_datetime(extracted_df['Date']).apply(
        lambda x: x.year)
month = pd.to_datetime(extracted_df['Date']).apply(
        lambda x: x.month)
day = pd.to_datetime(extracted_df['Date']).apply(
        lambda x: x.day)

extracted_df.insert(loc=2, column='year', value=year)
extracted_df.insert(loc=3, column='month', value=month)
extracted_df.insert(loc=4, column='day', value=day)

extracted_df

extracted_df[extracted_df['day']==6].groupby(
    ['year','month']).sum()['Confirmed']
```

Date __/__/____

3. filtered_df = extracted_df[(extracted_df['Date'] >= '2020-04-01') & (extracted_df['Date'] <= '2021-03-31')]

filtered_df

```
plt.subplot(2,2,1)
plt.plot(filtered_df[filtered_df['State/UnionTerritory'] == "Karnataka"]['Date'], filtered_df[filtered_df['State/UnionTerritory'] == "Karnataka"]['Cured'])
plt.title("Karnataka")

plt.subplot(2,2,2)
plt.plot(filtered_df[filtered_df['State/UnionTerritory'] == "Gujarat"]['Date'], filtered_df[filtered_df['State/UnionTerritory'] == "Gujarat"]['Cured'])
plt.title("Gujarat")

plt.subplot(2,2,3)
plt.plot(filtered_df[filtered_df['State/UnionTerritory'] == "Haryana"]['Date'], filtered_df[filtered_df['State/UnionTerritory'] == "Haryana"]['Cured'])
plt.title("Haryana")

plt.subplot(2,2,4)
plt.plot(filtered_df[filtered_df['State/UnionTerritory'] == "UttarPradesh"]['Date'], filtered_df[filtered_df['State/UnionTerritory'] == "UttarPradesh"]['Cured'])
plt.title("Uttar Pradesh")

plt.show()
```

Yoyo

Date. __/__/____

4.
```
filtered_df_2020 = extracted_df[(extracted_df
    ['year'] == 2020) & (extracted_df['month']
    == 5) & (extracted_df['State/UnionTerritory']
    .isin([['Karnataka', 'Delhi', 'Madhya Pradesh']]))]
    .groupby(['year', 'State/UnionTerritory'].as_index
    = False).sum()

filtered_df_2021 = extracted_df[(extracted_df
    ['year'] == 2021) & (extracted_df['month']
    == 5) & (extracted_df['State/UnionTerritory']
    .isin([['Karnataka', 'Delhi', 'Madhya Pradesh']]))]
    .groupby(['year', 'State/UnionTerritory'].as_index
    = False).sum()

plt.figure(figsize = (15,5))
plt.bar(filtered_df_2020['State/UnionTerritory'],
    filtered_df_2020['Deaths'], color = 'r')
plt.bar(filtered_df_2021['State/UnionTerritory'],
    filtered_df_2021['Deaths'], color = 'b',
    bottom = filtered_df_2020['Deaths'])
plt.xlabel("States")
plt.ylabel("Deaths due to Covid-19")
plt.legend(["May 2020", "May 2021"])
plt.title("Deaths due to Covid-19 in the
    months of May 2020 and May 2021 for the
    states Karnataka, Delhi, and Madhya Pradesh")
plt.show()
```

Date. __/__/____

5.
```
up_data = extracted_df[extracted_df['State/UnionTerritory'] == "Uttar Pradesh"]
up_data

up_data = up_data.groupby(['year', 'month'], as_index = False).sum()
up_data['period'] = up_data["year"].astype(str) + "-" + up_data["month"].astype(str)
up_data

correlation = up_data["Confirmed"].corr(up_data["Deaths"])
correlation

from matplotlib.pyplot import figure
figure(figsize = (15, 10), dpi = 80)

plt.plot(up_data['period'], up_data['Confirmed'], label = 'Confirmed')
plt.plot(up_data['period'], up_data["Deaths"], label = 'Deaths')
plt.legend()
plt.title("Correlation : " + correlation.astype(str))
plt.show()
```