

## DAY 3 LAB EXPERIMENTS

NAME: GUTTI CHAITANYA

Reg No : 192424185

Date : 18/12/2025

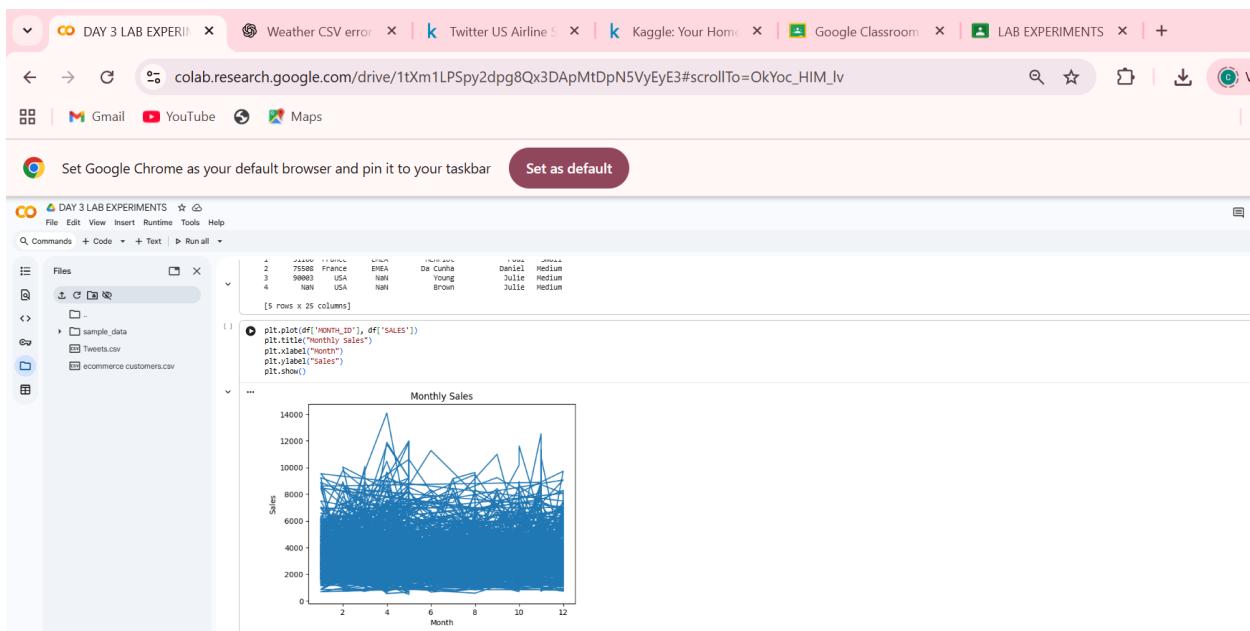
**EXP\_11 You have been tasked with creating a simple plot to show the sales of a product over time.**

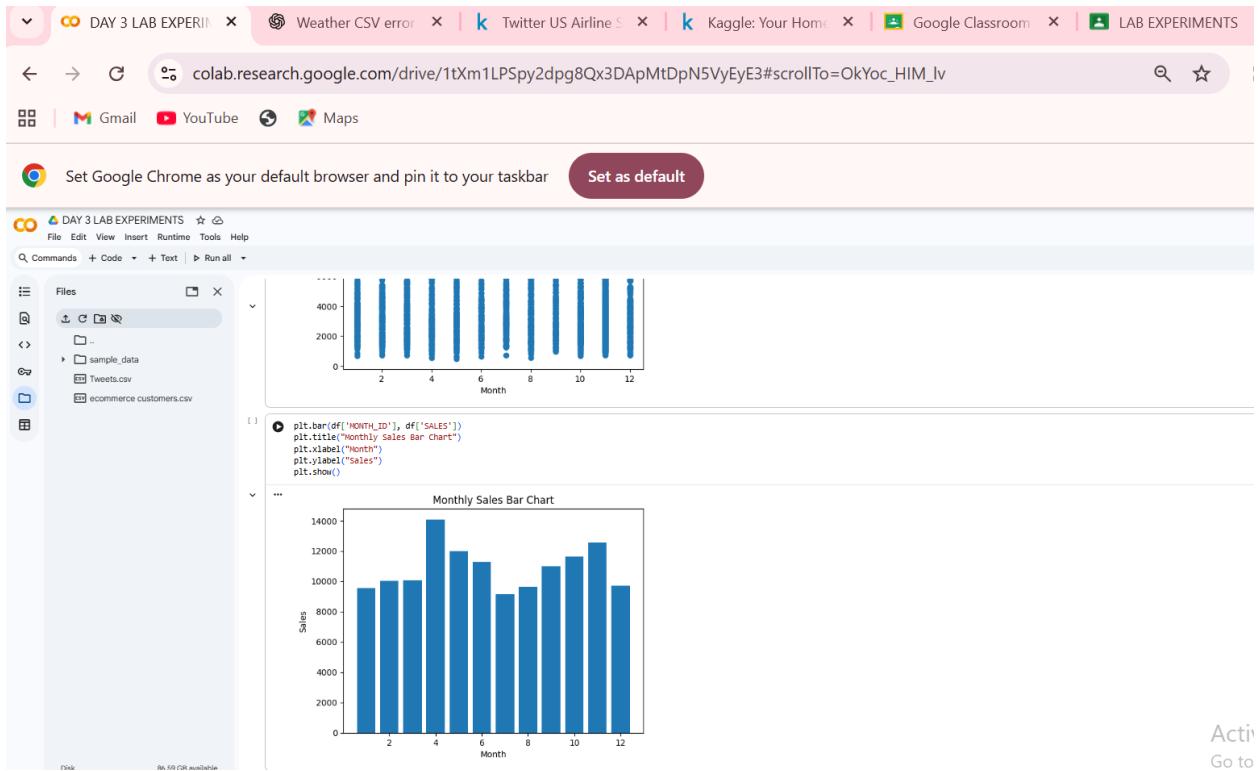
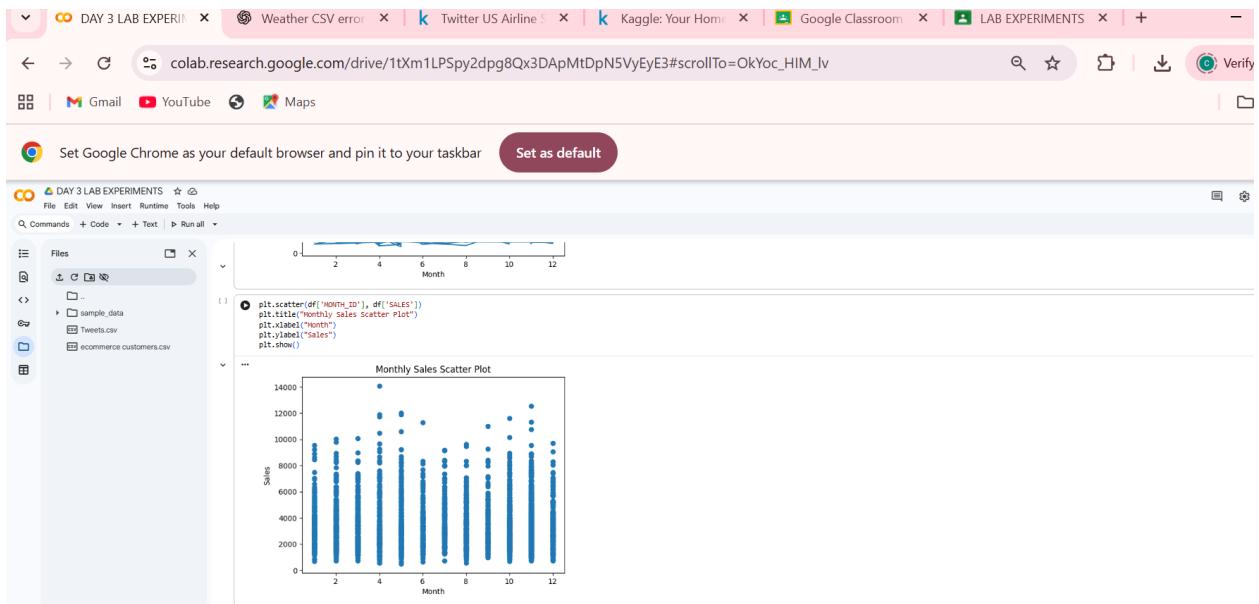
Google Colab interface showing the code for EXP\_11. The code imports pandas and matplotlib, reads a CSV file named 'monthly\_sales.csv', and prints the first few rows of the DataFrame. The DataFrame has columns: ORDERNUMBER, QUANTITYORDERED, PRICEEACH, ORDERLINENUMBER, SALES, ORDERDATE, STATUS, QTY\_ID, MONTH\_ID, YEAR\_ID, ADDRESSLINE1, ADDRESSLINE2, CITY, STATE, POSTCODE, COUNTRY, TERRITORY, CONTACTLASTNAME, CONTACTFIRSTNAME, DEALSIZE. The output shows monthly sales data from January 2003 to December 2003.

```
import pandas as pd
import matplotlib.pyplot as plt

df=pd.read_csv("monthly_sales.csv",encoding="latin1")
print(df.head())

[5 rows x 25 columns]
```





## EXP\_12 To Develop a Python program that generates line plots and scatter plots to visualize the temperature and rainfall data.

Set Google Chrome as your default browser and pin it to your taskbar [Set as default](#)

DAY 3 LAB EXPERIMENTS

```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv(
    "weather_data.csv",
    encoding="latin1",
    on_bad_lines="skip"
)

print(df.head())

Location Date_Time Temperature_C Humidity_pct \
0 San Diego 2024-01-18 21:12:46 18.683001 41.195754
1 San Diego 2024-01-17 15:22:10 8.73414 56.315107
2 San Diego 2024-01-17 15:22:10 11.53599 56.315175
3 Philadelphia 2024-02-26 17:32:39 -8.628976 54.874474
4 San Antonio 2024-04-29 13:23:51 39.086213 72.899980

Precipitation_mm Wind_Speed_mph
0 0.000000 23.5549
1 9.111623 27.715161
2 4.607511 28.723251
3 3.18372 26.367303
4 9.598282 29.698622
/tmppython-input-1239356151.py:3: DtypeWarning: Columns (2,3,4) have mixed types. Specify dtype option on import or set low_memory=False.
df = pd.read_csv(
```

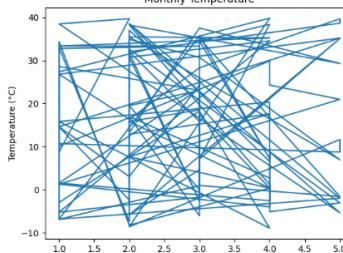
```
df["Date_Time"] = pd.to_datetime(df["Date_Time"], errors="coerce")
df["Month"] = df["Date_Time"].dt.month

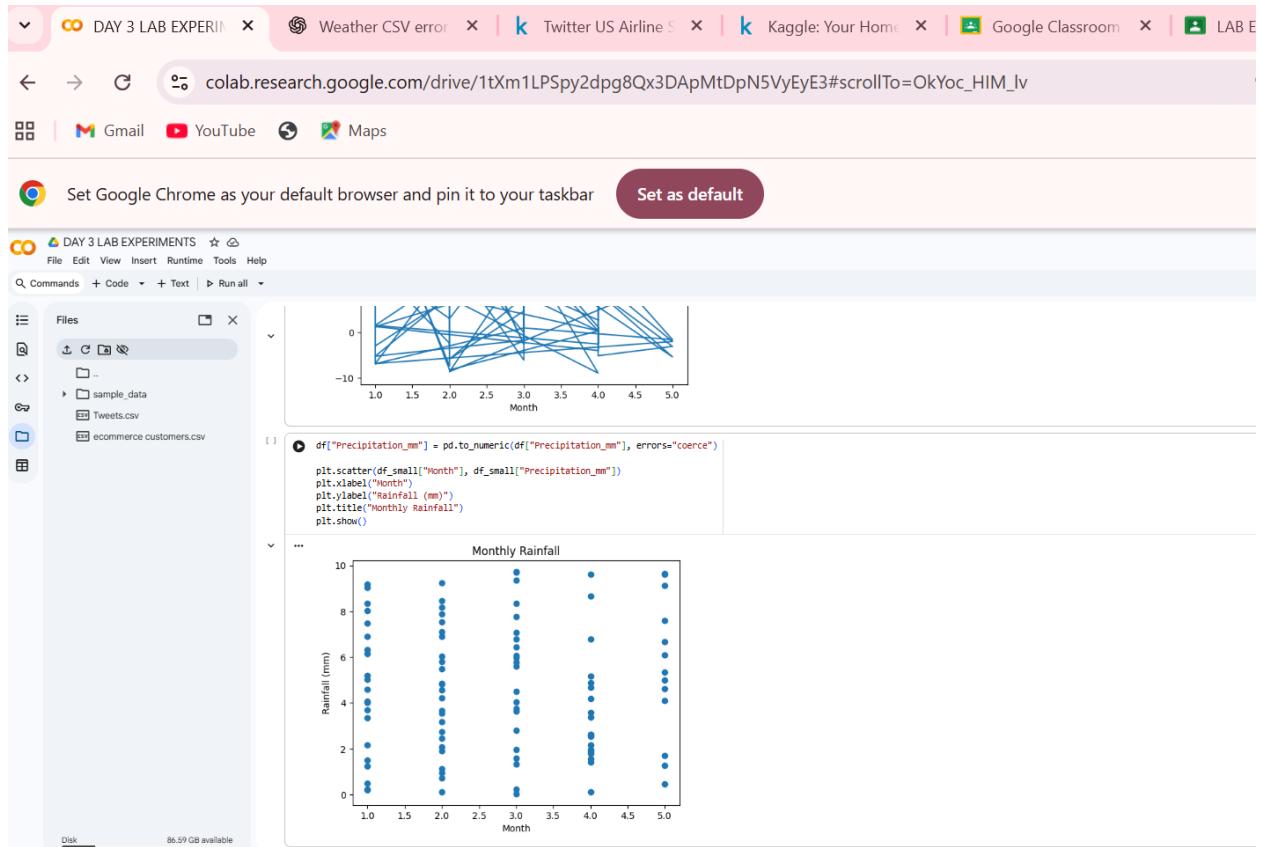
df["Date_Time"] = pd.to_datetime(df["Date_Time"], errors="coerce")
df["Month"] = df["Date_Time"].dt.month

df["Temperature_C"] = pd.to_numeric(df["Temperature_C"], errors="coerce")

df_small = df.head(100)

plt.plot(df_small["Month"], df_small["Temperature_C"])
plt.xlabel("Month")
plt.ylabel("Temperature (°C)")
plt.title("Monthly Temperature")
plt.show()
```





## **EXP\_13 To Develop a Python program that reads the text document, processes the text, and generates a frequency distribution of the words.**

The screenshot shows a Google Colab notebook interface. At the top, there are several tabs: DAY 3 LAB EXPERIMENTS, Weather CSV error, Twitter US Airline S, Kaggle: Your Home, Google Classroom, and LAB EXPERIMENT. Below the tabs, the URL is colab.research.google.com/drive/1tXm1LPSpy2dpg8Qx3DApMtDpN5VyEyE3#scrollTo=OkYoc\_HIM\_lv. The toolbar includes back, forward, search, and other navigation icons. Below the toolbar, there are links for Gmail, YouTube, and Maps. A message says "Set Google Chrome as your default browser and pin it to your taskbar" with a "Set as default" button. The main workspace shows a code cell with the following content:

```
from collections import Counter
with open("sample_data.txt", "r") as file:
    text = file.read()
words = text.lower().split()
word_freq = Counter(words)
for i in word_freq:
    print(f'{i}: {word_freq[i]}')
```

Below the code cell, the output is displayed as a list of words and their frequencies:

- welcome: 1
- to: 2
- natural: 1
- language: 1
- processing: 1
- it: 1
- is: 1
- one: 1
- of: 2
- the: 1
- most: 1
- exciting: 1
- research: 1
- areas: 1
- as: 1
- today: 1
- we: 1
- will: 1
- see: 1
- how: 1
- python: 1
- can: 1
- be: 1
- used: 1
- work: 1
- with: 1
- text: 1
- files.: 1

At the bottom left, it says "Disk 86.59 GB available".

## EXP\_14 To Develop a code in python to find the frequency distribution of the ages of the customers who have made a purchase in the past month.

The screenshot shows a Google Colab notebook interface. At the top, there are several tabs: DAY 3 LAB EXPERIMENTS, Weather CSV error, Twitter US Airline S, Kaggle: Your Home, and a file icon. Below the tabs, the URL is colab.research.google.com/drive/1tXm1LPSpy2dpg8Qx3DApMtDpN5VyEyE3#scrollTo=OkYoc\_HIN. The toolbar includes back, forward, search, and other navigation icons. Below the toolbar, there are links to Gmail, YouTube, and Maps.

A message bar at the top says "Set Google Chrome as your default browser and pin it to your taskbar" with a "Set as default" button.

The main workspace shows a code cell with the following Python code:

```
import pandas as pd
df=pd.read_csv("ecommerce_customers.csv")
print(df.head())
print(df.columns)

CustomerID  Genre  Age  Annual Income (k$)  Spending Score (1-100)
0           1  Male   19            15              39
1           2  Male   21            15              81
2           3 Female  20            16               6
3           4 Female  23            16              77
4           5 Female  31            17              40
Index(['CustomerID', 'Genre', 'Age', 'Annual Income (k$)', 'Spending Score (1-100)'],
      dtype='object')

age_frequency = df["Age"].value_counts()
print(age_frequency)
```

The code cell output displays the frequency distribution of ages:

| Age | Frequency |
|-----|-----------|
| 32  | 11        |
| 35  | 9         |
| 31  | 8         |
| 19  | 8         |
| 38  | 7         |
| 49  | 7         |
| 23  | 6         |
| 40  | 6         |
| 36  | 6         |
| 27  | 6         |
| 47  | 6         |
| 38  | 6         |
| 50  | 5         |
| 20  | 5         |
| 21  | 5         |
| 29  | 5         |
| 48  | 5         |
| 34  | 4         |
| 67  | 4         |
| 54  | 4         |
| 24  | 4         |
| 59  | 4         |

DAY 3 LAB EXPERIMENT Weather CSV error Twitter US Airline Kaggle: Your Home Good

colab.research.google.com/drive/1tXm1LPSpy2dpg8Qx3DApMtDpN5VyEyE3#scrollTo=OkYoc\_HIM\_I

Gmail YouTube Maps

Set Google Chrome as your default browser and pin it to your taskbar **Set as default**

DAY 3 LAB EXPERIMENTS

File Edit View Insert Runtime Tools Help

Commands + Code + Text Run all

Files

- ..
- sample\_data
- Tweets.csv
- ecommerce customers.csv

| 49 | 7 |
|----|---|
| 23 | 6 |
| 48 | 6 |
| 36 | 6 |
| 27 | 6 |
| 47 | 6 |
| 38 | 6 |
| 50 | 5 |
| 28 | 5 |
| 21 | 5 |
| 29 | 5 |
| 48 | 5 |
| 34 | 5 |
| 67 | 4 |
| 54 | 4 |
| 24 | 4 |
| 59 | 4 |
| 18 | 4 |
| 28 | 4 |
| 22 | 3 |
| 25 | 3 |
| 46 | 3 |
| 43 | 3 |
| 39 | 3 |
| 60 | 3 |
| 45 | 3 |
| 37 | 3 |
| 33 | 3 |
| 68 | 3 |
| 58 | 2 |
| 41 | 2 |
| 52 | 2 |
| 57 | 2 |
| 65 | 2 |
| 53 | 2 |
| 51 | 2 |
| 42 | 2 |
| 70 | 2 |
| 44 | 2 |
| 26 | 2 |
| 63 | 2 |
| 66 | 2 |
| 64 | 1 |
| 69 | 1 |
| 55 | 1 |
| 56 | 1 |

Disk 86.59 GB available

Name: count, dtype: int64

## EXP\_15 To Develop a Python program that calculates the frequency distribution of likes among the posts.

The screenshot shows two sessions in Google Colab. The top session is titled "EXP\_15 To Develop a Python program that calculates the frequency distribution of likes among the posts." It displays the following code and its output:

```
import pandas as pd
df=pd.read_csv("Tweets.csv",encoding="latin1")
print(df.head())
...[redacted]
likes_frequency = df["retweet_count"].value_counts()
```

The output shows the first few rows of the CSV file and the resulting frequency distribution of retweet counts:

| retweet_count | count |
|---------------|-------|
| 0             | 13873 |
| 1             | 640   |
| 2             | 66    |
| 3             | 22    |
| 4             | 17    |
| 5             | 5     |
| 7             | 3     |
| 6             | 3     |
| 22            | 2     |
| 18            | 1     |
| 15            | 1     |
| 31            | 1     |
| 11            | 1     |
| 8             | 1     |
| 9             | 1     |
| 28            | 1     |
| 32            | 1     |
| 44            | 1     |

The bottom session is also titled "EXP\_15 To Develop a Python program that calculates the frequency distribution of likes among the posts." It displays the following code and its output:

```
print("The Frequency Distribution of Likes are:")
print(likes_frequency)
```

The output shows the same frequency distribution of retweet counts as the top session:

| retweet_count | count |
|---------------|-------|
| 0             | 13873 |
| 1             | 640   |
| 2             | 66    |
| 3             | 22    |
| 4             | 17    |
| 5             | 5     |
| 7             | 3     |
| 6             | 3     |
| 22            | 2     |
| 18            | 1     |
| 15            | 1     |
| 31            | 1     |
| 11            | 1     |
| 8             | 1     |
| 9             | 1     |
| 28            | 1     |
| 32            | 1     |
| 44            | 1     |