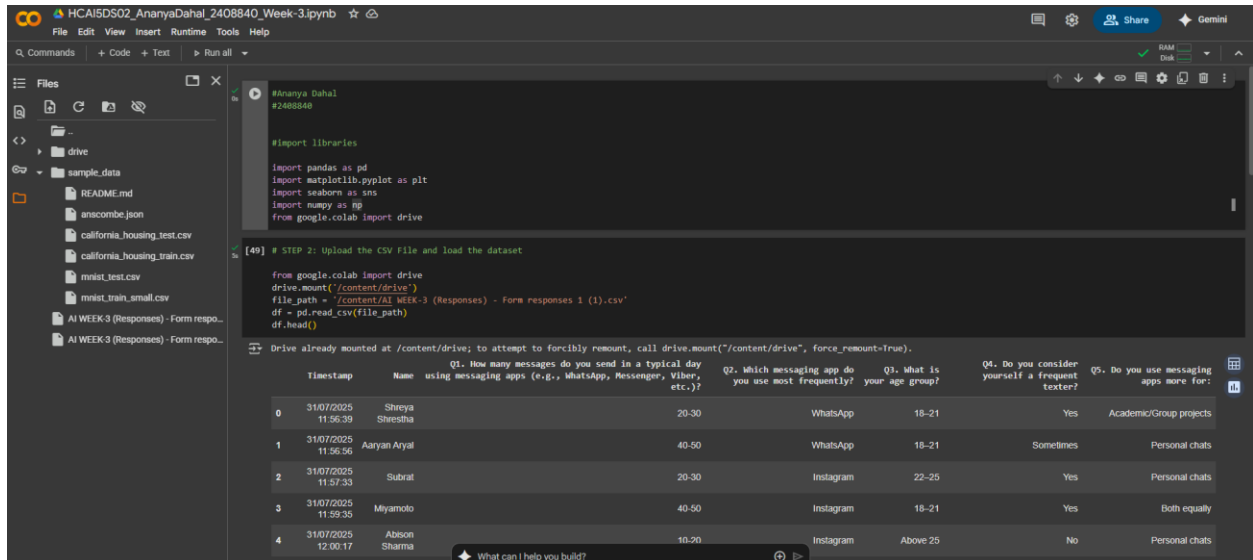


Week-3

Workshop -3



The screenshot shows a Jupyter Notebook titled "HCAISDS02_AnanyaDahal_2408840_Week-3.ipynb". The code in the notebook is as follows:

```
#Ananya Dahal
#2408840

#Import libraries

import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
from google.colab import drive

[40] # STEP 2: Upload the CSV File and load the dataset

from google.colab import drive
drive.mount('/content/drive')
file_path = '/content/AI WEEK-3 (Responses) - Form responses 1 (1).csv'
df = pd.read_csv(file_path)
df.head()
```

The output of the code shows a message: "Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount(\"/content/drive\", force_remount=True).". Below this, a table of survey responses is displayed:

	Timestamp	Name	Q1. How many messages do you send in a typical day using messaging apps (e.g., WhatsApp, Messenger, Viber, etc.)?	Q2. Which messaging app do you use most frequently?	Q3. What is your age group?	Q4. Do you consider yourself a frequent texter?	Q5. Do you use messaging apps more for:
0	31/07/2025 11:56:39	Shreya Shrestha	20-30	WhatsApp	18-21	Yes	Academic/Group projects
1	31/07/2025 11:56:56	Aaryan Aryal	40-50	WhatsApp	18-21	Sometimes	Personal chats
2	31/07/2025 11:57:33	Subrat	20-30	Instagram	22-25	Yes	Personal chats
3	31/07/2025 11:59:35	Miyamoto	40-50	Instagram	18-21	Yes	Both equally
4	31/07/2025 12:00:17	Abison Shama	10-20	Instagram	Above 25	No	Personal chats



The screenshot shows the same Jupyter Notebook with additional code for data analysis:

```
[50] # Rename the column name for easier use

df = df.rename(columns={
    df.columns[2]: "Messages_Per_Day"
})

[51] #STEP 4: Create frequency Table

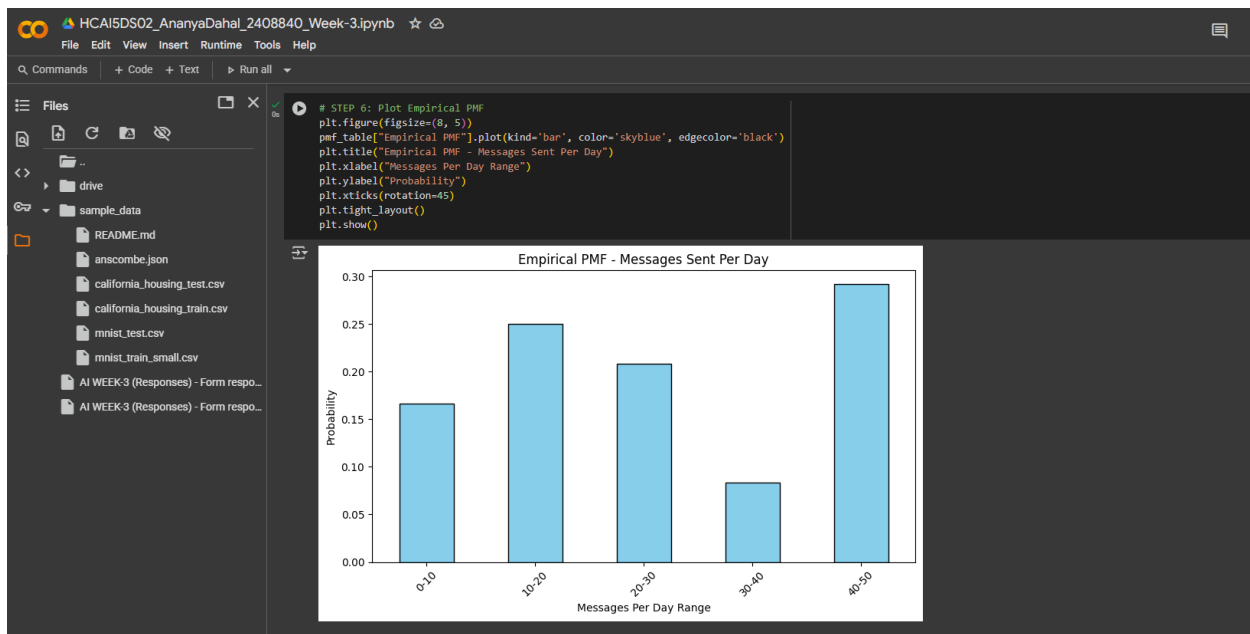
frequency_table = df["Messages_Per_Day"].value_counts().sort_index()
total_responses = frequency_table.sum()

[52] #Compute Empirical PMF

empirical_pmf = frequency_table / total_responses

[53] # Combine into a DataFrame

pmf_table = pd.DataFrame({
    "frequency": frequency_table,
    "Empirical PMF": empirical_pmf
})
```



HCIA5DS02_AnanyaDahal_2408840_Week-3.ipynb

File Edit View Insert Runtime Tools Help

Q Commands + Code + Text ▶ Run all

Files

- drive
- sample_data
 - README.md
 - anscombe.json
 - california_housing_test.csv
 - california_housing_train.csv
 - mnist_test.csv
 - mnist_train_small.csv
 - AI WEEK-3 (Responses) - Form respo...
 - AI WEEK-3 (Responses) - Form respo...

```
[55] # STEP 7: Estimate Mean and Variance using Midpoints
midpoints = {
    "0-10": 5,
    "10-20": 15,
    "20-30": 25,
    "30-40": 35,
    "40-50": 45
}

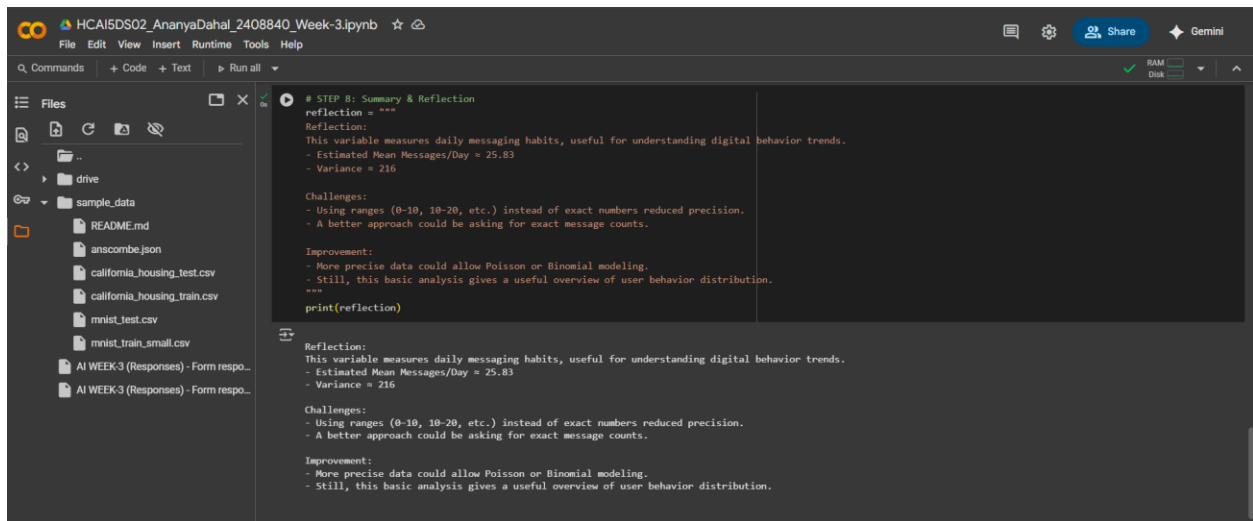
[56] # Add midpoints to the PMF table
pmf_table["Midpoint"] = pmf_table.index.map(midpoints)

[57] # Estimate sample mean
mean_estimate = (pmf_table["Midpoint"] * pmf_table["Empirical PMF"]).sum()

# Estimate sample variance
E_X2 = ((pmf_table["Midpoint"]**2) * pmf_table["Empirical PMF"]).sum()
variance_estimate = E_X2 - (mean_estimate**2)

print(f"Estimated Mean: {mean_estimate:.2f}")
print(f"Estimated Variance: {variance_estimate:.2f}")

Estimated Mean: 25.83
Estimated Variance: 215.97
```



Link:

https://docs.google.com/spreadsheets/d/1yMn_a5SyY9SO6obBIDt6sGbVcD7kUDoolf9thAWNfik/edit?usp=sharing