

localhost:8888/notebooks/nlp-project%201.ipynb

```
jupyter nlp-project 1 Last Checkpoint: 13 minutes ago
```

File Edit View Run Kernel Settings Help Trusted JupyterLab Python 3 (ipykernel)

```
[1]: import numpy as np
import pandas as pd
print("Libraries loaded")

Libraries loaded

[2]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

from statsmodels.tsa.seasonal import seasonal_decompose
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
from statsmodels.tsa.stattools import adfuller

sns.set(style="whitegrid")
```

```
[3]: # *****
# 1. Load dataset
# *****
```

```
try:
    df = pd.read_csv(r"C:\Users\Shenuke\Downloads\archive (2)\retail_sales_dataset.csv")
except:
    print("Error: file not found.")
    exit()

print("Dataset Preview:")
print(df.head())
```

```
Dataset Preview:
   Transaction ID Date Customer ID Gender Age Product Category \
0           1 2023-11-24 CUST001   Male  38      Beauty
1           2 2023-02-27 CUST002   Female 26      Clothing
2           3 2023-05-15 CUST003   Male  30 Electronics
3           4 2023-05-21 CUST004   Male  37      Clothing
4           5 2023-05-06 CUST005   Male  38      Beauty
```

```
   Quantity Price per unit Total Amount
0         1          100        100
1         2          500       1000
2         1          30         30
3         1          500       500
4         2          50        100
```

```
[4]: # convert Date column to datetime
df["Date"] = pd.to_datetime(df["Date"])
```

```
[5]: # convert Date column to datetime
df["Date"] = pd.to_datetime(df["Date"])

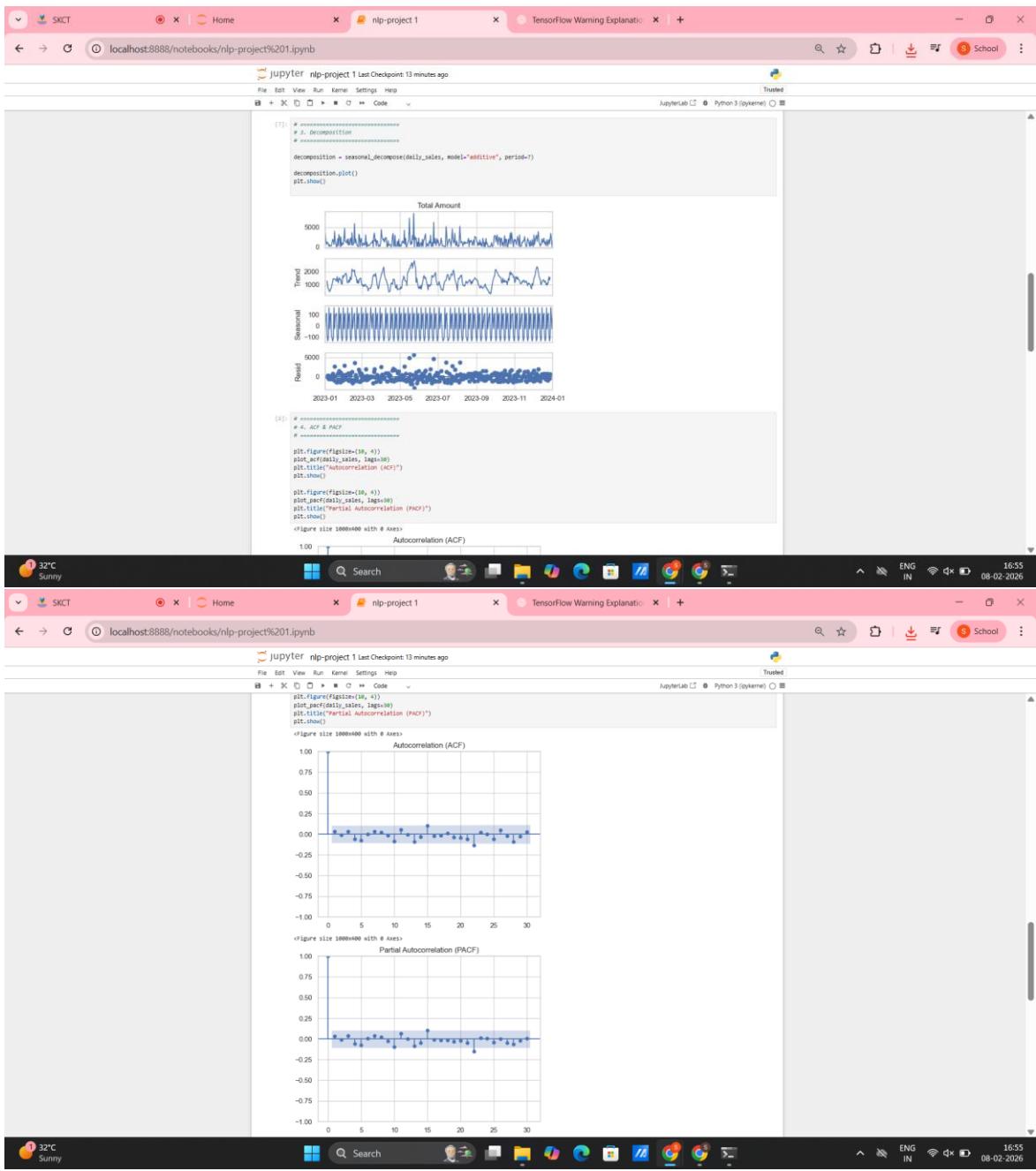
# sort by date
df = df.sort_values("Date")

# Create daily sales time series
daily_sales = df.groupby("Date")["Total Amount"].sum()

print(daily_sales.head())
```

```
Date
2023-01-01    3600
2023-01-02    1765
2023-01-03    1200
2023-01-04    1240
2023-01-05    1100
Name: Total Amount, dtype: int64
```

```
[6]: plt.figure(figsize=(12,5))
plt.plot(daily_sales)
plt.title("Daily Retail Sales Trend")
plt.xlabel("Date")
plt.ylabel("Total Sales")
plt.show()
```



SKCT

localhost:8888/notebooks/nlp-project%201.ipynb

jupyter nlp-project 1 Last Checkpoint: 13 minutes ago

File Edit View Run Kernel Settings Help Trusted JupyterLab Python 3 (ipykernel)

```
[1]: # =====#
# S: stationarity test (ADF)
# =====#
result = adfuller(daily_sales)
print("ADF statistic:", result[0])
print("p-value:", result[1])
else:
    print("The series is NOT stationary")
ADF statistic: -17.970363453121
p-value: 2.71e-16
The series is stationary
```

```
[10]: diff_sales = daily_sales.diff().dropna()
plt.figure(figsize=(12,3))
plt.plot(diff_sales)
plt.title("Differenced Sales Series")
plt.show()

# Test again
result = adfuller(diff_sales)
print("New p-value:", result[1])
```

Differenced Sales Series

```
[11]: print(result)
Key observations:
- Trend shows overall sales movement over time.
- Seasonal patterns indicate weekly buying behavior.
- Adfuller test checks for stationarity.
- ADF test checks stationarity.
- Differencing stabilizes the time series.
***
```

```
[12]:
```

32°C Sunny

localhost:8888/notebooks/nlp-project%201.ipynb

jupyter nlp-project 1 Last Checkpoint: 13 minutes ago

File Edit View Run Kernel Settings Help Trusted JupyterLab Python 3 (ipykernel)

Differenced Sales Series

```
[11]: print(result)
Key observations:
- Trend shows overall sales movement over time.
- Seasonal patterns indicate weekly buying behavior.
- Adfuller test checks for stationarity.
- ADF test checks stationarity.
- Differencing stabilizes the time series.
***
```

```
[12]:
```

ENG IN 1655 08-02-2026

32°C Sunny

localhost:8888/notebooks/nlp-project%201.ipynb

jupyter nlp-project 1 Last Checkpoint: 13 minutes ago

File Edit View Run Kernel Settings Help Trusted JupyterLab Python 3 (ipykernel)

Differenced Sales Series

```
[11]: print(result)
Key observations:
- Trend shows overall sales movement over time.
- Seasonal patterns indicate weekly buying behavior.
- Adfuller test checks for stationarity.
- ADF test checks stationarity.
- Differencing stabilizes the time series.
***
```

```
[12]:
```

ENG IN 1656 08-02-2026