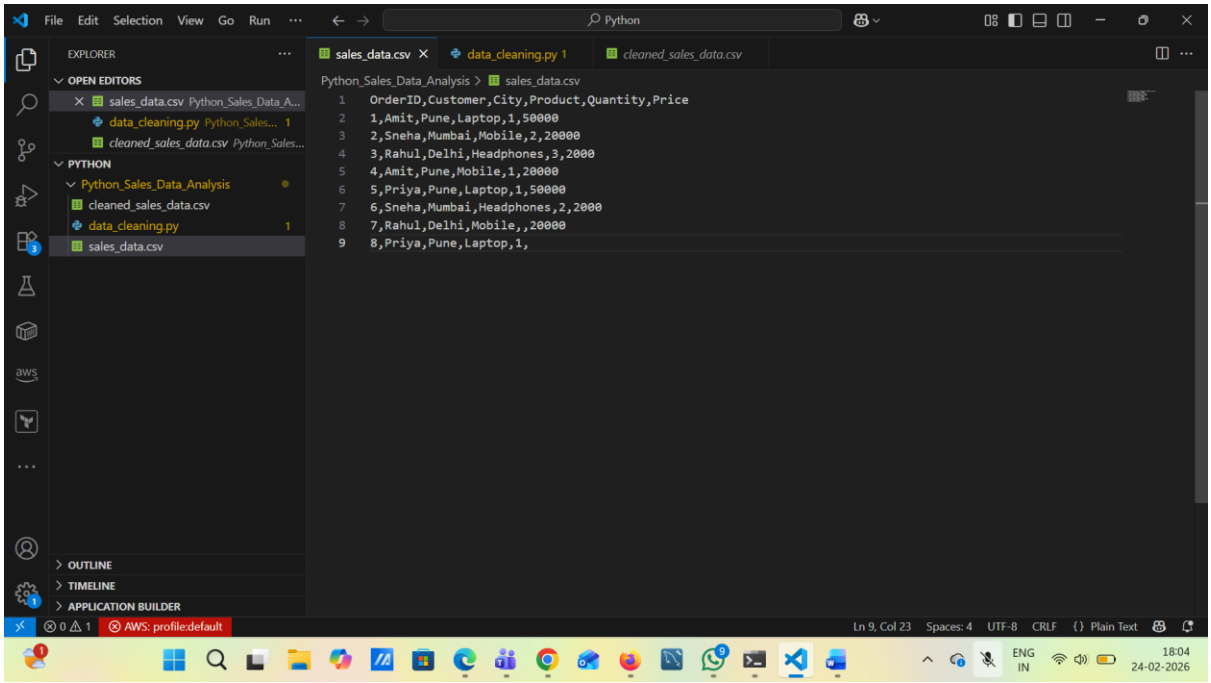
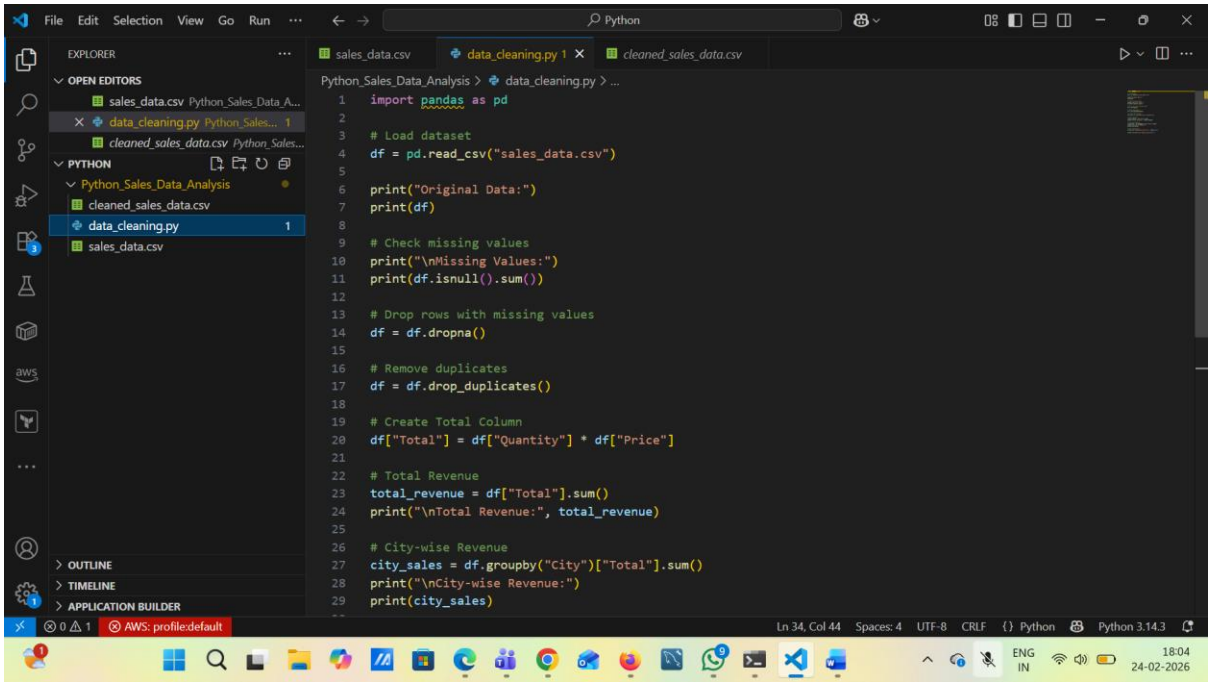
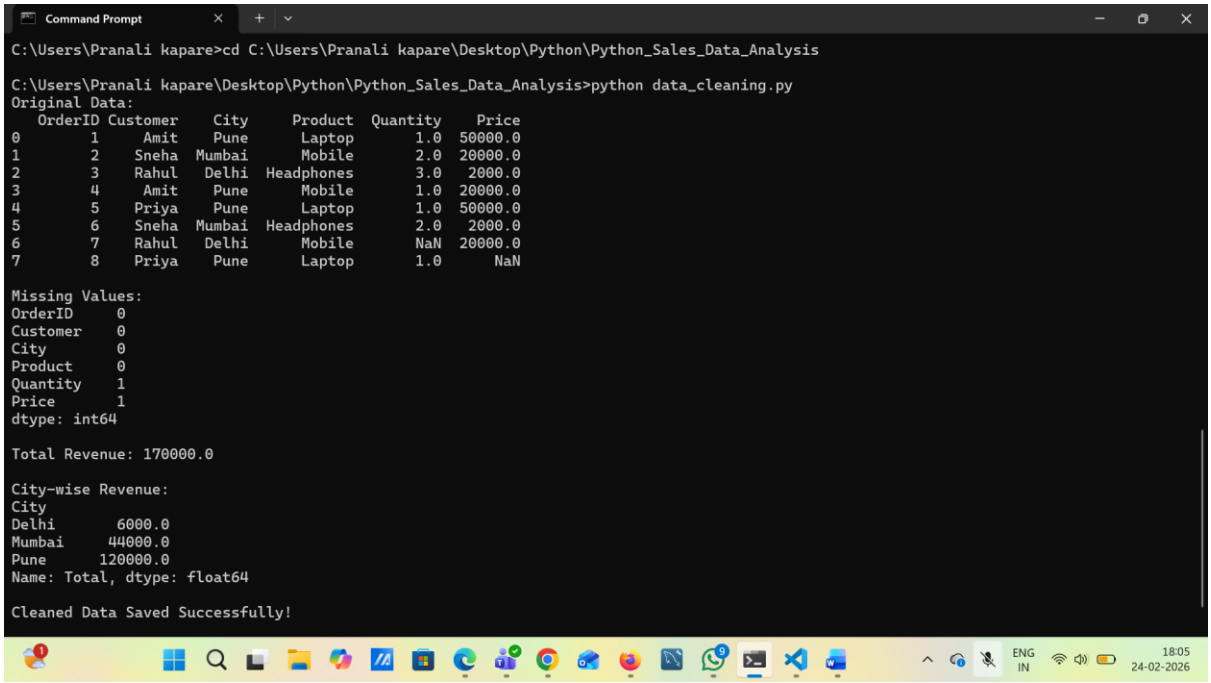
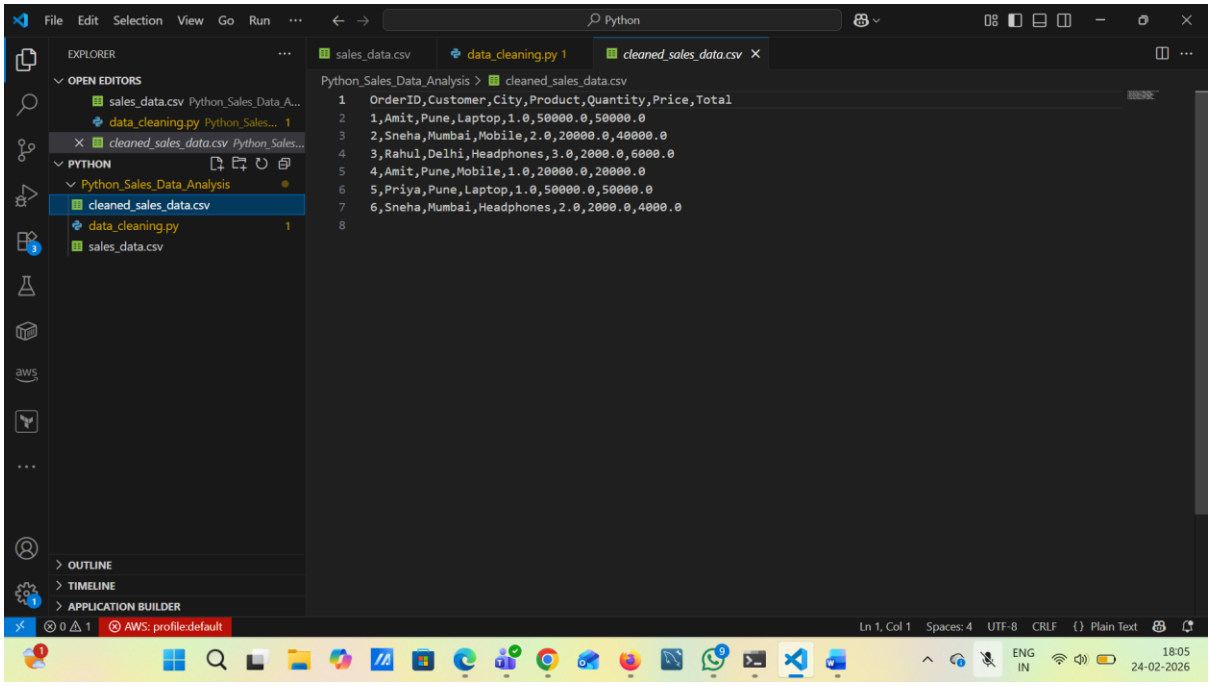

File Edit Selection View Go Run
Python
EXPLORER
OPEN EDITORS
sales_data.csv Python_Sales_Data_A...
data_cleaning.py Python_Sales... 1
cleaned_sales_data.csv Python_Sales...
PYTHON
Python_Sales_Data_Analysis
cleaned_sales_data.csv
data_cleaning.py 1
sales_data.csv
OUTLINE
TIMELINE
APPLICATION BUILDER
sales_data.csv
data_cleaning.py 1
cleaned_sales_data.csv
Python_Sales_Data_Analysis > sales_data.csv
1 OrderID,Customer,City,Product,Quantity,Price
2 1,Amit,Pune,Laptop,1,50000
3 2,Sneha,Mumbai,Mobile,2,20000
4 3,Rahul,Delhi,Headphones,3,2000
5 4,Amit,Pune,Mobile,1,20000
6 5,Priya,Pune,Laptop,1,50000
7 6,Sneha,Mumbai,Headphones,2,2000
8 7,Rahul,Delhi,Mobile,,20000
9 8,Priya,Pune,Laptop,1,
0 1
AWS: profile:default
Ln 9, Col 23 Spaces: 4 UTF-8 CRLF { } Plain Text
ENG
IN
18:04
24-02-2026


File Edit Selection View Go Run
Python
EXPLORER
OPEN EDITORS
sales_data.csv Python_Sales_Data_A...
data_cleaning.py Python_Sales... 1
cleaned_sales_data.csv Python_Sales...
PYTHON
Python_Sales_Data_Analysis
cleaned_sales_data.csv
data_cleaning.py 1
sales_data.csv
OUTLINE
TIMELINE
APPLICATION BUILDER
sales_data.csv
data_cleaning.py 1
cleaned_sales_data.csv
Python_Sales_Data_Analysis > data_cleaning.py > ...
1 import pandas as pd
2
3 # Load dataset
4 df = pd.read_csv("sales_data.csv")
5
6 print("Original Data:")
7 print(df)
8
9 # Check missing values
10 print("\nMissing Values:")
11 print(df.isnull().sum())
12
13 # Drop rows with missing values
14 df = df.dropna()
15
16 # Remove duplicates
17 df = df.drop_duplicates()
18
19 # Create Total Column
20 df["Total"] = df["Quantity"] * df["Price"]
21
22 # Total Revenue
23 total_revenue = df["Total"].sum()
24 print("\nTotal Revenue:", total_revenue)
25
26 # City-wise Revenue
27 city_sales = df.groupby("City")["Total"].sum()
28 print("\nCity-wise Revenue:")
29 print(city_sales)
0 1
AWS: profile:default
Ln 34, Col 44 Spaces: 4 UTF-8 CRLF { } Python Python 3.14.3
ENG
IN
18:04
24-02-2026


Command Prompt
C:\Users\Pranali kapare>cd C:\Users\Pranali kapare\Desktop\Python\Python_Sales_Data_Analysis
C:\Users\Pranali kapare\Desktop\Python\Python_Sales_Data_Analysis>python data_cleaning.py
Original Data:
OrderID Customer City Product Quantity Price
0 1 Amit Pune Laptop 1.0 50000.0
1 2 Sneha Mumbai Mobile 2.0 20000.0
2 3 Rahul Delhi Headphones 3.0 2000.0
3 4 Amit Pune Mobile 1.0 20000.0
4 5 Priya Pune Laptop 1.0 50000.0
5 6 Sneha Mumbai Headphones 2.0 2000.0
6 7 Rahul Delhi Mobile NaN 20000.0
7 8 Priya Pune Laptop 1.0 NaN
Missing Values:
OrderID 0
Customer 0
City 0
Product 0
Quantity 1
Price 1
dtype: int64
Total Revenue: 170000.0
City-wise Revenue:
City
Delhi 6000.0
Mumbai 44000.0
Pune 120000.0
Name: Total, dtype: float64
Cleaned Data Saved Successfully!


File Edit Selection View Go Run
Python
EXPLORER
OPEN EDITORS
sales_data.csv Python_Sales_Data_A...
data_cleaning.py Python_Sales... 1
cleaned_sales_data.csv Python_Sales...
PYTHON
Python_Sales_Data_Analysis
cleaned_sales_data.csv
data_cleaning.py 1
sales_data.csv
OUTLINE
TIMELINE
APPLICATION BUILDER
sales_data.csv
data_cleaning.py 1
cleaned_sales_data.csv
Python_Sales_Data_Analysis > cleaned_sales_data.csv
1 OrderID,Customer,City,Product,Quantity,Price,Total
2 1,Amit,Pune,Laptop,1.0,50000.0,50000.0
3 2,Sneha,Mumbai,Mobile,2.0,20000.0,40000.0
4 3,Rahul,Delhi,Headphones,3.0,2000.0,6000.0
5 4,Amit,Pune,Mobile,1.0,20000.0,20000.0
6 5,Priya,Pune,Laptop,1.0,50000.0,50000.0
7 6,Sneha,Mumbai,Headphones,2.0,2000.0,4000.0
8
AWS: profile:default
Ln 1, Col 1 Spaces: 4 UTF-8 CRLF Plain Text