

Python programming Lab(23CP301P)

Name: Krishika Vansh

Semester: V

Roll No: 23BCP448

Faculty: Mr. Davinder Singh

Division: VII Batch G13

Branch: Computer Engineering



School of Technology

November 2025

LAB4:**Experiment No: 4****Advanced Operations on csv file using Pandas****Code:**

```
import os

import pandas as pd

# === Paths ===

base_dir = "."

product_names_file = os.path.join(base_dir, "product_names.csv")

sales_dir = os.path.join(base_dir, "sales")

summary_file = os.path.join(base_dir, "sales_summary.csv")

# === Step 1: Load product names ===

product_df = pd.read_csv(product_names_file)

product_dict = dict(zip(product_df["ProductID"],
                        product_df["Product_Name"]))

# === Step 2: Gather sales CSV files ===

sales_files = [os.path.join(sales_dir, f) for f in os.listdir(sales_dir) if
                f.endswith(".csv")]

# === Step 3: Process sales data ===

product_sales = {}

months_set = set()
```

```
for file in sales_files:
```

```
    df = pd.read_csv(file)
```

```
    df['Month'] = pd.to_datetime(df['Date']).dt.to_period('M')
```

```
    for _, row in df.iterrows():
```

```
        pid = row["ProductID"]
```

```
        qty = int(row["Quantity"])
```

```
        product_sales[pid] = product_sales.get(pid, 0) + qty
```

```
    months_set.update(df['Month'].unique())
```

```
# === Step 4: Calculate statistics ===
```

```
months_count = len(months_set)
```

```
summary_data = []
```

```
for pid, total_qty in product_sales.items():
```

```
    avg_qty = total_qty / months_count if months_count else 0
```

```
    pname = product_dict.get(pid, "Unknown Product")
```

```
    summary_data.append([pid, pname, total_qty, round(avg_qty, 2)])
```

```
# === Step 5: Sort & Save Top 5 to CSV ===
```

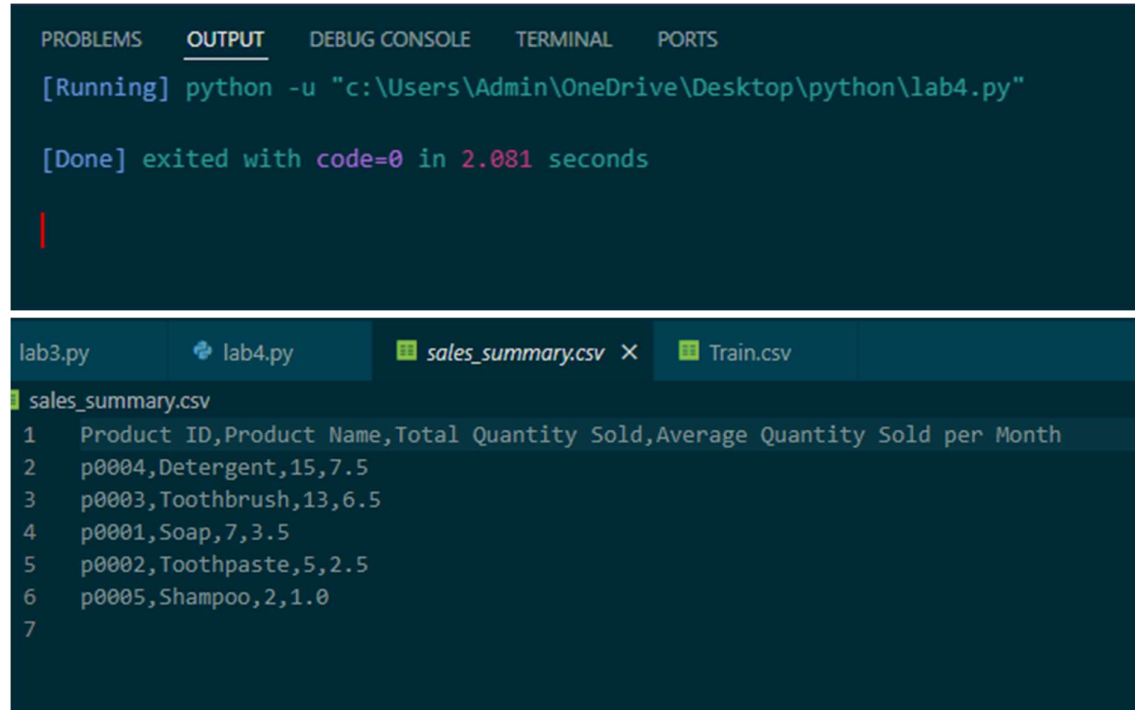
```
summary_df = pd.DataFrame(summary_data, columns=[
```

```
    "Product ID", "Product Name", "Total Quantity Sold", "Average Quantity  
Sold per Month"
```

```
])
```

```
summary_df = summary_df.sort_values(by="Total Quantity Sold",  
ascending=False).head(5)
```

```
summary_df.to_csv(summary_file, index=False)
```

Output:

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS
[Running] python -u "c:\Users\Admin\OneDrive\Desktop\python\lab4.py"
[Done] exited with code=0 in 2.081 seconds

lab3.py  lab4.py  sales_summary.csv  Train.csv
sales_summary.csv
1  Product ID,Product Name,Total Quantity Sold,Average Quantity Sold per Month
2  p0004,Detergent,15,7.5
3  p0003,Toothbrush,13,6.5
4  p0001,Soap,7,3.5
5  p0002,Toothpaste,5,2.5
6  p0005,Shampoo,2,1.0
7
```