MIT | Academy of Engineering

(An Autonomous Institute Affiliated to Savitribai Phule Pune University)

EDS Assignment

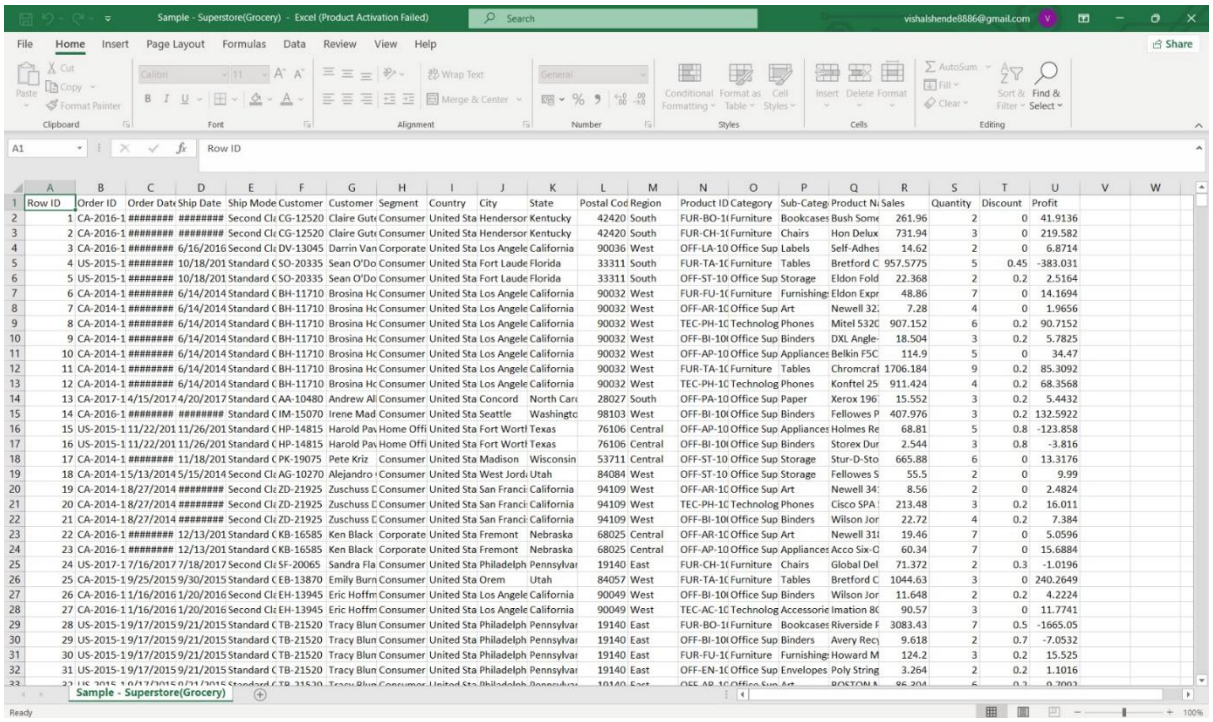
Name: - Vishal Shende

Class: - CS7

Roll No.: - CS7-65

PRN: - 202401110034

Sample-Superstore Grocery DATASET



Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer	Customer	Segment	Country	City	State	Postal Cod	Region	Product ID	Category	Sub-Categ	Product Na	Sales	Quantity	Discount	Profit
1	CA-2016-1	########	########	Second Cla	CG-12520	Claire Gut	Consumer	United Sta	Hendersor	Kentucky	42420	South	FUR-BO-1(	Furniture	Bookcases	Bush Some	261.96	2	0	41.9136
2	CA-2016-1	########	########	Second Cla	CG-12520	Claire Gut	Consumer	United Sta	Hendersor	Kentucky	42420	South	FUR-CH-1(	Furniture	Chairs	Hon Delux	731.94	3	0	219.582
3	CA-2016-1	########	6/16/2016	Second Cla	DV-13045	Darrin Van	Corporate	United Sta	Los Angele	California	90036	West	OFF-LA-10	Office Sup	Labels	Self-Adhes	14.62	2	0	6.8714
4	US-2015-1	########	10/18/201	Standard C	SO-20335	Sean O'Do	Consumer	United Sta	Fort Laude	Florida	33311	South	FUR-TA-1(	Furniture	Tables	Bretford C	957.5775	5	0.45	-383.031
5	US-2015-1	########	10/18/201	Standard C	SO-20335	Sean O'Do	Consumer	United Sta	Fort Laude	Florida	33311	South	OFF-ST-10	Office Sup	Storage	Eldon Fold	22.368	2	0.2	2.5164
6	CA-2014-1	########	6/14/2014	Standard C	BH-11710	Brosina Ho	Consumer	United Sta	Los Angele	California	90032	West	FUR-FU-1(	Furniture	Furnishing	Eldon Expr	48.86	7	0	14.1694
7	CA-2014-1	########	6/14/2014	Standard C	BH-11710	Brosina Ho	Consumer	United Sta	Los Angele	California	90032	West	OFF-AR-1(	Office Sup	Art	Newell 32	7.28	4	0	1.9656
8	CA-2014-1	########	6/14/2014	Standard C	BH-11710	Brosina Ho	Consumer	United Sta	Los Angele	California	90032	West	TEC-PH-1(	Technolog	Phones	Mitel 532(	907.152	6	0.2	90.7152
9	CA-2014-1	########	6/14/2014	Standard C	BH-11710	Brosina Ho	Consumer	United Sta	Los Angele	California	90032	West	OFF-BI-10(	Office Sup	Binders	DXL Angle-	18.504	3	0.2	5.7825
10	CA-2014-1	########	6/14/2014	Standard C	BH-11710	Brosina Ho	Consumer	United Sta	Los Angele	California	90032	West	OFF-AP-10	Office Sup	Appliances	Belkin F5C	114.9	5	0	34.47
11	CA-2014-1	########	6/14/2014	Standard C	BH-11710	Brosina Ho	Consumer	United Sta	Los Angele	California	90032	West	FUR-TA-1(	Furniture	Tables	Chromcraf	1706.184	9	0.2	85.3092
12	CA-2014-1	########	6/14/2014	Standard C	BH-11710	Brosina Ho	Consumer	United Sta	Los Angele	California	90032	West	TEC-PH-1(	Technolog	Phones	Konftel 25	911.424	4	0.2	68.3568
13	CA-2017-1	4/15/2017	4/20/2017	Standard C	AA-10480	Andrew All	Consumer	United Sta	Concord	North Car	28027	South	OFF-PA-10	Office Sup	Paper	Xerox 196	15.552	3	0.2	5.4432
14	CA-2016-1	########	########	Standard C	IM-15070	Irene Mad	Consumer	United Sta	Seattle	Washingto	98103	West	OFF-BI-10(	Office Sup	Binders	Fellowes P	407.976	3	0.2	132.5922
15	US-2015-1	11/22/201	11/26/201	Standard C	HP-14815	Harold Pa	Home Offi	United Sta	Fort Worth	Texas	76106	Central	OFF-AP-10	Office Sup	Appliances	Holmes Re	68.81	5	0.8	-123.858
16	US-2015-1	11/22/201	11/26/201	Standard C	HP-14815	Harold Pa	Home Offi	United Sta	Fort Worth	Texas	76106	Central	OFF-BI-10(	Office Sup	Binders	Storex Dur	2.544	3	0.8	-3.816
17	CA-2014-1	########	11/18/201	Standard C	PK-19075	Pete Kriz	Consumer	United Sta	Madison	Wisconsin	53711	Central	OFF-ST-10	Office Sup	Storage	Stur-D-Sto	665.88	6	0	13.3176
18	CA-2014-1	5/13/2014	5/15/2014	Second Cla	AG-10270	Alejandro	Consumer	United Sta	West Jordi	Utah	84084	West	OFF-ST-10	Office Sup	Storage	Fellowes S	55.5	2	0	9.99
19	CA-2014-1	8/27/2014	########	Second Cla	ZD-21925	Zuschuss D	Consumer	United Sta	San Franci	California	94109	West	OFF-AR-1(	Office Sup	Art	Newell 34	8.56	2	0	2.4824
20	CA-2014-1	8/27/2014	########	Second Cla	ZD-21925	Zuschuss D	Consumer	United Sta	San Franci	California	94109	West	TEC-PH-1(	Technolog	Phones	Cisco SPA	213.48	3	0.2	16.011
21	CA-2014-1	8/27/2014	########	Second Cla	ZD-21925	Zuschuss D	Consumer	United Sta	San Franci	California	94109	West	OFF-BI-10(	Office Sup	Binders	Wilson Jor	22.72	4	0.2	7.384
22	CA-2016-1	########	12/13/201	Standard C	KB-16585	Ken Black	Corporate	United Sta	Fremont	Nebraska	68025	Central	OFF-AR-1(	Office Sup	Art	Newell 31	19.46	7	0	5.0596
23	CA-2016-1	########	12/13/201	Standard C	KB-16585	Ken Black	Corporate	United Sta	Fremont	Nebraska	68025	Central	OFF-AP-10	Office Sup	Appliances	Acco Six-O	60.34	7	0	15.6884
24	US-2017-1	7/16/2017	7/18/2017	Second Cla	SF-20065	Sandra Fla	Consumer	United Sta	Philadelph	Pennsylvar	19140	East	FUR-CH-1(	Furniture	Chairs	Global Del	71.372	2	0.3	-1.0196
25	CA-2015-1	9/25/2015	9/30/2015	Standard C	EB-13870	Emily Burn	Consumer	United Sta	Orem	Utah	84057	West	FUR-TA-1(	Furniture	Tables	Bretford C	1044.63	3	0	240.2649
26	CA-2016-1	1/16/2016	1/20/2016	Second Cla	EH-13945	Eric Hoffm	Consumer	United Sta	Los Angele	California	90049	West	OFF-BI-10(	Office Sup	Binders	Wilson Jor	11.648	2	0.2	4.2224
27	CA-2016-1	1/16/2016	1/20/2016	Second Cla	EH-13945	Eric Hoffm	Consumer	United Sta	Los Angele	California	90049	West	TEC-AC-1(	Technolog	Accessorie	Imation 8(	90.57	3	0	11.7741
28	US-2015-1	9/17/2015	9/21/2015	Standard C	TB-21520	Tracy Blun	Consumer	United Sta	Philadelph	Pennsylvar	19140	East	FUR-BO-1(	Furniture	Bookcases	Riverside F	3083.43	7	0.5	-1665.05
29	US-2015-1	9/17/2015	9/21/2015	Standard C	TB-21520	Tracy Blun	Consumer	United Sta	Philadelph	Pennsylvar	19140	East	OFF-BI-10(	Office Sup	Binders	Avery Rec	9.618	2	0.7	-7.0532
30	US-2015-1	9/17/2015	9/21/2015	Standard C	TB-21520	Tracy Blun	Consumer	United Sta	Philadelph	Pennsylvar	19140	East	FUR-FU-1(	Furniture	Furnishing	Howard M	124.2	3	0.2	15.525
31	US-2015-1	9/17/2015	9/21/2015	Standard C	TB-21520	Tracy Blun	Consumer	United Sta	Philadelph	Pennsylvar	19140	East	OFF-EN-1(	Office Sup	Envelopes	Poly String	3.264	2	0.2	1.1016

Display of data in Python

Welcome | panadas.py | Untitled-1 | Sample - Superstore.csv

panadas.py > ...

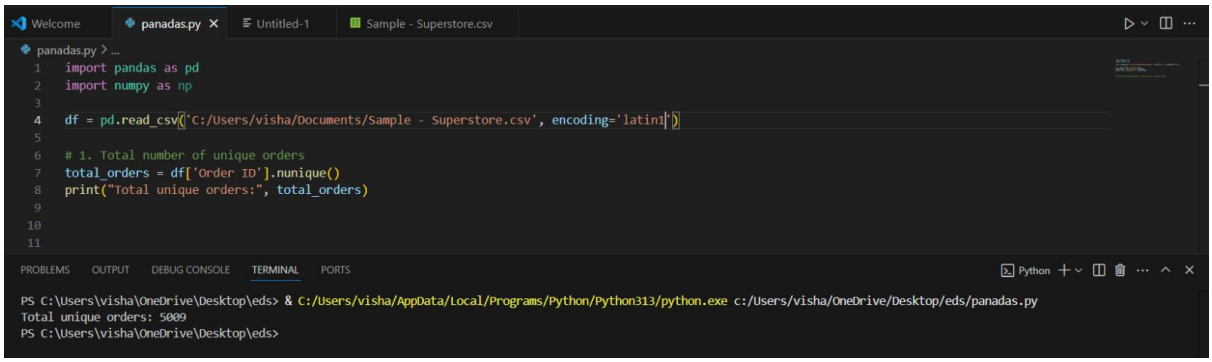
```python
# Import necessary libraries
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')
pd.set_option('display.max_rows', None)# Show all rows

pd.set_option('display.max_columns', None)# Show all columns
pd.set_option('display.width', None)# Don't cut lines
pd.set_option('display.max_colwidth', None)# Full content of each column
print(df)
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
  90045     West  OFF-AP-10000055  Office Supplies   Appliances                                                                                     Belkin F9S820V06
 8 Outlet Surge     64.9600        2     0.00    19.4880
9652    9653  CA-2016-107104  11/26/2016  11/30/2016  Standard Class    MS-17365     Maribeth Schnelling     Consumer  United States        Los Angeles           California
  90045     West  FUR-FU-10002937         Furniture  Furnishings                                                   GE 48" Fluorescent Tube, Cool White Energy Saver, 3
4 Watts, 30/Box     595.3800        6     0.00   297.6900
9653    9654  CA-2017-161893   1/23/2017   1/27/2017  Standard Class    HP-14815          Harold Pawlan  Home Office  United States             Tucson             Arizona
  85705     West  TEC-AC-10001874        Technology  Accessories                                                           Logitech Wireless Anywhere Mouse MX
 for PC and Mac     95.9840        2     0.20    11.9980
9654    9655  CA-2017-161893   1/23/2017   1/27/2017  Standard Class    HP-14815          Harold Pawlan  Home Office  United States             Tucson             Arizona
  85705     West  OFF-BI-10004506  Office Supplies      Binders                                                    Wilson Jones data.warehouse D-Ring Binder
s with DublLock      4.9380        2     0.70    -3.6212
9655    9656  CA-2014-156160   9/22/2014   9/29/2014  Standard Class    AS-10090      Adam Shillingsburg     Consumer  United States      New York City            New York
  10035     East  FUR-FU-10001876         Furniture  Furnishings                                                                                    Computer R
oom Manger, 14"     97.4400        3     0.00    35.0784
9656    9657  CA-2014-156160   9/22/2014   9/29/2014  Standard Class    AS-10090      Adam Shillingsburg     Consumer  United States      New York City            New York
  10035     East  OFF-BI-10003355  Office Supplies      Binders                                                                       Cardinal Holdit Busine
ss Card Pockets      3.9840        1     0.20     1.3944
9657    9658  CA-2014-156160   9/22/2014   9/29/2014  Standard Class    AS-10090      Adam Shillingsburg     Consumer  United States      New York City            New York
  10035     East  OFF-AR-10003876  Office Supplies          Art                                                 Avery Hi-Liter GlideStik Fluorescent Highligh
ter, Yellow Ink     13.0400        4     0.00     5.7376
9658    9659  CA-2014-156160   9/22/2014   9/29/2014  Standard Class    AS-10090      Adam Shillingsburg     Consumer  United States      New York City            New York
  10035     East  FUR-CH-10004983         Furniture       Chairs                             Office Star - Mid Back Dual function Ergonomic High Back Chair with 2-Way
Adjustable Arms    579.5280        4     0.10    83.7096
9659    9660  CA-2014-100860   3/26/2014   3/30/2014    Second Class    CS-12505          Cindy Stewart     Consumer  United States             Pomona          California
  91767     West  OFF-LA-10001982  Office Supplies       Labels                                                 Smead Alpha-Z Color-Coded Name Labels First Let
ter Starter Set     18.7500        5     0.00     9.0000
9660    9661  CA-2016-160717    6/6/2016   6/11/2016  Standard Class    ME-17320          Maria Etezadi  Home Office  United States      Santa Barbara          California
  93101     West  TEC-PH-10001459        Technology       Phones                                                                                       Samsung
Galaxy Mega 6.3   3023.9280        9     0.20   226.7946
9661    9662  CA-2016-160717    6/6/2016   6/11/2016  Standard Class    ME-17320          Maria Etezadi  Home Office  United States      Santa Barbara          California
  93101     West  TEC-AC-10002134        Technology  Accessories                                                        Rosewill 107 Normal Keys USB Wired St
andard Keyboard     26.9600        2     0.00     3.7744
9662    9663  CA-2016-160717    6/6/2016   6/11/2016  Standard Class    ME-17320          Maria Etezadi  Home Office  United States      Santa Barbara          California
  93101     West  TEC-PH-10001760        Technology       Phones                                                                              Bose SoundLink Bl
uetooth Speaker    477.6000        3     0.20   161.1900
9663    9664  CA-2017-132619   8/31/2017    9/5/2017  Standard Class    MS-17770        Maxwell Schwartz     Consumer  United States      San Francisco          California
  94109     West  OFF-ST-10001511  Office Supplies      Storage                                                                  Space Solutions Commercial
 Steel Shelving    193.9500        3     0.00     9.6975
9664    9665  US-2015-120957   12/7/2015   12/7/2015        Same Day    KH-16690        Kristen Hastings    Corporate  United States          Riverside          California
  92503     West  OFF-PA-10003953  Office Supplies        Paper
      Xerox 218     12.9600        2     0.00     6.2208
9665    9666  CA-2017-157448  11/16/2017  11/20/2017  Standard Class    LC-16885          Lena Creighton     Consumer  United States        Los Angeles          California
  90049     West  FUR-FU-10003268         Furniture  Furnishings                                                    Eldon Radial Chair Mat for Low to Medi
um Pile Carpets    119.9400        3     0.00    23.9880
9666    9667  CA-2017-157448  11/16/2017  11/20/2017  Standard Class    LC-16885          Lena Creighton     Consumer  United States        Los Angeles          California
  90049     West  FUR-FU-10001847         Furniture  Furnishings                                                                Eldon Image Series Black D
```

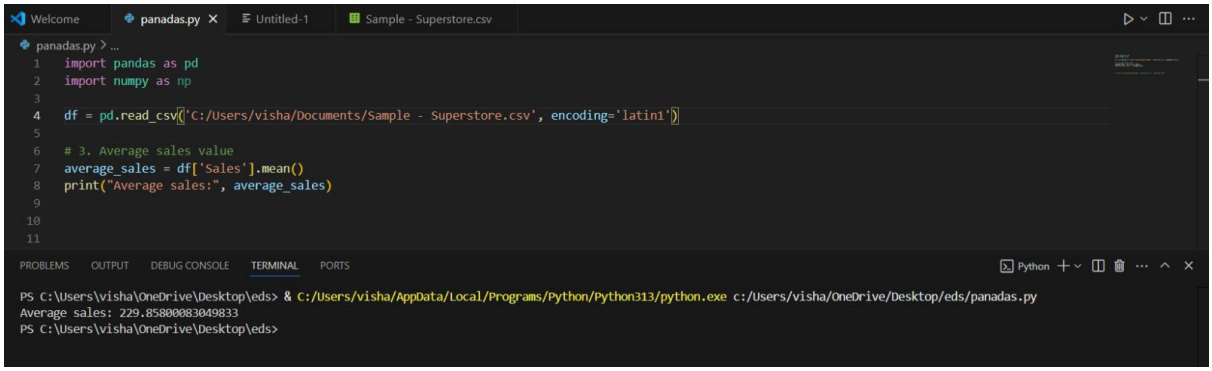
Ln 12, Col 1 Spaces: 4 UTF-8 CRLF {} Python 3.13.2 64-bit Go Live

1. Find the total number of unique orders.

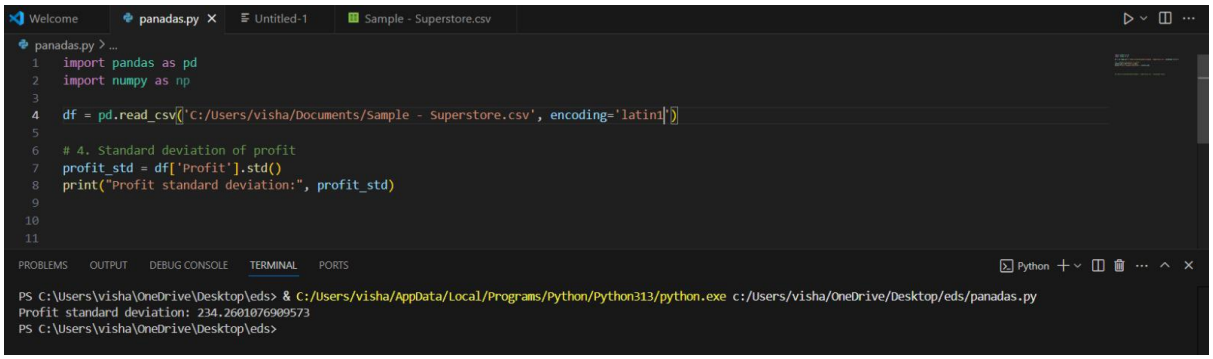


2.Find the customer with the highest number of orders.

3. Calculate the average sales value across all orders.



4. Find the standard deviation of profit overall.



```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')

# 4. Standard deviation of profit
profit_std = df['Profit'].std()
print("Profit standard deviation:", profit_std)
```

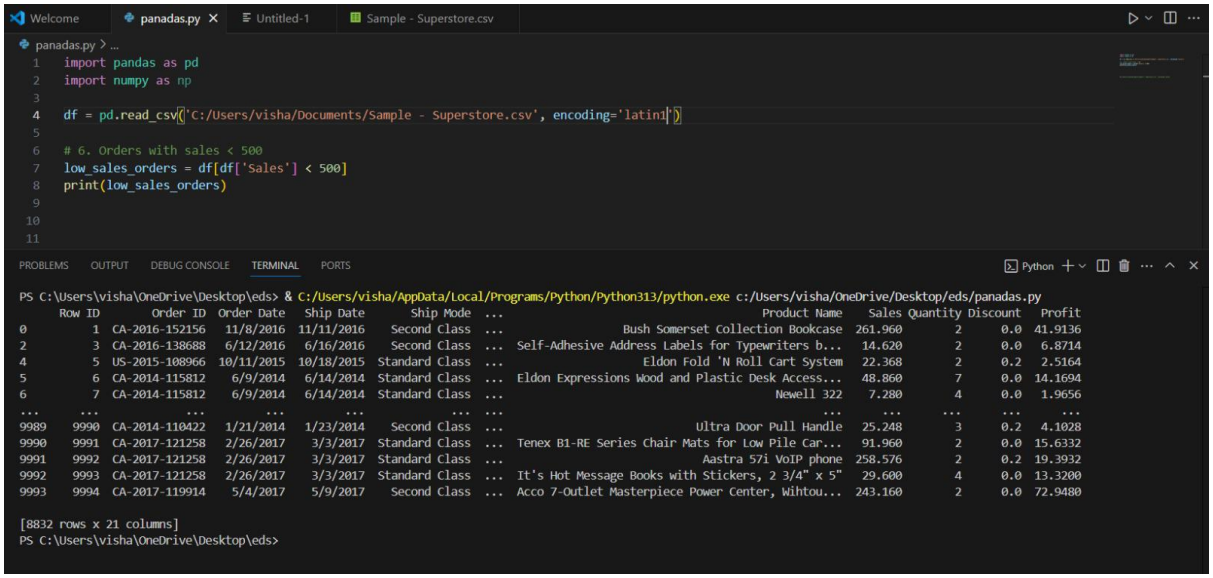
```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
Profit standard deviation: 234.2601076909573
PS C:\Users\visha\OneDrive\Desktop\eds>
```

5. Find the number of orders with discount greater than 30%

```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')

# 5. Number of orders with discount > 30%
high_discount_orders = len(df[df['Discount'] > 0.3])
print("Orders with discount > 30%:", high_discount_orders)
```

```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
Orders with discount > 30%: 1166
PS C:\Users\visha\OneDrive\Desktop\eds>
```

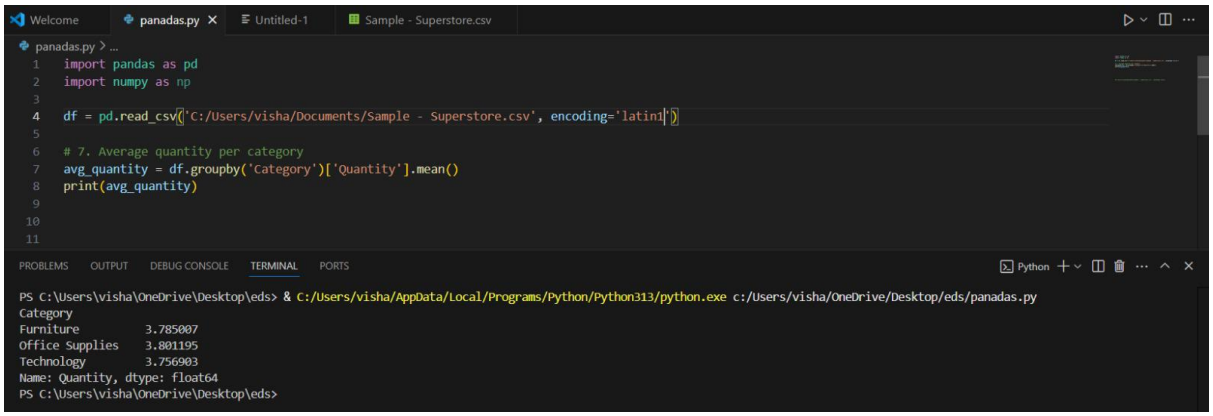
6. List order where sales value is less than Rs 500.



```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')

# 6. Orders with sales < 500
low_sales_orders = df[df['Sales'] < 500]
print(low_sales_orders)
```

```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
      Row ID        Order ID  Order Date   Ship Date       Ship Mode  ...                                       Product Name   Sales  Quantity  Discount   Profit
0          1  CA-2016-152156   11/8/2016  11/11/2016    Second Class  ...                      Bush Somerset Collection Bookcase  261.960         2       0.0  41.9136
2          3  CA-2016-138688   6/12/2016   6/16/2016    Second Class  ...  Self-Adhesive Address Labels for Typewriters b...  14.620         2       0.0   6.8714
4          5  US-2015-108966  10/11/2015  10/18/2015  Standard Class  ...                         Eldon Fold 'N Roll Cart System  22.368         2       0.2   2.5164
5          6  CA-2014-115812    6/9/2014   6/14/2014  Standard Class  ...  Eldon Expressions Wood and Plastic Desk Access...  48.860         7       0.0  14.1694
6          7  CA-2014-115812    6/9/2014   6/14/2014  Standard Class  ...                                    Newell 322   7.280         4       0.0   1.9656
...      ...             ...         ...         ...            ...  ...                                           ...     ...       ...       ...      ...
9989    9990  CA-2014-110422   1/21/2014   1/23/2014    Second Class  ...                               Ultra Door Pull Handle  25.248         3       0.2   4.1028
9990    9991  CA-2017-121258   2/26/2017   3/3/2017  Standard Class  ...  Tenex B1-RE Series Chair Mats for Low Pile Car...  91.960         2       0.0  15.6332
9991    9992  CA-2017-121258   2/26/2017   3/3/2017  Standard Class  ...                                  Aastra 57i VoIP phone  258.576         2       0.2  19.3932
9992    9993  CA-2017-121258   2/26/2017   3/3/2017  Standard Class  ...  It's Hot Message Books with Stickers, 2 3/4" x 5"  29.600         4       0.0  13.3200
9993    9994  CA-2017-119914    5/4/2017    5/9/2017    Second Class  ...  Acco 7-Outlet Masterpiece Power Center, Wihtou...  243.160         2       0.0  72.9480

[8832 rows x 21 columns]
PS C:\Users\visha\OneDrive\Desktop\eds>
```

7. Calculate the average quantity sold per category.



```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')

# 7. Average quantity per category
avg_quantity = df.groupby('Category')['Quantity'].mean()
print(avg_quantity)
```

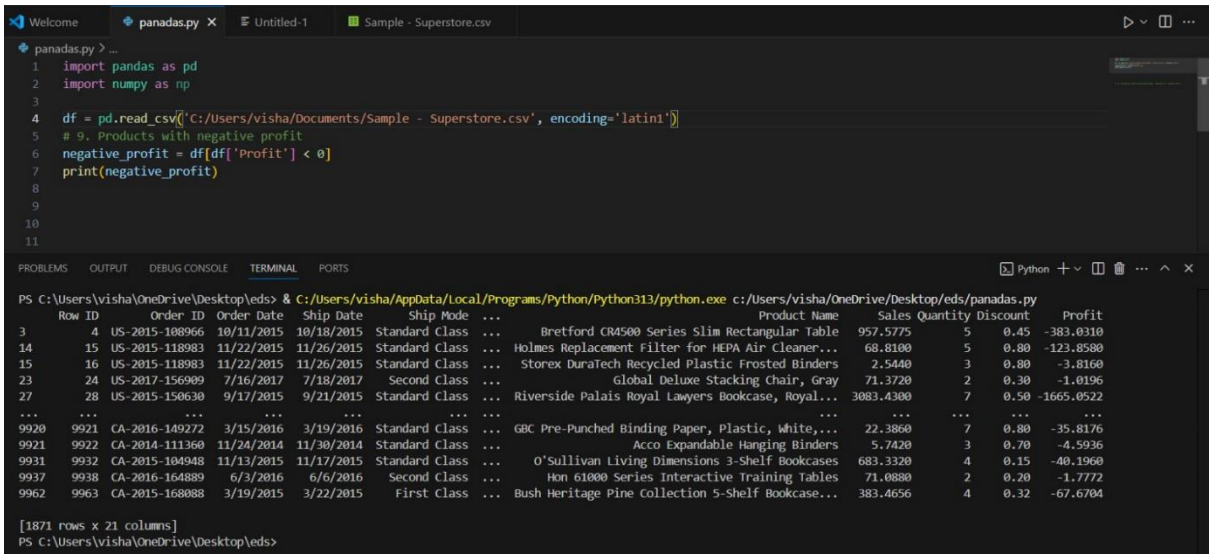
```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
Category
Furniture          3.785007
Office Supplies    3.801195
Technology         3.756903
Name: Quantity, dtype: float64
PS C:\Users\visha\OneDrive\Desktop\eds>
```

8. Find the number of unique product sold in each category.

```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')

# 8. Unique products sold in each category
unique_products = df.groupby('Category')['Product ID'].nunique()
print(unique_products)
```

```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
Category
Furniture          375
Office Supplies    1083
Technology         404
Name: Product ID, dtype: int64
PS C:\Users\visha\OneDrive\Desktop\eds>
```

9. Find product when profit is negative.



```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')
# 9. Products with negative profit
negative_profit = df[df['Profit'] < 0]
print(negative_profit)
```

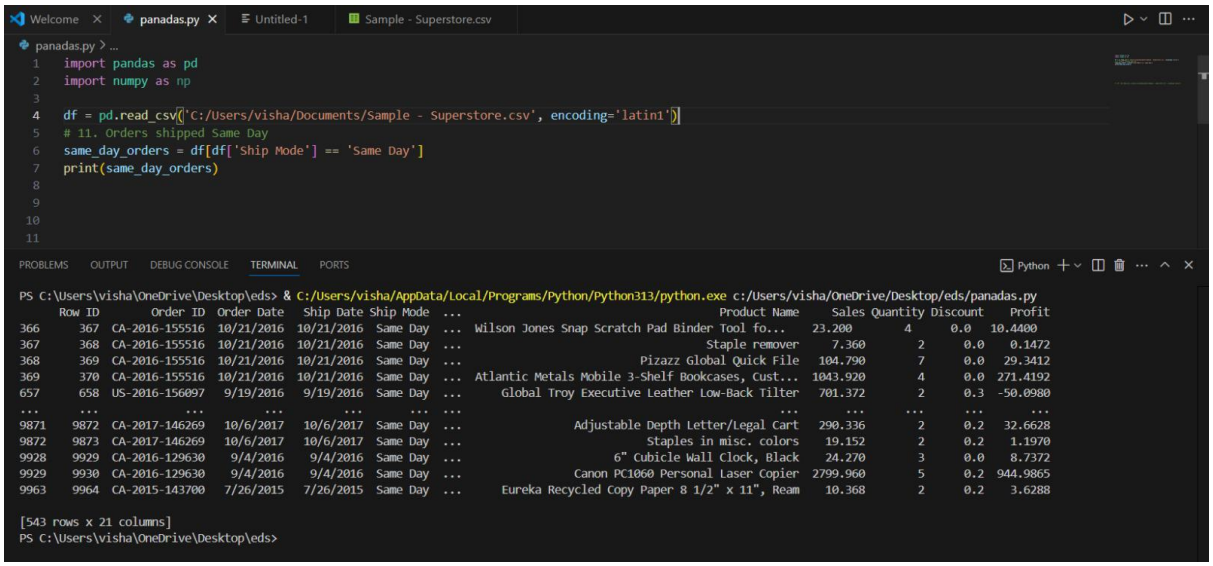
```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
      Row ID        Order ID  Order Date   Ship Date       Ship Mode  ...                                       Product Name     Sales  Quantity  Discount     Profit
3          4  US-2015-108966  10/11/2015  10/18/2015  Standard Class  ...      Bretford CR4500 Series Slim Rectangular Table  957.5775         5      0.45  -383.0310
14        15  US-2015-118983  11/22/2015  11/26/2015  Standard Class  ...  Holmes Replacement Filter for HEPA Air Cleaner...   68.8100         5      0.80  -123.8580
15        16  US-2015-118983  11/22/2015  11/26/2015  Standard Class  ...   Storex DuraTech Recycled Plastic Frosted Binders    2.5440         3      0.80    -3.8160
23        24  US-2017-156909   7/16/2017   7/18/2017    Second Class  ...                Global Deluxe Stacking Chair, Gray   71.3720         2      0.30    -1.0196
27        28  US-2015-150630   9/17/2015   9/21/2015  Standard Class  ...  Riverside Palais Royal Lawyers Bookcase, Royal...  3083.4300         7      0.50 -1665.0522
...      ...             ...         ...         ...             ...  ...                                            ...        ...       ...      ...        ...
9920    9921  CA-2016-149272   3/15/2016   3/19/2016  Standard Class  ...  GBC Pre-Punched Binding Paper, Plastic, White,...   22.3860         7      0.80   -35.8176
9921    9922  CA-2014-111360  11/24/2014  11/30/2014  Standard Class  ...                    Acco Expandable Hanging Binders    5.7420         3      0.70    -4.5936
9931    9932  CA-2015-104948  11/13/2015  11/17/2015  Standard Class  ...     O'Sullivan Living Dimensions 3-Shelf Bookcases  683.3320         4      0.15   -40.1960
9937    9938  CA-2016-164889   6/3/2016    6/6/2016    Second Class  ...       Hon 61000 Series Interactive Training Tables   71.0880         2      0.20    -1.7772
9962    9963  CA-2015-168088   3/19/2015   3/22/2015     First Class  ...  Bush Heritage Pine Collection 5-Shelf Bookcase...  383.4656         4      0.32   -67.6704

[1871 rows x 21 columns]
PS C:\Users\visha\OneDrive\Desktop\eds>
```

10. Find the maximum sales recorded in a single order.

```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')
# 10. Maximum sales in an order
max_sales = df['Sales'].max()
print("Maximum sales:", max_sales)
```

```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
Maximum sales: 22638.48
PS C:\Users\visha\OneDrive\Desktop\eds>
```

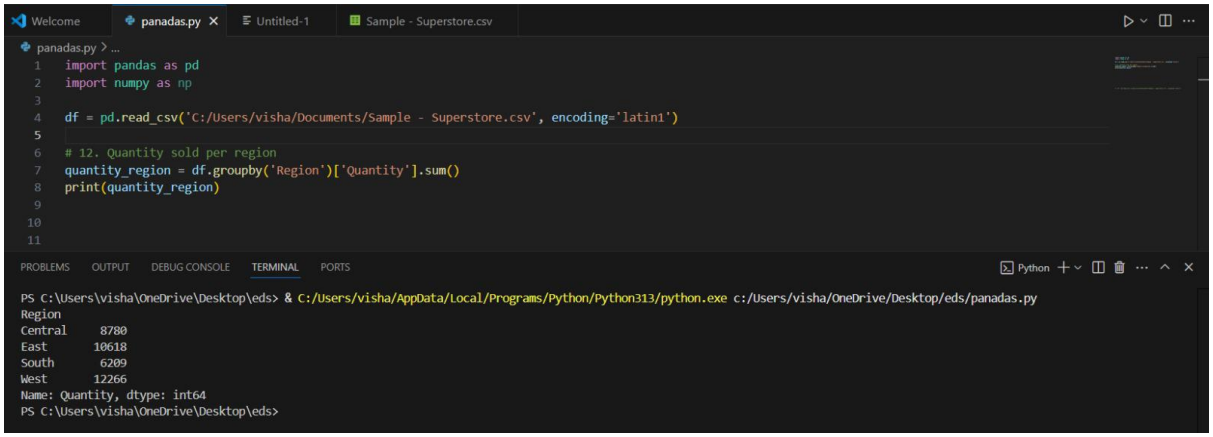
11. List all order where ship mode is “Same Day”.



```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')
# 11. Orders shipped Same Day
same_day_orders = df[df['Ship Mode'] == 'Same Day']
print(same_day_orders)
```

```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
      Row ID        Order ID  Order Date   Ship Date Ship Mode  ...                                       Product Name     Sales Quantity Discount    Profit
366      367  CA-2016-155516  10/21/2016  10/21/2016  Same Day  ...  Wilson Jones Snap Scratch Pad Binder Tool fo...    23.200        4      0.0   10.4400
367      368  CA-2016-155516  10/21/2016  10/21/2016  Same Day  ...                                    Staple remover     7.360        2      0.0    0.1472
368      369  CA-2016-155516  10/21/2016  10/21/2016  Same Day  ...                              Pizazz Global Quick File   104.790        7      0.0   29.3412
369      370  CA-2016-155516  10/21/2016  10/21/2016  Same Day  ...  Atlantic Metals Mobile 3-Shelf Bookcases, Cust...  1043.920        4      0.0  271.4192
657      658  US-2016-156097   9/19/2016   9/19/2016  Same Day  ...      Global Troy Executive Leather Low-Back Tilter   701.372        2      0.3  -50.0980
...      ...             ...         ...         ...       ...  ...                                               ...       ...      ...       ...       ...
9871    9872  CA-2017-146269   10/6/2017   10/6/2017  Same Day  ...               Adjustable Depth Letter/Legal Cart   290.336        2      0.2   32.6628
9872    9873  CA-2017-146269   10/6/2017   10/6/2017  Same Day  ...                          Staples in misc. colors    19.152        2      0.2    1.1970
9928    9929  CA-2016-129630    9/4/2016    9/4/2016  Same Day  ...                6" Cubicle Wall Clock, Black    24.270        3      0.0    8.7372
9929    9930  CA-2016-129630    9/4/2016    9/4/2016  Same Day  ...               Canon PC1060 Personal Laser Copier  2799.960        5      0.2  944.9865
9963    9964  CA-2015-143700   7/26/2015   7/26/2015  Same Day  ...      Eureka Recycled Copy Paper 8 1/2" x 11", Ream    10.368        2      0.2    3.6288

[543 rows x 21 columns]
PS C:\Users\visha\OneDrive\Desktop\eds>
```

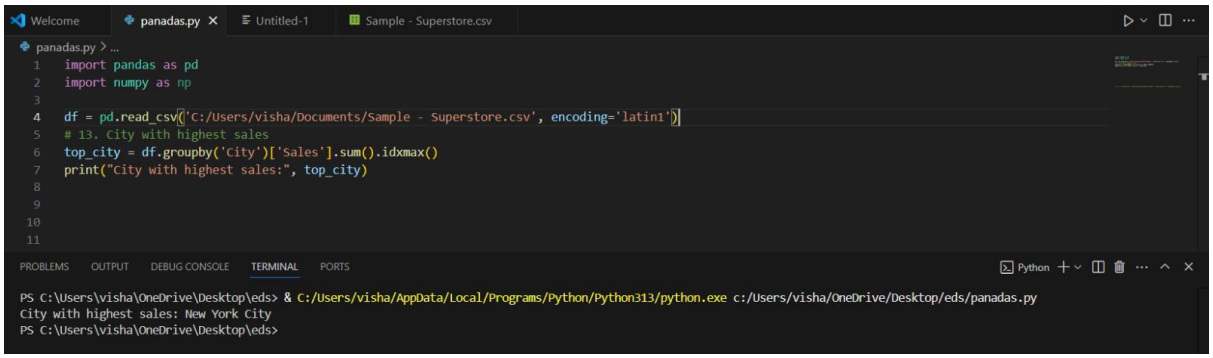
12. Find the total quantity of product sold per region.



```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')

# 12. Quantity sold per region
quantity_region = df.groupby('Region')['Quantity'].sum()
print(quantity_region)
```

```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
Region
Central     8780
East       10618
South       6209
West       12266
Name: Quantity, dtype: int64
PS C:\Users\visha\OneDrive\Desktop\eds>
```

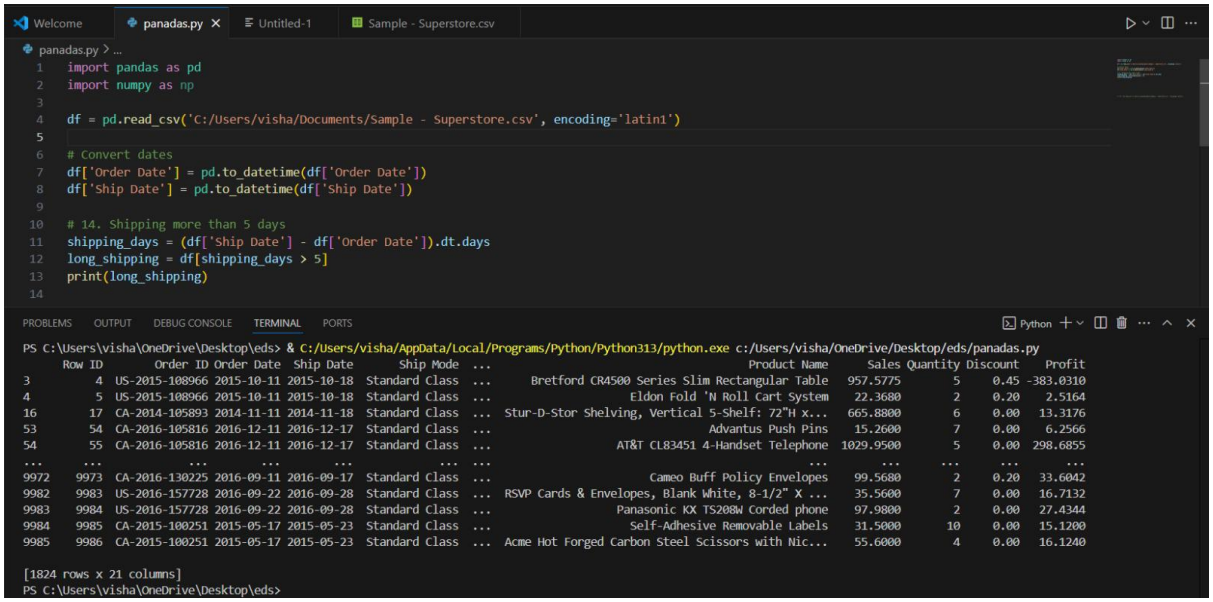
13. Find the city with highest total sales.



```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')
# 13. City with highest sales
top_city = df.groupby('City')['Sales'].sum().idxmax()
print("City with highest sales:", top_city)
```

```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
City with highest sales: New York City
PS C:\Users\visha\OneDrive\Desktop\eds>
```

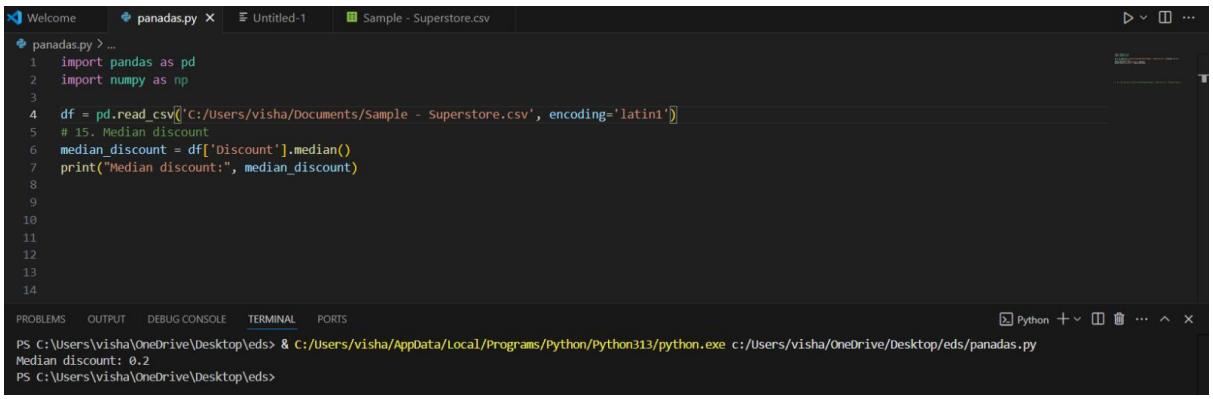
14. Find orders where shipping took more than 5 days.



```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')

# Convert dates
df['Order Date'] = pd.to_datetime(df['Order Date'])
df['Ship Date'] = pd.to_datetime(df['Ship Date'])

# 14. Shipping more than 5 days
shipping_days = (df['Ship Date'] - df['Order Date']).dt.days
long_shipping = df[shipping_days > 5]
print(long_shipping)
```

```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
      Row ID        Order ID Order Date  Ship Date       Ship Mode  ...                                       Product Name      Sales Quantity Discount    Profit
3          4  US-2015-108966 2015-10-11 2015-10-18  Standard Class  ...      Bretford CR4500 Series Slim Rectangular Table   957.5775        5     0.45 -383.0310
4          5  US-2015-108966 2015-10-11 2015-10-18  Standard Class  ...                          Eldon Fold 'N Roll Cart System    22.3680        2     0.20    2.5164
16        17  CA-2014-105893 2014-11-11 2014-11-18  Standard Class  ...  Stur-D-Stor Shelving, Vertical 5-Shelf: 72"H x...   665.8800        6     0.00   13.3176
53        54  CA-2016-105816 2016-12-11 2016-12-17  Standard Class  ...                                  Advantus Push Pins    15.2600        7     0.00    6.2566
54        55  CA-2016-105816 2016-12-11 2016-12-17  Standard Class  ...                AT&T CL83451 4-Handset Telephone  1029.9500        5     0.00  298.6855
...      ...             ...        ...        ...             ...  ...                                             ...       ...      ...      ...       ...
9972    9973  CA-2016-130225 2016-09-11 2016-09-17  Standard Class  ...                          Cameo Buff Policy Envelopes    99.5680        2     0.20   33.6042
9982    9983  US-2016-157728 2016-09-22 2016-09-28  Standard Class  ...  RSVP Cards & Envelopes, Blank White, 8-1/2" X ...    35.5600        7     0.00   16.7132
9983    9984  US-2016-157728 2016-09-22 2016-09-28  Standard Class  ...                       Panasonic KX TS208W Corded phone    97.9800        2     0.00   27.4344
9984    9985  CA-2015-100251 2015-05-17 2015-05-23  Standard Class  ...                          Self-Adhesive Removable Labels    31.5000       10     0.00   15.1200
9985    9986  CA-2015-100251 2015-05-17 2015-05-23  Standard Class  ...  Acme Hot Forged Carbon Steel Scissors with Nic...    55.6000        4     0.00   16.1240

[1824 rows x 21 columns]
PS C:\Users\visha\OneDrive\Desktop\eds>
```

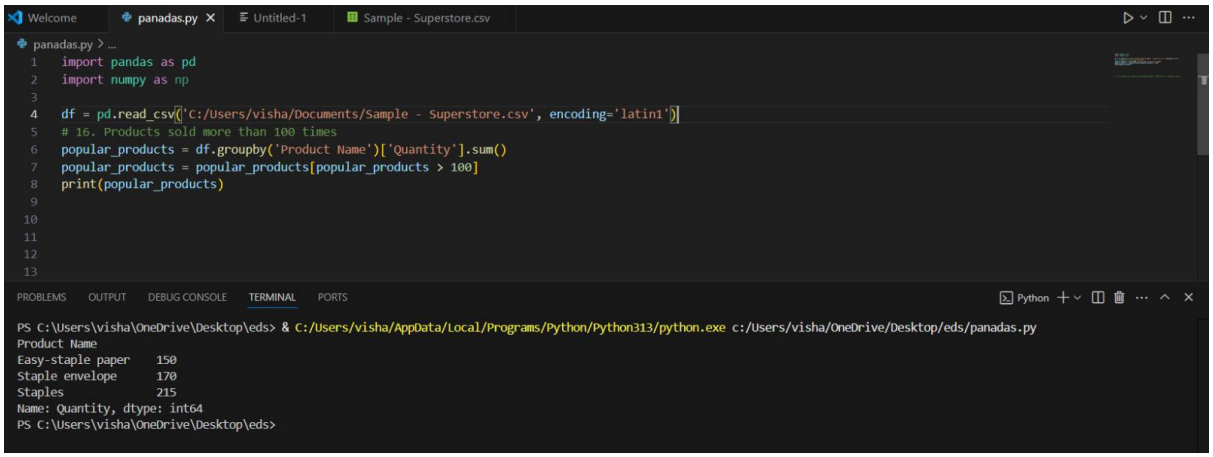
15. Calculate the median discount across all orders.



```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')
# 15. Median discount
median_discount = df['Discount'].median()
print("Median discount:", median_discount)
```

```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
Median discount: 0.2
PS C:\Users\visha\OneDrive\Desktop\eds>
```

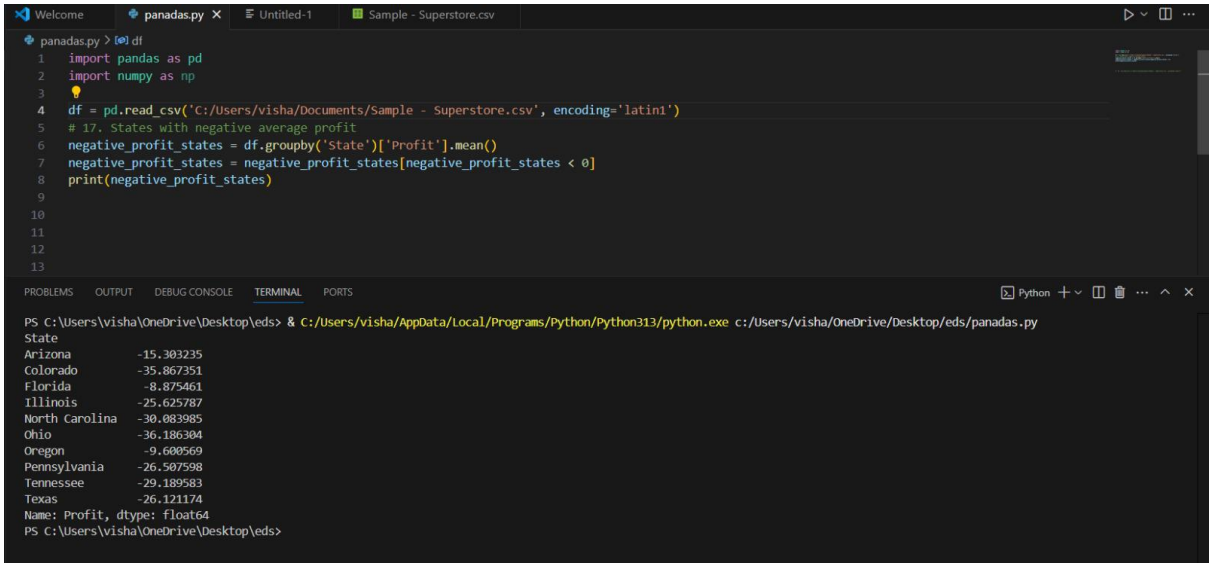
16. Find product sold more than 100 times.



```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')
# 16. Products sold more than 100 times
popular_products = df.groupby('Product Name')['Quantity'].sum()
popular_products = popular_products[popular_products > 100]
print(popular_products)
```

```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
Product Name
Easy-staple paper    150
Staple envelope      170
Staples              215
Name: Quantity, dtype: int64
PS C:\Users\visha\OneDrive\Desktop\eds>
```

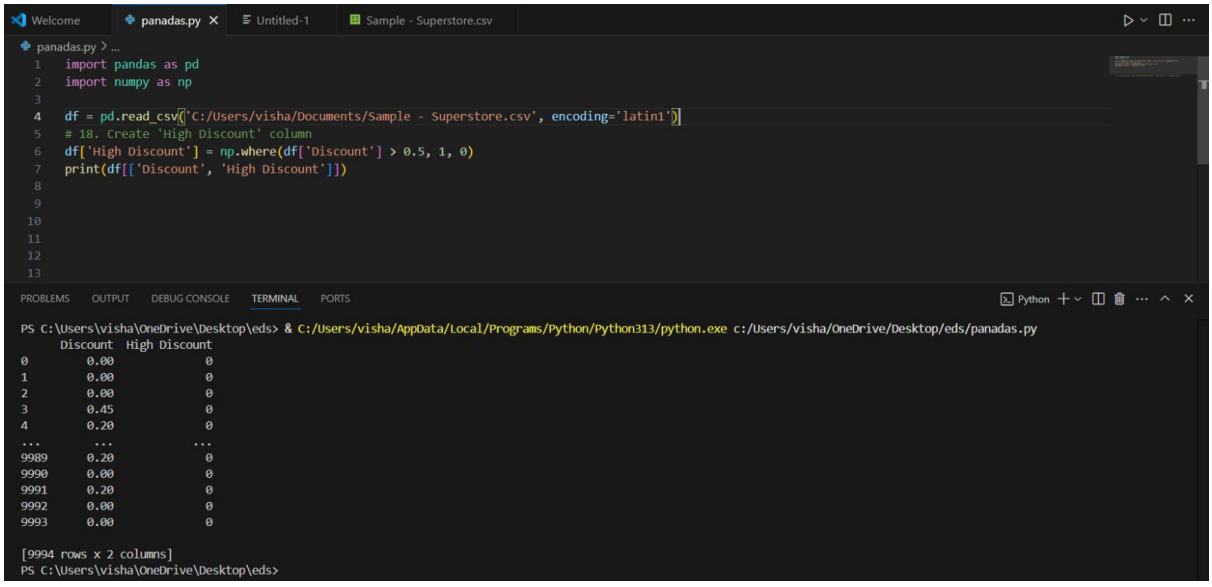
17. Find states where average profit is negative.



```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')
# 17. States with negative average profit
negative_profit_states = df.groupby('State')['Profit'].mean()
negative_profit_states = negative_profit_states[negative_profit_states < 0]
print(negative_profit_states)
```

```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
State
Arizona          -15.303235
Colorado         -35.867351
Florida           -8.875461
Illinois         -25.625787
North Carolina   -30.083985
Ohio             -36.186304
Oregon            -9.600569
Pennsylvania     -26.507598
Tennessee        -29.189583
Texas            -26.121174
Name: Profit, dtype: float64
PS C:\Users\visha\OneDrive\Desktop\eds>
```

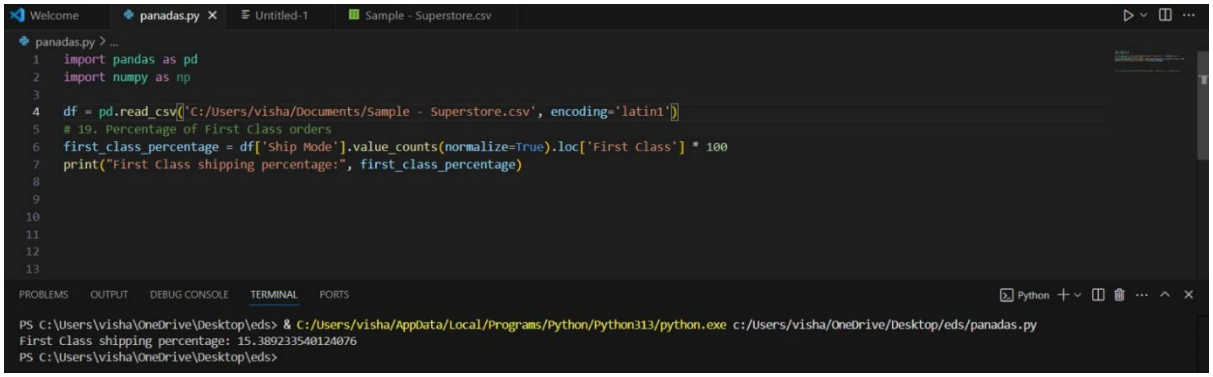
18. Create a new column ‘High Discount’ (Discount>50%).



```
import pandas as pd
import numpy as np

df = pd.read_csv('C:/Users/visha/Documents/Sample - Superstore.csv', encoding='latin1')
# 18. Create 'High Discount' column
df['High Discount'] = np.where(df['Discount'] > 0.5, 1, 0)
print(df[['Discount', 'High Discount']])
```

```
PS C:\Users\visha\OneDrive\Desktop\eds> & C:/Users/visha/AppData/Local/Programs/Python/Python313/python.exe c:/Users/visha/OneDrive/Desktop/eds/panadas.py
      Discount  High Discount
0         0.00              0
1         0.00              0
2         0.00              0
3         0.45              0
4         0.20              0
...        ...            ...
9989      0.20              0
9990      0.00              0
9991      0.20              0
9992      0.00              0
9993      0.00              0

[9994 rows x 2 columns]
PS C:\Users\visha\OneDrive\Desktop\eds>
```

19. Find the percentage of orders shipping in “First Class” mode



20. Create a new column ‘Fast shipping’ (Order Date = Ship Date).