In [1]:
```
import pandas as pd
```

In [98]:
```
path = 'SuperMarket Analysis.csv'
market = pd.read_csv(path)
market.columns = market.columns.str.strip()
```

Q1. Load the (SuperMarket Analysis.csv) dataset and display the first 5 rows.

In [99]:
```
market.head()
```

Out[99]:

	Invoice ID	Branch	City	Customer type	Gender	Product line	Unit price	Quant
0	750-67-8428	Alex	Yangon	Member	Female	Health and beauty	74.69	
1	226-31-3081	Giza	Naypyitaw	Normal	Female	Electronic accessories	15.28	
2	631-41-3108	Alex	Yangon	Normal	Female	Home and lifestyle	46.33	
3	123-19-1176	Alex	Yangon	Member	Female	Health and beauty	58.22	
4	373-73-7910	Alex	Yangon	Member	Female	Sports and travel	86.31	

Q2. Display the dataset shape (rows and columns).

In [100…
```
market.shape
```

Out[100… (1000, 17)

Q3. List all column names.

In [101…
```
market.columns
```

Out[101…
```
Index(['Invoice ID', 'Branch', 'City', 'Customer type', 'Gender',
       'Product line', 'Unit price', 'Quantity', 'Tax 5%', 'Sale
s', 'Date',
       'Time', 'Payment', 'cogs', 'gross margin percentage', 'gros
s income',
       'Rating'],
      dtype='object')
```

Q4. Identify categorical and numerical columns.

In [102…
```
categorical_col = market.select_dtypes(include='object').columns
numerical_col = market.select_dtypes(exclude='object').columns
print("categorical columns -> " + categorical_col)
print()
print("numerical columns -> " + numerical_col)
```

```
Index(['categorical columns -> Invoice ID', 'categorical columns ->
Branch',
       'categorical columns -> City', 'categorical columns -> Custom
er type',
       'categorical columns -> Gender', 'categorical columns -> Prod
uct line',
       'categorical columns -> Date', 'categorical columns -> Time',
       'categorical columns -> Payment'],
      dtype='object')

Index(['numerical columns -> Unit price', 'numerical columns -> Quan
tity',
       'numerical columns -> Tax 5%', 'numerical columns -> Sales',
       'numerical columns -> cogs',
       'numerical columns -> gross margin percentage',
       'numerical columns -> gross income', 'numerical columns -> Ra
ting'],
      dtype='object')
```

Q5. Check for missing values in each column.

In [103…
```
market.isnull().sum()
```

Out[103…
```
Invoice ID                0
Branch                    0
City                      0
Customer type             0
Gender                    0
Product line              0
Unit price                0
Quantity                  0
Tax 5%                    0
Sales                     0
Date                      0
Time                      0
Payment                   0
cogs                      0
gross margin percentage   0
gross income              0
Rating                    0
dtype: int64
```

Q6. Display the data types of each column.

In [104…
```
market.dtypes
```

Out[104…
```
Invoice ID                 object
Branch                     object
City                       object
Customer type              object
Gender                     object
Product line               object
Unit price                float64
Quantity                    int64
Tax 5%                    float64
Sales                     float64
Date                       object
Time                       object
Payment                    object
cogs                      float64
gross margin percentage   float64
gross income              float64
Rating                    float64
dtype: object
```

Q7. Show summary statistics for numerical columns (use pandas method).

In [105…
```
market.describe(exclude='object')
```

Out [105…

	Unit price	Quantity	Tax 5%	Sales	cogs	g
count	1000.000000	1000.000000	1000.000000	1000.000000	1000.00000	1.
mean	55.672130	5.510000	15.379369	322.966749	307.58738	4
std	26.494628	2.923431	11.708825	245.885335	234.17651	(
min	10.080000	1.000000	0.508500	10.678500	10.17000	4
25%	32.875000	3.000000	5.924875	124.422375	118.49750	4
50%	55.230000	5.000000	12.088000	253.848000	241.76000	4
75%	77.935000	8.000000	22.445250	471.350250	448.90500	4
max	99.960000	10.000000	49.650000	1042.650000	993.00000	4

Q8. Filter rows where Sales > 500

In [112…

```
print(market[market['Sales'] > 500])
```

```
     Invoice ID Branch       City Customer type  Gender  \
0    750-67-8428   Alex     Yangon        Member  Female
4    373-73-7910   Alex     Yangon        Member  Female
5    699-14-3026   Giza  Naypyitaw        Member  Female
7    315-22-5665   Giza  Naypyitaw        Member  Female
14   829-34-3910   Alex     Yangon        Member  Female
..           ...    ...        ...           ...     ...
988  267-62-7380   Giza  Naypyitaw        Member    Male
989  430-53-4718  Cairo   Mandalay        Member    Male
991  602-16-6955  Cairo   Mandalay        Normal  Female
996  303-96-2227  Cairo   Mandalay        Normal  Female
999  849-09-3807   Alex     Yangon        Member  Female

               Product line  Unit price  Quantity   Tax 5%      Sale
s  \
0         Health and beauty       74.69         7  26.1415   548.971
5
4         Sports and travel       86.31         7  30.2085   634.378
5
5    Electronic accessories       85.39         7  29.8865   627.616
5
7        Home and lifestyle       73.56        10  36.7800   772.380
0
14        Health and beauty       71.38        10  35.6900   749.490
0
..                      ...         ...       ...      ...
...
988  Electronic accessories       82.34        10  41.1700   864.570
0
989       Health and beauty       75.37         8  30.1480   633.108
0
991       Sports and travel       76.60        10  38.3000   804.300
```

```
0
996      Home and lifestyle       97.38        10  48.6900  1022.490
0
999     Fashion accessories       88.34         7  30.9190   649.299
0

          Date         Time      Payment    cogs  gross margin perce
ntage  \
0     1/5/2019   1:08:00 PM      Ewallet  522.83                 4.7
61905
4     2/8/2019  10:37:00 AM      Ewallet  604.17                 4.7
61905
5    3/25/2019   6:30:00 PM      Ewallet  597.73                 4.7
61905
7    2/24/2019  11:38:00 AM      Ewallet  735.60                 4.7
61905
14   3/29/2019   7:21:00 PM         Cash  713.80                 4.7
61905
..         ...          ...          ...     ...
...
988  3/29/2019   7:12:00 PM      Ewallet  823.40                 4.7
61905
989  1/28/2019   3:46:00 PM  Credit card  602.96                 4.7
61905
991  1/24/2019   6:10:00 PM      Ewallet  766.00                 4.7
61905
996   3/2/2019   5:16:00 PM      Ewallet  973.80                 4.7
61905
999  2/18/2019   1:28:00 PM         Cash  618.38                 4.7
61905

     gross income  Rating
0         26.1415     9.1
4         30.2085     5.3
5         29.8865     4.1
7         36.7800     8.0
14        35.6900     5.7
..            ...     ...
988       41.1700     4.3
989       30.1480     8.4
991       38.3000     6.0
996       48.6900     4.4
999       30.9190     6.6

[227 rows x 17 columns]
```

Q9. Filter sales in City = “Yangon” and Sales > 200.

In [113…

```
sales_yangon = market[(market['City'] == "Yangon") & (market['Sales
print(sales_yangon)
```

```
     Invoice ID Branch    City Customer type  Gender            Pro
duct line  \
```

```
0    750-67-8428   Alex  Yangon        Member  Female       Health a
nd beauty
2    631-41-3108   Alex  Yangon        Normal  Female      Home and
lifestyle
3    123-19-1176   Alex  Yangon        Member  Female       Health a
nd beauty
4    373-73-7910   Alex  Yangon        Member  Female       Sports a
nd travel
6    355-53-5943   Alex  Yangon        Member  Female  Electronic ac
cessories
..           ...    ...     ...           ...     ...
...
976  221-25-5073   Alex  Yangon        Normal  Female      Food and
beverages
981  809-46-1866   Alex  Yangon        Normal    Male       Health a
nd beauty
982  139-32-4183   Alex  Yangon        Member  Female       Sports a
nd travel
990  886-18-2897   Alex  Yangon        Normal  Female      Food and
beverages
999  849-09-3807   Alex  Yangon        Member  Female     Fashion ac
cessories

     Unit price  Quantity   Tax 5%     Sales       Date         Time
\
0         74.69         7  26.1415  548.9715   1/5/2019   1:08:00 PM
2         46.33         7  16.2155  340.5255   3/3/2019   1:23:00 PM
3         58.22         8  23.2880  489.0480  1/27/2019   8:33:00 PM
4         86.31         7  30.2085  634.3785   2/8/2019  10:37:00 AM
6         68.84         6  20.6520  433.6920  2/25/2019   2:36:00 PM
..          ...       ...      ...       ...        ...          ...
976       74.66         4  14.9320  313.5720   3/4/2019  10:39:00 AM
981       58.15         4  11.6300  244.2300  1/23/2019   5:44:00 PM
982       97.48         9  43.8660  921.1860  3/14/2019   2:19:00 PM
990       56.56         5  14.1400  296.9400  3/22/2019   7:06:00 PM
999       88.34         7  30.9190  649.2990  2/18/2019   1:28:00 PM

         Payment    cogs  gross margin percentage  gross income  Rat
ing
0        Ewallet  522.83                 4.761905       26.1415
9.1
2    Credit card  324.31                 4.761905       16.2155
7.4
3        Ewallet  465.76                 4.761905       23.2880
8.4
4        Ewallet  604.17                 4.761905       30.2085
5.3
6        Ewallet  413.04                 4.761905       20.6520
5.8
..           ...     ...                      ...           ...
...
976         Cash  298.64                 4.761905       14.9320
8.5
981         Cash  232.60                 4.761905       11.6300
8.4
982      Ewallet  877.32                 4.761905       43.8660
```

```
7.4
990  Credit card  282.80                4.761905       14.1400
4.5
999         Cash  618.38                4.761905       30.9190
6.6

[204 rows x 17 columns]
```

Q10. Sort all orders by Sales in descending order.

In [114…

```
market.sort_values(by = 'Sales', ascending=False)
```

Out [114…

	Invoice ID	Branch	City	Customer type	Gender	Product line	Unit price	Qua
350	860-79-0874	Giza	Naypyitaw	Member	Female	Fashion accessories	99.30	
167	687-47-8271	Alex	Yangon	Normal	Male	Fashion accessories	98.98	
557	283-26-5248	Giza	Naypyitaw	Member	Female	Food and beverages	98.52	
699	751-41-9720	Giza	Naypyitaw	Normal	Male	Home and lifestyle	97.50	
996	303-96-2227	Cairo	Mandalay	Normal	Female	Home and lifestyle	97.38	
...	...	...	...	...	...	...	...	
402	236-86-3015	Giza	Naypyitaw	Member	Male	Home and lifestyle	13.98	
443	192-98-7397	Giza	Naypyitaw	Normal	Male	Fashion accessories	12.78	
223	279-62-1445	Giza	Naypyitaw	Member	Female	Fashion accessories	12.54	
629	308-39-1707	Alex	Yangon	Normal	Female	Fashion accessories	12.09	
822	784-21-9238	Giza	Naypyitaw	Member	Male	Sports and travel	10.17	

1000 rows × 17 columns

Q11. Sort by Date (ascending) and then Time (ascending).

In [116…

```
print(market.sort_values(by = ['Date', 'Time'], ascending=[True, Tr
```

```
     Invoice ID Branch       City Customer type  Gender  \
17   765-26-6951   Alex     Yangon        Member  Female
970  746-04-1077  Cairo   Mandalay        Member  Female
839  271-77-8740   Giza  Naypyitaw        Member  Female
```

```
523  133-14-7229   Giza  Naypyitaw       Normal    Male
567  651-88-7328   Alex     Yangon       Normal  Female
..           ...    ...        ...          ...     ...
122  219-22-9386  Cairo   Mandalay       Member  Female
45   132-32-9879  Cairo   Mandalay       Member  Female
73   841-35-6630   Giza  Naypyitaw       Member  Female
234  157-13-5295   Alex     Yangon       Member    Male
326  815-11-1168   Alex     Yangon       Member    Male

               Product line  Unit price  Quantity  Tax 5%    Sales
Date  \
17        Sports and travel       72.61         6  21.783  457.443
1/1/2019
970      Food and beverages       84.63        10  42.315  888.615
1/1/2019
839       Sports and travel       29.22         6   8.766  184.086
1/1/2019
523       Health and beauty       62.87         2   6.287  132.027
1/1/2019
567     Fashion accessories       65.74         9  29.583  621.243
1/1/2019
..                      ...         ...       ...     ...      ...
...
122       Sports and travel       99.96         9  44.982  944.622
3/9/2019
45   Electronic accessories       93.96         4  18.792  394.632
3/9/2019
73   Electronic accessories       75.91         6  22.773  478.233
3/9/2019
234       Health and beauty       51.94        10  25.970  545.370
3/9/2019
326      Food and beverages       99.78         5  24.945  523.845
3/9/2019

            Time      Payment    cogs  gross margin percentage  gros
s income  \
17   10:39:00 AM  Credit card  435.66                4.761905
21.783
970  11:36:00 AM  Credit card  846.30                4.761905
42.315
839  11:40:00 AM      Ewallet  175.32                4.761905
8.766
523  11:43:00 AM         Cash  125.74                4.761905
6.287
567   1:55:00 PM         Cash  591.66                4.761905
29.583
..           ...          ...     ...                     ...
...
122   5:26:00 PM  Credit card  899.64                4.761905
44.982
45    6:00:00 PM         Cash  375.84                4.761905
18.792
73    6:21:00 PM         Cash  455.46                4.761905
22.773
234   6:24:00 PM      Ewallet  519.40                4.761905
25.970
```

```
326   7:09:00 PM         Cash  498.90                4.761905
24.945

     Rating
17      6.9
970     9.0
839     5.0
523     5.0
567     7.7
..      ...
122     4.2
45      9.5
73      8.7
234     6.5
326     5.4

[1000 rows x 17 columns]
```

Q12. Sort by Unit price and Quantity.

In [146…
```
print(market.sort_values(by=['Unit price', 'Quantity'], ascending=[
```

```
     Invoice ID Branch       City Customer type  Gender         Pro
duct line  \
944  333-23-2632   Alex     Yangon        Member    Male    Health a
nd beauty
572  239-48-4278   Alex     Yangon        Member    Male   Food and
beverages
784  516-77-6464   Giza  Naypyitaw        Member  Female    Health a
nd beauty
822  784-21-9238   Giza  Naypyitaw        Member    Male    Sports a
nd travel
881  115-38-7388   Giza  Naypyitaw        Member  Female  Fashion ac
cessories
..           ...    ...        ...           ...     ...
...
283  667-92-0055   Alex     Yangon        Member    Male    Health a
nd beauty
494  437-53-3084  Cairo   Mandalay        Normal    Male  Fashion ac
cessories
930  641-62-7288  Cairo   Mandalay        Normal    Male   Home and
lifestyle
983  148-41-7930   Giza  Naypyitaw        Normal    Male    Health a
nd beauty
122  219-22-9386  Cairo   Mandalay        Member  Female    Sports a
nd travel

     Unit price  Quantity   Tax 5%     Sales       Date         Time
\
944       10.08         7   3.5280   74.0880  3/28/2019   8:14:00 PM
572       10.13         7   3.5455   74.4555  3/10/2019   7:35:00 PM
784       10.16         5   2.5400   53.3400  2/24/2019   1:08:00 PM
822       10.17         1   0.5085   10.6785   2/7/2019   2:15:00 PM
881       10.18         8   4.0720   85.5120  3/30/2019  12:51:00 PM
```

```
..          ...       ...      ...       ...        ...          ...
283       99.83         6  29.9490  628.9290   3/4/2019   3:02:00 PM
494       99.89         2   9.9890  209.7690  2/26/2019  11:48:00 AM
930       99.92         6  29.9760  629.4960  3/24/2019   1:33:00 PM
983       99.96         7  34.9860  734.7060  1/23/2019  10:33:00 AM
122       99.96         9  44.9820  944.6220   3/9/2019   5:26:00 PM

          Payment    cogs  gross margin percentage  gross income  Rat
ing
944          Cash   70.56                 4.761905        3.5280
4.2
572       Ewallet   70.91                 4.761905        3.5455
8.3
784       Ewallet   50.80                 4.761905        2.5400
4.1
822          Cash   10.17                 4.761905        0.5085
5.9
881   Credit card   81.44                 4.761905        4.0720
9.5
..           ...     ...                      ...           ...
...
283       Ewallet  598.98                 4.761905       29.9490
8.5
494       Ewallet  199.78                 4.761905        9.9890
7.1
930       Ewallet  599.52                 4.761905       29.9760
7.1
983          Cash  699.72                 4.761905       34.9860
6.1
122   Credit card  899.64                 4.761905       44.9820
4.2

[1000 rows x 17 columns]
```

Q13. Calculate the total sales in “Sales” per Branch.

In [120…
```
market.groupby('Branch')['Sales'].sum()
```

Out[120…
```
Branch
Alex     106200.3705
Cairo    106197.6720
Giza     110568.7065
Name: Sales, dtype: float64
```

Q14. Calculate average Sales per City

In [121…
```
market.groupby('City')['Sales'].mean()
```

Out[121…
```
City
Mandalay     319.872506
Naypyitaw    337.099715
Yangon       312.354031
Name: Sales, dtype: float64
```

Q15. Find the quantities sold per product line.

In [126…
```
market.groupby('Product line')['Quantity'].sum()
```

Out[126…
```
Product line
Electronic accessories    971
Fashion accessories       902
Food and beverages        952
Health and beauty         854
Home and lifestyle        911
Sports and travel         920
Name: Quantity, dtype: int64
```

Q16. Calculate average gross income per Gender.

In [128…
```
market.groupby('Gender')['gross income'].mean()
```

Out[128…
```
Gender
Female    16.234829
Male      14.240749
Name: gross income, dtype: float64
```

Q17. Count number of sales per Payment method.

In [131…
```
market.groupby('Payment')['Sales'].count()
```

Out[131…
```
Payment
Cash           344
Credit card    311
Ewallet        345
Name: Sales, dtype: int64
```

Q18. Find maximum Sales per Branch.

In [133…
```
market.groupby('Branch')['Sales'].max()
```

Out[133…
```
Branch
Alex     1039.29
Cairo    1022.49
Giza     1042.65
Name: Sales, dtype: float64
```

Q19. Find minimum Unit price per Product line.

In [134…
```
market.groupby('Product line')['Unit price'].min()
```

Out[134…
```
Product line
Electronic accessories    10.56
Fashion accessories       10.18
Food and beverages        10.13
Health and beauty         10.08
Home and lifestyle        10.53
Sports and travel         10.17
Name: Unit price, dtype: float64
```

Q20. Find the sum of gross income per Product line and Branch.

In [136…
```
market.groupby(['Product line', 'Branch'])['gross income'].sum()
```

Out[136…
```
Product line            Branch
Electronic accessories  Alex       872.2435
                        Cairo      811.9735
                        Giza       903.2845
Fashion accessories     Alex       777.7385
                        Cairo      781.5865
                        Giza      1026.6700
Food and beverages      Alex       817.2905
                        Cairo      724.5185
                        Giza      1131.7550
Health and beauty       Alex       599.8930
                        Cairo      951.4600
                        Giza       791.2060
Home and lifestyle      Alex      1067.4855
                        Cairo      835.6745
                        Giza       661.6930
Sports and travel       Alex       922.5095
                        Cairo      951.8190
                        Giza       750.5680
Name: gross income, dtype: float64
```

21. What is the total quantities sold in Product line: “Electronic accessories”?

In [140…
```
market[market['Product line'] == 'Electronic accessories']['Quantit
```

Out[140…
```
971
```

22. What is the average Sales for female customers?

In [142…
```
market[market['Gender'] == 'Female']['Sales'].mean()
```

Out[142…
```
340.9314141856392
```

23. What is the most expensive Unit price among Customer type members only?

In [145…
```
market[market['Customer type'] == 'Member']['Unit price'].max()
```

Out[145…
```
99.96
```

24. How many orders with Rating >= 9 ?

In [150…
```
market[market['Rating'] >= 9]['Invoice ID'].count()
```

Out[150…
```
166
```

25. What is the total Sales for Payment “Credit card” in Branch C?

In [152…
```
market[(market['Payment'] == 'Credit card') & (market['Branch'] ==
```

Out[152…
```
37344.8565
```

In [ ]:
```

```