

lab-1

April 13, 2024

```
[19]: import pandas as pd
import numpy as np
```

```
[20]: df = pd.read_csv("supermarket_sales - Sheet1.csv")
```

```
[21]: df.head()
```

```
[21]: Invoice ID Branch      City Customer type Gender \
0  750-67-8428      A      Yangon      Member  Female
1  226-31-3081      C  Naypyitaw      Normal  Female
2  631-41-3108      A      Yangon      Normal   Male
3  123-19-1176      A      Yangon      Member   Male
4  373-73-7910      A      Yangon      Normal   Male
```

```
Product line Unit price Quantity Tax 5% Total Date \
0 Health and beauty      74.69         7 26.1415 548.9715 1/5/2019
1 Electronic accessories    15.28         5  3.8200  80.2200 3/8/2019
2 Home and lifestyle       46.33         7 16.2155 340.5255 3/3/2019
3 Health and beauty       58.22         8 23.2880 489.0480 1/27/2019
4 Sports and travel       86.31         7 30.2085 634.3785 2/8/2019
```

```
Time Payment cogs gross margin percentage gross income Rating
0 13:08 Ewallet 522.83          4.761905      26.1415      9.1
1 10:29 Cash    76.40          4.761905       3.8200      9.6
2 13:23 Credit card 324.31          4.761905      16.2155      7.4
3 20:33 Ewallet 465.76          4.761905      23.2880      8.4
4 10:37 Ewallet 604.17          4.761905      30.2085      5.3
```

```
[22]: df.head(10)
```

```
[22]: Invoice ID Branch      City Customer type Gender \
0  750-67-8428      A      Yangon      Member  Female
1  226-31-3081      C  Naypyitaw      Normal  Female
2  631-41-3108      A      Yangon      Normal   Male
3  123-19-1176      A      Yangon      Member   Male
4  373-73-7910      A      Yangon      Normal   Male
5  699-14-3026      C  Naypyitaw      Normal   Male
```

| | | | | | |
|---|-------------|---|-----------|--------|--------|
| 6 | 355-53-5943 | A | Yangon | Member | Female |
| 7 | 315-22-5665 | C | Naypyitaw | Normal | Female |
| 8 | 665-32-9167 | A | Yangon | Member | Female |
| 9 | 692-92-5582 | B | Mandalay | Member | Female |

| | Product line | Unit price | Quantity | Tax 5% | Total | Date \ |
|---|------------------------|------------|----------|---------|----------|-----------|
| 0 | Health and beauty | 74.69 | 7 | 26.1415 | 548.9715 | 1/5/2019 |
| 1 | Electronic accessories | 15.28 | 5 | 3.8200 | 80.2200 | 3/8/2019 |
| 2 | Home and lifestyle | 46.33 | 7 | 16.2155 | 340.5255 | 3/3/2019 |
| 3 | Health and beauty | 58.22 | 8 | 23.2880 | 489.0480 | 1/27/2019 |
| 4 | Sports and travel | 86.31 | 7 | 30.2085 | 634.3785 | 2/8/2019 |
| 5 | Electronic accessories | 85.39 | 7 | 29.8865 | 627.6165 | 3/25/2019 |
| 6 | Electronic accessories | 68.84 | 6 | 20.6520 | 433.6920 | 2/25/2019 |
| 7 | Home and lifestyle | 73.56 | 10 | 36.7800 | 772.3800 | 2/24/2019 |
| 8 | Health and beauty | 36.26 | 2 | 3.6260 | 76.1460 | 1/10/2019 |
| 9 | Food and beverages | 54.84 | 3 | 8.2260 | 172.7460 | 2/20/2019 |

| | Time | Payment | cogs | gross margin percentage | gross income | Rating |
|---|-------|-------------|--------|-------------------------|--------------|--------|
| 0 | 13:08 | Ewallet | 522.83 | 4.761905 | 26.1415 | 9.1 |
| 1 | 10:29 | Cash | 76.40 | 4.761905 | 3.8200 | 9.6 |
| 2 | 13:23 | Credit card | 324.31 | 4.761905 | 16.2155 | 7.4 |
| 3 | 20:33 | Ewallet | 465.76 | 4.761905 | 23.2880 | 8.4 |
| 4 | 10:37 | Ewallet | 604.17 | 4.761905 | 30.2085 | 5.3 |
| 5 | 18:30 | Ewallet | 597.73 | 4.761905 | 29.8865 | 4.1 |
| 6 | 14:36 | Ewallet | 413.04 | 4.761905 | 20.6520 | 5.8 |
| 7 | 11:38 | Ewallet | 735.60 | 4.761905 | 36.7800 | 8.0 |
| 8 | 17:15 | Credit card | 72.52 | 4.761905 | 3.6260 | 7.2 |
| 9 | 13:27 | Credit card | 164.52 | 4.761905 | 8.2260 | 5.9 |

[23]: df.tail()

[23]:

| | Invoice ID | Branch | City | Customer type | Gender | Product line \ |
|-----|-------------|--------|-----------|---------------|--------|---------------------|
| 995 | 233-67-5758 | C | Naypyitaw | Normal | Male | Health and beauty |
| 996 | 303-96-2227 | B | Mandalay | Normal | Female | Home and lifestyle |
| 997 | 727-02-1313 | A | Yangon | Member | Male | Food and beverages |
| 998 | 347-56-2442 | A | Yangon | Normal | Male | Home and lifestyle |
| 999 | 849-09-3807 | A | Yangon | Member | Female | Fashion accessories |

| | Unit price | Quantity | Tax 5% | Total | Date | Time | Payment \ |
|-----|------------|----------|---------|-----------|-----------|-------|-----------|
| 995 | 40.35 | 1 | 2.0175 | 42.3675 | 1/29/2019 | 13:46 | Ewallet |
| 996 | 97.38 | 10 | 48.6900 | 1022.4900 | 3/2/2019 | 17:16 | Ewallet |
| 997 | 31.84 | 1 | 1.5920 | 33.4320 | 2/9/2019 | 13:22 | Cash |
| 998 | 65.82 | 1 | 3.2910 | 69.1110 | 2/22/2019 | 15:33 | Cash |
| 999 | 88.34 | 7 | 30.9190 | 649.2990 | 2/18/2019 | 13:28 | Cash |

| | cogs | gross margin percentage | gross income | Rating |
|-----|-------|-------------------------|--------------|--------|
| 995 | 40.35 | 4.761905 | 2.0175 | 6.2 |

| | | | | |
|-----|--------|----------|---------|-----|
| 996 | 973.80 | 4.761905 | 48.6900 | 4.4 |
| 997 | 31.84 | 4.761905 | 1.5920 | 7.7 |
| 998 | 65.82 | 4.761905 | 3.2910 | 4.1 |
| 999 | 618.38 | 4.761905 | 30.9190 | 6.6 |

[24]: df.tail(10)

[24]:

| | Invoice ID | Branch | City | Customer type | Gender | \ |
|-----|-------------|--------|-----------|---------------|--------|---|
| 990 | 886-18-2897 | A | Yangon | Normal | Female | |
| 991 | 602-16-6955 | B | Mandalay | Normal | Female | |
| 992 | 745-74-0715 | A | Yangon | Normal | Male | |
| 993 | 690-01-6631 | B | Mandalay | Normal | Male | |
| 994 | 652-49-6720 | C | Naypyitaw | Member | Female | |
| 995 | 233-67-5758 | C | Naypyitaw | Normal | Male | |
| 996 | 303-96-2227 | B | Mandalay | Normal | Female | |
| 997 | 727-02-1313 | A | Yangon | Member | Male | |
| 998 | 347-56-2442 | A | Yangon | Normal | Male | |
| 999 | 849-09-3807 | A | Yangon | Member | Female | |

| | Product line | Unit price | Quantity | Tax 5% | Total | \ |
|-----|------------------------|------------|----------|---------|-----------|---|
| 990 | Food and beverages | 56.56 | 5 | 14.1400 | 296.9400 | |
| 991 | Sports and travel | 76.60 | 10 | 38.3000 | 804.3000 | |
| 992 | Electronic accessories | 58.03 | 2 | 5.8030 | 121.8630 | |
| 993 | Fashion accessories | 17.49 | 10 | 8.7450 | 183.6450 | |
| 994 | Electronic accessories | 60.95 | 1 | 3.0475 | 63.9975 | |
| 995 | Health and beauty | 40.35 | 1 | 2.0175 | 42.3675 | |
| 996 | Home and lifestyle | 97.38 | 10 | 48.6900 | 1022.4900 | |
| 997 | Food and beverages | 31.84 | 1 | 1.5920 | 33.4320 | |
| 998 | Home and lifestyle | 65.82 | 1 | 3.2910 | 69.1110 | |
| 999 | Fashion accessories | 88.34 | 7 | 30.9190 | 649.2990 | |

| | Date | Time | Payment | cogs | gross margin percentage | \ |
|-----|-----------|-------|-------------|--------|-------------------------|---|
| 990 | 3/22/2019 | 19:06 | Credit card | 282.80 | 4.761905 | |
| 991 | 1/24/2019 | 18:10 | Ewallet | 766.00 | 4.761905 | |
| 992 | 3/10/2019 | 20:46 | Ewallet | 116.06 | 4.761905 | |
| 993 | 2/22/2019 | 18:35 | Ewallet | 174.90 | 4.761905 | |
| 994 | 2/18/2019 | 11:40 | Ewallet | 60.95 | 4.761905 | |
| 995 | 1/29/2019 | 13:46 | Ewallet | 40.35 | 4.761905 | |
| 996 | 3/2/2019 | 17:16 | Ewallet | 973.80 | 4.761905 | |
| 997 | 2/9/2019 | 13:22 | Cash | 31.84 | 4.761905 | |
| 998 | 2/22/2019 | 15:33 | Cash | 65.82 | 4.761905 | |
| 999 | 2/18/2019 | 13:28 | Cash | 618.38 | 4.761905 | |

| | gross income | Rating |
|-----|--------------|--------|
| 990 | 14.1400 | 4.5 |
| 991 | 38.3000 | 6.0 |
| 992 | 5.8030 | 8.8 |

```

993      8.7450      6.6
994      3.0475      5.9
995      2.0175      6.2
996     48.6900      4.4
997      1.5920      7.7
998      3.2910      4.1
999     30.9190      6.6

```

```
[25]: df.index
```

```
[25]: RangeIndex(start=0, stop=1000, step=1)
```

```
[26]: df.info
```

```
[26]: <bound method DataFrame.info of Invoice ID Branch City Customer type
Gender \
```

```

0    750-67-8428      A    Yangon      Member  Female
1    226-31-3081      C  Naypyitaw      Normal  Female
2    631-41-3108      A    Yangon      Normal   Male
3    123-19-1176      A    Yangon      Member   Male
4    373-73-7910      A    Yangon      Normal   Male
..      ...      ...      ...      ...      ...
995  233-67-5758      C  Naypyitaw      Normal   Male
996  303-96-2227      B  Mandalay      Normal  Female
997  727-02-1313      A    Yangon      Member   Male
998  347-56-2442      A    Yangon      Normal   Male
999  849-09-3807      A    Yangon      Member  Female

```

```

Product line  Unit price  Quantity  Tax 5%  Total \
0    Health and beauty    74.69         7  26.1415  548.9715
1  Electronic accessories    15.28         5   3.8200   80.2200
2    Home and lifestyle    46.33         7  16.2155  340.5255
3    Health and beauty    58.22         8  23.2880  489.0480
4    Sports and travel     86.31         7  30.2085  634.3785
..      ...      ...      ...      ...      ...
995  Health and beauty    40.35         1   2.0175   42.3675
996  Home and lifestyle    97.38        10  48.6900  1022.4900
997  Food and beverages    31.84         1   1.5920   33.4320
998  Home and lifestyle    65.82         1   3.2910   69.1110
999  Fashion accessories    88.34         7  30.9190   649.2990

```

```

Date    Time    Payment  cogs  gross margin percentage \
0  1/5/2019  13:08    Ewallet  522.83      4.761905
1  3/8/2019  10:29      Cash   76.40      4.761905
2  3/3/2019  13:23  Credit card  324.31      4.761905
3  1/27/2019  20:33    Ewallet  465.76      4.761905
4  2/8/2019  10:37    Ewallet  604.17      4.761905

```

```

..      ...      ...      ...      ...
995  1/29/2019  13:46      Ewallet  40.35      4.761905
996   3/2/2019  17:16      Ewallet  973.80      4.761905
997   2/9/2019  13:22      Cash    31.84      4.761905
998  2/22/2019  15:33      Cash    65.82      4.761905
999  2/18/2019  13:28      Cash   618.38      4.761905

```

```

      gross income  Rating
0          26.1415    9.1
1           3.8200    9.6
2          16.2155    7.4
3          23.2880    8.4
4          30.2085    5.3
..      ...      ...
995          2.0175    6.2
996         48.6900    4.4
997          1.5920    7.7
998          3.2910    4.1
999         30.9190    6.6

```

```
[1000 rows x 17 columns]>
```

```
[27]: df.columns
```

```
[27]: Index(['Invoice ID', 'Branch', 'City', 'Customer type', 'Gender',
          'Product line', 'Unit price', 'Quantity', 'Tax 5%', 'Total', 'Date',
          'Time', 'Payment', 'cogs', 'gross margin percentage', 'gross income',
          'Rating'],
          dtype='object')
```

```
[28]: df.shape
```

```
[28]: (1000, 17)
```

```
[29]: df.dtypes
```

```
[29]: Invoice ID      object
      Branch      object
      City        object
      Customer type object
      Gender      object
      Product line  object
      Unit price    float64
      Quantity      int64
      Tax 5%        float64
      Total         float64
      Date          object
```

```

Time                object
Payment             object
cogs                float64
gross margin percentage float64
gross income        float64
Rating              float64
dtype: object

```

```
[30]: df.columns.values
```

```
[30]: array(['Invoice ID', 'Branch', 'City', 'Customer type', 'Gender',
        'Product line', 'Unit price', 'Quantity', 'Tax 5%', 'Total',
        'Date', 'Time', 'Payment', 'cogs', 'gross margin percentage',
        'gross income', 'Rating'], dtype=object)
```

```
[31]: df.describe()
```

```
[31]:
```

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

| | gross margin percentage | gross income | Rating |
|-------|-------------------------|--------------|-------------|
| count | 1.000000e+03 | 1000.000000 | 1000.000000 |
| mean | 4.761905e+00 | 15.379369 | 6.97270 |
| std | 6.131498e-14 | 11.708825 | 1.71858 |
| min | 4.761905e+00 | 0.508500 | 4.00000 |
| 25% | 4.761905e+00 | 5.924875 | 5.50000 |
| 50% | 4.761905e+00 | 12.088000 | 7.00000 |
| 75% | 4.761905e+00 | 22.445250 | 8.50000 |
| max | 4.761905e+00 | 49.650000 | 10.00000 |

```
[32]: df.iloc[0:3]
```

```
[32]:
```

| | Invoice ID | Branch | City | Customer type | Gender \ |
|---|-------------|--------|-----------|---------------|----------|
| 0 | 750-67-8428 | A | Yangon | Member | Female |
| 1 | 226-31-3081 | C | Naypyitaw | Normal | Female |
| 2 | 631-41-3108 | A | Yangon | Normal | Male |

| | Product line | Unit price | Quantity | Tax 5% | Total | Date \ |
|---|------------------------|------------|----------|---------|----------|----------|
| 0 | Health and beauty | 74.69 | 7 | 26.1415 | 548.9715 | 1/5/2019 |
| 1 | Electronic accessories | 15.28 | 5 | 3.8200 | 80.2200 | 3/8/2019 |

| | | | | | | |
|---|--------------------|-------|---|---------|----------|----------|
| 2 | Home and lifestyle | 46.33 | 7 | 16.2155 | 340.5255 | 3/3/2019 |
|---|--------------------|-------|---|---------|----------|----------|

| | Time | Payment | cogs | gross margin | percentage | gross income | Rating |
|---|-------|-------------|--------|--------------|------------|--------------|--------|
| 0 | 13:08 | Ewallet | 522.83 | | 4.761905 | 26.1415 | 9.1 |
| 1 | 10:29 | Cash | 76.40 | | 4.761905 | 3.8200 | 9.6 |
| 2 | 13:23 | Credit card | 324.31 | | 4.761905 | 16.2155 | 7.4 |

```
[33]: df.iloc[0:3,0:3]
```

```
[33]: Invoice ID Branch      City
0  750-67-8428      A      Yangon
1  226-31-3081      C  Naypyitaw
2  631-41-3108      A      Yangon
```