```
In [3]: import pandas as pd
In [4]: | df= pd.read_csv("supermarket_sales - Sheet1.csv")
In [5]:
         df.head()
Out[5]:
             Invoice
                                                             Product
                                        Customer
                                                                       Unit
                                   City
                      Branch
                                                  Gender
                                                                             Quantity
                                                                                      Tax 5%
                  ID
                                             type
                                                                 line
                                                                      price
             750-67-
                                                            Health and
          0
                           Α
                                Yangon
                                          Member
                                                   Female
                                                                      74.69
                                                                                   7 26.1415 548.
                8428
                                                               beauty
             226-31-
                                                            Electronic
                              Naypyitaw
                                           Normal
                                                   Female
                                                                      15.28
                                                                                       3.8200
                                                                                               80.
                3081
                                                           accessories
             631-41-
                                                            Home and
                                                                      46.33
                                                                                   7 16.2155 340.
                                Yangon
                                           Normal
                                                     Male
                3108
                                                              lifestyle
                                                            Health and
              123-19-
           3
                           Α
                                Yangon
                                          Member
                                                     Male
                                                                      58.22
                                                                                   8 23.2880
                                                                                              489.
                                                               beauty
                1176
             373-73-
                                                            Sports and
                                Yangon
                                           Normal
                                                     Male
                                                                      86.31
                                                                                   7 30.2085 634.
                7910
                                                                travel
In [6]:
         df.shape
Out[6]: (1000, 17)
In [7]:
         df.isnull().sum()
Out[7]: Invoice ID
                                          0
          Branch
                                          0
         City
                                          0
          Customer type
                                          0
          Gender
                                          0
          Product line
                                         0
                                         0
         Unit price
          Quantity
                                         0
          Tax 5%
                                          0
          Total
                                         0
         Date
                                          0
                                          0
          Time
          Payment
                                          0
                                         0
          cogs
                                         0
          gross margin percentage
          gross income
                                         0
                                          0
          Rating
```

dtype: int64

```
In [8]:
         df.dtypes
 Out[8]: Invoice ID
                                       object
          Branch
                                       object
          City
                                       object
          Customer type
                                       object
                                       object
          Gender
          Product line
                                       object
          Unit price
                                      float64
          Quantity
                                        int64
          Tax 5%
                                      float64
          Total
                                      float64
          Date
                                       object
          Time
                                       object
          Payment
                                       object
          cogs
                                      float64
                                      float64
          gross margin percentage
          gross income
                                      float64
                                      float64
          Rating
          dtype: object
 In [9]: |df['Quantity'] = df['Quantity'].astype(object)
In [10]: df.dtypes
Out[10]: Invoice ID
                                       object
          Branch
                                       object
                                       object
          City
          Customer type
                                       object
          Gender
                                       object
          Product line
                                       object
          Unit price
                                      float64
                                       object
          Quantity
          Tax 5%
                                      float64
          Total
                                      float64
          Date
                                       object
          Time
                                       object
          Payment
                                       object
                                      float64
          cogs
                                      float64
          gross margin percentage
          gross income
                                      float64
                                      float64
          Rating
          dtype: object
In [11]: |df['Gender'].value_counts()
Out[11]: Gender
          Female
                    501
                    499
          Male
          Name: count, dtype: int64
In [12]: | df['Gender'] = df['Gender'].map({'Male':0, 'Female':1})
```

In [13]: df.head()

Out[13]:

	Invoice ID	Branch	City	Customer type	Gender	Product line	Unit price	Quantity	Tax 5%	
0	750-67- 8428	А	Yangon	Member	1	Health and beauty	74.69	7	26.1415	548.
1	226-31- 3081	С	Naypyitaw	Normal	1	Electronic accessories	15.28	5	3.8200	80.
2	631-41- 3108	Α	Yangon	Normal	0	Home and lifestyle	46.33	7	16.2155	340.
3	123-19- 1176	Α	Yangon	Member	0	Health and beauty	58.22	8	23.2880	489.
4	373-73- 7910	Α	Yangon	Normal	0	Sports and travel	86.31	7	30.2085	634.
4										•

In []: