```
import numpy as np
import pandas as pd
import matplotlib as mpl
import matplotlib.pyplot as plt
import seaborn as sns
import sklearn
import sklearn.preprocessing as preprocessing
import sklearn.model_selection as model_selection
import sklearn.ensemble as ensemble
import sklearn.metrics as metrics
import sklearn.tree as tree
import sklearn.linear_model as linear_model
import sklearn.neighbors as neighbors
import sklearn.svm as svm
import sklearn.naive_bayes as naive_bayes
import sklearn.decomposition as decomposition
import sklearn.feature_selection as feature_selection
import sklearn.cluster as cluster
import sklearn.manifold as manifold
import sklearn.mixture as mixture
import sklearn.gaussian_process as gaussian_process
import sklearn.neural_network as neural_network
import sklearn.pipeline as pipeline
import sklearn.model_selection as model_selection
import sklearn.metrics as metrics
import sklearn.preprocessing as preprocessing
import sklearn.ensemble as ensemble
import sklearn.tree as tree
import sklearn.linear_model as linear_model
import sklearn.svm as svm
import sklearn.naive_bayes as naive_bayes
import sklearn.neighbors as neighbors
import sklearn.decomposition as decomposition
import sklearn.cluster as cluster
import sklearn.manifold as manifold
import sklearn.mixture as mixture
import sklearn.gaussian_process as gaussian_process
import sklearn.neural_network as neural_network
import sklearn.pipeline as pipeline
import sklearn.feature_selection as feature_selection
```

```
education  marital status  income  total bill  total visits  total purchases  age  tenure  spent
0  Graduate   Single   60182.0   1.0   1.0   1.0   26   1.0   1.0
1  Graduate   Single   82754.0   0.0   0.0   0.0   26   0.0   0.0
2  Graduate   Married  19514.0   2.0   1.0   2.0   94   1.0   2.0
3  Graduate   Married  62513.0   1.0   1.0   1.0   16   1.0   1.0
4  Graduate   Married  55635.0   1.0   1.0   1.0   34   1.0   1.0
5  Graduate   Married  33929.0   2.0   1.0   2.0   32   1.0   2.0
6  Graduate   Married  62187.0   1.0   1.0   1.0   19   1.0   1.0
7  Graduate   Married  30096.0   2.0   1.0   2.0   74   1.0   2.0
8  Graduate   Married  48432.0   1.0   1.0   1.0   79   1.0   1.0
9  Graduate   Single   18744.0   2.0   1.0   2.0   65   1.0   2.0
10  Graduate   Married  34421.0   2.0   1.0   2.0   81   1.0   2.0
11  Graduate   Married  62499.0   1.0   1.0   1.0   79   1.0   1.0
12  Graduate   Married  37284.0   2.0   1.0   2.0   78   1.0   2.0
13  Graduate   Married  35860.0   2.0   1.0   2.0   72   1.0   2.0
14  Graduate   Married  35860.0   2.0   1.0   2.0   72   1.0   2.0
15  Graduate   Married  35860.0   2.0   1.0   2.0   72   1.0   2.0
16  Graduate   Married  35860.0   2.0   1.0   2.0   72   1.0   2.0
```

```
Gender                  0
Customer type           0
Age                     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
City                    0
Branch                  0
dtype: int64
```

```
df[['Tax 5%', 'cogs', 'gross income', 'Sales']].sample(5)
```

	Tax 5%	cogs	gross income	Sales
293	2.0895	39.79	41.7895	831.579
7	16.1865	308.57	324.7565	6457.413
338	1.5855	30.21	31.7955	632.109
62	3.7905	72.25	76.0405	1511.691
43	2.4255	46.22	48.6455	967.275

FINAL ANALYSIS

```
plotnine.0.10.1+gg.0.5.5.9dev.patchwork.from_list('total_gross', 'Rating', 'Payment', 'Gender')
```

```
fig,axs = plt.subplots(2,2,figsize=(12,8))
sns.boxplot(data=df,x='Gender',y='Rating',ax=axs[0,0])
sns.boxplot(data=df,x='Payment',y='Rating',ax=axs[0,1])
sns.boxplot(data=df,x='Gender',y='total_gross',ax=axs[1,0])
sns.boxplot(data=df,x='Payment',y='total_gross',ax=axs[1,1])
```

```
sns.boxplot(data=df,x='Payment',y='Rating',hue='Gender',palette='Set1')
```

```
fig,axs = plt.subplots(2,2,figsize=(12,8))
sns.boxplot(data=df,x='Gender',y='Rating',hue='Payment',palette='Set1',ax=axs[0,0])
sns.boxplot(data=df,x='Payment',y='Rating',hue='Gender',palette='Set1',ax=axs[0,1])
sns.boxplot(data=df,x='Gender',y='total_gross',hue='Payment',palette='Set1',ax=axs[1,0])
sns.boxplot(data=df,x='Payment',y='total_gross',hue='Gender',palette='Set1',ax=axs[1,1])
```

```
sns.scatplot(data=df,x='Income',y='Spending',hue='Gender',palette='Set1')
```

```
sns.scatplot(data=df,x='Age',y='Spending',hue='Gender',palette='Set1')
```

```
sns.scatplot(data=df,x='Income',y='Age',hue='Gender',palette='Set1')
```



```
sns.scatplot(data=df,x='Spending',y='Rating',hue='Gender',palette='Set1')
```



```
sns.scatplot(data=df,x='Spending',y='Rating',hue='Payment',palette='Set1')
```



```
sns.scatplot(data=df,x='Spending',y='Rating',hue='Gender',palette='Set1')
```



```
sns.scatplot(data=df,x='Income',y='Rating',hue='Gender',palette='Set1')
```



```
sns.scatplot(data=df,x='Age',y='Rating',hue='Gender',palette='Set1')
```



```
sns.scatplot(data=df,x='Income',y='Spending',hue='Gender',palette='Set1')
```



```
sns.scatplot(data=df,x='Age',y='Spending',hue='Gender',palette='Set1')
```


Spending
Age