

```
[1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
df=pd.read_csv("C:\Users\ELC07\Desktop\supermarket.csv")
df

In [1]: Invoice ID Branch City Customer type Gender Product line Unit price Quantity Tax 5% Total Date Time Payment coogs gross margin percentage gross income Rating
0 74.69 7.0 26.1415 548.9715 522.83 4.761905 26.1415 9.1
1 15.28 5.0 3.8200 80.2200 76.40 4.761905 3.8200 9.6
2 46.33 7.0 16.2155 340.5265 324.31 4.761905 16.2155 7.4
3 58.22 8.0 23.2880 489.0480 465.76 4.761905 23.2880 8.4
4 86.31 7.0 30.2085 634.3785 604.17 4.761905 30.2085 5.3
... ..
995 40.35 1.0 2.0175 42.3675 40.35 4.761905 2.0175 6.2
996 97.38 10.0 48.6900 1022.4900 973.80 4.761905 48.6900 4.4
997 31.84 1.0 1.5920 33.4320 31.84 4.761905 1.5920 7.7
998 65.82 1.0 3.2910 69.1110 65.82 4.761905 3.2910 4.1
999 88.34 7.0 30.9190 649.2990 618.38 4.761905 30.9190 6.6
1000 rows x 17 columns

In [4]: df["branch"].value_counts()
Out[2]: Branch
A 340
B 132
C 328
Name: count, dtype: int64

In [3]: sns.countplot(x=df["Payment"])
plt.title("Count Vs Method of Payment")
plt.xlabel("Method of Payment")
plt.ylabel("Count")
Out[3]: Text(0, 0.5, 'Count')

In [4]: plt.figure(figsize=(14,6))
plt.style.use("classic")
ax = sns.countplot(x="Payment", hue="Branch", data=df, palette="tab20")
plt.title("Payment distribution in all branches")
ax.set_xlabel(xlabel = "Payment Method", fontsize=5)
ax.set_ylabel(ylabel = "People count", fontsize=16)
Out[4]: Text(0, 0.5, 'People count')

In [5]: plt.figure(figsize=(14,6))
plt.title("Payment distribution in all branches")
ax = sns.countplot(x="Payment", hue="Branch", data=df, palette="tab20")
plt.title("Payment distribution in all branches")
ax.set_xlabel(xlabel = "Payment Method", fontsize=5)
ax.set_ylabel(ylabel = "People count", fontsize=16)
Out[4]: Text(0, 0.5, 'People count')

In [5]: sns.set_style("darkgrid")
sns.scatterplot(x=df["Rating"], y=df["gross income"])
plt.title("Gross income Vs Rating")
Out[5]: Text(0.5, 1.0, 'Gross income Vs Rating')

In [6]: plt.figure(figsize=(8,4))
ax = sns.boxplot(x="branch", y="Rating", data=df, palette="magma")
ax.set_title("Rating distribution between branches", fontsize=25)
ax.set_xlabel(xlabel = "Branches", fontsize=16)
ax.set_ylabel(ylabel = "Rating distribution", fontsize=16)
plt.grid()
Out[6]: Text(0.5, 1.0, 'Rating distribution between branches')

In [7]: sns.boxplot(x=df["Branch"], y=df["gross income"])
plt.title("gross income Vs Branch")
Out[7]: Text(0.5, 1.0, 'gross income Vs Branch')

In [8]: sns.boxplot(x=df["Gender"], y=df["gross income"])
plt.title("gross income Vs gender")
Out[8]: Text(0.5, 1.0, 'gross income Vs gender')

In [9]: sns.boxplot(x=df["Gender"], y=df["gross income"])
plt.title("gross income Vs gender")
Out[9]: Text(0.5, 1.0, 'gross income Vs gender')

In [10]: df.groupby(df.index).mean(numeric_only=True)
Out[10]:
Unit price Quantity Tax 5% Total coogs gross margin percentage gross income Rating
0 74.69 7.0 26.1415 548.9715 522.83 4.761905 26.1415 9.1
1 15.28 5.0 3.8200 80.2200 76.40 4.761905 3.8200 9.6
2 46.33 7.0 16.2155 340.5265 324.31 4.761905 16.2155 7.4
3 58.22 8.0 23.2880 489.0480 465.76 4.761905 23.2880 8.4
4 86.31 7.0 30.2085 634.3785 604.17 4.761905 30.2085 5.3
... ..
995 40.35 1.0 2.0175 42.3675 40.35 4.761905 2.0175 6.2
996 97.38 10.0 48.6900 1022.4900 973.80 4.761905 48.6900 4.4
997 31.84 1.0 1.5920 33.4320 31.84 4.761905 1.5920 7.7
998 65.82 1.0 3.2910 69.1110 65.82 4.761905 3.2910 4.1
999 88.34 7.0 30.9190 649.2990 618.38 4.761905 30.9190 6.6
1000 rows x 8 columns

In [16]: cat=df[["Product line","gross income"]].groupby(["Product line",as_index=False).sum().sort_values(yo="gross income",ascending=False)
plt.figure(figsize=(28,8))
sns.barplot(x="Product line", y="gross income", data=cat)
plt.title("gross income Vs Product line")
Out[16]: Text(0.5, 1.0, 'gross income Vs Product line')

In [17]: plt.figure(figsize=(16,6))
plt.title("Total Monthly transaction by Gender")
sns.countplot(x=df["Product line"], hue=df["Gender"])
Out[17]: 
```