

```
▶ import pandas as pd  
df=pd.read_csv("/content/Mall_Customers.csv")
```

```
▶ df.head()
```

CustomerID Gender Age Annual Income (k\$) Spending Score (1-100)

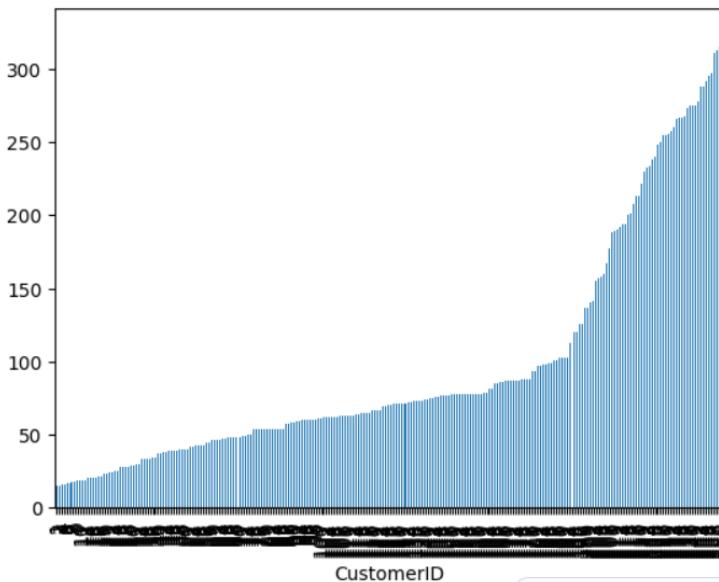
	CustomerID	Gender	Age	Annual Income (k\$)	Spending Score (1-100)
0	1	Male	19	15	39
1	2	Male	21	15	81
2	3	Female	20	16	6
3	4	Female	23	16	77
4	5	Female	31	17	40

Next steps: [Generate code with df](#) [View recommended plots](#) [New interactive sheet](#)

```
[4] import matplotlib.pyplot as plt
```

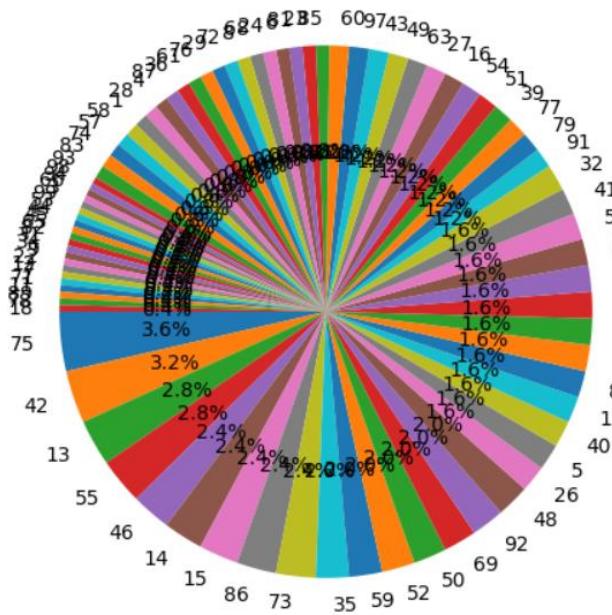
```
▶ plt.figure(figsize=(8,5))  
df.dropna(subset=["CustomerID"]).plot(kind="bar",x="CustomerID",y="Annual Income (k$)",legend=False)  
plt.title("Mall Customers")  
plt.ylabel("Annual Income")  
plt.xticks(rotation=90)  
plt.show()
```

<Axes: xlabel='CustomerID'>



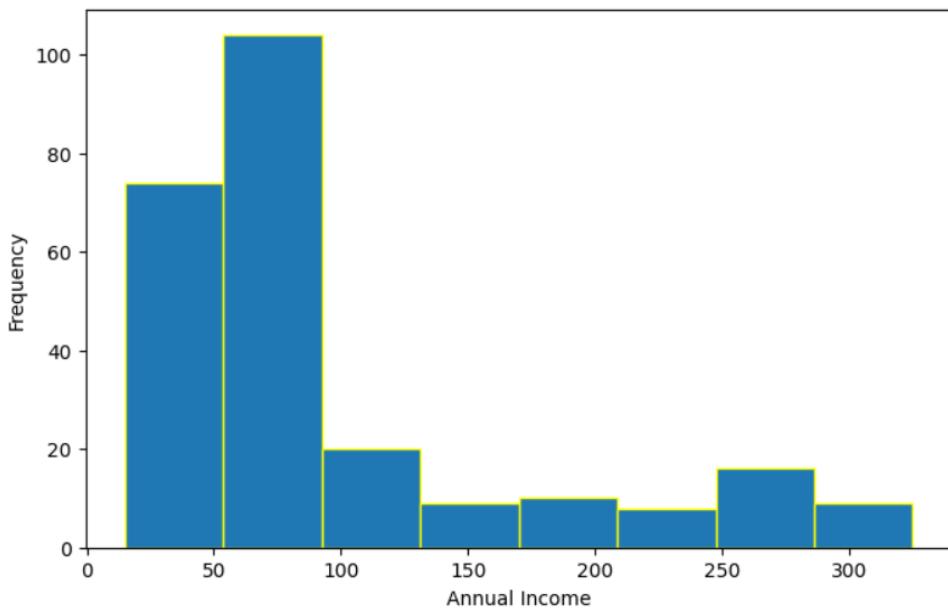
```
▶ plt.figure(figsize=(6,6))
df["Spending Score (1-100)"].value_counts().plot(kind="pie", autopct="%1.1f%%", startangle=180)
plt.title("Spending Score")
plt.ylabel("")
plt.show()
```

→ Spending Score



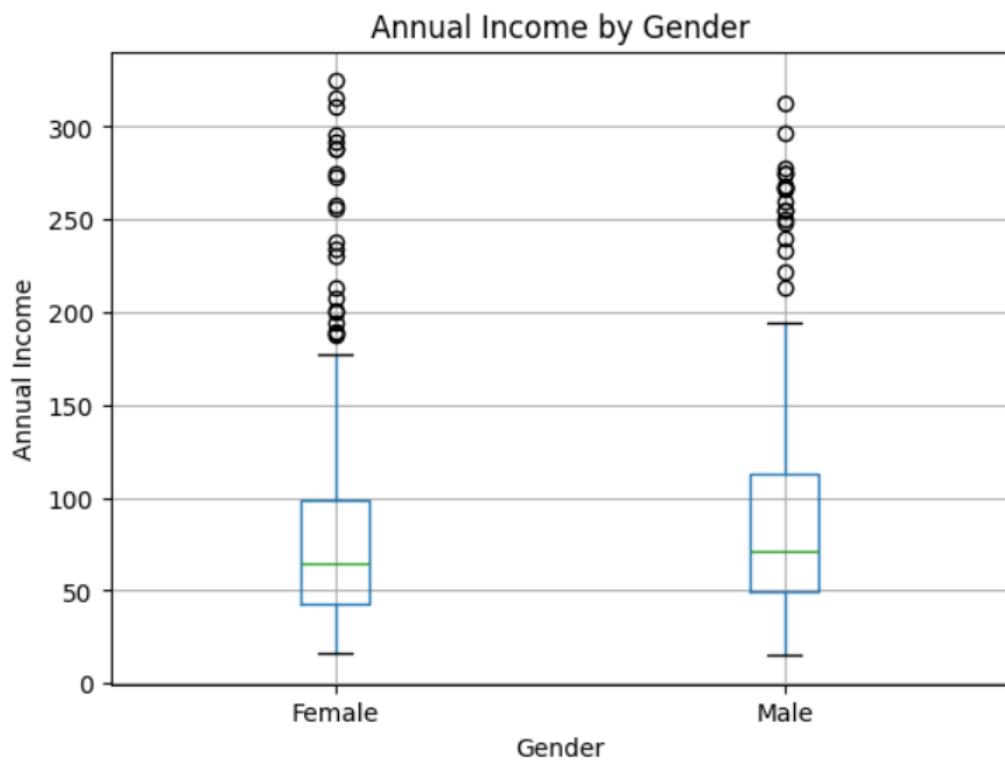
```
▶ plt.figure(figsize=(8,5))
df["Annual Income (k$)"].plot(kind="hist", bins=8, edgecolor="yellow")
plt.title("Annual Income Distribution")
plt.xlabel("Annual Income")
plt.show()
```

→ Annual Income Distribution



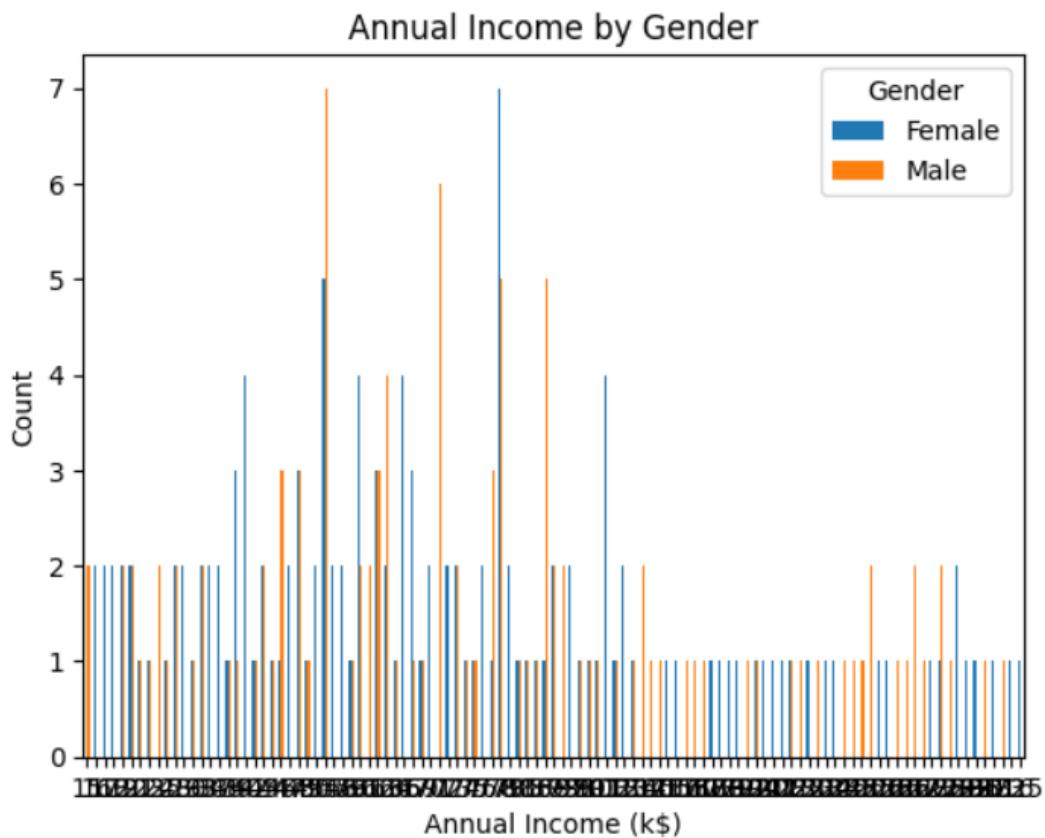
```
▶ plt.figure(figsize=(8,5))
df.boxplot(column="Annual Income (k$)",by="Gender")
plt.title("Annual Income by Gender")
plt.suptitle("")
plt.ylabel("Annual Income")
plt.show()
```

☞ <Figure size 800x500 with 0 Axes>



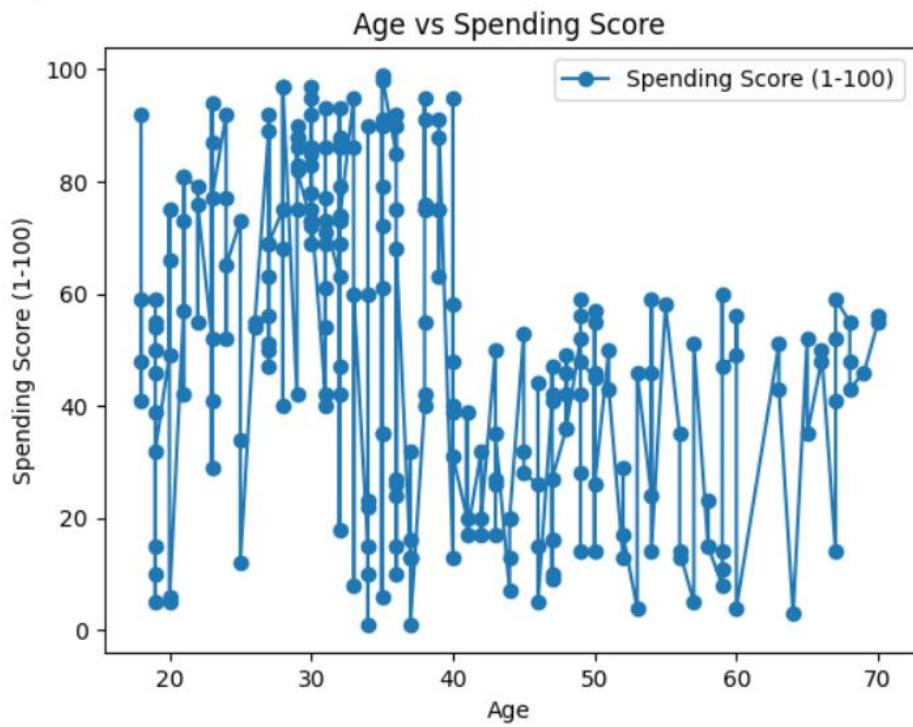
```
▶ plt.figure(figsize=(8,5))
pd.crosstab(df["Annual Income (k$)"],df["Gender"]).plot(kind="bar")
plt.title("Annual Income by Gender")
plt.ylabel("Count")
plt.xticks(rotation=0)
plt.show()
```

→ <Figure size 800x500 with 0 Axes>

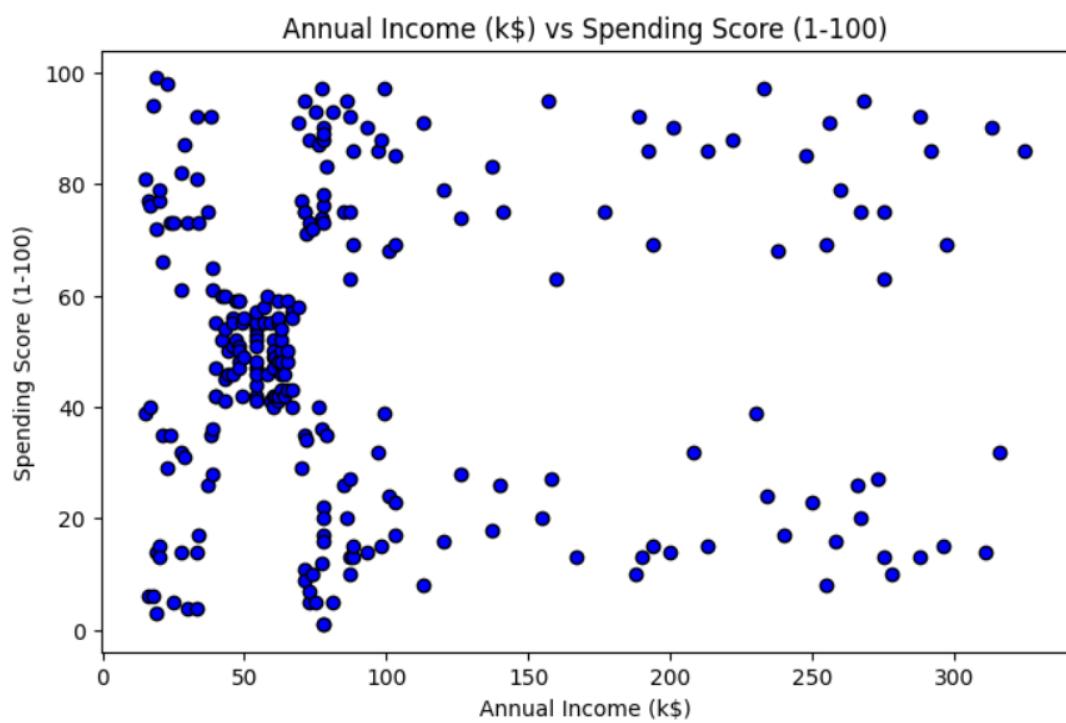


```
plt.figure(figsize=(10,5))
df.sort_values("Age").plot(x="Age", y="Spending Score (1-100)", kind="line", marker="o")
plt.title("Age vs Spending Score")
plt.ylabel("Spending Score (1-100)")
plt.show()
```

Figure size 1000x500 with 0 Axes

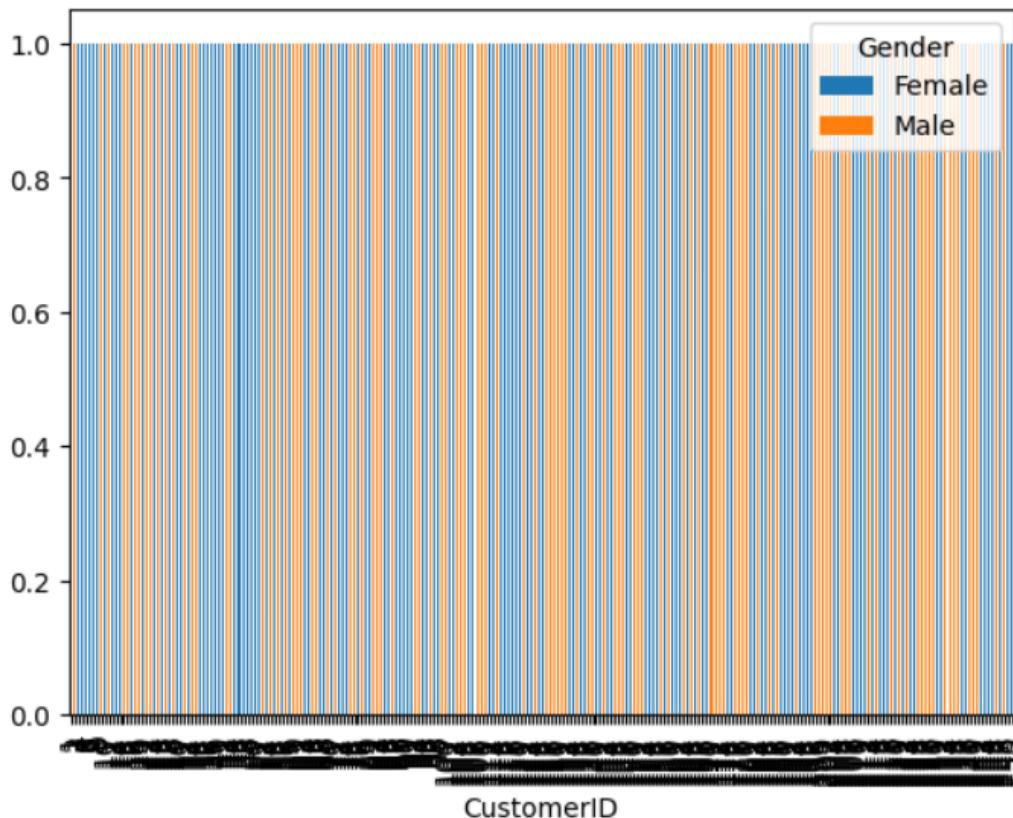


```
▶ plt.figure(figsize=(8,5))
plt.scatter(df["Annual Income (k$)"],df["Spending Score (1-100)"],c="blue",edgecolor="k")
plt.title("Annual Income (k$) vs Spending Score (1-100)")
plt.xlabel("Annual Income (k$)")
plt.ylabel("Spending Score (1-100)")
plt.show()
```



```
▶ plt.figure(figsize=(10,5))
pd.crosstab(df["CustomerID"],df["Gender"]).plot(kind="bar",stacked=True)
plt.title("Gender Distribution by CustomerID")
plt.ylabel("Count")
plt.xticks(rotation=90)
plt.show()
```

→ <Axes: xlabel='CustomerID'>



```
▶ plt.figure(figsize=(6,5))
corr = df[["Annual Income (k$)","Age"]].corr()
plt.imshow(corr, cmap="coolwarm", interpolation="none", aspect="auto")
plt.colorbar()
plt.xticks(range(len(corr)), corr.columns, rotation=45)
plt.yticks(range(len(corr)), corr.columns)
plt.title("Correlation Heatmap")
plt.show()
```

Correlation Heatmap

