DWM Lab Assignment -3 ¶

Bharath Kumar Reddy Sanikommu

AP19110010116/CSE-D

- Problem: Import a dataset or create a dataset and then
- Use matplot library to plot graphs.
    1. Histogram
    2. Pie Chart
    3. Box plot Chart
    4. Scatter plot
    5. Column Chart

In [13]:
```
import matplotlib.pyplot as plt
from matplotlib import pyplot as plt

df = pd.read_csv("Social_Network_Ads.csv")
df
```

Out[13]:

	User ID	Gender	Age	EstimatedSalary	Purchased
0	15624510	Male	19	19000	0
1	15810944	Male	35	20000	0
2	15668575	Female	26	43000	0
3	15603246	Female	27	57000	0
4	15804002	Male	19	76000	0
...	...	...	...	...	...
395	15691863	Female	46	41000	1
396	15706071	Male	51	23000	1
397	15654296	Female	50	20000	1
398	15755018	Male	36	33000	0
399	15594041	Female	49	36000	1

400 rows × 5 columns

Histogram

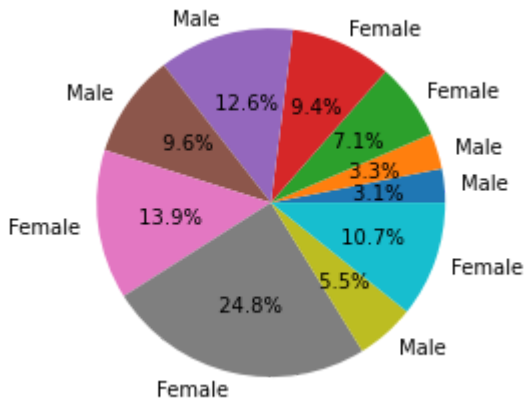
- A histogram shows the frequency on the vertical axis and the horizontal axis is another dimension. Usually it has bins where every bin has as minimum and maximum value. Each bin has a frequency between x and infinite

In [30]:
```
plt.hist(df["Gender"])
plt.ylabel("Age of the Employee")
plt.show()
```

Pie chart

- A pie chart can only display one series of data. Pie chart shows the size of items in one data series, proportional to the sum of the items.
- Matplot API has pie() function that generates a pie diagram representing data in a array. The fractional area of each wedge is given by x/sum(x)

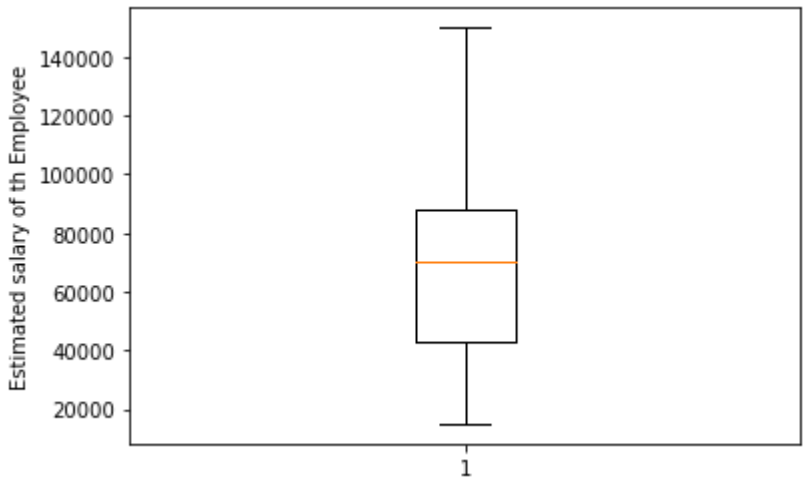
In [14]:
```
plt.pie(df["EstimatedSalary"][:10],labels = df["Gender"][:10], autopct = "% 1.1f%%"
plt.show()
```



Box Plot Chart

- A box plot represents the summary of a set of data containing the minimum, first quartile, median, third quartile and maximum.
- It is also called as whisker plot

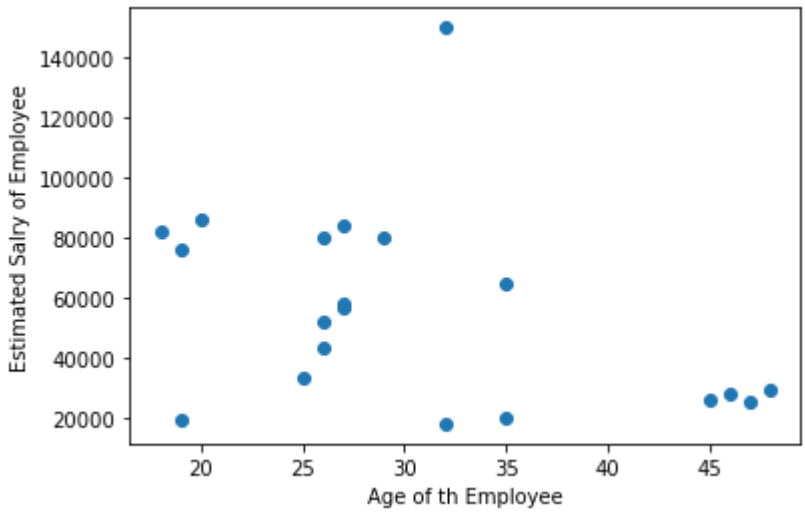
In [31]:
```
plt.boxplot(df["EstimatedSalary"])
plt.ylabel("Estimated salary of th Employee")
plt.show()
```



Scatter Plot

- A scatter plot is a type of plot that shows the data as a collection of points. The position of a point depends on its two dimensional value, where each vaue is a position on either the horizontal or vertical dimensions

In [16]:
```
plt.scatter(df["Age"][:20],df["EstimatedSalary"][:20])
plt.xlabel("Age of th Employee")
plt.ylabel("Estimated Salry of Employee")
plt.show()
```



Column Chart or Bar graph

- Column chart or commonly called as bar chart is a chart that presents categorical data with rectangular bars to the values that they represent. The bars can be plotted vertically or horizontally.
- A bar chart shows comparisions among discrete categories. One axis of the chart shows the specific categories being compared and the other axis represents a measured value.

In [32]:

```
plt.bar(df["Gender"],df["Age"])
plt.ylabel("Age of th Employee")
plt.show()
```