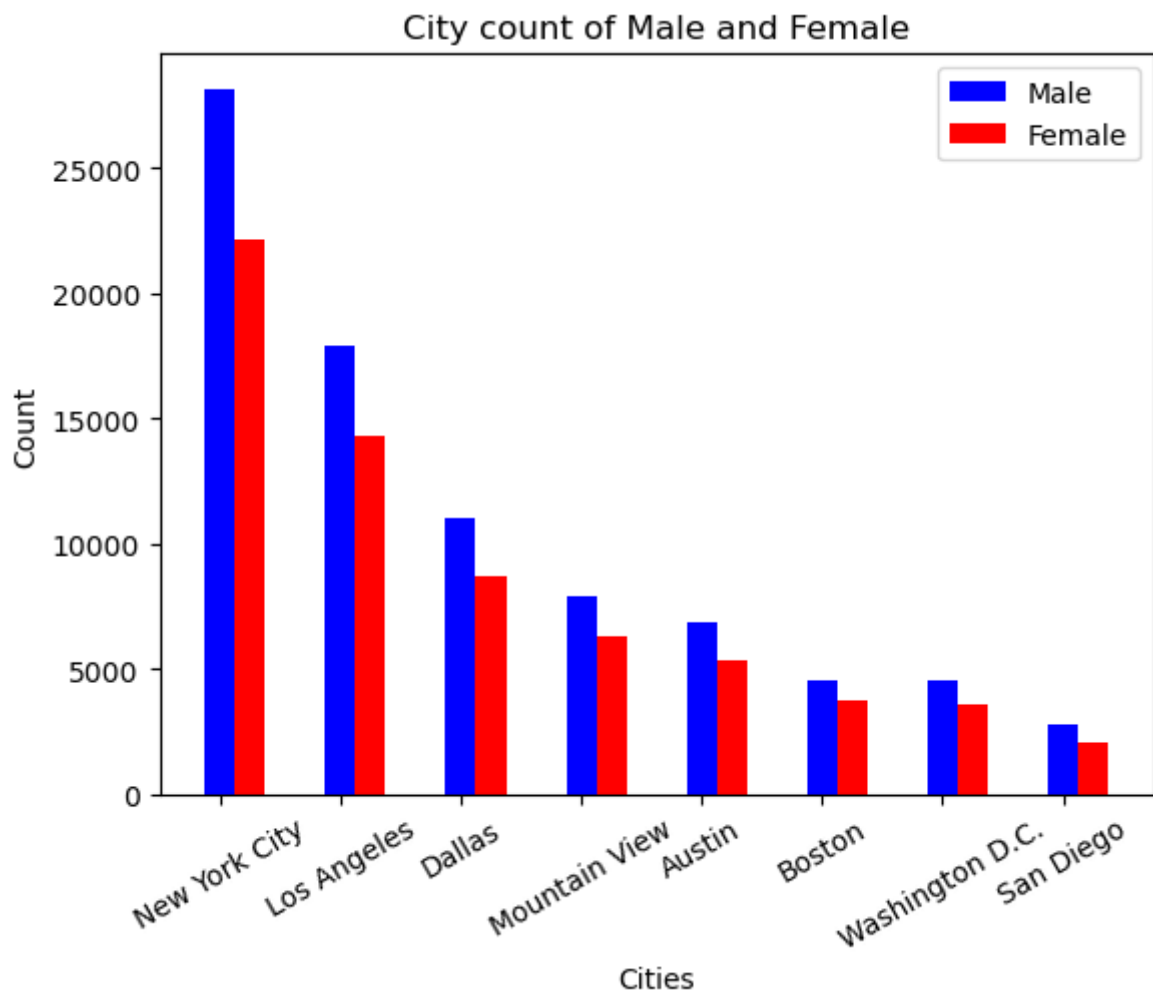


Lab1

```
In [107... df = pd.read_csv('toy_dataset.csv')
```

```
In [108... city_cnt = df['City'].value_counts()  
gender_cnt = df['Gender'].value_counts()  
city_cnt_male = df[df['Gender']=='Male']['City'].value_counts()  
city_cnt_female = df[df['Gender']=='Female']['City'].value_counts()
```

```
In [109... xlabel = list(city_cnt_male.index)  
  
fig = plt.figure()  
X = np.arange(8)  
plt.bar(X, city_cnt_male, color = 'b', width = 0.25)  
plt.bar(X+0.25, city_cnt_female, color = 'r', width = 0.25)  
plt.xticks(X, xlabel, rotation=30)  
plt.legend(labels=['Male', 'Female'])  
plt.title("City count of Male and Female")  
plt.ylabel("Count")  
plt.xlabel("Cities")  
plt.show()
```



Homework 1

```
In [90]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
In [91]: fishdf = pd.read_csv("Fish.csv")
```

```
In [92]: fishdf.head()
```

```
Out[92]:
```

	Species	Weight	Length1	Length2	Length3	Height	Width
0	Bream	242.0	23.2	25.4	30.0	11.5200	4.0200
1	Bream	290.0	24.0	26.3	31.2	12.4800	4.3056
2	Bream	340.0	23.9	26.5	31.1	12.3778	4.6961
3	Bream	363.0	26.3	29.0	33.5	12.7300	4.4555
4	Bream	430.0	26.5	29.0	34.0	12.4440	5.1340

```
In [93]: fishdf.info
```

```
Out[93]: <bound method DataFrame.info of
Height Width Species Weight Length1 Length2 Length3
0 Bream 242.0 23.2 25.4 30.0 11.5200 4.0200
1 Bream 290.0 24.0 26.3 31.2 12.4800 4.3056
2 Bream 340.0 23.9 26.5 31.1 12.3778 4.6961
3 Bream 363.0 26.3 29.0 33.5 12.7300 4.4555
4 Bream 430.0 26.5 29.0 34.0 12.4440 5.1340
.. ... ..
154 Smelt 12.2 11.5 12.2 13.4 2.0904 1.3936
155 Smelt 13.4 11.7 12.4 13.5 2.4300 1.2690
156 Smelt 12.2 12.1 13.0 13.8 2.2770 1.2558
157 Smelt 19.7 13.2 14.3 15.2 2.8728 2.0672
158 Smelt 19.9 13.8 15.0 16.2 2.9322 1.8792

[159 rows x 7 columns]>
```

```
In [94]: fishdf["Species"].value_counts()
```

```
Out[94]: Perch      56
Bream      35
Roach      20
Pike       17
Smelt      14
Parkki     11
Whitefish   6
Name: Species, dtype: int64
```

```
In [95]: fishdf["Species"].value_counts().index
```

```
Out[95]: Index(['Perch', 'Bream', 'Roach', 'Pike', 'Smelt', 'Parkki', 'Whitefish'], dtype='object')
```

```
In [96]: fishIndex = list(fishdf["Species"].value_counts().index)
fishIndex
```

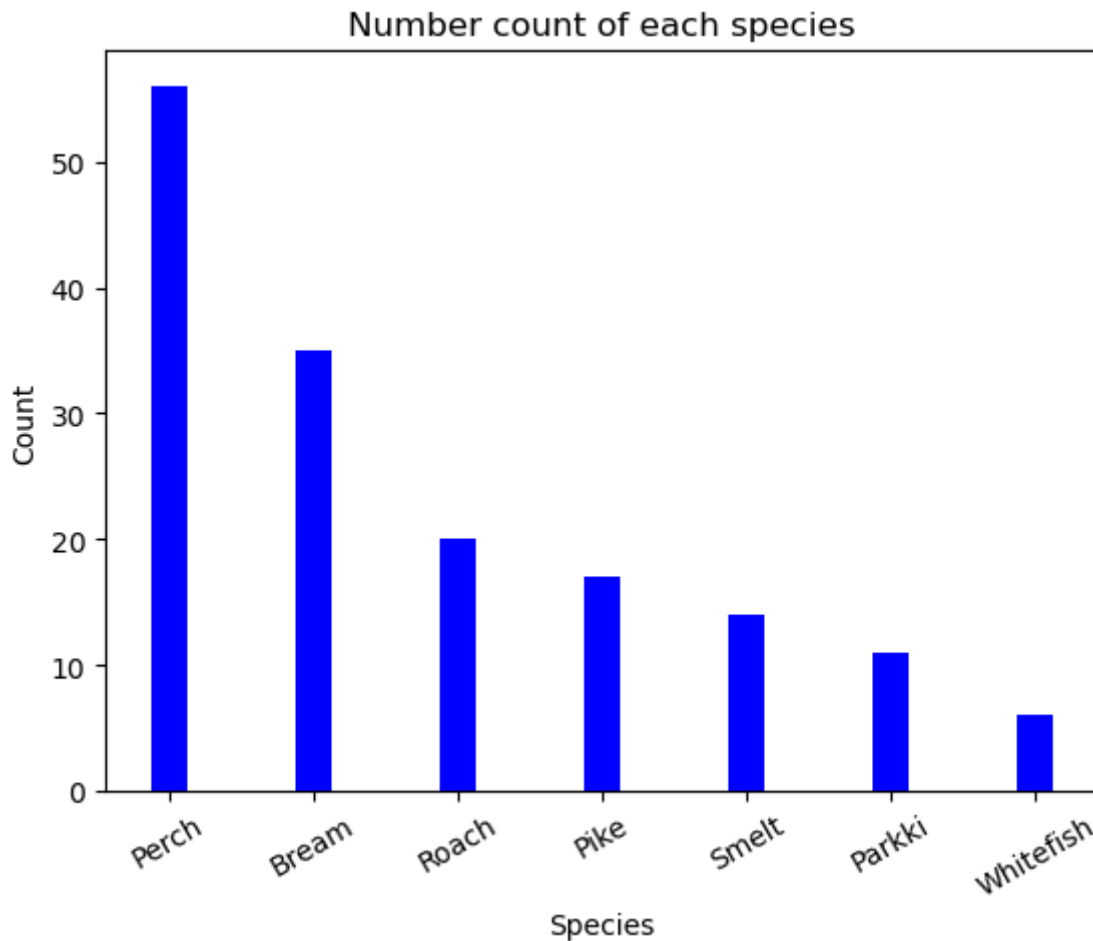
```
Out[96]: ['Perch', 'Bream', 'Roach', 'Pike', 'Smelt', 'Parkki', 'Whitefish']
```

1.Bar Graph

```
In [97]: # Bar Graph

fig1 = plt.figure()

X = np.arange(len(fishIndex))
plt.bar(X, fishdf["Species"].value_counts(), color = 'b', width = 0.25)
plt.xticks(X, fishIndex, rotation=30)
plt.title("Number count of each species")
plt.ylabel("Count")
plt.xlabel("Species")
plt.show()
```



```
In [98]: # fishdf_sorted = fishdf[fishdf["Weight"]].sort_values(ascending=True)
fishdf_sorted = fishdf.sort_values(by="Weight")
```

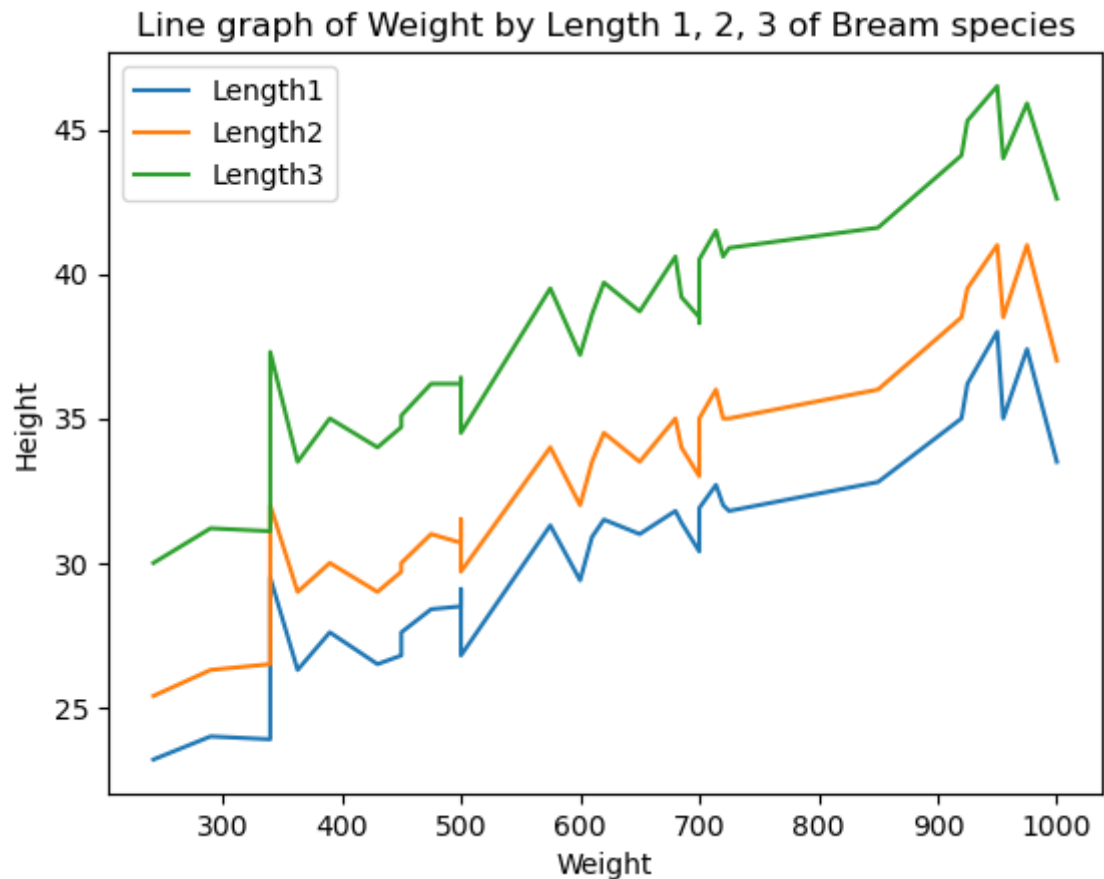
2. Line Graph

```
In [99]: # Line Graph
fig2 = plt.figure()

Bream_sorted = fishdf_sorted[fishdf_sorted["Species"]=="Bream"]
# Bream_sorted

plt.plot(Bream_sorted.iloc[:,1:2], Bream_sorted.iloc[:,2:5])
plt.legend(labels=["Length1", "Length2", "Length3"])
```

```
plt.title("Line graph of Weight by Length 1, 2, 3 of Bream species")
plt.xlabel("Weight")
plt.ylabel("Height")
plt.show()
```

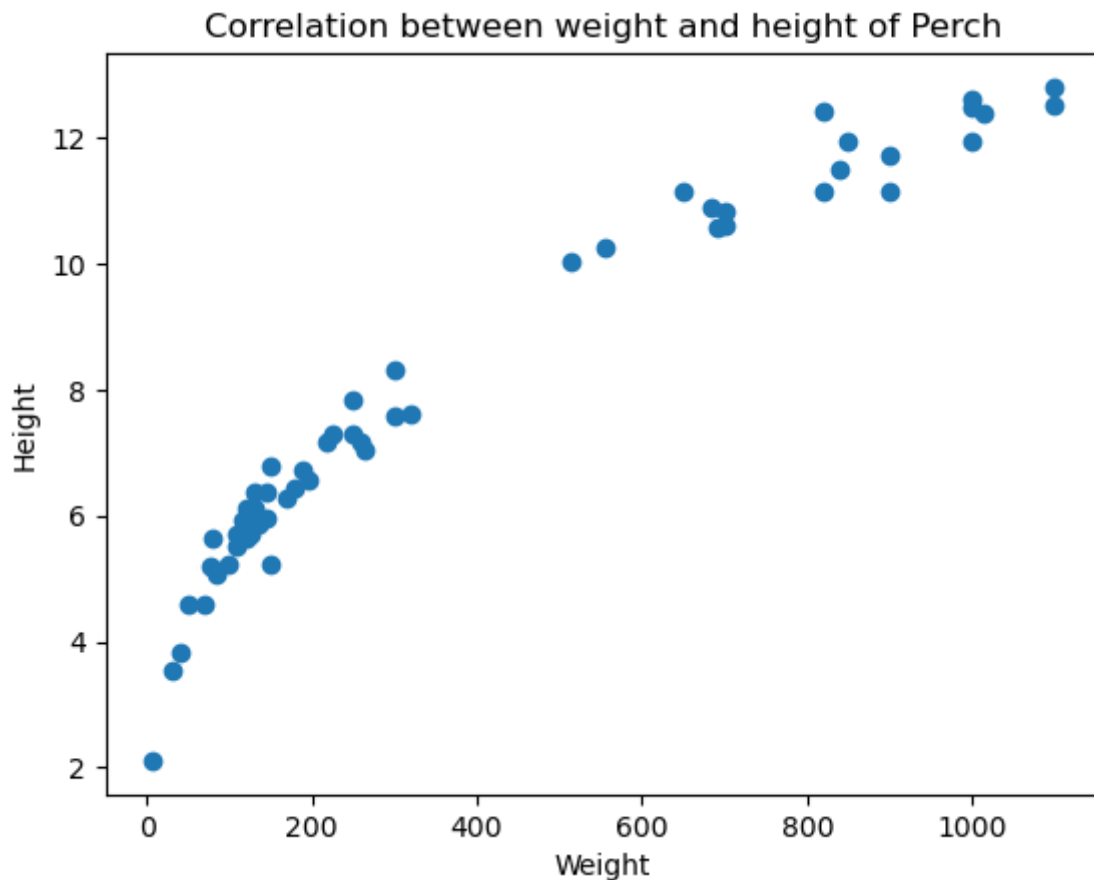


3. Scatter Graph

```
In [100... # Scatter Grpah

fig3 = plt.figure()

plt.scatter(fishdf[fishdf['Species']=="Perch"]["Weight"], fishdf[fishdf['Species']=="Perch"]["Height"])
plt.title("Correlation between weight and height of Perch")
plt.xlabel("Weight")
plt.ylabel("Height")
plt.show()
```



```
In [101...] fishdf["Species"].value_counts()
```

```
Out[101]: Perch      56
          Bream     35
          Roach     20
          Pike      17
          Smelt     14
          Parkki    11
          Whitefish  6
          Name: Species, dtype: int64
```

```
In [102...] Perch_weight = fishdf[fishdf["Species"]=="Perch"]["Weight"]
          Bream_weight = fishdf[fishdf["Species"]=="Bream"]["Weight"]
          Roach_weight = fishdf[fishdf["Species"]=="Roach"]["Weight"]
          Pike_weight = fishdf[fishdf["Species"]=="Pike"]["Weight"]
          Smelt_weight = fishdf[fishdf["Species"]=="Smelt"]["Weight"]
          Parkki_weight = fishdf[fishdf["Species"]=="Parkki"]["Weight"]
          Whitefish_weight = fishdf[fishdf["Species"]=="Whitefish"]["Weight"]
```

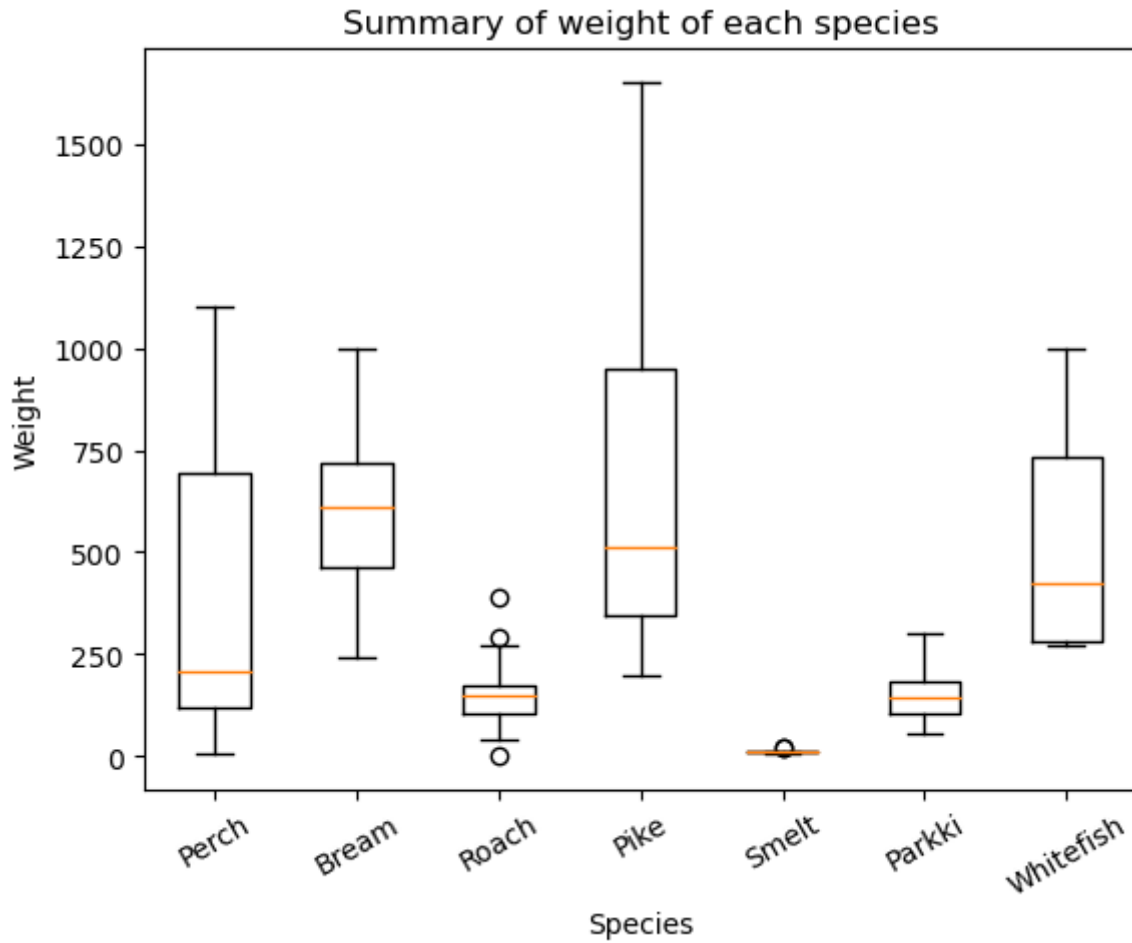
```
In [103...] species_data = [Perch_weight, Bream_weight, Roach_weight, Pike_weight, Smelt_weight, Parkki_weight, Whitefish_weight]
```

4. Box Plot

```
In [104...] # Box plot

fig4 = plt.figure()
X1 = np.arange(1,8)
plt.boxplot(species_data)
plt.xticks(X1, fishIndex, rotation=30)
plt.title("Summary of weight of each species")
```

```
plt.xlabel("Species")
plt.ylabel("Weight")
plt.show()
```



5. Pie Chart

```
In [105... # Pie chart

cnt = []
for i in fishdf.Species.value_counts():
    cnt.append(i)

plt.pie(cnt, labels=fishIndex, autopct='%1.1f%%')
plt.title("Percentage of each species in Fish.csv file")
plt.show()
```

Percentage of each species in Fish.csv file

