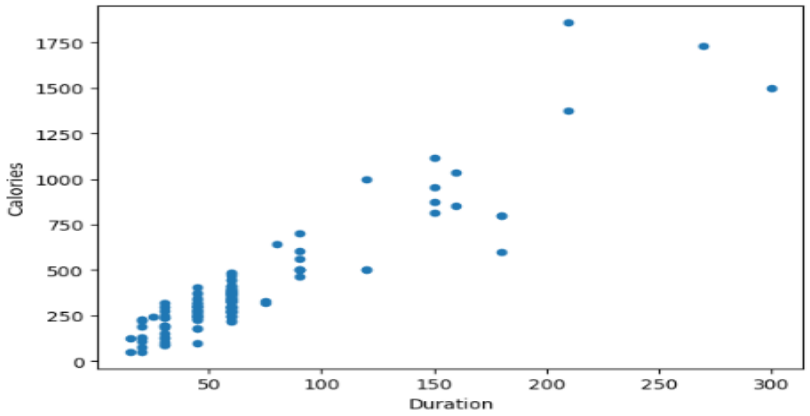
YOUTUBE LINK: ***https://youtu.be/kxajGlChVRs***

GITHUB: ***https://github.com/Harishwar-reddi/ICP-4***

Q1

```
[33] import pandas as pd

     # Read File
     url = 'https://drive.google.com/uc?id={}'.format('19jhAQZno1RpcFpJAimU0KDYtq-QDnLi5')
     df = pd.read_csv(url)

     description = df.describe()
     print(description)
     #nulls
     null_values = df.isnull().sum()
     print(null_values)

     df.fillna(df.mean(numeric_only=True), inplace=True)

     aggregated_data = df[['Duration', 'Calories']].agg(['min', 'max', 'count', 'mean'])
     print(aggregated_data)

     filtered_data = df[(df['Calories'] > 500) & (df['Calories'] < 1000)]

     filtered_data_2 = df[(df['Calories'] > 500) & (df['Pulse'] < 100)]

     df_modified = df.drop(columns=['Maxpulse'])

     #drop
     df.drop(columns=['Maxpulse'], inplace=True)

     #convert
     df['Calories'] = df['Calories'].astype(int)

     #plot
     df.plot.scatter(x='Duration', y='Calories')
```

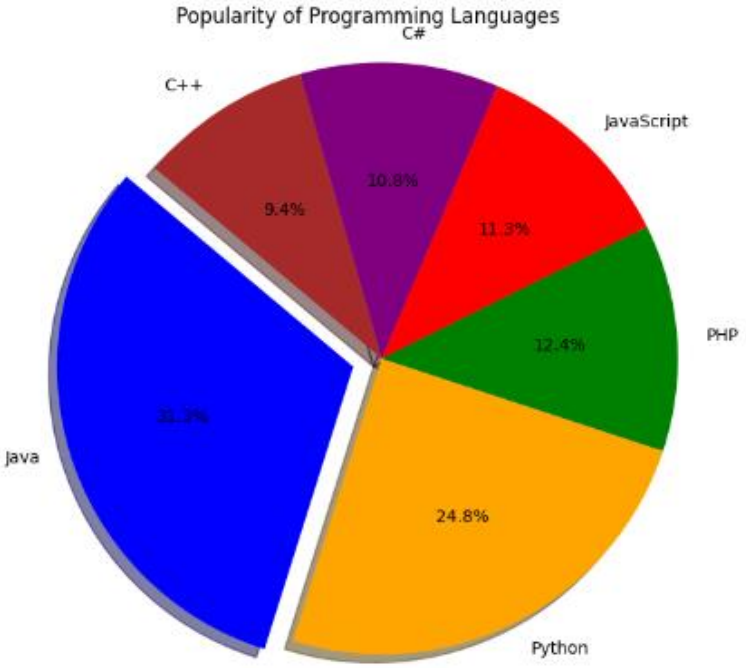
```
         Duration       Pulse    Maxpulse     Calories
count  169.000000  169.000000  169.000000   164.000000
mean    63.846154  107.461538  134.047337   375.790244
std     42.299949   14.510259   16.450434   266.379919
min     15.000000   80.000000  100.000000    50.300000
25%     45.000000  100.000000  124.000000   250.925000
50%     60.000000  105.000000  131.000000   318.600000
75%     60.000000  111.000000  141.000000   387.600000
max    300.000000  159.000000  184.000000  1860.400000
Duration    0
Pulse       0
Maxpulse    0
Calories    5
dtype: int64
          Duration     Calories
min      15.000000    50.300000
max     300.000000  1860.400000
count   169.000000   169.000000
mean     63.846154   375.790244
<Axes: xlabel='Duration', ylabel='Calories'>
```



Q2

1

```
import matplotlib.pyplot as plt

# Data
languages = ['Java', 'Python', 'PHP', 'JavaScript', 'C#', 'C++']
popularity = [22.2, 17.6, 8.8, 8, 7.7, 6.7]

# Create pie chart
colors = ['blue', 'orange', 'green', 'red', 'purple', 'brown']
explode = (0.1, 0, 0, 0, 0, 0)

plt.figure(figsize=(10, 7))
plt.pie(popularity, explode=explode, labels=languages, colors=colors, autopct='%1.1f%%', shadow=True, startangle=140)
plt.title('Popularity of Programming Languages')
plt.axis('equal')
plt.show()
```



2

```
import matplotlib.pyplot as plt

# Data
math_marks = [88, 92, 80, 89, 100, 80, 60, 100, 80, 34]
science_marks = [35, 79, 79, 48, 100, 88, 32, 45, 20, 30]
marks_range = [10, 20, 30, 40, 50, 60, 70, 80, 90, 100]

# Create scatter plot
plt.figure(figsize=(10,6))
plt.scatter(marks_range, math_marks, label='Math Marks', color='red')
plt.scatter(marks_range, science_marks, label='Science Marks', color='green')

# Title and labels
plt.title('Maths and Science Marks', fontsize=15)
plt.xlabel('Marks Range', fontsize=13)
plt.ylabel('Marks Scored', fontsize=13)

# Legend
plt.legend(loc='upper right')

# Show the plot
plt.tight_layout()
plt.show()
```