1. **Produce a line plot showing multiple lines with proper labels and legend. Describe what conclusions you can draw from this plot.**

**CODE:**

# importing appropriate libraries

import pandas as pds

import matplotlib.pyplot as pylt

# To read csv file

movie = pds.read\_csv("movies\_marvel\_dc.csv", encoding = 'latin1')

# To assign variable

mname = movie["Original Title"]

Mbud = movie["Budget"]

Musa = movie["Gross USA"]

Mww = movie["Gross Worldwide"]

**# To plot line graph**

pylt.figure(figsize=(8,8))

pylt.plot(mname,Mbud, label ='Budget') # It will produce 1st line

pylt.plot(mname, Musa, label = 'Gross USA') # It will produce 2nd line

pylt.plot(mname, Mww, label = 'Gross Worldwide') # It will produce 3rd line

pylt.legend()

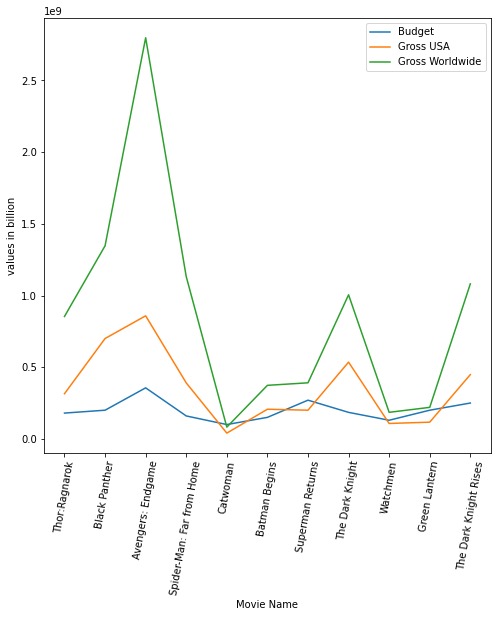
pylt.xticks(rotation=80) # will rotate the text of x axis

pylt.xlabel('Movie Name')

pylt.ylabel('values in billion')

pylt.show()

**OUTPUT:**



*Figure 1: Multiple Line Plot*

**CONCLUSIONS:**

The program and graph illustrate the budget, gross USA and gross worldwide of the movies of Marvel and DC. It clearly shows the budget of the movies and how much is the gross of the movies in the USA and worldwide. It compares the all movies which are listed. We can depict that the movie avengers: endgame has the Highest gross among all and followed by Black Panther and Spider-man: Far from Home. The movie Catwoman has the least gross among all followed by watchmen. Even the movie Catwoman has its budget more than its gross.

I used the line graph for this as in it we can easily compare budget, gross USA, gross worldwide of the all different movies in one single graph.

**Dataset Link:** <https://www.kaggle.com/datasets/leonardopena/marvel-vs-dc>

**2. Produce graphs using two other visualization methods. Explain why you picked this type of graph and describe what conclusions you can draw.**

**CODE 1 for pie chart:**

**#To plot pie chart**

car = pds.read\_csv("car\_data.csv")

hpower = car["horsepower"]

cname = car["car name"]

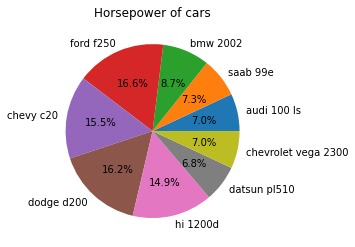
pylt.figure()

pylt.pie(hpower, labels=cname, autopct='%1.1f%%') “““ It will plot the pie chart and also give percentage to portion as we used autopct”””

pylt.title("Horsepower of cars") # used to give title to pie chart

pylt.show()

**OUTPUT:**



*Figure 2: Pie Chart*

**CONCLUSIONS:**

The program and graph below depict the horsepower of 9 cars. Among all 9 cars, the Ford F250, Dodge D200, and Chevy C20 have the highest horsepower while the Datsun PL510, Chevrolet Vega 2300, and Audi 100 have the lowest horsepower. There is not much difference between the horsepower of the car Ford F250 and Dodge D200. Audi 100 and Chevrolet Vega 2300 have the same horsepower.

I used a pie chart for this as it allows us to compare the horsepower of cars effectively because it is small data set.

**Dataset Link:** <https://www.kaggle.com/datasets/raghupalem/auto-mpg-data-set>

**CODE 2 for histogram:**

**#To plot histogram**

stp= pds.read\_csv("StudentsPerformance.csv")

pylt.figure(1)

pylt.hist(stp["reading score"],density=True, label="reading score")

pylt.hist(stp["writing score"], density=True, alpha=0.7, label="writing score")

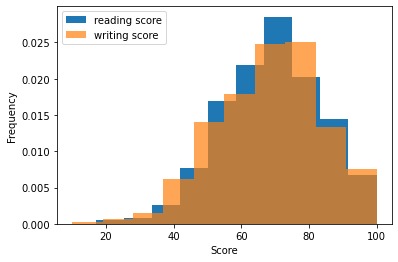
pylt.legend()

pylt.xlabel('Score')

pylt.ylabel('Frequency')

pylt.show()

**OUTPUT:**



*Figure 3: Histogram Plot*

**CONCLUSIONS:**

The program and histogram shows the scores of students in reading and writing. Scores are little bit similar, but the reading score is bit higher than writing score and it is between 60 and 80. Most of the students’ scores lie in this range. And only a few students’ scores is less than 20.

I used the histogram for this data because it is a large data set, and it allows me to depict any trends more easily.

**Dataset Link:** <https://www.kaggle.com/datasets/adithyabshetty100/student-performance>

**Link of my Repository**: - <https://github.com/hetanshipatel/Data-Visulization-ADS-assignment-1>