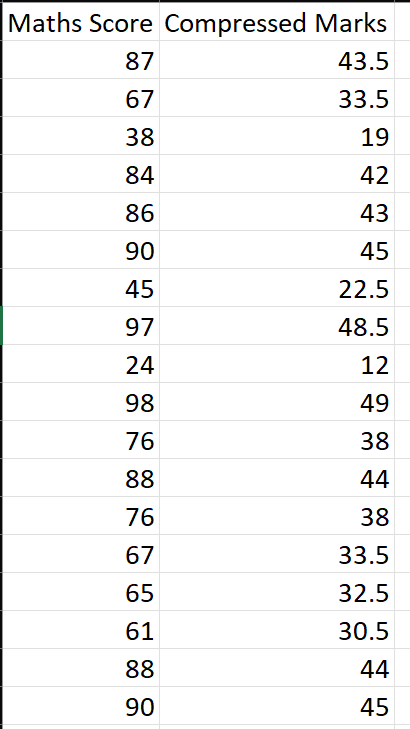
Name: Tirth Patel

Roll No: 16010421075

Batch A2

Dataset



import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

df = pd.read\_csv("dataset.csv")

c1 = df["Maths Score"]

c2 = df["Compressed Marks"]

correlation = c1.corr(c2)

def relation():

    if correlation == 1.0:

        print("Relation: Linear")

    else:

        print("Relation: Not Linear")

def dfmean1():

    maths\_marks = df["Maths Score"]

    total1 = maths\_marks.sum()

    count1 = len(maths\_marks.axes[0])

    mean1 = total1/count1

    print("x(mean) =", mean1)

def dfmean2():

    compressed\_marks = df["Compressed Marks"]

    total2 = compressed\_marks.sum()

    count2 = len(compressed\_marks.axes[0])

    mean2 = total2/count2

    print("y(mean) =", mean2)

print("\n")

relation()

dfmean1()

dfmean2()

def estimate\_coef(x, y):

    n = np.size(x)

    m\_x = np.mean(x)

    m\_y = np.mean(y)

    SS\_xy = np.sum(y\*x) - n\*m\_y\*m\_x

    SS\_xx = np.sum(x\*x) - n\*m\_x\*m\_x

    global W\_0

    global W\_1

    W\_1 = SS\_xy / SS\_xx

    W\_0 = m\_y - W\_1\*m\_x

    return (W\_0, W\_1)

def plot\_regression\_line(x, y, b):

    plt.scatter(x, y, color="m",

    marker="o", s=30)

    y\_pred = b[0] + b[1]\*x

    plt.plot(x, y\_pred, color="g")

    plt.xlabel('x')

    plt.ylabel('y')

    plt.show()

def main():

    x = df["Maths Score"]

    y = df["Compressed Marks"]

    b = estimate\_coef(x, y)

    print("Estimated coefficients:\nW\_0 = {} \

    \nb\_1 = {}".format(b[0], b[1]))

    plot\_regression\_line(x, y, b)

if \_\_name\_\_ == "\_\_main\_\_":

    main()

y = int(input("Input the value for Maths Score: "))

x = y\*W\_1 +W\_0

print("The missing value in the Compressed Marks column will be:", x)

