**Problem - 5**

**Student Name: Milan Sharma UID: 23MAI10003**

**Branch: ME – CSE - AIML Section/Group: MAI – 1 (A)**

**Semester: 1st Date of Performance: 06 Nov 2023**

**Subject Name: Artificial Intelligence Subject Code: 23 CSH 621**

1. **Aim of the Experiment :**

You are given the scores of N students in three different subjects - Mathematics,\*Physics\* and Chemistry; all of which have been graded on a scale of 0 to 100. Your task is to compute the Pearson product-moment correlation coefficient between the scores of different pairs of subjects (Mathematics and Physics, Physics and Chemistry, Mathematics and Chemistry) based on this data. This data is based on the records of the CBSE K-12 Examination - a national school leaving examination in India, for the year 2013.

1. **Objective of the Experiment :**

Pearson product-moment correlation coefficient

This is a measure of linear correlation described well [on this Wikipedia page](http://en.wikipedia.org/wiki/Pearson_product-moment_correlation_coefficient). The formula, in brief, is given by:  
http://upload.wikimedia.org/math/c/a/6/ca68fbe94060a2591924b380c9bc4e27.png

where x and y denote the two vectors between which the correlation is to be measured.

The first row contains an integer N.  
This is followed by N rows containing three tab-space ('\t') separated integers, M P C corresponding to a candidate's scores in Mathematics, Physics and Chemistry respectively.  
Each row corresponds to the scores attained by a unique candidate in these three subjects.

1. **Code :**

def pearson\_correlation(x, y, n):

    mean\_x = sum([i for i in x]) / len(x)

    mean\_y = sum([i for i in y]) / len(y)

    sum\_xy = sum([x[i] \* y[i] for i in range(n)])

    Sx = pow(sum([pow(i - mean\_x, 2) for i in x]) / (n-1), 0.5)

    Sy = pow(sum([pow(i - mean\_y, 2) for i in y]) / (n-1), 0.5)

    corr = (sum\_xy - n \* mean\_x \* mean\_y) / ((n-1) \* Sx \* Sy)

    return corr

n = int(input())

data = [list(map(float, input().split())) for i in range(n)]

math = [data[i][0] for i in range(n)]

physics = [data[i][1] for i in range(n)]

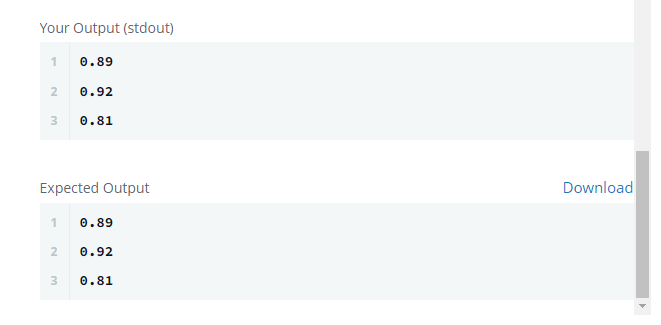
chem = [data[i][2] for i in range(n)]

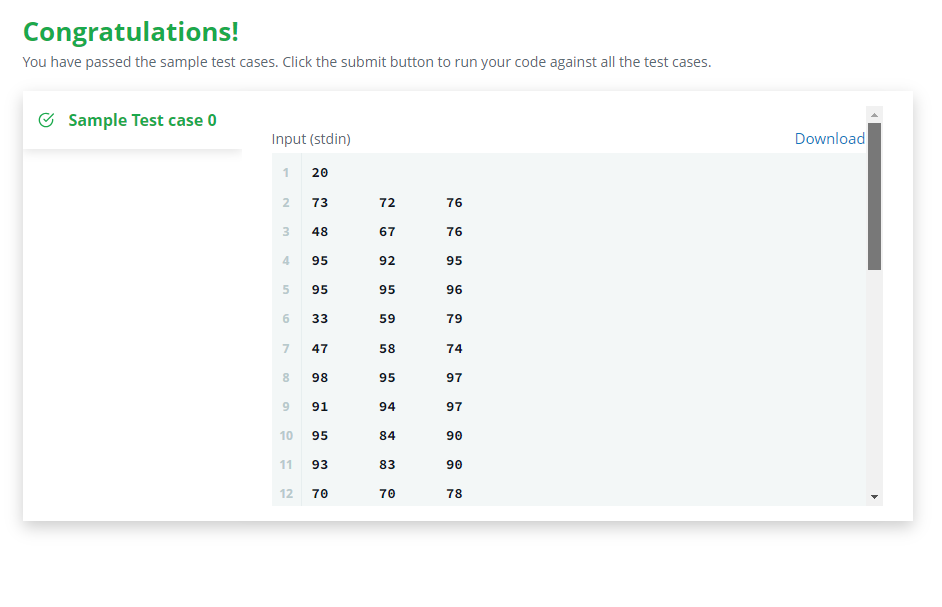
print("%.2f" % pearson\_correlation(math, physics, n))

print("%.2f" % pearson\_correlation(physics, chem, n))

print("%.2f" % pearson\_correlation(chem, math, n))

1. **Output:**

****

****

**Learning outcomes (What I have learnt):**

1. **Learned about numpy and pandas libraries.**
2. **Learned to use time constraints in python.**
3. **Learned how to use mean mode median in python using libraries.**
4. **Learned about Pearson correlation.**