**LAB 04**

**CORRELATION COEFFICIENT**

**Aim: Finding the correlation coefficient between 2 data vectors**

**New-Term used:**

1. **c(): This function is used to create vectors by combining multiple values. It is used to create the vectors x and y.**
2. **cov(): This function is used to calculate the covariance between two numeric vectors. It is used to calculate the covariance between the vectors x and y, stored in the variable coff.**
3. **sd(): This function is used to calculate the standard deviation of a numeric vector. It is used to calculate the standard deviations of the vectors x and y in the calculation of the correlation coefficient.**

**Input:**

**x = c(17,18,19,19,20,20,21,21,22,23)**

**xsum = sum(x)**

**xsum**

**y = c(12,16,14,11,15,19,22,16,15,20)**

**ysum = sum(y);**

**ysum**

**n = length(x);**

**n**

**x2 = x^2;**

**x2sum = sum(x2);**

**x2sum**

**y2 = y^2;**

**y2sum = sum(y2);**

**y2sum**

**xy = x\*y;**

**xysum = sum(xy);**

**xysum**

**x2**

**y2**

**xy**

**#Rxy = cov(x,y)/sigx\*sigy**

**#finding correlation coefficient**

**coff = cov(x,y)/(sd(x)\*sd(y));**

**coff**

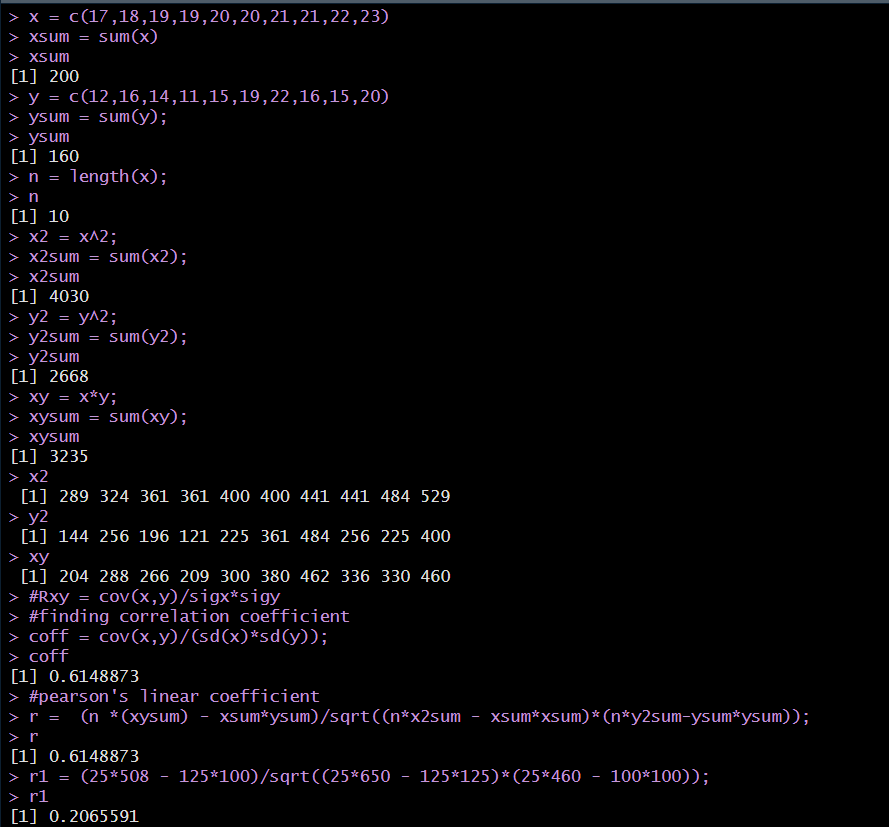
**#pearson's linear coefficient**

**r = (n \*(xysum) - xsum\*ysum)/sqrt((n\*x2sum - xsum\*xsum)\*(n\*y2sum-ysum\*ysum));**

**r**

**r1 = (25\*508 - 125\*100)/sqrt((25\*650 - 125\*125)\*(25\*460 - 100\*100));**

**r1**

**OUTPUT:**

**Inference:**

**Hence the correlation coefficien tof a given data can be calculated using R.**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***