

# EXCEL ASSIGNMENT

## SUBMITTED BY:-

NAME:- SURESH KUMAR PRAJAPATI  
ENROLL.NO. :-CUSB2302222008  
MASTER IN DATA SCIENCE AND APPLIED STATISTICS

## SUBMITTED BY:-

PROFESSOR:- Dr. INDRAJEET KUMAR  
DATA SCIENCE FACULTY

### DESCRIPTIVE STATISTICS:-

Mean(average):-

$$\text{Mean} = \Sigma x/n$$

Median = middle value (for odd case)

Median =  $((n/2)^{\text{th}} \text{ term} + (n/2+1)^{\text{th}} \text{ term})/2$  (for even case)

Mode = The mode is the value that appears most frequently in the data set.

Range = Max-Min

$$\text{Variance} = \sigma^2 = \Sigma (X_i - \text{mean})^2/n$$

$$\text{Standard Deviation } (\sigma) = (\sigma^2)^{1/2}$$

$$\text{Interquartile Range(IQR)} = Q_3 - Q_1$$

$$\text{Coefficient of variation(CV)} = (\sigma/\text{mean}) * 100\%$$

$$\text{Kurtosis} = [\Sigma (X_i - \text{mean})^4 / (1/n \Sigma (X_i - \text{mean})^2)] - 3$$

$$\text{Skewness} = 1/n \Sigma (X_i - \text{mean})^3 / (1/n \Sigma (X_i - \text{mean})^2)^{3/2}$$

where n is the number of observation.

	Column 1	y.GNP	y.Armed.Forces	y.Year
Mean	8.5	387.6984375	260.66875	1954.5
Standard Error	1.19023807142381	24.848734448822	17.3979901108097	1.19023807142381
Median	8.5	381.427	271.75	1954.5
First Quartile	4.75	317.881	229.8	1950.75
Third Quartile	12.25	454.0855	306.075	1958.25
Variance	22.6666666666	9879.35365	4843.040958333	22.66666666

	6667	932917	33	666667
StandardDeviation	4.76095228569523	99.394937795288	69.591960443239	4.76095228569523
Kurtosis	-1.2	-1.07202105801688	-0.835442289946648	-1.2
Skewness	0	0.0279793231224011	-0.447652802591812	0
Range	15	320.605	213.8	15
Minimum	1	234.289	145.6	1947
Maximum	16	554.894	359.4	1962
Sum	136	6203.175	4170.7	31272
Count	16	16	16	16

$R = \frac{\sum(X_i - \text{mean})(Y_i - \text{mean})}{[\sum(X_i - \text{mean})^2 \sum(Y_i - \text{mean})^2]^{1/2}}$   
 measures linear relationships

Correlations	Column 1	Column 2	Column 3	Column 4
Column 1	1			
Column 2	0.994337974201009	1		
Column 3	0.417245149834945	0.340842943122361	1	
Column 4	1	0.994337974201009	0.417245149834945	1

$\text{Cov}(X, Y) = \frac{\sum(X_i - \text{mean})(Y_i - \text{mean})}{n-1}$

covariance measures the direction of the relationship (positive or negative),

Covariances	Column 1	Column 2	Column 3	Column 4
Column 1	21.25			
Column 2	389.15953333 3333	8205.79237 832889		
Column 3	129.603125	1978.92669 244445	4540.3508 984375	
Column 4	21.25	389.159533 333333	129.60312 5	21.25

The general formula for a simple linear regression, which involves one independent variable, is given by:  $Y = \beta_0 + \beta_1 X + \epsilon$

**Prediction:** Building models to predict the values of the dependent variable based on new values of the independent variables.

- 1. Understanding Relationships:** Analyzing the strength and direction of relationships between variables.
- 2. Variable Selection:** Identifying which independent variables are most important in predicting the dependent variable.
- 3. Trend Analysis:** Examining trends in data and making predictions about future values.
- 4. Model Evaluation:** Assessing the goodness of fit of the model to the observed data.

REGRESSION						
Regression Model	Linear					
LINEST raw output						
0.56720088187908 1	18.5830577 461206					
0.28658917530002 6	11.0518791 531133					
0.43927367128507 3	16.6355785 697159					
3.91700593310644	5					
1084.00191396833	1383.71237 174595					

Regression Statistics						
R <sup>2</sup>	0.43927367 1285073					
Standard Error	16.6355785 697159					
Count of X variables	1					
Observations	7					
Adjusted R <sup>2</sup>	0.32712840 5542087					
Analysis of Variance (ANOVA)						
	df	SS	MS	F	Significance F	
Regression	1	1084.00191396 833	1084.00191 396833	3.91700593 310644	0.1046927833 77419	
Residual	5	1383.71237174 595	276.742474 34919			
Total	6	2467.71428571 429				
Confidence level	0.95					
	Coefficients	Standard Error	t-Statistic	P-value	Lower 95%	Upper 95%
Intercept	18.5830577 461206	11.0518791531 133	1.68143873 893932	0.15350644 864799	- 9.8267020545 2013	46.99281 7546761 3
CP	0.56720088 1879081	0.28658917530 0026	1.97914272 681544	0.10469278 3377419	- 0.1695000464 37158	1.303901 8101953 2
CP	Predicted Y	SP	Residual			

10	24.2550665 649114	34	9.74493343 508862
35	38.4350886 118884	45	6.56491138 81116
34	37.8678877 300093	60	22.1321122 699907
30	35.5990842 02493	32	- 3.59908420 2493
10	24.2550665 649114	20	- 4.25506656 491139
23	31.6286780 293394	5	- 26.6286780 293394
80	63.9591282 96447	60	- 3.95912829 644703



