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):-

Mean=$\Sigma x/n$

Median = middle value (for odd case)

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

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

Range = Max-Min

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

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

Interquartile Range(IQR) = $Q_3 - Q_1$

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

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

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

where n is the number of observation.

	Column 1	y.GNP	y.Armed.Forces	y.Year
Mean	8.5	387.698437 5	260.66875	1954.5
Standard Error	1.1902380714 2381	24.8487344 48822	17.39799011080 97	1.19023807 142381
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	22.666666666 6667	9879.35365 932917	4843.040958333 33	22.6666666 666667
StandardDeviation	4.7609522856 9523	99.3949377 95288	69.59196044323 9	4.76095228 569523

Kurtosis	-1.2	-1.07202105801688	-0.835442289946648	-1.2
Skewness	0	-0.02797932031224011	-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

$R = \Sigma(Xi\text{-}mean)(Yi\text{-}mean)/[(\Sigma(Xi\text{-}mean)^2\Sigma(Y_i\text{-}mean)^2]^{1/2}$

measures linear relationships

Correlations	Column 1	Column 2	Column 3	Column 4
Column 1	1			
Column 2	0.994337974201009	1		
Column 3	0.41724514983494561	0.340842943122361	1	
Column 4	1	0.994337974201009	0.41724514983494561	1

$Cov(X,Y) = \Sigma(X_i\text{-}mean)(Yi\text{-}mean)/n\text{-}1$

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	389.15953333333	8205.792378328		
Column 3	129.603125	1978.926692444	4540.3508984375	

		45		
		389.159		
		5333333		
Column 4	21.25	33	129.603125	21.25

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

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

0.567200881879081	18.5830577461206
0.286589175300026	11.0518791531133
0.439273671285073	16.6355785697159
3.91700593310644	5
1084.00191396833	1383.71237174595

Regression Statistics

R^2	0.43927367 1285073
Standard Error	16.6355785 697159
Count of X variables	1
Observations	7
Adjusted R^2	0.32712840 5542087

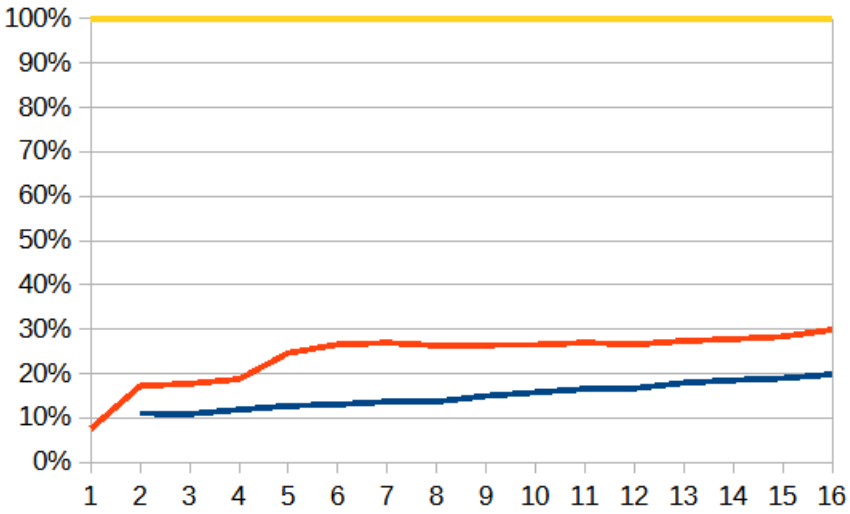
Analysis of Variance (ANOVA)

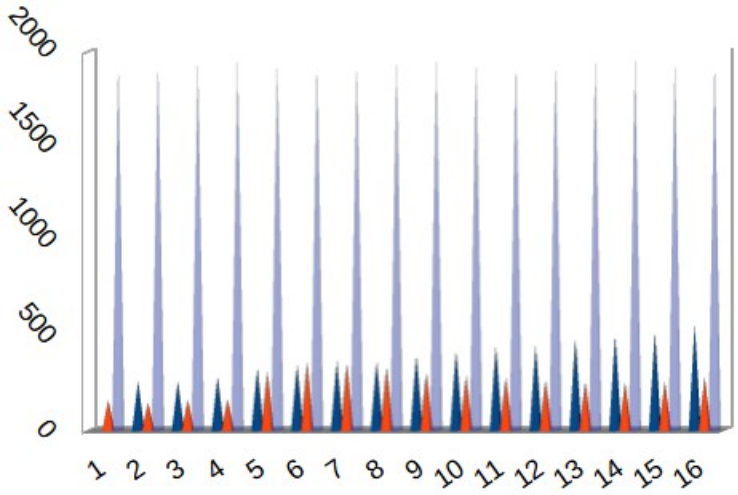
	df	SS	MS	F	Significance F
Regression	1	1084.001913961833	1084.001913961833	3.917005933106443	0.1046927833774190
Residual	5	1383.712371745950	276.7424743491900		
Total	6	2467.714285714290			

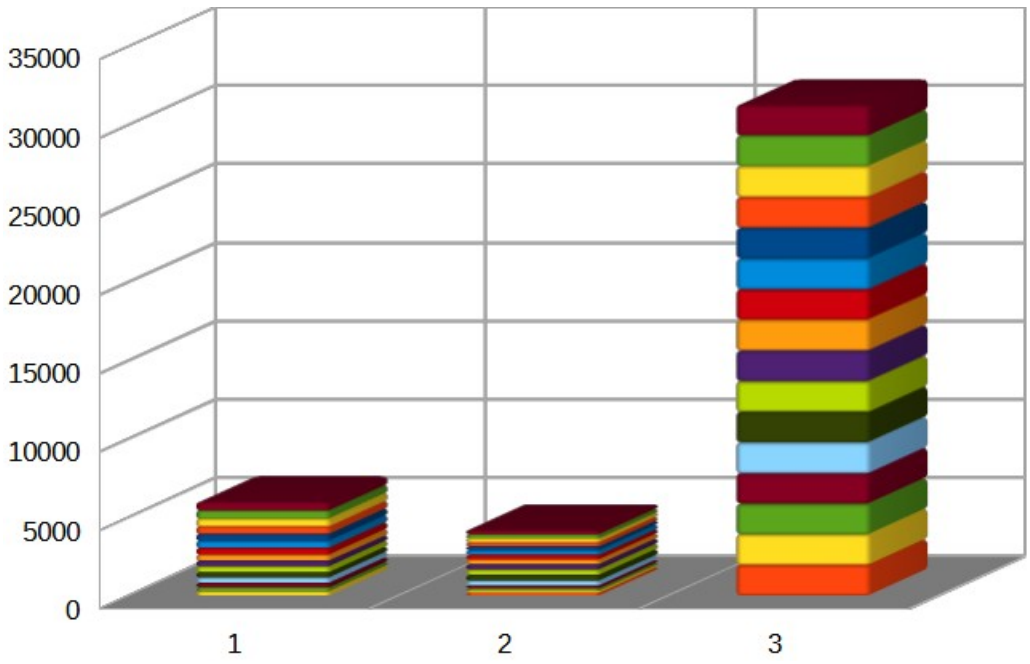
Confidence level 0.95

	Coefficients	Standard Error	t-Statistic	P-value	Lower 95%	Upper 95%
Intercept	18.58305774612060	11.05187915311330	1.681438738939320	0.1535064448647990	-9.826702054520130	46.99281754675870
CP	0.5672008818790810	0.2865891753000260	1.979142726815440	0.1046927833774190	-0.1695000464371580	1.303901810195320

CP	Predicted Y	SP	Residual
10	24.25506651491140	34	9.744933435085080
35	38.43508861188840	45	6.564911381081160
34	37.86788773000930	60	22.13211226999070
30	35.59908420249300	32	-3.599084202493000
10	24.25506651491140	20	-4.255066564911390
23	31.62867802933940	5	-26.62867802933940
80	63.95912829644700	60	-3.959128296447030


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16


2000
1500
1000
500
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16


35000
30000
25000
20000
15000
10000
5000
0
1
2
3