

# SPSS Report

Safalya Pal

A90555919003 Sem-5

STAT305

---

[27/8/2021]

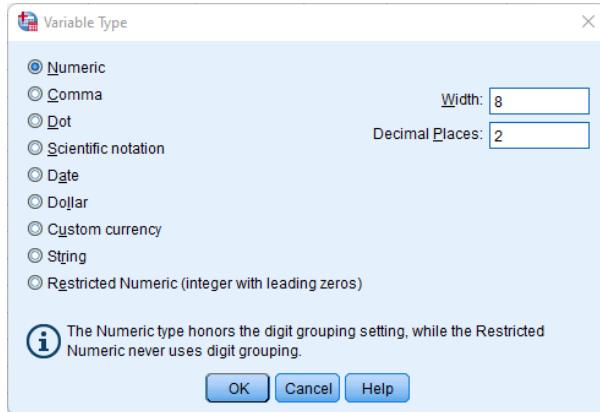
## Variable View



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Var1	Numeric	8	2		None	None	8	Right	Scale	Input
2											
3											

Name - Name of the variable.

Type - The types of values the variable can take.



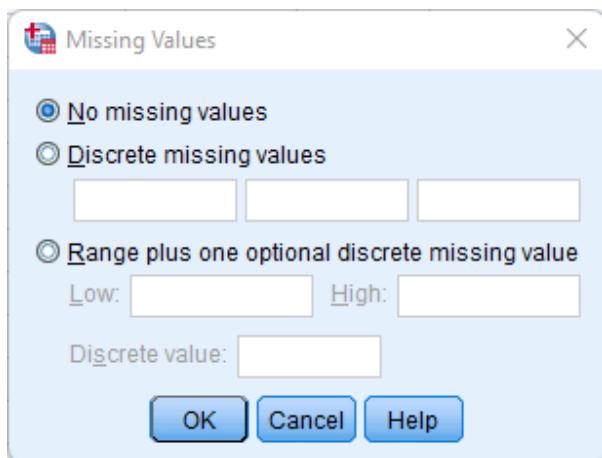
*Variable type selector popup window*

**Width** - Specifies the column width for the display of variables in the Data Editor.

**Decimals** - No. of decimal places permitted by the variable.

**Label** - Description for the variable.

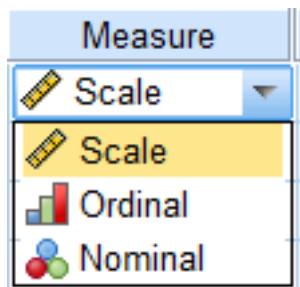
**Values** - Values which can be taken by the variable.



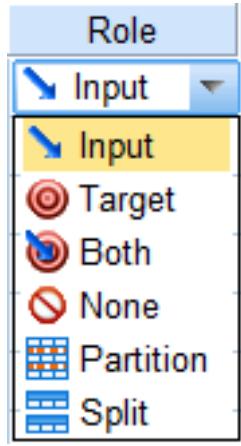
Missing values editor

**Align** - Alignment of the values in the cell of the Data Editor view.

**Measure** - SPSS measurement levels are limited to nominal(i.e. categorical), ordinal(i.e. ordered like 1st, 2nd, 3rd...), or scale. Essentially, a scale variable is a measurement variable.



**Role** - Specifies the variable's role in analysis.



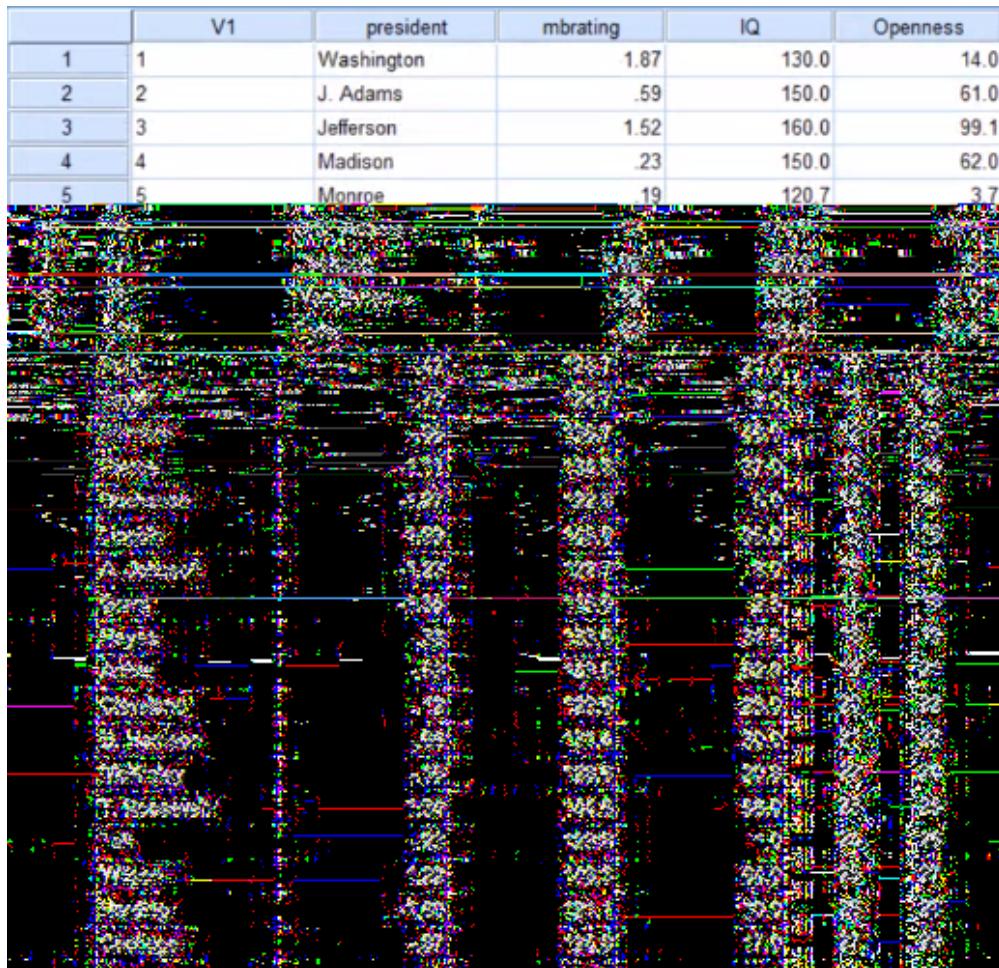
## Classwork

### Variable View

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	President	String	25	0		None	None	8	Left	Nominal	Input
2	MB_Rating	Numeric	25	2	M B Rating of p...	None	None	8	Right	Scale	Input
3	IQ	Numeric	25	1		None	None	8	Right	Scale	Input

## Data View

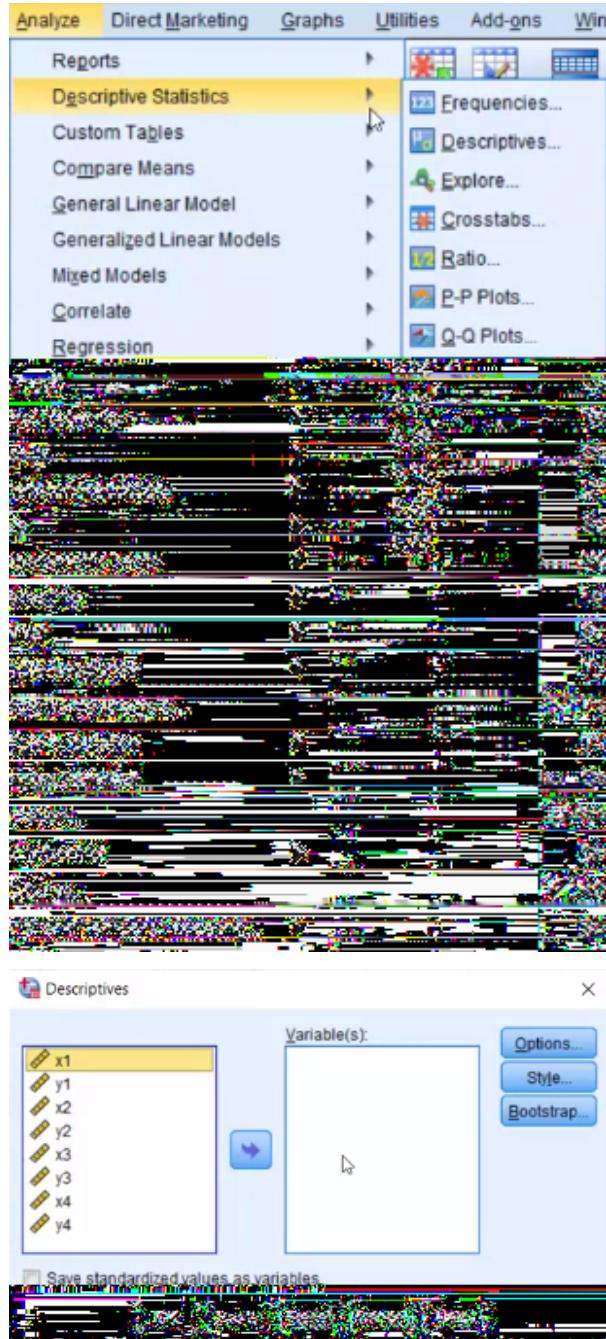
	V1	president	mbrating	IQ	Openness
1	1	Washington	1.87	130.0	14.0
2	2	J. Adams	.59	150.0	61.0
3	3	Jefferson	1.52	160.0	99.1
4	4	Madison	.23	150.0	62.0
5	5	Monroe	.19	120.7	3.7



[1/9/2021]

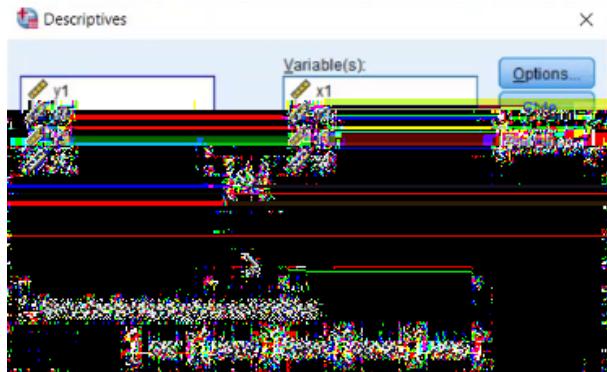
## Descriptive Analysis

The Descriptives menu lies under Analyze → Descriptive Statistics → Descriptives...

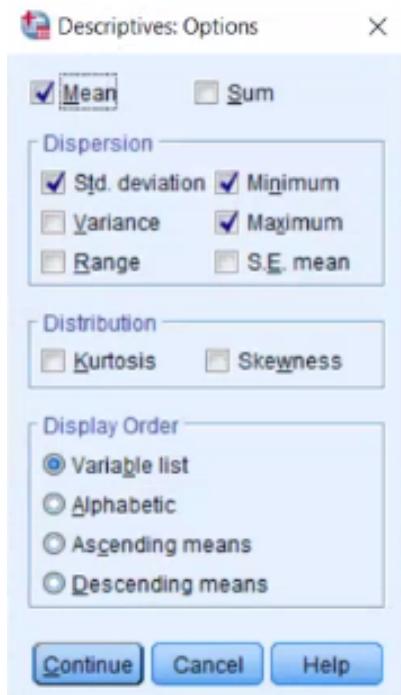


The descriptives popup window

Once the window is open, select the variables you want to perform the analysis on and click the arrow in the middle to move them to the variables side.



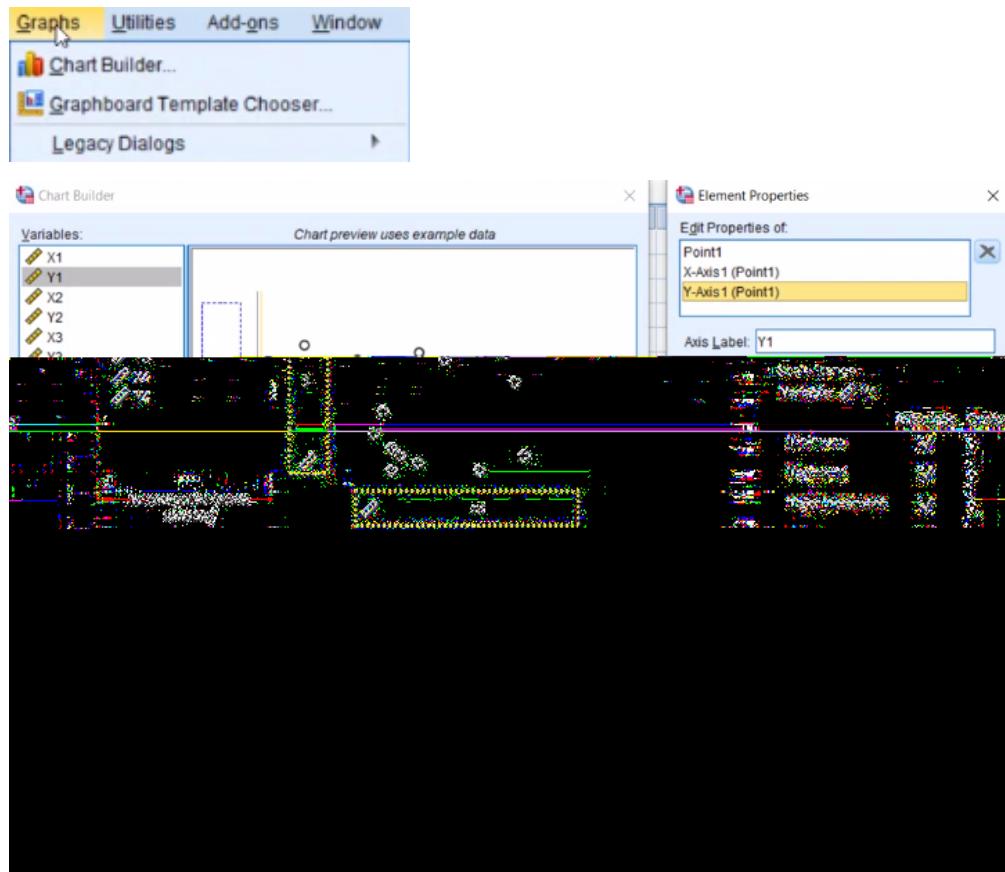
The descriptives window after moving the variables



Options available for descriptive analysis

## Chart Builder

To build a graph, Chart Builder is used which is available under Graphs → Chart Builder



*Chart Builder Window*

**Variables** - A list of variables which can be used to build the graph.

**Gallery** - Types of plots for the user to choose from.

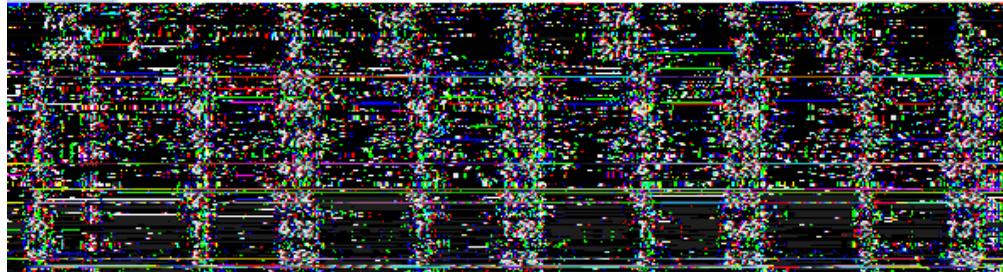
**Element Properties** - This window allows the variables to be modified in the graph.

Scatterplot   Alternate   Method   Graphs   →   Legacy Dialogs   →  
Scatter/Dot...

## Classwork

### Data View

	x1	y1	x2	y2	x3	y3	x4	y4
1	10	8.04	10	9.14	10	7.46	8	6.58
2	8	6.95	8	8.14	8	6.77	8	5.76



## Output View

### Descriptives

[DataSet1] C:\Users\Abhirup Moitra\Pictures\anscombe.sav

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
X1	11	4.00	14.00	9.0000	3.31662	11.000	.000	.661	-1.200	1.279
X2	11	4.00	14.00	9.0000	3.31662	11.000	.000	.661	-1.200	1.279
X3	11	4.00	14.00	9.0000	3.31662	11.000	.000	.661	-1.200	1.279
X4	11	8.00	19.00	9.0000	3.31662	11.000	3.317	.661	11.000	1.279
Valid N (listwise)	11									

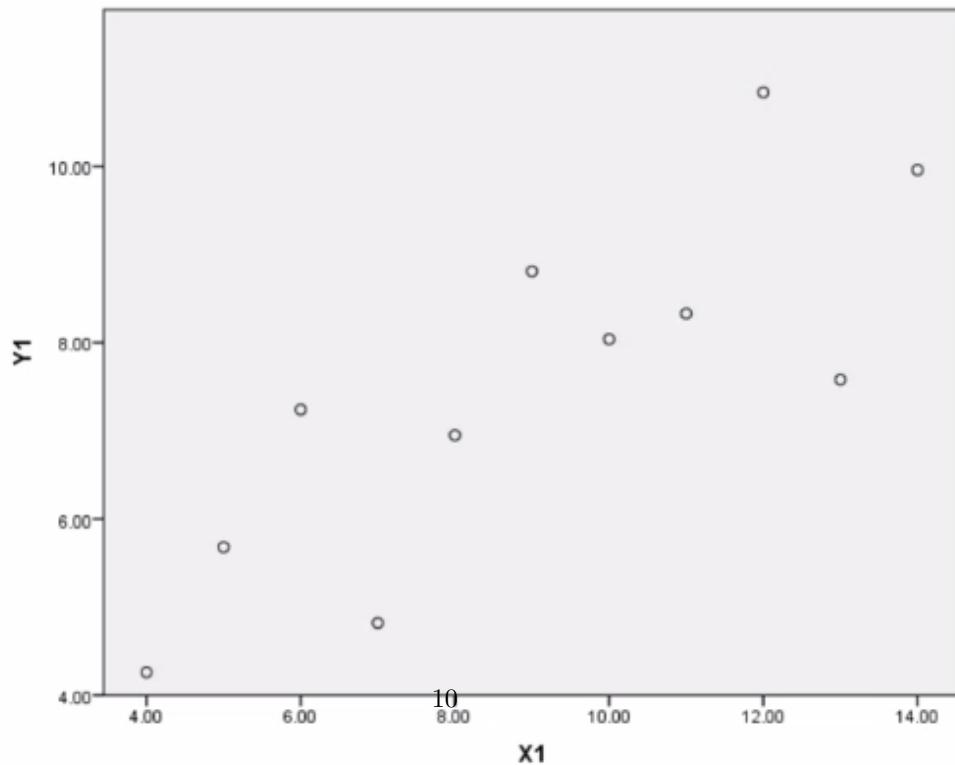
DESCRIPTIVES VARIABLES=Y1 Y2 Y3 Y4  
/STATISTICS=MEAN STDDEV VARIANCE MIN MAX KURTOSIS SKEWNESS.

### Descriptives

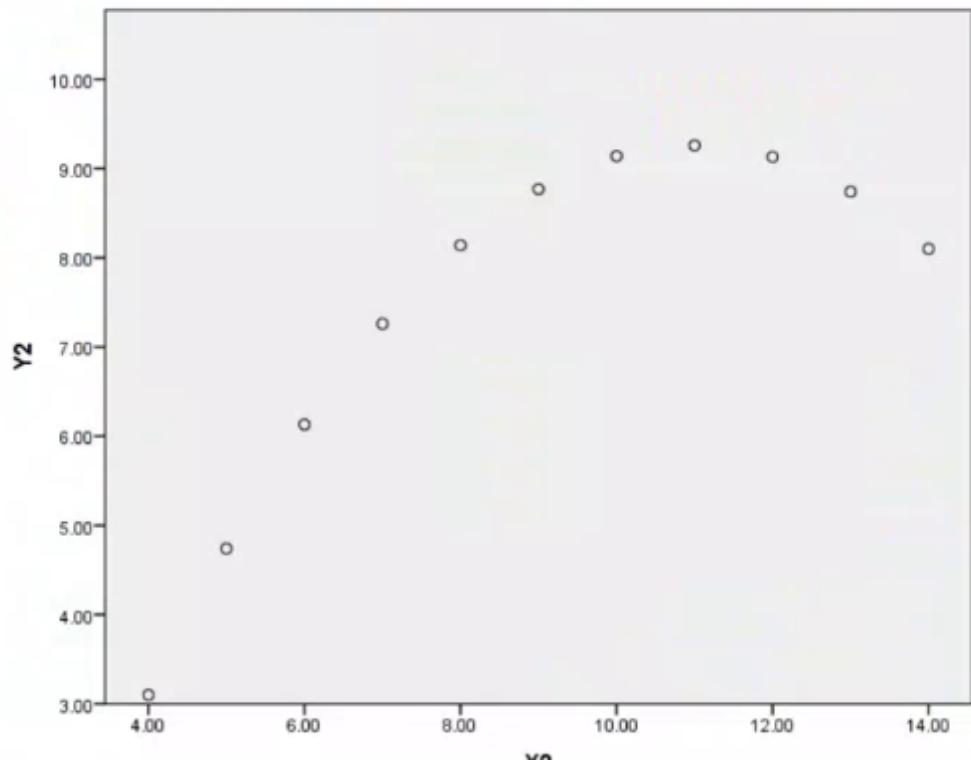
Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness	Kurtosis
Y1	11	4.00	14.00	9.0000	3.31662	11.000	.000	.661
Y2	11	4.00	14.00	9.0000	3.31662	11.000	.000	.661
Y3	11	4.00	14.00	9.0000	3.31662	11.000	.000	.661
Y4	11	8.00	19.00	9.0000	3.31662	11.000	3.317	.661

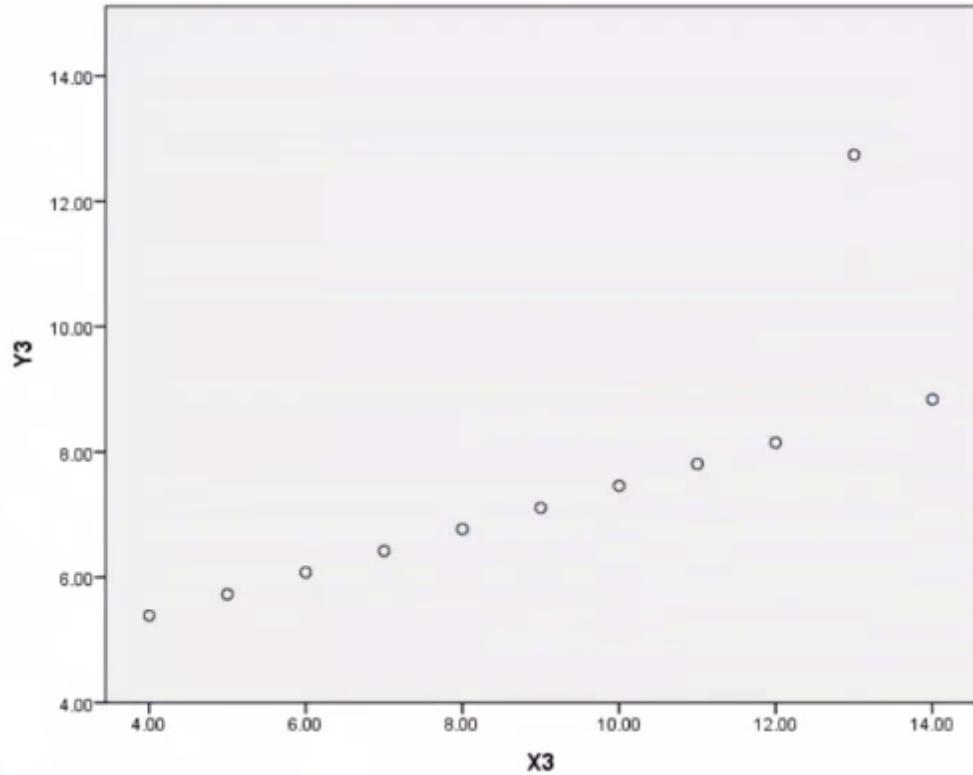
### GGraph



## GGraph



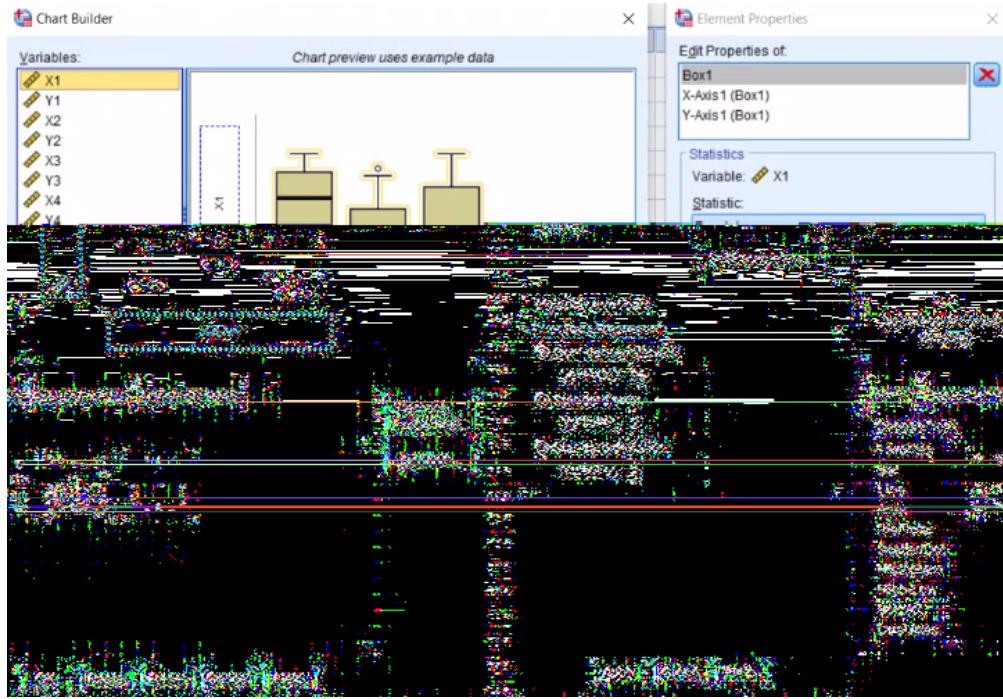
## **GGraph**



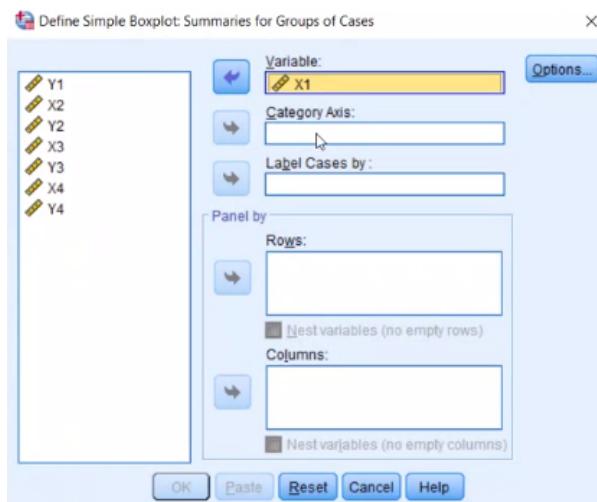
[3/9/2021]

## **Outlier Detection with Boxplots**

The Boxplot option is available in Chart Builder under Gallery



### Alternate Method Graphs → Legacy Dialogs → Boxplot



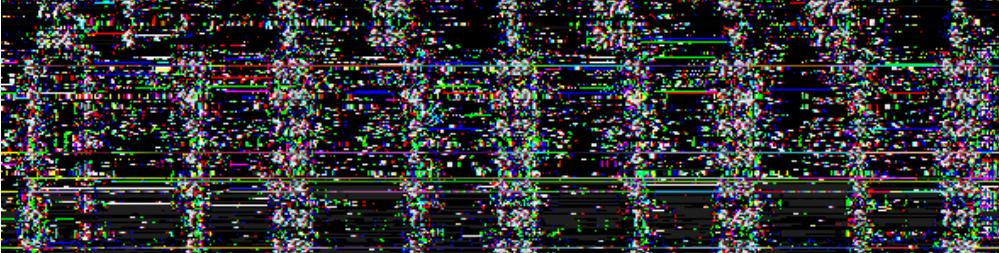
*Legacy Boxplot window*

Panel by - Separates into different boxes based on variables in this option

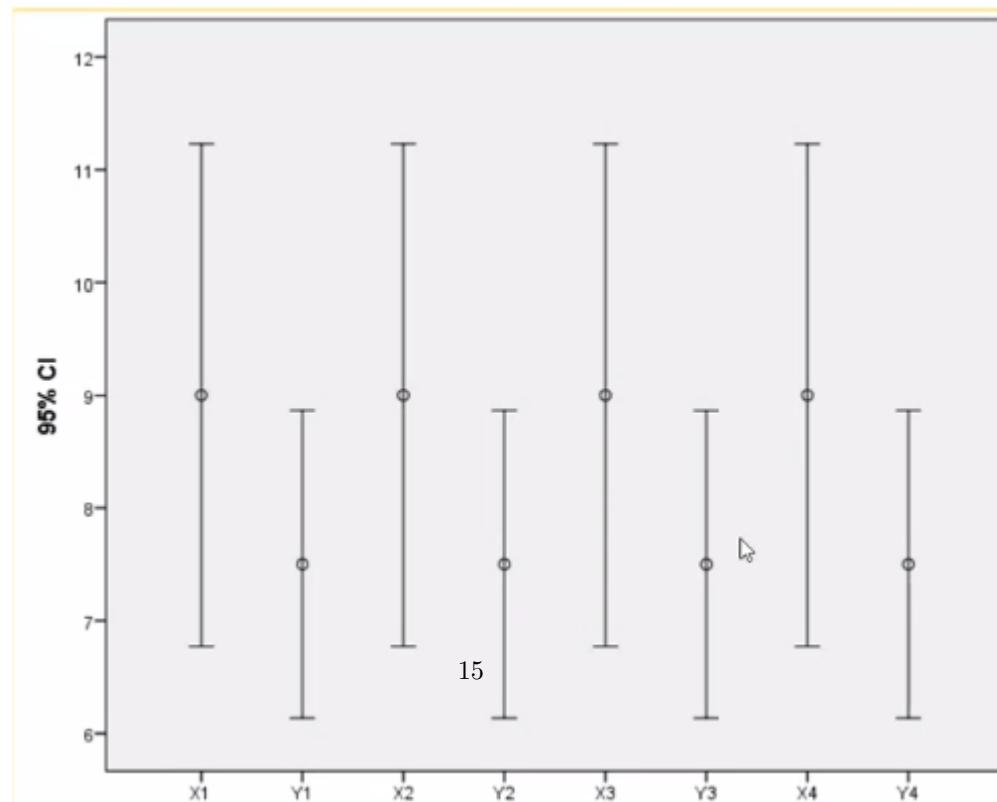
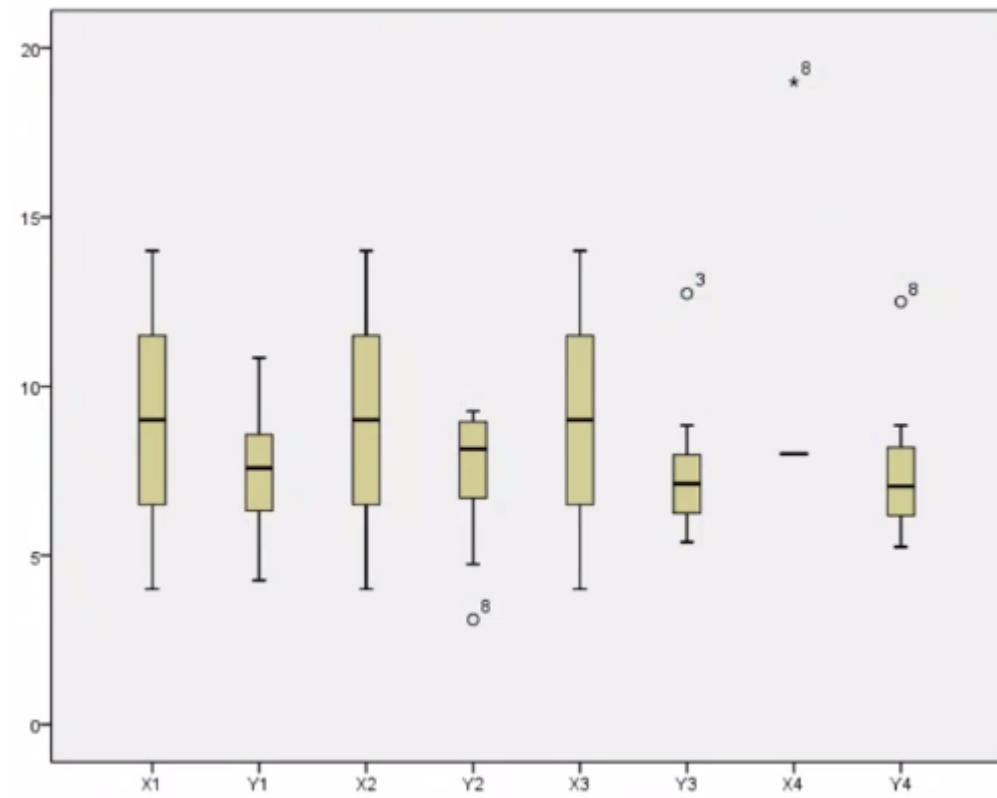
## Classwork 1

### Data View

	x1	y1	x2	y2	x3	y3	x4	y4
1	10	8.04	10	9.14	10	7.46	8	6.58
2	8	6.95	8	8.14	8	6.77	8	5.76



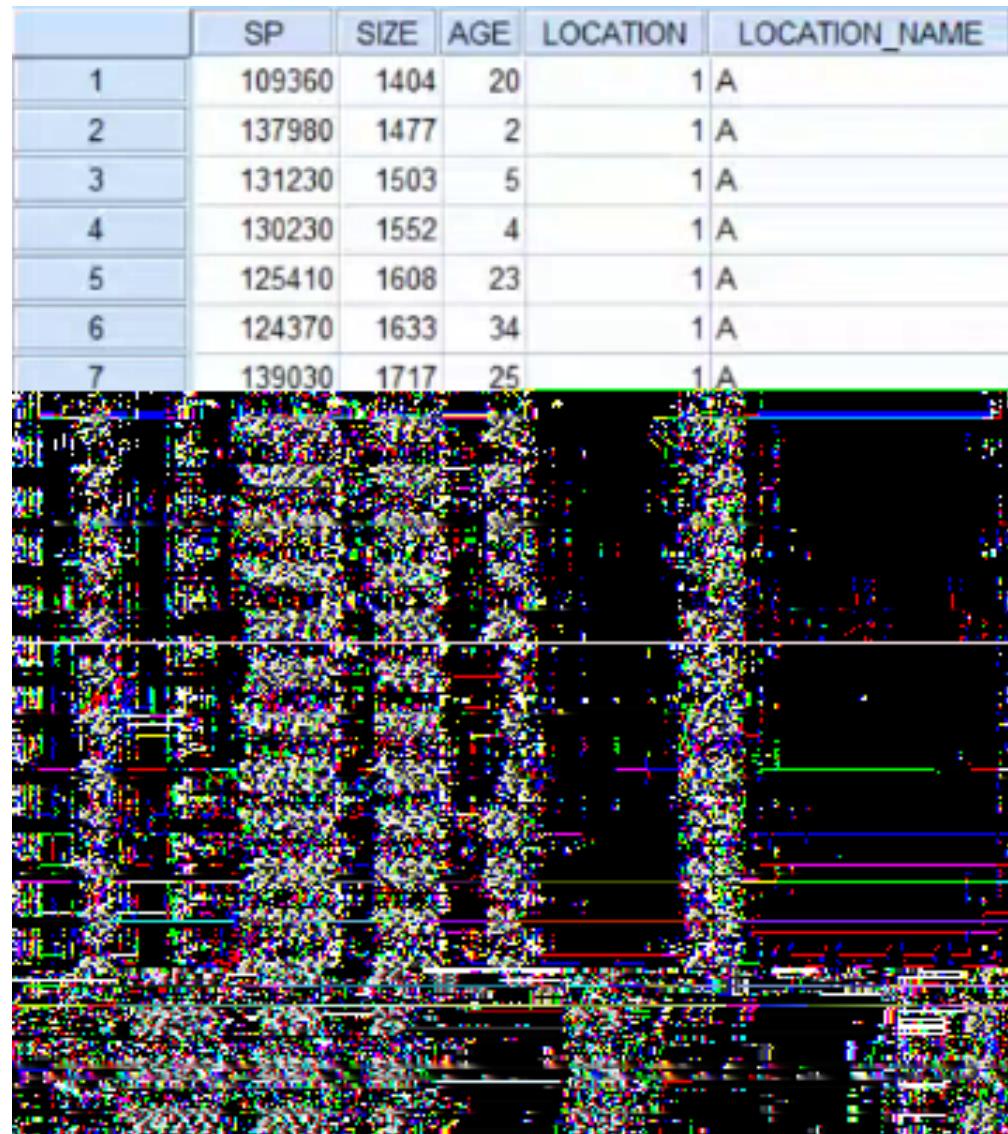
## Output View



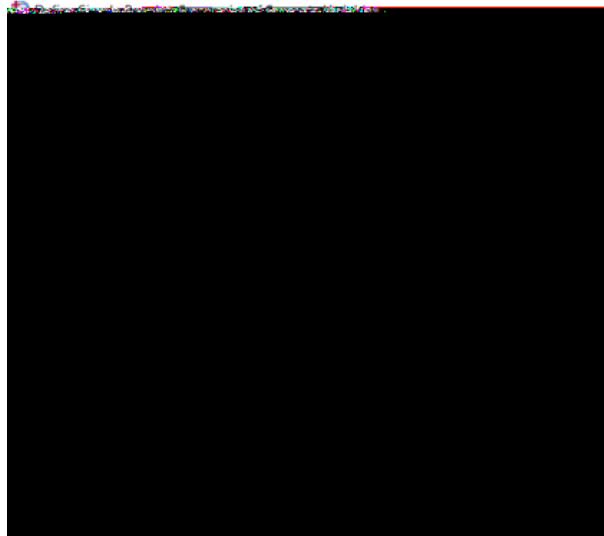
*95% Confidence interval plot*

## Classwork 2

### Data View



Boxplot Window



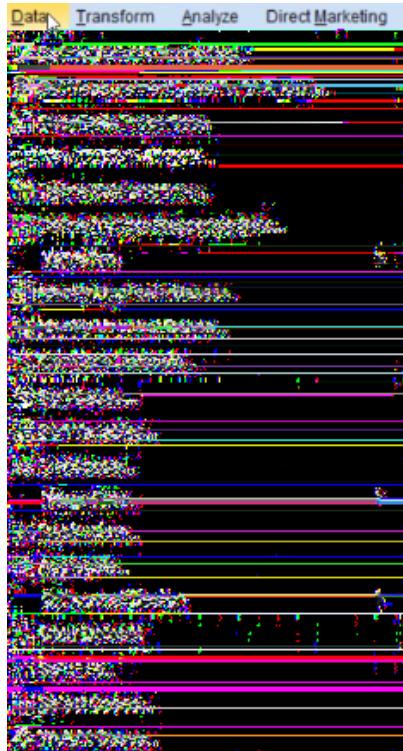
Output View



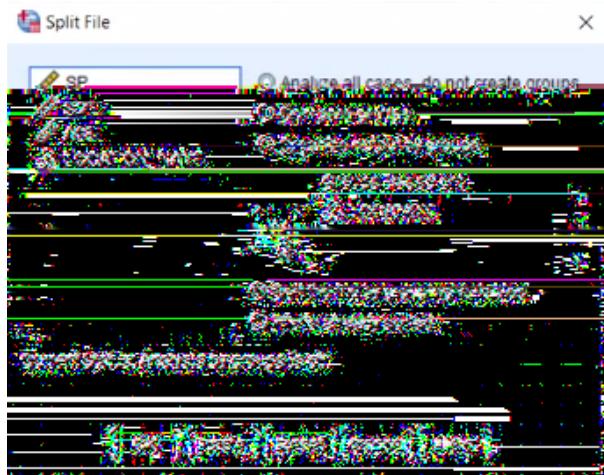
[8/9/2021]

## Splitting Files into Groups

To split the data w.r.t some variable, we use **Split File** under **Data**—>**Split File...**



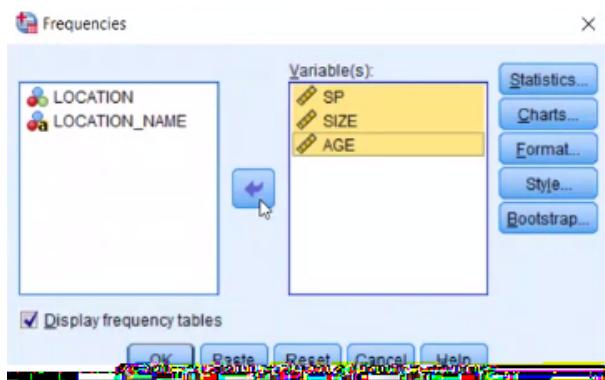
To split the file based on a variable, the variables must be in the **Group Based on:** dialogue.



## Frequencies for Descriptive Statistics

To display the descriptive statistics of a dataset in a vertical tabular form, an alternative method exists i.e. by using the **Frequencies...** dialogue.

The Frequencies window is under **Analyze**—>**Descriptive Statistics**—>**Frequencies...**



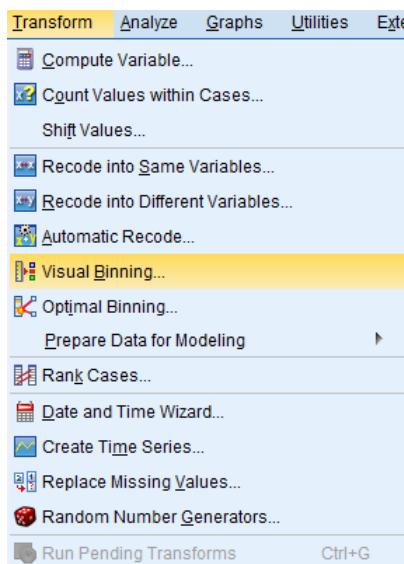
Variable selector window

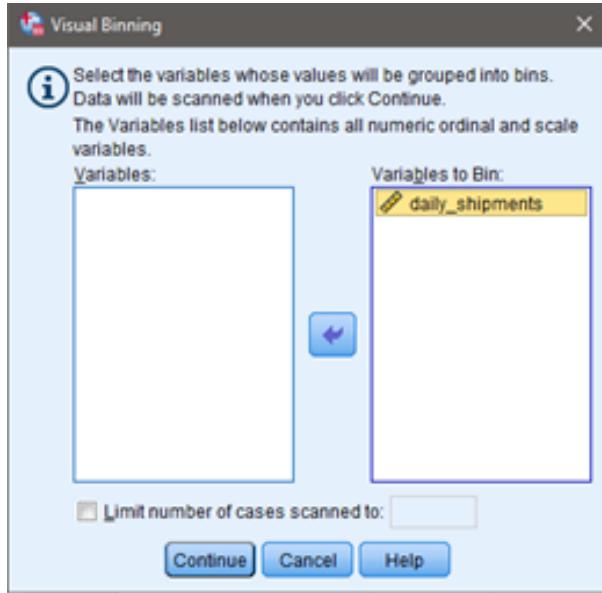


Stats selector window

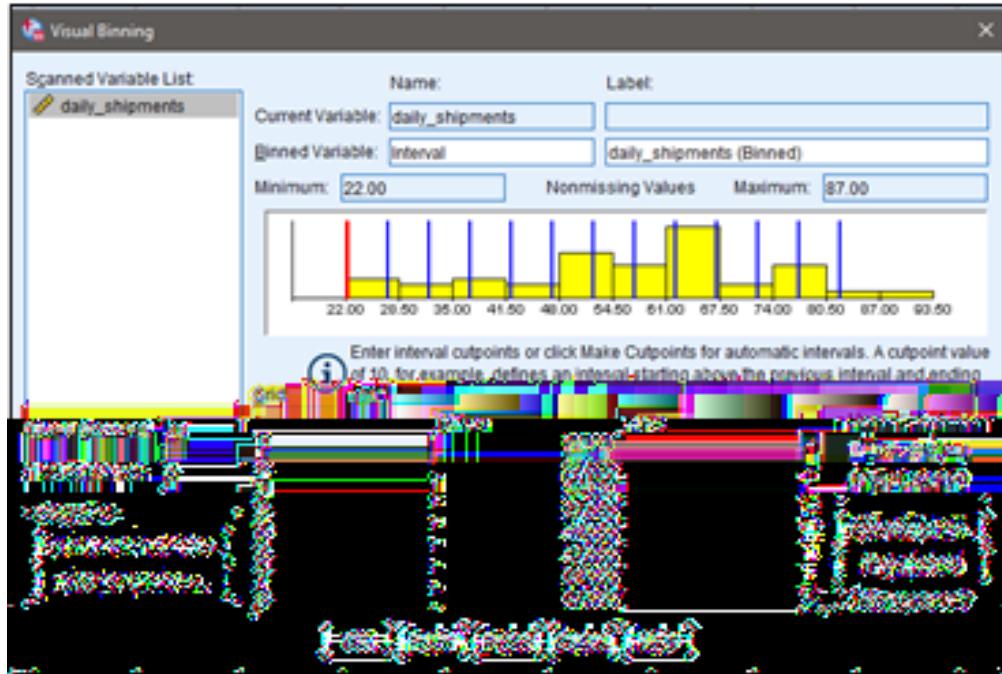
## Visual Binning

To create class intervals for continuous data, we use the **Visual Binning** dialogue under **Transform**—>**Visual Binning**



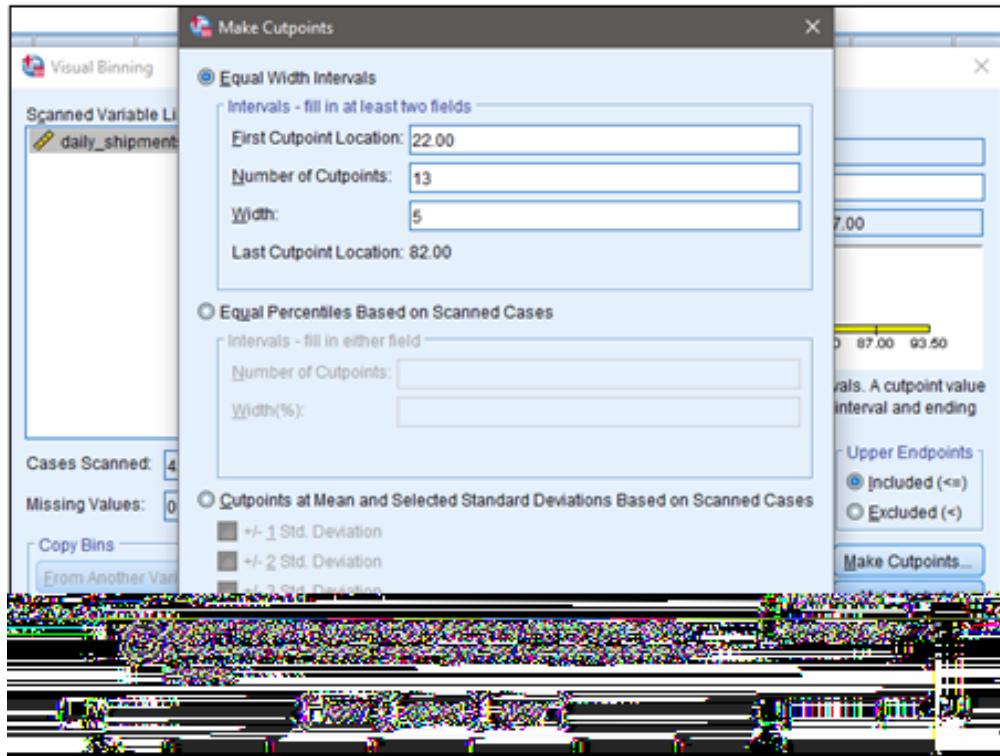


Variable Selector Window



Visual Binning Window

To specify the cutpoint properties, use the Make Cutpoint options



## Classwork 1

### Variable View

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	daily_shipm...	Numeric	8	2		None	None	18	Right	Scale	Input
2											

## Data View

	daily_shipments	var
1	22.00	
2	65.00	
3	65.00	
4	67.00	
5	55.00	
6	50.00	
7	65.00	
8	77.00	
9	73.00	
10	30.00	
11	62.00	
12	54.00	
13	48.00	
14	65.00	
15	79.00	
16	60.00	
17	63.00	
18	45.00	
19	51.00	
20	68.00	
21	79.00	
22	63.00	
23	33.00	
24	44.00	

Data View Variable View

## Post Binning Data View

	daily_shipments	Interval	var
1	22.00	1	
2	65.00	10	
3	65.00	10	
4	67.00	10	
5	55.00	8	
6	50.00	7	
7	65.00	10	
8	77.00	12	
9	79.00	12	
10	30.00	3	
11	62.00	9	
12	54.00	8	
13	48.00	7	
14	65.00	10	
15	79.00	13	
16	60.00	9	
17	63.00	10	
18	45.00	6	
19	51.00	7	
20	68.00	11	
21	79.00	13	
22	83.00	14	
23	33.00	4	
24	xx.00	6	

#

[Data View](#) [Variable View](#)

## Classwork 2

### Variable View

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Sales	Numeric	8	2		None	None	18	Right	Scale	Input
2											

## Data View

	Sales	var
1	13.00	
2	19.00	
3	22.00	
4	14.00	
5	13.00	
6	16.00	
7	19.00	
8	21.00	
9	23.00	
10	11.00	
11	27.00	
12	25.00	
13	17.00	
14	17.00	
15	16.00	
16	20.00	
17	23.00	
18	19.00	
19	26.00	
20	20.00	
21	22.00	
22	16.00	
23	28.00	
24	25.00	

### Post Binning Data View

	Sales	interval	var
1	13.00	2	
2	19.00	2	
3	22.00	3	
4	14.00	2	
5	13.00	2	
6	16.00	2	
7	19.00	2	
8	21.00	3	
9	23.00	3	
10	11.00	2	
11	27.00	3	
12	25.00	3	
13	17.00	2	
14	17.00	2	
15	13.00	2	
16	20.00	2	
17	23.00	3	
18	24.00	2	
19	24.00	3	
20	21.00	2	
21	23.00	3	
22	23.00	2	
23	18.00	2	
24	24.00	2	

[10/9/2021]

### Classwork 1

#### Variable View

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Number	8	2	None	None	None	2	Right	Scale	Input
2	String	8	0	Second contact	None	None	2	Right	Text	Output
3	Number	8	2	None	None	None	2	Right	Scale	Output
4	Number	8	2	None	None	None	2	Right	Scale	Output
5	String	8	0	Gender	None	None	2	Right	Text	Output

## Data View

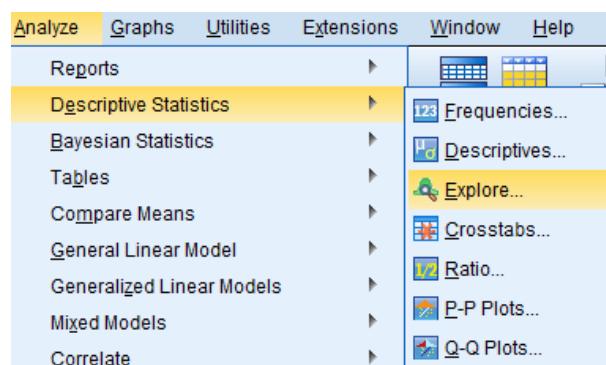
	rollno	name	subject1	subject2	subject3
1	1	Aditi	89.00	87.00	78.00
2	2	Anushka	56.00	76.00	98.00
3	3	Ritika	78.00	75.00	89.00
4	4	Rishika	67.00	88.00	76.00
5	5	Aish	75.00	99.00	56.00
6	6	Laxmi	67.00	79.00	56.00
7	7	Safalya	89.00	68.00	46.00
8	8	Shoumit	78.00	77.00	76.00
9	9	Abhirup	80.00	66.00	65.00
10	10	Shougata	97.00	96.00	76.00

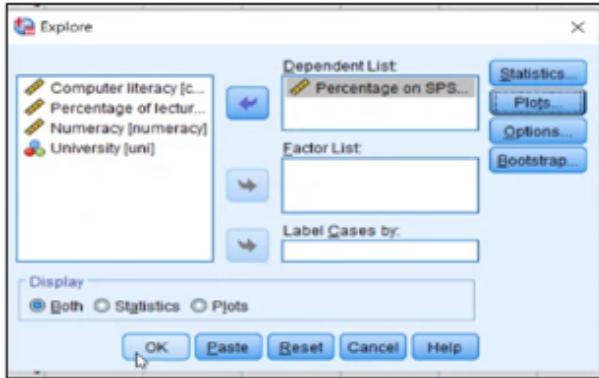
## Output View

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
subject1	10	56.00	97.00	77.6000	12.20382
subject2	10	66.00	99.00	81.1000	11.10005
subject3	10	46.00	98.00	71.6000	15.93180
Valid N (listwise)	10				

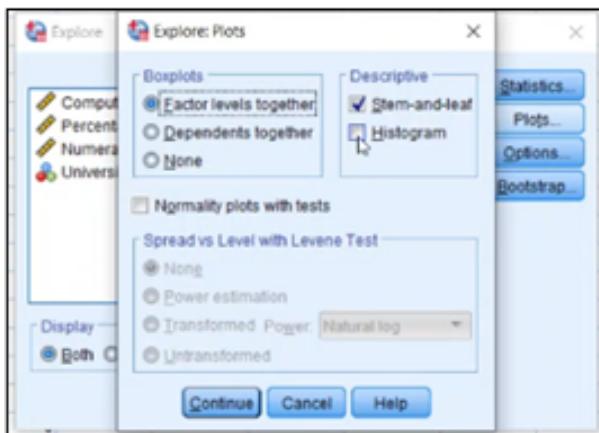
## Normality Check

To check if the data is normal or not, the **Explore** option under **Analyze**→**Descriptive Statistics**→**Explore**





Explore Window



Plot Selector Window

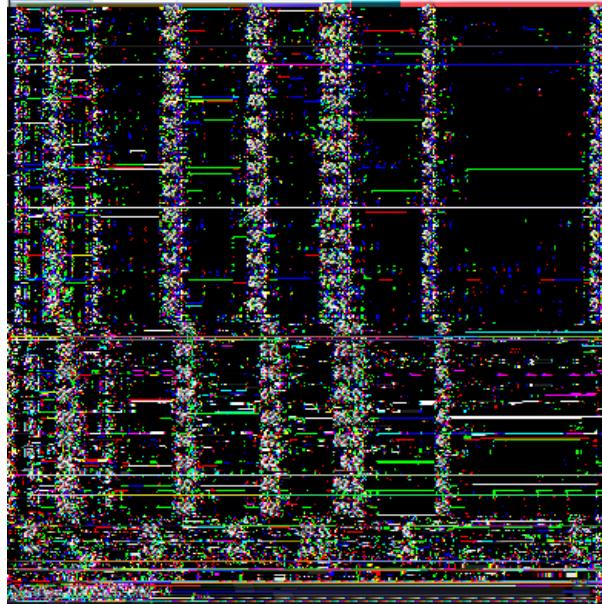
## Classwork 2

### Variable View

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Computer literacy	Scale	8	2	Computer literacy	1	1	1	Right	Continuous	Dependent Variable
2	Percentage on SPS	Scale	8	2	Percentage on SPS	1	1	1	Right	Continuous	Dependent Variable
3	Percentage of lecturer	Scale	8	2	Percentage of lecturer	1	1	1	Right	Continuous	Independent Variable
4	Numeracy	Scale	8	2	Numeracy	1	1	1	Right	Continuous	Independent Variable
5	University	Categorical	8	0	University	1	1	1	Right	Continuous	Independent Variable

## Data View

	exam	computer	lectures	numeracy	uni
1	18	54	75.0	7	0
2	30	47	8.5	1	0
3	40	58	69.5	6	0



## Output View

Explore

[DataSet1] D:\AUK\_STAT\SEM 5\spss\_dataset\Data File - SPSSExam.sav

Case Processing Summary

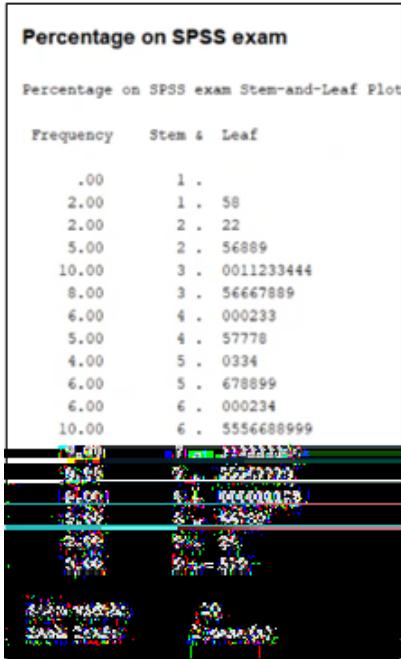
	Cases					
	Valid		Missing		Total	
N	Percent	N	Percent	N	Percent	
Percentage on SPSS exam	100	100.0%	0	0.0%	100	100.0%

Descriptives

Percentage on SPSS exam	Mean	Statistic	Std. Error
Mean	58.10	2.132	
95% Confidence Interval for Mean	Lower Bound	53.87	
	Upper Bound	62.33	
5% Trimmed Mean		58.12	
Median		60.00	
Variance		454.354	
Std. Deviation		21.316	
Minimum		15	
Maximum		99	
Range		84	
Interquartile Range		37	
Skewness		-1.07	.241
Kurtosis		-1.105	.478

Tests of Normality

	Kolmogorov-Smirnov*			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Percentage on SPSS exam	.102	100	.012	.961	100	.005

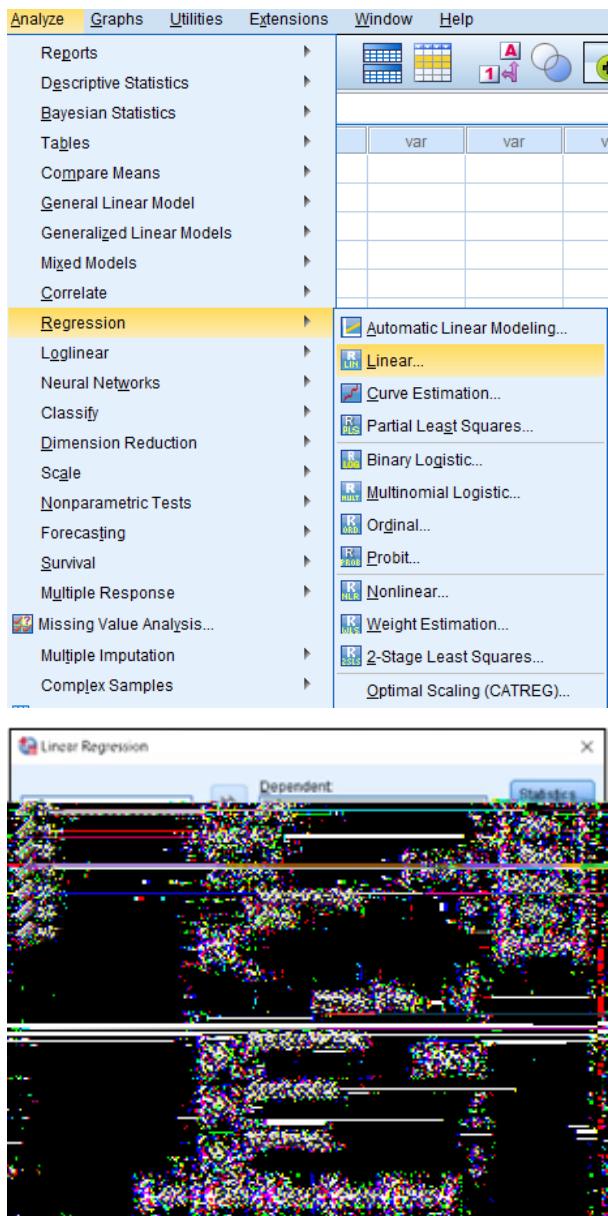


If the significance value is greater than 0.05 then the data is normal.

[16/9/2021]

## Linear Regression

To fit a linear regression model to your data, use the Linear Regression dialogue under Analyze → Regression → Linear



Linear Regression Window

## Classwork

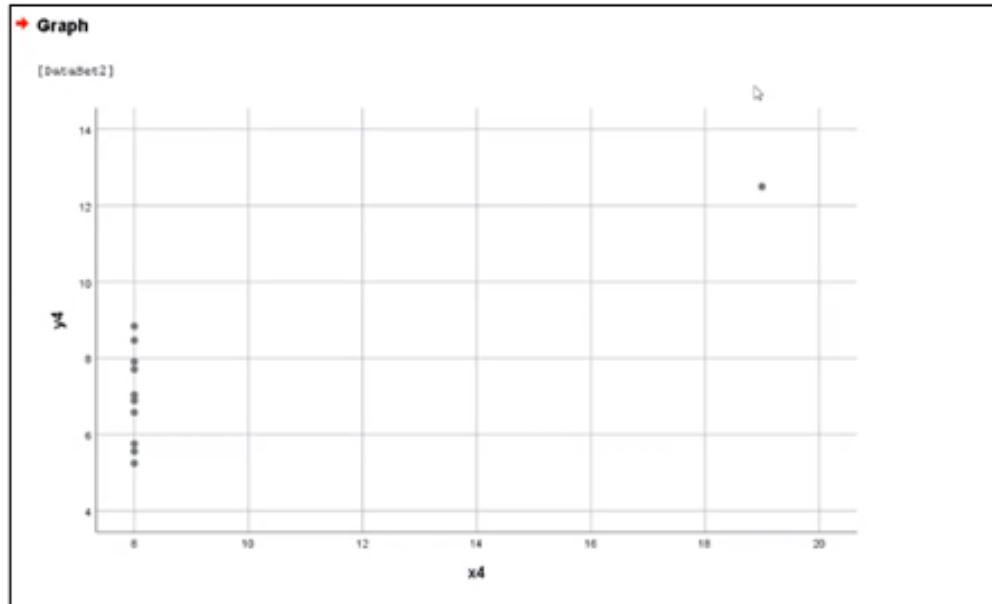
### Variable View

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	x1	Numeric	8	0		None	None	8	Right	Scale	Input
2	y1	Numeric	8	0		None	None	8	Right	Scale	Input
3	x2	Numeric	8	0		None	None	8	Right	Scale	Input
4	y2	Numeric	8	0		None	None	8	Right	Scale	Input
5	x3	Numeric	8	0		None	None	8	Right	Scale	Input
6	y3	Numeric	8	0		None	None	8	Right	Scale	Input
7	x4	Numeric	8	0		None	None	8	Right	Scale	Input
8	y4	Numeric	8	0		None	None	8	Right	Scale	Input

### Data View

	x1	y1	x2	y2	x3	y3	x4	y4
1	10	8	10	9	10	7	8	7
2	8	7	8	8	8	7	8	6
3	13	8	13	9	13	13	8	8
4	9	9	9	9	9	7	8	9
5	11	8	11	9	11	8	8	8
6	14	10	14	8	14	9	8	7
7	6	7	6	6	6	6	8	5
8	4	4	4	3	4	5	19	13
9	12	11	12	9	12	8	8	6
10	7	5	7	7	7	6	8	8
11	5	6	5	5	5	6	8	7

## Scatter Output View



The coordinate with value of  $Y4 > 18$  is an outlier. Hence we remove that coordinate

## Processed Data View

	x1	y1	x2	y2	x3	y3	x4	y4
1	10	8	10	9	10	7	8	7
2	8	7	8	8	8	7	8	6
3	13	8	13	9	13	13	8	8
4	9	9	9	9	9	7	8	9

A large, distorted image of a DNA double helix structure, likely a watermark or background image.

## Output View

Model Summary					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	
1	1.000 <sup>a</sup>	1.000	1.000	.003	

a. Predictors: (Constant), x3

ANOVA <sup>a</sup>						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	11.023	1	11.023	1160687.631	.000 <sup>b</sup>
	Residual	.000	8	.000		
	Total	11.023	9			

a. Dependent Variable: y3  
b. Predictors: (Constant), x3

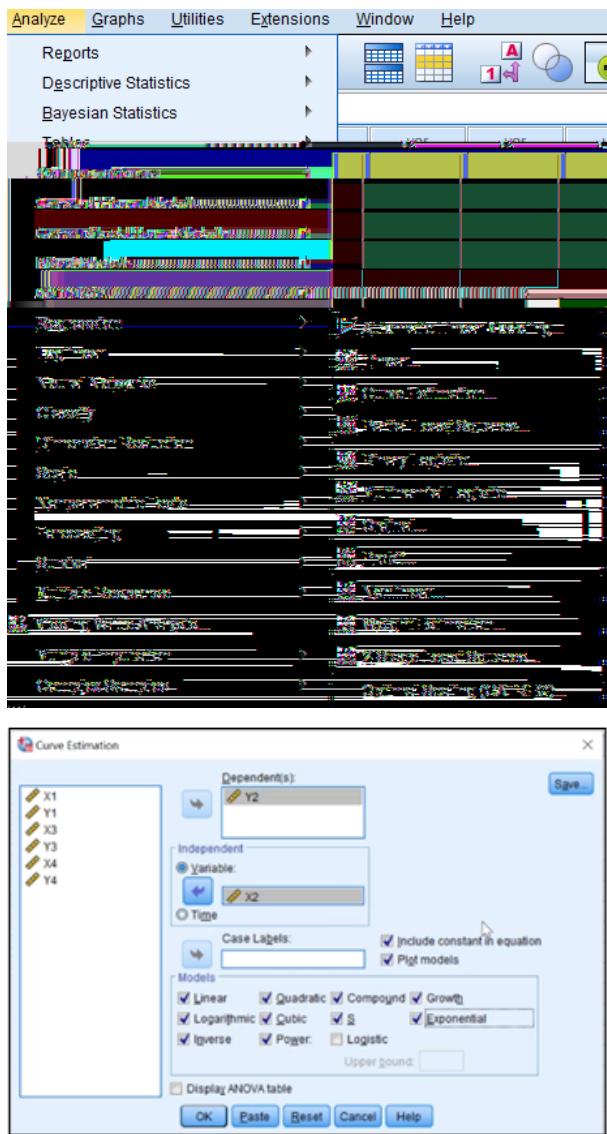
Coefficients <sup>a</sup>						
Model	Unstandardized Coefficients			Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.	
1	(Constant)	4.006	.003		1369.808	.000
	x3	.345	.000	1.000	1077.352	.000

a. Dependent Variable: y3

[20/9/2021]

## Curve Fitting

To fit a curve to your data, use the Curve Estimate window under Analyze → Regression → Curve Estimate



*Curve Estimate Window*

**Models** - Specify the type of curve(s) to be fitted to the data

## Classwork

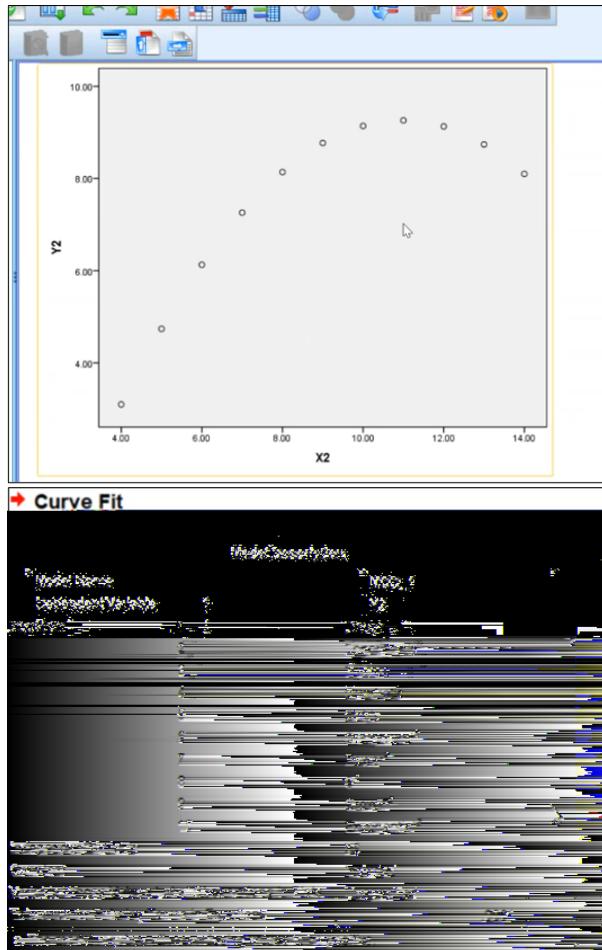
### Variable View

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
	X1	Numeric	8	2		None	None	8	Right	Unknown	Input
	Y1	Numeric	8	2		None	None	8	Right	Unknown	Input
	X2	Numeric	8	2		None	None	8	Right	Unknown	Input
	Y2	Numeric	8	2		None	None	8	Right	Unknown	Input
	X3	Numeric	8	2		None	None	8	Right	Unknown	Input
	Y3	Numeric	8	2		None	None	8	Right	Unknown	Input
	X4	Numeric	8	2		None	None	8	Right	Unknown	Input
	Y4	Numeric	8	2		None	None	8	Right	Unknown	Input

### Data View

	X1	Y1	X2	Y2	X3	Y3	X4	Y4
	10.00	8.04	10.00	9.14	10.00	7.46	8.00	6.58
	8.00	6.95	8.00	8.14	8.00	6.77	8.00	5.76
	13.00	7.58	13.00	8.74	13.00	12.74	8.00	7.71
	9.00	8.81	9.00	8.77	9.00	7.11	8.00	8.84
	11.00	8.33	11.00	9.26	11.00	7.81	8.00	8.47
	14.00	9.96	14.00	8.10	14.00	8.84	8.00	7.04
	6.00	7.24	6.00	6.13	6.00	6.08	8.00	5.25
	4.00	4.26	4.00	3.10	4.00	5.39	19.00	12.50
	12.00	10.84	12.00	9.13	12.00	8.15	8.00	5.56
	7.00	4.82	7.00	7.26	7.00	6.42	8.00	7.91
	5.00	5.68	5.00	4.74	5.00	5.73	8.00	6.89

## Output View



Case Processing Summary	
	N
Total Cases	11
Excluded Cases <sup>a</sup>	0
Forecasted Cases	0
Newly Created Cases	0

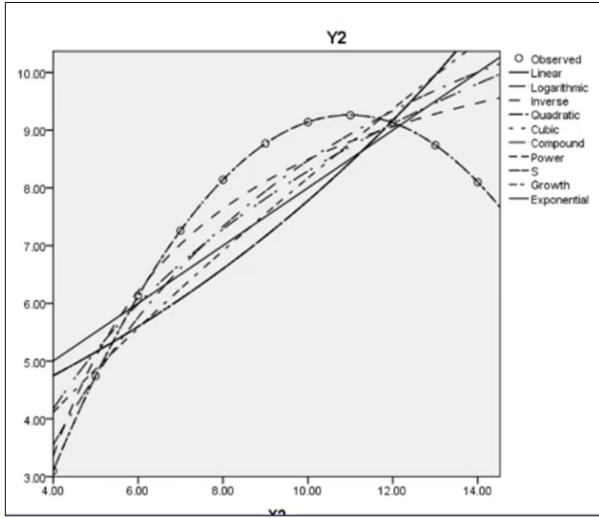
a. Cases with a missing value in any variable are excluded from the analysis.

Variable Processing Summary		Variables	
		Dependent	Independent
		Y2	X2
Number of Positive Values		11	11
Number of Zeros		0	0
Number of Negative Values		0	0
Number of Missing Values	User-Missing	0	0
	System-Missing	0	0

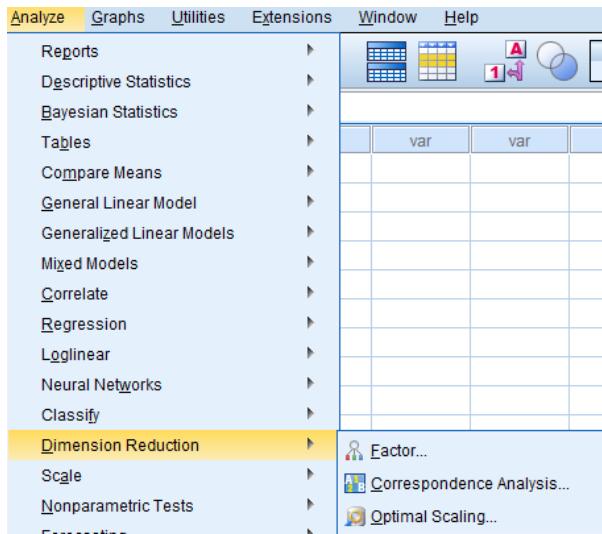
Model Summary and Parameter Estimates								
Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	Constant	b1	b2
Linear	.666	17.966	1	9	.002	3.001	.500	
Logarithmic	.812	38.830	1	9	.000	-2.070	4.499	
Inverse	.910	90.673	1	9	.000	11.915	-34.238	
Quadratic	1.000	7378133.000	2	8	.000	-5.996	2.781	-.127
Cubic	1.000	4303910.917	3	7	.000	-5.996	2.781	-.127



[21/09/2021]

## Factor Analysis

To perform factor analysis on your data, use Factor... under Analyze → Dimension Reduction → Factor...

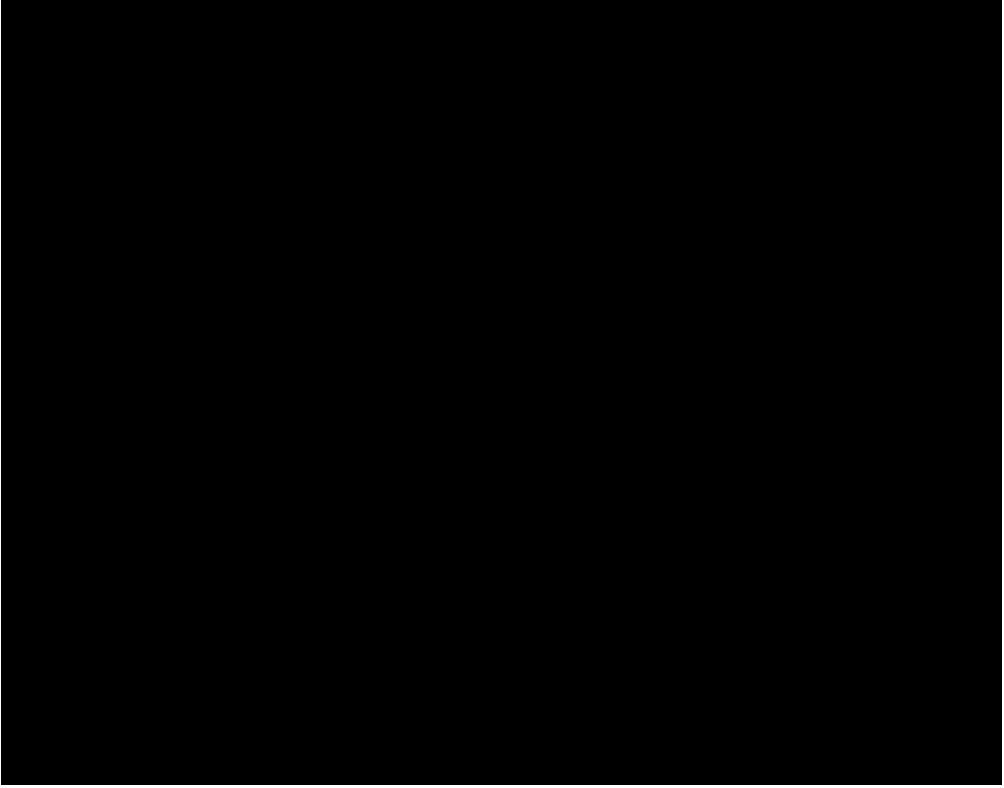


## Classwork

### Data View

	X1	Y1	X2	Y2	X3	Y3	X4	Y4	
	10.00	8.04	10.00	9.14	10.00	7.46	8.00	6.58	
	8.00	6.95	8.00	8.14	8.00	6.77	8.00	5.76	
	13.00	7.58	13.00	8.74	13.00	12.74	8.00	7.71	
	9.00	8.81	9.00	8.77	9.00	7.11	8.00	8.84	
	11.00	8.33	11.00	9.26	11.00	7.81	8.00	8.47	
	14.00	9.96	14.00	8.16	14.00	8.84	8.00	7.03	
	6.00	7.24	5.00	6.43	6.00	6.08	8.00	6.93	
	4.00	4.76	4.00	4.02	4.00	4.00	4.00	3.89	
	12.00	10.84	12.00	9.73	12.00	8.74	10.00	6.86	
	7.00	4.82	7.00	7.26	7.00	6.42	8.00	7.91	
	5.00	5.68	5.00	4.74	5.00	5.73	8.00	6.89	

## Output View



Coefficients <sup>a</sup>								
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		
	B	Std. Error	Beta			Lower Bound	Upper Bound	
1	(Constant)	3.000	1.125	2.667	.026	.456	5.544	
	X1	.500	.118	.816	4.241	.002	.233	.767

a. Dependent Variable: Y1

Descriptive Statistics													
	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis		
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error	
Y1	11	6.58	4.26	10.84	7.5009	.61254	2.03157	4.127	-.065	.661	-.535	1.279	
Y2	11	6.16	3.10	9.26	7.5009	.61257	2.03166	4.128	-1.316	.661	.846	1.279	
Y3	11	7.35	5.39	12.74	7.5000	.61220	2.03042	4.123	1.855	.661	4.384	1.279	
Y4	11	7.25	5.25	12.50	7.5009	.61224	2.03058	4.123	1.507	.661	3.151	1.279	
Valid N (listwise)	11												

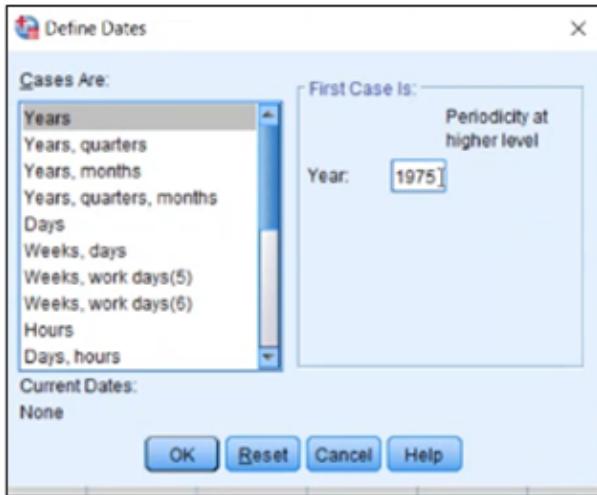
```

NEW FILE.
DATASET NAME DataSet4 WINDOW=FRONT.
DATASET ACTIVATE DataSet1.
DATASET CLOSE DataSet4.
DATASET ACTIVATE DataSet3.
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS CI(95) R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT Y1
  /METHOD=ENTER X1.
  
```

[22/09/2021]

## Date and Time

To define a Date and Time column, use Data → Define Date and Time



Deefine Data and Time window

## Classwork

### Variable View

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	USA	Numerical	8	2		None	None	8	Right	Scale	Input
2	Canada	Numerical	8	2		None	None	8	Right	Scale	Input
3	YEAR	Text	4	0	YEAR	None	None	4	Right	Scale	Input
4	DATE	Date	8	0	DATE	None	None	7	Left	Scale	Input

### Data View

	USA	Canada	YEAR	DATE
1	61.79	41.75	1975	1975
2	64.37	47.99	1976	1976
3	70.25	58.12	1977	1977
4	79.69	73.31	1978	1978
5	91.52	84.77	1979	1979
6	100.00	100.00	1980	1980
7	105.91	110.56	1981	1981
8	111.09	115.49	1982	1982
9	114.21	117.72	1983	1983
10	118.56	119.79	1984	1984
11	124.70	124.66	1985	1985
12	133.50	132.35	1986	1986
13	144.67	144.75	1987	1987
14	153.78	162.98	1988	1988
15	162.55	196.04	1989	1989
16	170.77	227.07	1990	1990
17	171.77	228.09	1991	1991
18	176.01	227.03	1992	1992

## Regression

[DataSet0]

### Variables Entered/Removed<sup>a</sup>

Model	Variables Entered	Variables Removed	Method
1	USA <sup>b</sup>	.	Enter

a. Dependent Variable: Canada

b. All requested variables entered.

### Model Summary<sup>b</sup>

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.980 <sup>a</sup>	.960	.959	22.29788

a. Predictors: (Constant), USA

b. Dependent Variable: Canada

### ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	347603.888	1	347603.888	699.129	.000 <sup>b</sup>
	Residual	14418.664	29	497.195		
	Total	362022.551	30			

a. Dependent Variable: Canada

b. Predictors: (Constant), USA

### Coefficients<sup>a</sup>

Model	Unstandardized Coefficients		Beta	t	Sig.	95.0% Confidence Interval for B		
	B	Std. Error				Lower Bound	Upper Bound	
1	(Constant)	-54.881	10.193		-5.384	.000	-75.728	-34.033
	USA	1.474	.056	.980	26.441	.000	1.360	1.588

a. Dependent Variable: Canada

### Residuals Statistics<sup>a</sup>

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	36.1939	459.2672	192.9848	107.64198	31
Residual	-167.8918	34.8284	0.0000	21.8340	31
Df Fit	1	29	0.000	1.000	29
Df Residuals	29	29	0.000	1.000	29

a. Dependent Variable: Canada

From the analysis, we see that  $\text{Canada\_price} = -54.881 + 1.474(\text{USA\_price})$