

SPSS Lab Work

1. Following is the 20 subjects related with socio-demographic and some health related variables

Sn	age	gender	bmi	bp-sy	bp-dy	occu	lit	smoke	alco
1	92	1	24	120	80	2	2	3	2
2	65	1	20	90	80	1	2	2	2
3	85	2	17	140	95	1	1	1	2
4	65	2	25	140	100	1	1	1	2
5	64	2	26	150	90	1	1	1	2
6	85	1	17	120	70	1	1	2	2
7	76	1	21	110	80	1	1	2	2
8	65	2	22	110	75	1	1	3	2
9	78	1	23	90	60	1	1	2	2
10	60	2	20	130	80	1	1	1	2
11	66	1	28	120	60	1	1	3	2
12	65	2	30	130	80	2	2	1	2
13	69	2	25	120	80	1	2	2	2
14	61	2	24	140	90	1	2	1	1
15	67	2	26	120	80	1	2	1	2
16	68	2	20	120	80	1	2	1	1
17	80	1	25	120	80	1	2	2	2
18	65	1	27	110	70	1	2	3	2
19	65	2	29	130	80	1	2	1	2
20	70	2	30	150	60	1	2	1	2

The details of the variables are as follows:

sn	Serial number
age	Age of persons
gender	1 for male ; 2 for female
bp-sy	Symboliz blood pressure
bp-dy	Diastolic blood pressure
occu	Occupation: 1 agriculture ; 2 others
lit	Literacy status: 1 literate; 2 illiterate
smoke	Smoking habit: 1 non smokers; 2 previous smoker; 3 current smokers
bmi	Body mass index
alco	Habit of using alcohol: 1 never; 2 past; 3 current

2. Entering data into the IBM SPSS Statistics

Variable View

The screenshot shows the IBM SPSS Statistics Data Editor in Variable View mode. The main window displays a table of variables with the following columns:

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	sn	Numeric	8	0		None	None	8	Left	Nominal	Input
2	age	Numeric	8	0		None	None	8	Left	Nominal	Input
3	gender	Numeric	8	0		{1, Male}	None	8	Left	Nominal	Input
4	bp_sy	Numeric	8	0		None	None	8	Left	Unknown	Input
5	bp_dy	Numeric	8	0		None	None	8	Left	Unknown	Input
6	occu	Numeric	8	0		None	None	8	Left	Unknown	Input
7	lit	Numeric	8	0		{1, agricultu}	None	8	Left	Unknown	Input
8	smoke	Numeric	8	0		{1, Literate}	None	8	Left	Unknown	Input
9	bmi	Numeric	8	0		{1, Non Sm.	None	8	Left	Unknown	Input
10	alco	Numeric	8	0		None	None	3	Left	Unknown	Input
11						{1, Never}	None	8	Left	Unknown	Input
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											
27											
28											
29											
30											
31											
32											
33											
34											
35											
36											
37											
38											

A 'Value Labels' dialog box is open over the Variable View table. It contains the following fields:

- Value: []
- Label: []
- Buttons: OK, Cancel, Help

In the dialog box, the 'Label' field contains the text "1 = 'Literate'" and "2 = 'Illiterate'".

Data View

20		sn	age	gender	bp_sy	bp_dy	occu	lit	smoke	bmi	alco	var
1	1	92	Male	120	80	other	illiterate	Current S...	24	Past		
2	2	65	Male	90	80	agriculture	illiterate	Previus S...	20	Past		
3	3	85	Female	140	95	agriculture	Literate	Non Smoker	17	Past		
4	4	65	Female	140	100	agriculture	Literate	Non Smoker	25	Past		
5	5	64	Female	150	90	agriculture	Literate	Non Smoker	26	Past		
6	6	85	Male	120	70	agriculture	Literate	Previus S...	17	Past		
7	7	76	Male	110	80	agriculture	Literate	Previus S...	21	Past		
8	8	65	Female	110	75	agriculture	Literate	Current S...	22	Past		
9	9	78	Male	90	60	agriculture	Literate	Previus S...	23	Past		
10	10	60	Female	130	80	other	illiterate	Non Smoker	20	Past		
11	11	66	Male	120	60	agriculture	illiterate	Current S...	28	Past		
12	12	65	Female	130	80	agriculture	illiterate	Non Smoker	30	Past		
13	13	69	Female	120	80	agriculture	illiterate	Previus S...	25	Past		
14	14	61	Female	140	90	agriculture	illiterate	Non Smoker	24	Never		
15	15	67	Female	120	80	agriculture	illiterate	Non Smoker	26	Past		
16	16	68	Female	120	80	agriculture	illiterate	Non Smoker	20	Never		
17	17	80	Male	120	80	agriculture	illiterate	Previus S...	25	Past		
18	18	65	Male	110	70	agriculture	illiterate	Current S...	27	Past		
19	19	65	Female	130	80	agriculture	illiterate	Non Smoker	29	Past		
20	20	70	Female	100	60	agriculture	illiterate	Non Smoker	30	Past		
21												
22												
23												
24												
25												
26												
27												
28												
29												
30												
31												
32												
33												
34												
35												
36												
37												

Data View Variable View

Computing mean, median, mode, standard deviation, variance, minimum value, maximum value, skewness and kurtosis of age:
 Question B 1.

→ Frequencies

Statistics

age		
N	Valid	20
	Missing	1
Mean		70.55
Median		66.50
Mode		65
Std. Deviation		8.959
Variance		80.261
Skewness		1.123
Std. Error of Skewness		.512
Kurtosis		.270
Std. Error of Kurtosis		.992
Minimum		60
Maximum		92

age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	60	1	4.8	5.0	5.0
	61	1	4.8	5.0	10.0
	64	1	4.8	5.0	15.0
	65	6	28.6	30.0	45.0
	66	1	4.8	5.0	50.0
	67	1	4.8	5.0	55.0
	68	1	4.8	5.0	60.0
	69	1	4.8	5.0	65.0
	70	1	4.8	5.0	70.0
	76	1	4.8	5.0	75.0
	78	1	4.8	5.0	80.0
	80	1	4.8	5.0	85.0
	85	2	9.5	10.0	95.0
	92	1	4.8	5.0	100.0
Total		20	95.2	100.0	
Missing	System	1	4.8		
Total		21	100.0		

Question B 2. Analyzing / Statisticizing the bmi according to question:

→ Frequencies

Statistics

bmi

N	Valid	20
	Missing	1
Mean		23.95
Median		24.50
Mode		20 ^a
Std. Deviation		3.927
Variance		15.418
Skewness		-.196
Std. Error of Skewness		.512
Kurtosis		-.740
Std. Error of Kurtosis		.992
Minimum		17
Maximum		30

a. Multiple modes exist. The smallest value is shown

bmi

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	17	2	9.5	10.0	10.0
	20	3	14.3	15.0	25.0
	21	1	4.8	5.0	30.0
	22	1	4.8	5.0	35.0
	23	1	4.8	5.0	40.0
	24	2	9.5	10.0	50.0
	25	3	14.3	15.0	65.0
	26	2	9.5	10.0	75.0
	27	1	4.8	5.0	80.0
	28	1	4.8	5.0	85.0
	29	1	4.8	5.0	90.0
	30	2	9.5	10.0	100.0
	Total	20	95.2	100.0	
Missing	System	1	4.8		
	Total	21	100.0		

Question B 3. Computing mean, median, mode, s.d., variance, minimum value, maximum value, skewness and kurtosis of bp_sy:

→ Frequencies

Statistics

bp_sy

N	Valid	20
	Missing	1
Mean		120.50
Median		120.00
Mode		120
Std. Deviation		16.051
Variance		257.632
Skewness		-.259
Std. Error of Skewness		.512
Kurtosis		-.101
Std. Error of Kurtosis		.992
Minimum		90
Maximum		150

bp_sy

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	90	2	9.5	10.0	10.0
	100	1	4.8	5.0	15.0
	110	3	14.3	15.0	30.0
	120	7	33.3	35.0	65.0
	130	3	14.3	15.0	80.0
	140	3	14.3	15.0	95.0
	150	1	4.8	5.0	100.0
	Total	20	95.2	100.0	
Missing	System	1	4.8		
Total		21	100.0		

Question B 4.

Computing mean, median, mode, standard deviation, variance, minimum value, maximum value, skewness and kurtosis of bp_dy:

→ Frequencies Statistics

bp_dy

N	Valid	20
	Missing	1
Mean		78.50
Median		80.00
Mode		80
Std. Deviation		10.773
Variance		116.053
Skewness		-.129
Std. Error of Skewness		.512
Kurtosis		.134
Std. Error of Kurtosis		.992
Minimum		60
Maximum		100

bp_dy

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	60	3	14.3	15.0	15.0
	70	2	9.5	10.0	25.0
	75	1	4.8	5.0	30.0
	80	10	47.6	50.0	80.0
	90	2	9.5	10.0	90.0
	95	1	4.8	5.0	95.0
	100	1	4.8	5.0	100.0
	Total	20	95.2	100.0	
Missing	System	1	4.8		
Total		21	100.0		

Question c. Generating frequency distribution for each categorical variable gender, occu, lit, smoke and alco:

Statistics

	gender	occu	lit	smoke	alco
N	Valid	20	20	20	20
	Missing	1	1	1	1

Frequency Table

gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	8	38.1	40.0	40.0
	Female	12	57.1	60.0	100.0
	Total	20	95.2	100.0	
Missing	System	1	4.8		
Total		21	100.0		

occu

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	agriculture	18	85.7	90.0	90.0
	other	2	9.5	10.0	100.0
	Total	20	95.2	100.0	
Missing	System	1	4.8		
Total		21	100.0		

lit

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Literate	7	33.3	35.0	35.0
	illiterate	13	61.9	65.0	100.0
	Total	20	95.2	100.0	
Missing	System	1	4.8		
Total		21	100.0		

smoke

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Non Smoker	10	47.6	50.0	50.0
	Previous Smoker	6	28.6	30.0	80.0
	Current Smoker	4	19.0	20.0	100.0
	Total	20	95.2	100.0	
Missing	System	1	4.8		
Total		21	100.0		

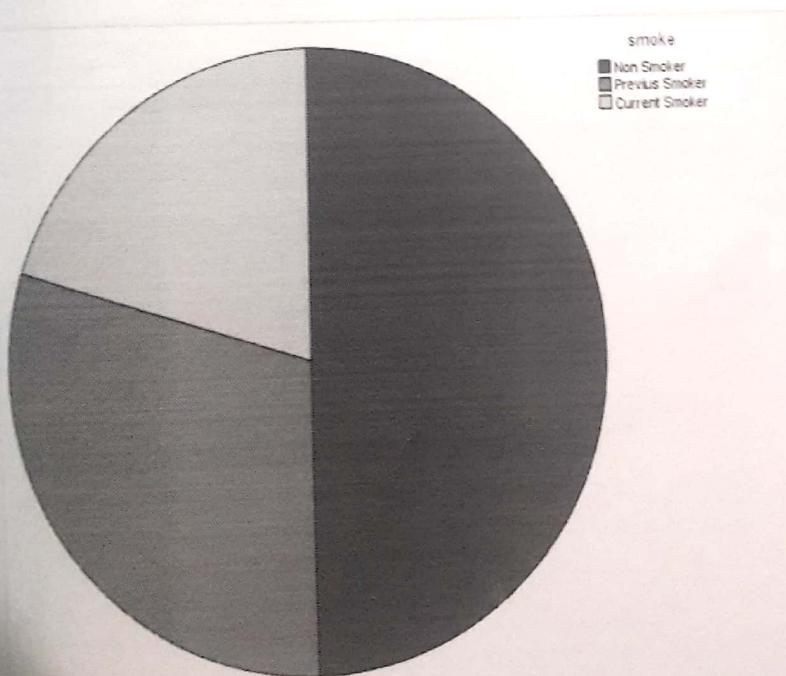
alco

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Never	2	9.5	10.0	10.0
	Past	18	85.7	90.0	100.0
	Total	20	95.2	100.0	
Missing	System	1	4.8		
	Total	21	100.0		

Question H.

Pie-chart of smoke :

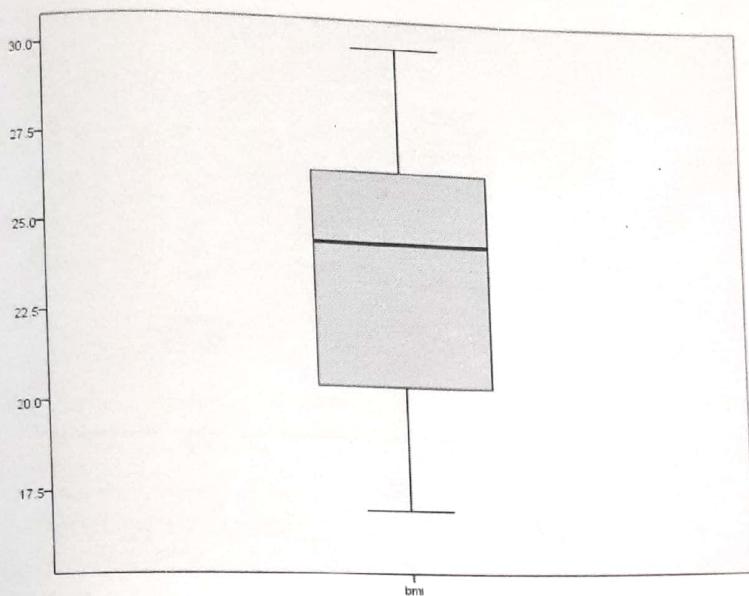
→ Graph



BoxPlot of BMI Only

Case Processing Summary

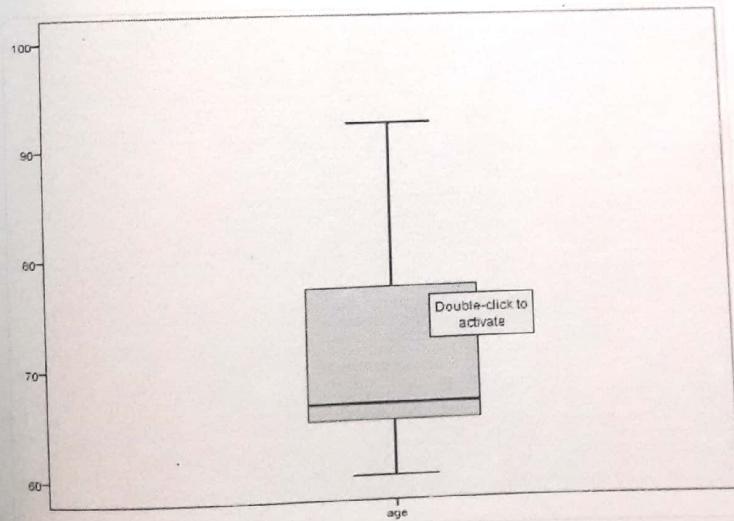
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
bmi	20	95.2%	1	4.8%	21	100.0%



Boxplots of Age Only

Case Processing Summary

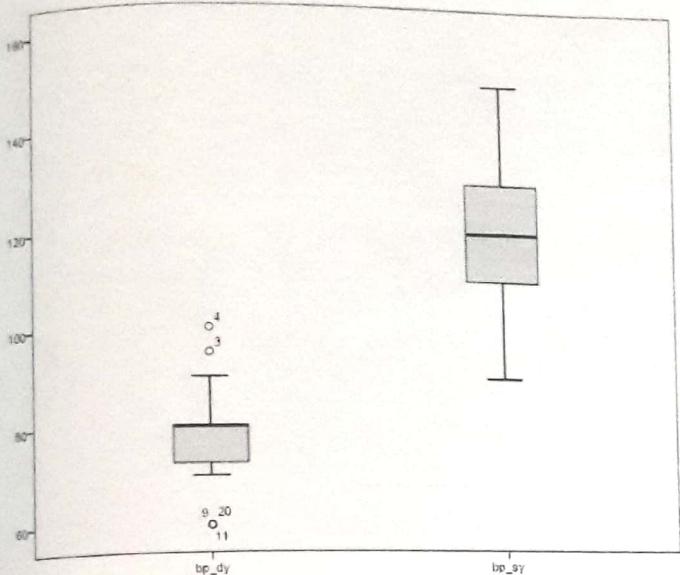
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
age	20	95.2%	1	4.8%	21	100.0%



Boxplot of bp_dy and bp_sy on single graph:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
bp_dy	20	95.2%	1	4.8%	21	100.0%
bp_sy	20	95.2%	1	4.8%	21	100.0%



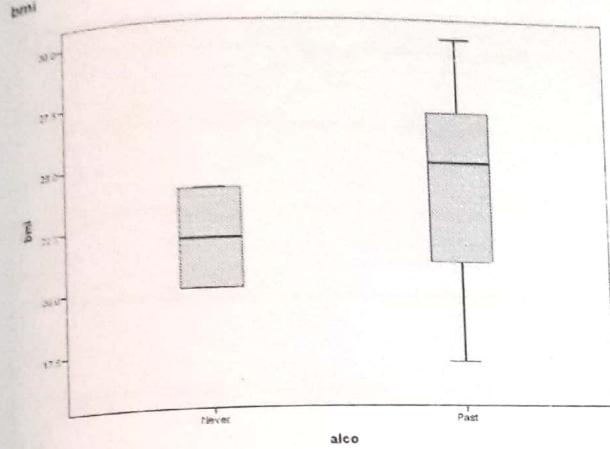
Question G>

Bonplot of bmi over alco:

Explore

alco

Case Processing Summary						
	Cases		Missing		Total	
	Valid	N Percent	N	Percent		
alco	2	100.0%	0	0.0%	2	100.0%
never	18	100.0%	0	0.0%	18	100.0%



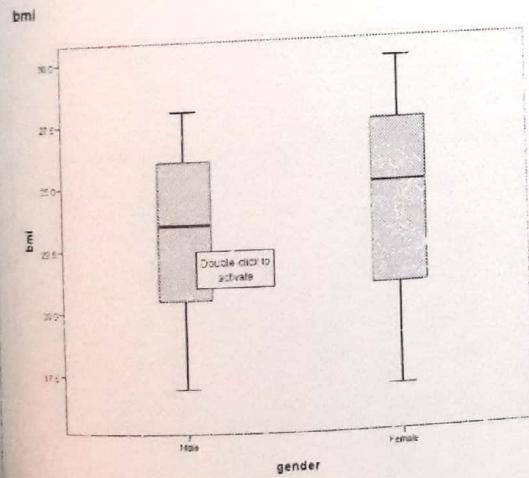
Question F.

Bonplot of bmi over gender:

Explore

gender

	Case Processing Summary					
	Cases		Missing		Total	
gender	Valid	N Percent	N	Percent	N	Percent
	Male	8	100.0%	0	0.0%	8
Female	12	100.0%	0	0.0%	12	100.0%



2. Compute the Karl Pearson's correlation coefficient using SPSS
for the following age of husband and wives (in years)
at the time of their marriage.

hus - age	wife - age
23	18
27	20
28	22
28	27
28	21
30	29
30	27
33	29
35	28
38	29

Question No 2

Karls pearson correlation:

→ Correlations

[DataSet0]

Correlations

		in years	in years
in years	Pearson Correlation	1	.801**
	Sig. (2-tailed)		.005
	N	10	10
in years	Pearson Correlation	.801**	1
	Sig. (2-tailed)	.005	
	N	10	10

**. Correlation is significant at the 0.01 level (2-tailed).

. fit a regression line of blood pressure on age of 20 patients using SPSS. The data is as follows:

age	bp
92	120
65	90
85	140
65	140
64	150
85	120
76	110
65	110
78	90
60	130
66	120
65	130
69	120
61	140
67	120
68	120
80	120
65	110
65	130
70	150

Q. Compute correlation between experience and performance. Also fit a regression line of performance on experience.

The data is as follows:

Experience	16	12	18	4	3	10	5	12
Performance	87	88	89	68	78	80	75	83

Question No 3

Regression line:

Model Description	
Model Name	
Dependent Variable	1
Equation	1
Independent Variable	Linear
Constant	bp
Variable Whose Values Label Observations in Plots	Included
	Unspecified

Case Processing Summary

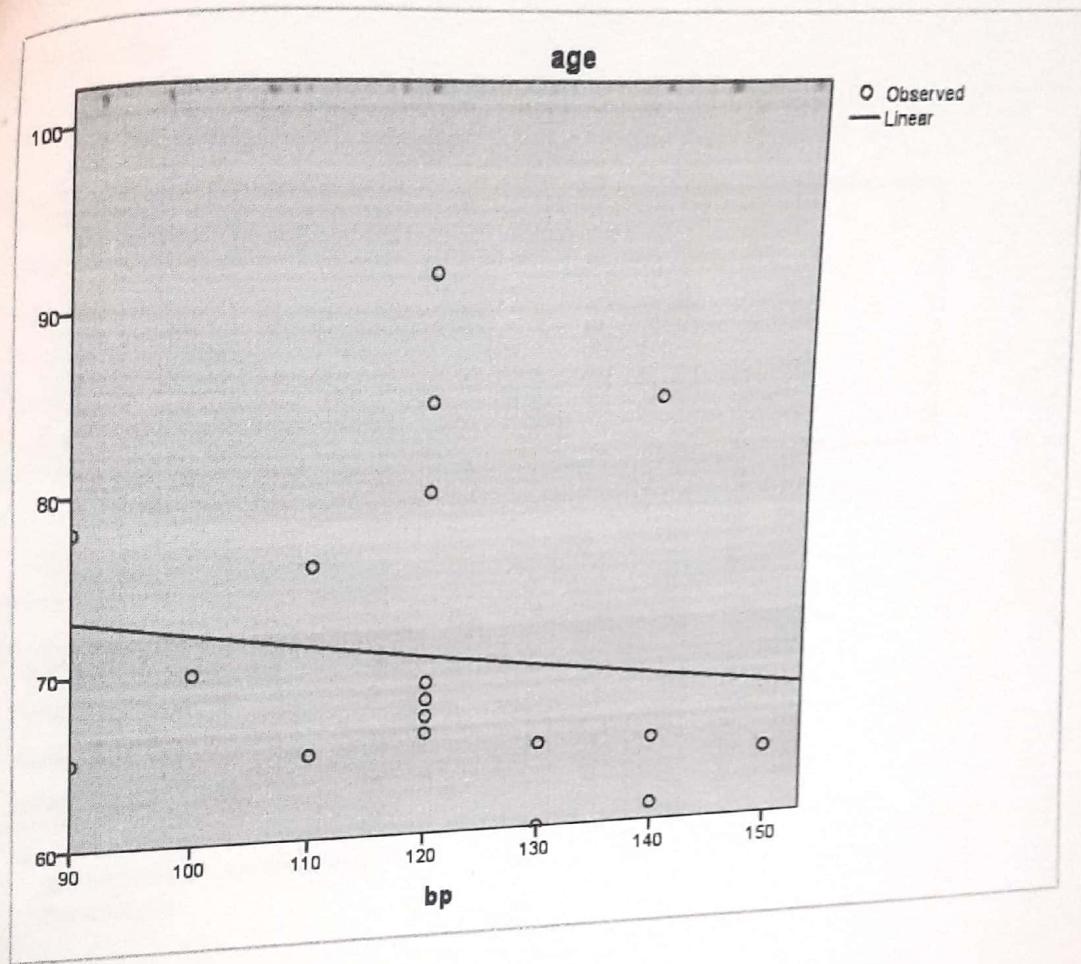
	N
Total Cases	20
Excluded Cases ^a	0
Forecasted Cases	0
Newly Created Cases	0

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

Variable Processing Summary

	Variables	
	Dependent	Independent
	age	bp
Number of Positive Values	20	20
Number of Zeros	0	0
Number of Negative Values	0	0
Number of Missing Values	User-Missing System-Missing	0 0

Model Summary and Parameter Estimates



Question No 4

Correlation:

→ Correlations

		Correlations	
		Experence	Performance
Experence	Pearson Correlation	1	.872**
	Sig. (2-tailed)		.005
	N	8	8
Performance	Pearson Correlation	.872**	1
	Sig. (2-tailed)	.005	
	N	8	8

**. Correlation is significant at the 0.01 level (2-tailed).

Regression line:

• Curve Fit

Model Description

Model Name	MOD_3
Dependent Variable	Experence
Equation	Linear
Independent Variable	Performance
Constant	Included
Variable Whose Values Label Observations in Plots	Unspecified

Case Processing Summary

	N
Total Cases	20
Excluded Cases ^a	12
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
	Experence	Performance
Number of Positive Values	8	8
Number of Zeros	0	0
Number of Negative Values	0	0
Number of Missing Values	User-Missing Values	0 12
	System-Missing	12

Model Summary and Parameter Estimates

Dependent Variable: Experence

Equation	Model Summary					Parameter Estimates	
	R Square	F	df1	df2	Sig.	Constant	b1
Linear	.760	19.050	1	6	.005	-44.367	.671

The independent variable is Performance.

