Unique paper code	:	61018313
Name of Paper	:	Statistical Data Analysis Using Software Packages
Name of Course	:	B. Voc.
Semester	:	III
Duration	:	2 Hours
Maximum marks	:	50

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. All questions are compulsory.

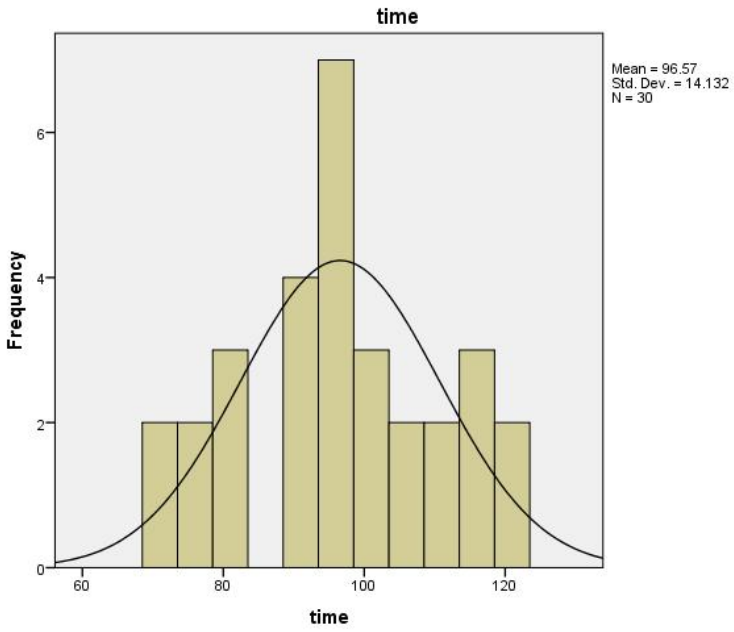
Q1.

a. A social researcher is interested in measuring the level of religiosity of a sample ofsenior citizens. Help her in establishing the levels of measurement for the following variables: pray (do you pray?), services (number of times you attend formal church services per year), money (donated to church), volunteer (hours per year of volunteer assistance), member (are you an official member of a church?), discuss (how many times each week do you discuss religious doctrine?), and times pray (how many times per week do you pray?). For input these data into the SPSS Variable View spreadsheet, what type of measurement the data have, give your answer in detail.

b. The highway patrol officer wants to set up an SPSS file to record trafficviolations. Shewishes to record data at the *nominal*, *ordinal*, and *scale*levels of measurement. The first item of interest (the largest source ofincome for the highway patrol) is speeding. Input three variables thatcould record speed at each level of measurement. The next item of interest is vehicle violations—in the same database set up a variable atthe correct level of measurement and with three categories if necessary.Impaired driving is another important violation. How would you measure and record information for this violation? Show all this in the samedatabase by recoding them and giving example of each types of variable.

Q2.

a. Interpret the following figure?



b. Interpret the result based on below table.

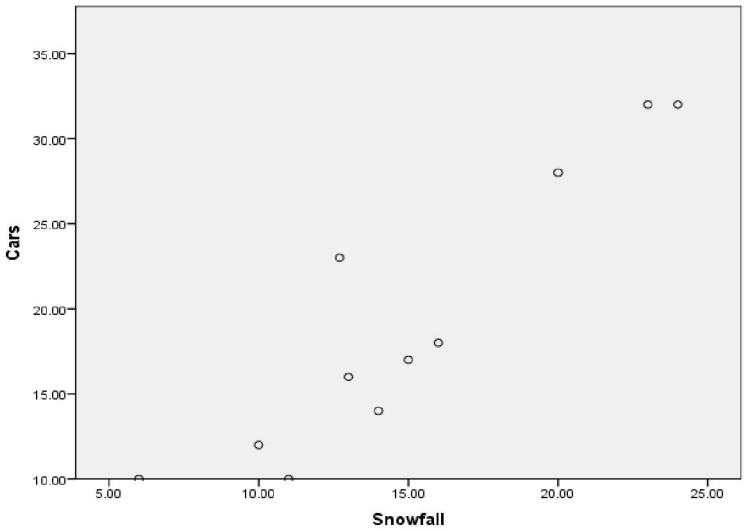
Statistics

Variable_1

N	Valid	18
	Missing	0
Mean		53.83
Median		53.50
Mode		51
Std. Deviation		8.893
Variance		79.088
Skewness		.074
Std. Error of Skewness		.536
Kurtosis		.136
Std. Error of Kurtosis		1.038
Range		35
Minimum		37
Maximum		72
Percentiles	25	49.50
	50	53.50
	75	58.25

Q3.

(a) Let's say you live on a little used back road that leads to the ski slopes.Over the years, you have noticed that there seems to be a correlationbetween the number of inches of snowfall and traffic on your road.You collect some data and now wish to analyze them using correlation and a test of significance. You also wish to visualize the data on agraph. Write the null and alternative hypotheses, and analyze the table and scatterplot given below. Theoutput are as follows:



Correlations

		Snowfall	Cars
Snowfall	Pearson Correlation	1	.920
	Sig. (2-tailed)		.000
	N	11	11
Cars	Pearson Correlation	.920	1
	Sig. (2-tailed)	.000	
	N	11	11

Correlations

			Snowfall	Cars
Spearman's rho	Snowfall	Correlation Coefficient	1.000	.881
		Sig. (2-tailed)	.	.000
		N	11	11

	Cars	Correlation Coefficient	.881	1.000
		Sig. (2-tailed)	.000	.
		N	11	11

(b)A community activist believed that there was a relationship betweenmembership in the police SWAT Team and prior military experience.He collected data from several police departments in an effort to support his belief. He found that there were 57 members of the SWATteam with prior military experience and 13 members with no priormilitary service. There were also 358 police personnel who had military experience but were not members of SWAT and another 413 withno military experience and not members of SWAT. You must write thenull and alternative hypotheses and interpret the results.

prior military * swat member Crosstabulation

			swat member		
			yes	no	Total
prior military	yes	Count	57	358	415
		Expected Count	34.5	380.5	415.0
	no	Count	13	413	426
		Expected Count	35.5	390.5	426.0
Total		Count	70	771	841
		Expected Count	70.0	771.0	841.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	31.442[a]	1	.000		
Continuity Correction[b]	30.058	1	.000		
Likelihood Ratio	33.631	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	31.405	1	.000		
N of Valid Cases	841				

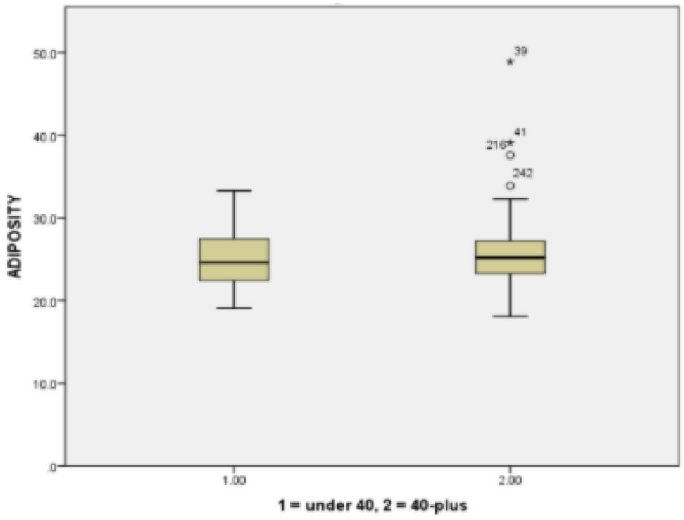
a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 34.54.

b. Computed only for a 2x2 table

Q4.

(a) Explain with example, how to generate report using report summaries in rows in SPSS.

(b) How we can draw a box plot in SPSS, interpret the following plot.



Q5. Trace metals in drinking water affects the flavor, and unusually high concentration can pose a health hazard. The article "Trace Metals of South Indian River" reports on a study in which six river locations were selected and zinc concentrations (mg/L) determined for both surface water and bottom water at each location. The following results were obtained after analyzing the data:

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Zinc Concentration in bottom water	.53617	6	.171326	.069944
	Zinc Concentration in surface water	.44450	6	.141770	.057877

	Paired Differences					
	Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
Zinc Concentration in bottom water - Zinc Concentration in surface water	.091667	.060688	.024776	3.700	5	.014

i. What is the parameter of interest used?

ii. Write the hypothesis used for reaching to these results.

iii. Interpret the result obtained in the above table.

iv. Draw inference about the difference between the two population means by calculating the confidence interval estimate.

6. Can you help the manager of a senior citizen center at the local librarydetermine if there was any merit to her idea that the patron's ageand the number of books checked out were related? Her thoughtwas that as an individual got older, more books would be checked out. She would like to be able to predict the number of books that would be checked out by looking at a person's age. The manager isespecially interested in the number of books checked out for those65 years of age. She selected a random sample of 24 senior patrons and collected details of the age and the number of books checked outduring a 4-week period. If you wish to help, select the correct statistical approach, write the null and alternative hypothesis, and interpret the results. Her outputs are as follows:

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.941[a]	.885	.880	.84146

a. Predictors: (Constant), Age

b. Dependent Variable: Books

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	120.381	1	120.381	170.015	.000[b]
	Residual	15.577	22	.708		
	Total	135.958	23			

a. Dependent Variable: Books

b. Predictors: (Constant), Age

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-52.337	4.417		-11.850	.000
	Age	.876	.067	.941	13.039	.000

a. Dependent Variable: Books