

- (v) Using the Count Values within Cases transformation, a user can specify the following combinations of values to count (select all that apply) :
- (a) Range of values
- (b) System-missing only
- (c) User-missing only
- (d) Individual values 1×5
3. (a) Explain the difference between the files with extension .sav and .sps in detail.
- (b) What is the difference between Sort cases and Rank cases ? Explain with examples how the two can be performed.
- (c) Name the command and sub-commands used to draw a simple random sample of size 20 from given 50 observations.
- (d) What the different types of data files that can be read in SPSS ? Can you read any ASCII file ?
- (e) Name the commands and sub-commands required to compute the median. 2×5

8/5/2018

This question paper contains 8 printed pages]

Roll No.

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

S. No. of Question Paper : 7209

Unique Paper Code : 32373901

HC

Name of the Paper : Statistical Data Analysis Using Software Packages

Name of the Course : B.Sc. (H) Statistics : SEC

Semester : IV

Duration : 2 Hours

Maximum Marks : 50

(Write your Roll No. on the top immediately on receipt of this question paper.)

All questions are compulsory.

1. Fill in the blanks :

- (i) Rank option is included under command.
- (ii) SPSS by default aligns the text to side of the cell.
- (iii) file is used for recalling the SPSS commands to run different procedures.
- (iv) The data editor has two views and
- (v) SPSS syntax file has extension name 1×5

2. Answer the following :

(i) If the null hypothesis is true, then $p < .05$ means :

- (a) the obtained result is not due to chance.
- (b) the obtained result is likely to occur by chance 95% of the time.
- (c) the obtained result is likely to occur by chance less than 5 times in a hundred.
- (d) the obtained result is likely to occur by chance 5 times in a hundred.

(ii) In SPSS, what is the "Data Viewer" ?

- (a) A table summarizing the frequencies of data for one variable.
- (b) A spreadsheet into which data can be entered.
- (c) A dialog box that allows you to choose a statistical test.
- (d) A screen in which variables can be defined and labeled.

(iii) In which sub-dialog box can the Chi-square test be found ?

- (a) Frequencies : Percentages
- (b) Crosstabs : Statistics
- (c) Bivariate : Pearson
- (d) Gender : Female

(iv) What is the advantage of a stem and leaf display over a histogram ?

- (a) The shape of the distribution can be seen more clearly
- (b) It is quicker to produce
- (c) No information from the data is lost
- (d) It can be used with all data

- (a) Which command has been executed ?
- (b) What coefficient has been computed between Protein intake (PROT_IN) and Calorie intake (CAL_IN) ?
- (c) How many observations are there in the study ?
- (d) What is the value of this coefficient ?
- (e) Comment on the significance of this coefficient. 5×6

- 4 Answer any *six* of the following :
- (a) Mention the rules that are to be followed while naming a variable.
 - (b) What is the function of the following commands ?
 - (i) Select Cases
 - (ii) Insert Variables
 - (iii) Cross Tabs
 - (iv) Descriptive
 - (v) Recode.
 - (c) How will you draw a systematic sample of size 5 from 50 observations ?
 - (d) What are the different statistics available under Frequency/Statistics command ?
 - (e) Given a 2×3 contingency table on health conditions and smoking habits, how will you run Chi-square test for independence of attributes ?
 - (f) A group of Sports Science students are selected from the population to investigate whether a 12-week plyometric-training programme improves their standing

long jump performance. In order to test whether this training improves performance, the students are tested for their long jump performance before they undertake a plyometric-training programme and then again at the end of the programme. The following output has been obtained :

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 Jump 1	2.4815	20	.16135	.03608
Jump 2	2.5155	20	.15982	.03574

Paired Samples Test

	Paired Differences					t	df	sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% confidence Interval of the difference				
				Lower	Upper			
Pair 1 Jump1-Jump2	-.03400	.03185	.00712	-.04891	-.01909	-4.773	19	.000

- (i) How many cases are under study ?
- (ii) What can be the possible null hypothesis that can be tested ?
- (iii) What is the value of the calculated test statistic and its degrees of freedom ?
- (iv) What conclusions do we draw about the null hypothesis ? Give reason.
- (g) Read the following outputs and analyse the result :

Controlling for.. WEIGHT

	PROT_IN	CAL_IN
PROT_IN	1.0000 (0) P= .	.9936 (12) P= .000
CAL_IN	.9936 (12) P= .000	1.0000 (0) P= .

(Coefficient / (D.F.) / 2-tailed Significance)