COVENANT UNIVERSITY
NIGERIA

TUTORIAL KIT
OMEGA SEMESTER

PROGRAMME:
DEMOGRAPHY AND SOCIAL STATISTICS

COURSE: DSS 325

DISCLAIMER

The contents of this document are intended for practice and leaning purposes at the undergraduate level. The materials are from different sources including the internet and the contributors do not in any way claim authorship or ownership of them. The materials are also not to be used for any commercial purpose.

DSS 325
STATISTICAL DATA PROCESSING

BY
DR. AMOO, E. O.

QUESTIONS

Q1. Summarise the fundamental operations in the two windows of SPSS

Q2. The table below presents the regression analysis illustrating the interrelationships between selected variables and exposure to the risk of HIV/AIDS among adolescent in university

Selected Variables	B	S.E.	t-stat	Sig.
Age Group				
10-14 years	0.545	0.394	1.913	0.000
15-19 years	0.106	0.164	0.414	0.001
20-24 years	0.545	0.394	1.913	0.000
Education Attainment				
No formal education	-0.70	0.21	-3.340	0.15
Primary education	-0.17	0.027	2.442	0.15
Secondary education	0.009	0.40	0.215	0.830
Tertiary education	0.933	0.109	8.573	0.00

Dependent Variable: Exposure to HIV/AIDS

In not more than three sentences each, interpret the implications of the following coefficients (B-coefficient, t –value and P-value) for all selected variables.

Q3. What 'Menu' are suitable for discovering <u>duplicated variables</u> and for <u>ranking of cases</u>, <u>recoding</u> and <u>plausible aggregation</u> of data.

Q4. Using the Measure DHS Dataset of 2013, test the hypothesis that that there is a significant positive association between the use of family planning and children ever born (CEB) in Nigeria. Set your confidence level at .05. Report the obtained value of the test statistic, the N, the df and probability level, and whether or not you can reject the null hypothesis of no association between the two variables.

Q5. Explain the logic of data analysis.

Q6. The table below presents the regression analysis illustrating the interrelationships between selected variables and exposure to the risk of HIV/AIDS in XY Community

Selected Variables	B	S.E.	t-stat	Sig.

Age at Marriage				
10-14 years	0.545	0.394	1.913	0.000
15-19 years	0.106	0.164	0.414	0.001
20-24 years	0.545	0.394	1.913	0.000
Education Attainment				
Unemployed	-0.70	0.21	-3.340	0.15
Self-Employed	-0.17	0.027	2.442	0.15
Civil Servant	0.009	0.40	0.215	0.830
Others (Labourers)	0.933	0.109	8.573	0.00

Dependent Variable: Exposure to HIV/AIDS

In not more than two sentences each, interpret the implications of the following coefficients (B-coefficient, t –value and P-value) for all selected variables.

Q7. In the logic of data, indicate what should be the focus of an evaluator or analyst

Q8. A social researcher conducted a survey on the topic "Socio-Economic factors influencing fertility in Ado-Odo\Ota Local Government Area".   Data were generated data on 10 variables namely: children ever born, gender, religious affiliation, occupation, contraceptive use, income, education, usual place of residence, age at marriage and respondent's age. You are required to use the information to answer the following questions:

a) Which variable would you use as the Dependent variable?
b) What do you consider as the appropriate measurement scale for your Dependent variable?
c) Formulate one relevant hypothesis for this study.
d) Indicate the appropriate statistical model using the hypothesis formulated
e) Provide the likely interpretation of your model parameters

Q9. What is data management?
  ➤ .

Q10. State the characteristics of the data require for Two-way Analysis of Variance as a statistical techniques.

Q11. What are the importance of data management?

Q12. Discuss two methods of analysing qualitative data

Q13. What factors can account for variations in hypothesis and analysis results?

Q14. Highlight importance of Post Hoc Test in Two-way Analysis of Variance

Q15. When can you use Binary Logistic regression in the analysis of social data?
Q16. Compare and Contrast One-way and Two-way ANOVA in data analysis.

Q17. State the characteristics of the data require for Regression Analysis as a statistical techniques.

Q18. State the characteristics of the data require for One-Way Analysis of Variance as a statistical techniques.

Q19. Identify four specific generic questions and the appropriate statistic for each of them

Q20. State the characteristics of the data require for Pearson Product Moment Correlation
ANSWER

Q1. Summarise the fundamental operations in the two windows of SPSS
Solution
The two prominent windows in SPSS are (1) Data view and (1) Variable View
Variable View
Functions of the window
  ➢ It is the first operation performed in standard data analysis procedures
  ➢ Candidate to mention how it can be accessed
  ➢ It is programming page
  ➢ Used to define the name of the variables, the types of variable, labels and values of the variables
  ➢ Without appropriate designing in variable view, no interpretable or meaningful result can be obtained.
Data view
  ➢ Described as a row and column cell-like tables
  ➢ Take only codes in numeric data and summary of response in qualitative data
  ➢ Mainly for data entry
  ➢ Mainly for viewing the data

Q3. What 'Menu' are suitable for discovering duplicated variables and for ranking of cases, recoding and plausible aggregation of data.
Solution

Menu for discovering duplicated variables
Selected Variable view window, click Data menu and sorting of variables. This display variables in specific order and duplicated ones are seen side-by-side
Menu for Ranking of cases
Select Transform in Variable view window and then click rank cases
Menu for 'Recoding' data
Select Transform in Variable view window and then click recoding into the same variable or different variable
Menu for Plausible aggregation of data
Select aggregate from the Data Menu in Variable view window

Q5. Explain the logic of data analysis.

Logic of data implies the art and science of following reason judgment, statistical sense or common sense in analysis processes.

Logic basically entails avoiding pitfall in data analysis processes.

It calls for accurate selection of appropriate statistics in analysing data

It considers the nature, quality of dataset in the choice of statistical technique(s)

It is to consider the study questions given the dataset and approaches to present the answer

Q7. In the logic of data, indicate what should be the focus of an evaluator or analyst

The research design: procedure used in data collection

The nature of the data: this deals with the form and quality of the data which governs the overall analysis and what can be inferred from them.

The kind of report expected to be written

The users of the result

Q9. What is data management?

➢ as a process of controlling the information generated during a research project. It is an integral part of the research process.

➢ Concerned with collection, organising, auditing, storing, archiving, sharing and ownership of data

➢ Data management is contingent upon: the types of data involved, process of data collection; storage facility and the usage of the data.

Q11. What are the importance of data management?

It helps in proper organisation of data files for easier access and analysis.

It helps to preserve the quality of your research. It enhance the publication f the result

In the long-run, management of data enhances accountability in data analysis.

Q13. What factors can account for variations in hypothesis and analysis results?

Examples of such factors include but not limited to the following:

- Target population: the peculiarity of the population studied, the number and how representative are they, etc.
- Data used: nature of the data, the timing of the data collection, etc.
- Analysis/analysis procedure: the types of analytic technique employed, are the right variables selected, the nature of the variable used, the manner in which data are edited, etc.
- Timing: The gap between the time of data collection and the time of analysis
- Interpretation of result

Q15. When can you use Binary Logistic regression in the analysis of social data?

Q17. State the characteristics of the data require for Regression Analysis as a statistical techniques.

Q19. Identify four specific generic questions and the appropriate statistic for each of them

Generic Types of Quantitative Questions

S/N	Generic Question	Specific Questions	Useful Statistics
1	What is a typical value of the variable?	• At the state level, how many pounds of soft drink bottles (per unit of population) were typically returned annually?	Measures of central Tendency.
2	How much spread is there among the cases?	• To what extent are two or more variables associated? • How similar are the individual states' return rates?	Measures of spread.
3	To what extent are there causal relationships among two or more variables?	• What factors cause high return rates: existence of state bottle bills? • State economic conditions? • State level of environmental awareness?	Measures of association