|  |  |  |  |
| --- | --- | --- | --- |
| **COURSE LEVEL** | **COURSE CODE** | **COURSE TITLE** | **CONTACT PERIODS** |
| **LH** | **Cl.H** | **TH** | **PH** | **CH** | **CU** |
| Yr1Sem2 | CDC9123 | Advanced Data Analysis | 30 | 00 | 30 | 15 | 60 | 4 |

**Course Description**

The course provides postgraduate students with advanced expertise in data analysis using statistical packages STATA, R and SPSS

**Course Justification/Rationale**

The course is designed to equip students with knowledge and skill to make an informative inquiry of the implication of data. The student builds onto the basic tools to more sophisticated and advanced tools to make sense of use of data generated in doctoral research study.

**Course Objectives**

By the end of this course, students will be able:

1. To develop advanced expertise in identification and application of statistical approaches to analyse data based on the objectives and hypotheses advanced
2. To train students in more advanced statistical data analysis using STATA and R Statistical packages
3. To provide an opportunity for the students to work independently on data collection and analysis tasks

**Learning Outcomes**

At completion of this course, learners will be able to:

1. Implement various stages of advanced statistical analysis using various statistical packages

2. Interpret correctly outputs from the statistical packages

3. Critically collate results and conclusions

4. Present the result of data analyses in form of written reports

5. Critically assess published quantitative data analyses

6. Work independently on practical data analysis problems

**Course Content**

1. Qualitative Data analysis (Manual & computer-assisted): Constant comparative method, categorisation, thematic analysis, theory generation (5hours)
2. Overview of Descriptive statistics and probability theory: Basic Probability Theory; The Normal Distribution; Data Description and Summarization; Elements of Statistical Inference (4hours)
3. Introduction & Data Analysis with STATA - Running STATA/R; Data and data types; Data screening; Variable and value labels; Recording, new variables; Descriptive statistics; T-test, confidence interval: one sample; two independence samples, two paired samples (5hours)
4. Study Design - Sample size and precision; Sample surveys: random, systematic, stratified, cluster, multistage; Experimental designs; Comparative studies: cohort studies, case-control studies (5hours)
5. Analysis of Cross-classified Data - Preparing a contingency table; Chi-square test; The STATA/R crosstabs command (4hours)
6. Non-Parametric Statistics - Ranks; One sample tests; Two sample tests; Comparison of several samples; Ranks correlation coefficients (5hours)
7. Bivariate Correlation - The correlation coefficient; Scatter diagrams; Plotting the scatter diagram with STATA/R; Significance of the correlation coefficient; The STATA/R correlation command; Connection with-and introduction to Regression (6hours)
8. Linear Regression - Introduction to regression; The regression coefficient; The intercept; The relation between regression and correlation coefficients; Significance of the regression coefficient; Simple regression diagnostics; The STATA/R regression procedure (6hours)
9. Multiple Regression - Partial regression coefficients; Model (variable) selection; Regression with dummy variables; The STATA/R regression procedure (4hours)
10. Binary and Multinomial Regression - Binary and polytomous response variables; Binary logistic regression; Logit and probit analysis; Multinomial regression; Use of STATA/R logistic regression and probit procedures (6hours)
11. Analysis of Variance - Comparison of several means: one-way ANOVA; Multiple classifications; Interactions; Adjusting for covariates; STATA/R Analysis of variance procedures (6hours)
12. Multivariate Analysis - The nature of multivariate data; Generalization of the t and F tests; Repeated measures analysis of variance; The STATA MANOVA procedure(6hours)

**Teaching – Learning Methods**

 Lectures/Discussion

 Group Demonstrations

 Class Presentation

 Self-Directed learning

**Teaching and Learning Facilities**

 Lecture facilities

 Class Rooms

 White boards / m

 arkers / cleaners

 LCD Projectorsvarious forms from the Uganda National Council for Science and Technology (UNCST)

**Assessment Strategies**

 Continuous Assessment Tests 40%

 End of Semester Examinations 60%

 Total 100%

**Recommended Reading and Study Materials**

1. Anscombe, F. 1973. Graphs in Statistical Analysis, The American Statistician, pp. 195-199

2. Barnett and Lewis .1994. Outliers in Statistical Data, 3rd. Ed., John Wiley and Sons.

3. Birnbaum, Z. W. and Saunders, S. C. 1958. A Statistical Model for Life Length of

4. Materials, Journal of the American Statistical Association, 53(281), pp. 151-160.

5. Bloomfield, Peter 1976. Fourier Analysis of Time Series, John Wiley and Sons.