

Introduction to Data Analysis – SPSS Without Tears

Synopsis

This short course introduces data analysis using IBM SPSS. Data analysis using basic bio-statistical principles and methods are covered in this short course. Students will have opportunity to analyse real life data and supports for results discussion and conclusion. The short course will cover: data entering, data labelling, data cleaning, data computing/transforming and data analysis (using commands on menus) including summary statistics, hypothesis test, 95% CI, ANOVA, non-parametric method, correlation analysis, linear regression analysis, relative risk (RR), odds ratio (OR), logistic regression analysis, chi-square test and Cox's regression analysis (analysis of hazard ratio).

This course is suitable for those who have basic concepts in biostatistics, e.g., classification of data; summarizing data using graphs and descriptive statistics, normal and t distributions, quantifying uncertainty in results from a sample; two samples independent t-tests, paired samples t-test, non-parametric tests, one-way ANOVA, simple and multiple linear regression, relative risk (RR), odds ratio (OR), simple and multiple logistic regression and chi-squared test.

Learning Objectives/Outcomes

Day 1: 9:00am to 5:00pm

- Be familiar with basic SPSS functions and its tools. These functions and tools will enable students to proficiently open and create SPSS data files.
- Presenting data using SPSS generated graphs and summary statistics: **descriptive statistics.**
- Conducting independent and paired samples t-tests to compare two groups.
- Conducting a one-way ANOVA to compare more than two groups where the test variable is collected on a continuous scale and the data in each group follows the normal distribution: **One-way ANOVA.**
- Analysing data when normality assumption for data does not hold, i.e., the data does not follow the normal distribution. The statistical methods to analyse such data are collectively known as Non-Parametric methods or distribution free method: **non-parametric tests.**
- Evaluating the association between an outcome and one or multiple exposures where outcome is continuous however, exposure could be numerical or categorical or a combination of both: **correlation and linear regression analysis.**

Day 2: 9:00am to 5:00pm

- Quantifying the risk of the disease (outcome) in the exposed group compared to the unexposed group: **relative risk and odds ratio**.
- Evaluating the association between an outcome and one or more exposures where outcome is categorical BINARY but exposure could be numerical or categorical or a combination of both: **logistic regression analysis**.
- Assessing the association between an outcome and an exposure where both of them are categorical, and both have two or more categories or a combination of both: **chi-square analysis**.
- Evaluating the association between an outcome and one or more exposures where outcome is categorical BINARY and time dependent but exposure could be numerical or categorical or a combination of both: **survival analysis**.
- Managing our data (entering, labelling, creating, cleaning, merging, etc.)