Semester 1 (Study Period 1) 2020 Statistics Short Courses-Fully online (no on-campus classes)

Course: STAA0001 - Basic Statistics

Duration-12 weeks: from 31st August 2020 to 29th November 2020

Canvas Site: STAA0001

Assumed Knowledge: None

Software used: IBM SPSS Statistics Version 26

Course Description:

This course develops students' understanding of a range of statistical methods along with their assumptions and limitations of their application. It enables students to develop the capacity to carry out independent statistical analysis of data using a standard statistical software package and also aims to develop students' abilities in effectively communicating the outcomes of statistical investigations.

Students who successfully complete this unit will be able to:

1. Effectively display information in datasets graphically.
2. Select an appropriate descriptive or inferential statistical technique based on the researcher's hypothesis, the level of measurement of the variables and testing of the appropriate assumptions to analyse the data.
3. Select appropriate IBM SPSS Statistics procedures, Java applets on the web and mathematical calculations, to obtain basic statistical test results, including confidence intervals and effect size statistics.
4. Explain the foundations of statistical inference, in particular the role of sampling distributions and the use of the normal distribution as a density curve.
5. Recognize when more advanced techniques are needed.
6. Write interpretive summary reports for both descriptive and inferential statistical analysis.

Online quizzes will be available for self-assessment.

Course: STAA0002 - Simple Linear Regression and ANOVA

Duration-12 weeks: from 31st August 2020 to 29th November 2020

Canvas Site: STAA0002

Assumed Knowledge: Basic Statistics (e.g. STAA0001)

Software used: IBM SPSS Statistics Version 26

Course Description:

This course of study aims to extend the ideas developed in Basic Statistics to include more advanced analyses, broaden the range of applications students are familiar with so that they will be able to carry out independent statistical investigations, and develop an awareness of the assumptions and limitations involved in the generalisation of results of such investigations.

Students who successfully complete this unit will be able to:

1. Choose the appropriate statistical analysis based on the researcher's hypothesis, the level of measurement of the variables and the testing of the appropriate assumptions.
2. Select appropriate statistical procedures, for example, using SPSS, Java applets on the web and mathematical calculations, to analyse data in a variety of contexts.
3. Relate the concepts of effect size, sample size, one or two tailed tests, level of significance and power of a statistical test.
4. Write interpretive summary reports for the inferential statistical techniques covered.
5. Judge when more advanced techniques are needed by comparing different statistical techniques for the variety of research questions.

Online quizzes will be available for self-assessment.

Course: STAA0003A - **Introduction to R**

Duration-6 weeks: from 31st August 2020 to 11th October 2020

Canvas Site: STAA0003

Assumed Knowledge: None

Software used: RStudio

Course Description:

In this course participants will learn how to install and configure R software. They will also learn how to read data into R, access R packages, and organise and comment R code. Furthermore, they will learn how to use R for effective data analysis and visualisation. Some of the most commonly used probability distributions will be introduced. Statistical data analysis will be conducted using working examples. After successfully completing this unit, students will be able to:

1. Arrange and consolidate large datasets.
2. Develop the ability to perform basic programming.
3. Visualise data using R packages.
4. Relate the basics of fundamental probability distributions to different types of data.
5. Formulate practical and user-friendly solutions to real life problems in the form of a statistical model in an RStudio software environment.

Online quizzes will be available for self-assessment.

Course: STAA0003B - **Using R   for Statistical Analysis**

Canvas Site: STAA0003

Duration: 5 weeks: from 12ᵗʰ October 2020 to 21ˢᵗ November 2020

Assumed Knowledge: Introduction to R (e.g. STAA0003A)

Software used: RStudio

Course Description:

This course of study aims to extend the ideas developed in Introduction to R. In this course you will learn key programming principles of R and how to develop and perform different types of statistical analyses and data visualisation. Participants expected to develop competence in programming in R - essential skill for a statistician or data scientist.

After successfully completing this unit, students will be able to:

1. Write R programs to conduct hypothesis testing and compare means.
2. Perform simple linear regression.
3. Analyse categorical data.

Online quizzes will be available for self-assessment.

Course: STAA0003C – Introduction to SAS

Canvas Site: STAA0003

Duration-6 weeks: from 31st August 2020 to 11th October 2020

Assumed Knowledge: None

Software used: SAS

Course Description:

This unit is an introduction to one of the leading statistical software packages SAS - a powerful tool for organising and analysing data. Students will use this statistical software to develop their knowledge in data management, data presentation, and statistical analysis.

After successfully completing this unit, students will be able to:

1. Demonstrate professionally relevant SAS programming and data management skills.
2. Write SAS programs to enter, load, and merge different types of data from multiple sources.
3. Visualise data graphically using SAS graphical procedures.
4. Summarise data and perform descriptive statistical analyses using SAS procedures.

Online quizzes will be available for self-assessment.

Course: STAA0004A - Survey Design

Canvas Site: STAA0004

Duration: 6 weeks: from 31st August 2020 to 11th October 2020

Assumed Knowledge: Basic Statistics (e.g. STAA0001)

Software used: None

Course Description:

Participants will acquire skills and knowledge in the collection of survey, observational, experimental and secondary data; developing a questionnaire, and writing of descriptive reports.

Topics will include:

- Introduction to survey research
- The basics of survey sampling
- How to collect survey data
- Making the most of secondary data
- Developing a questionnaire
- Introduction to scale development
- Coding and cleaning survey data

Online quizzes will be available for self-assessment.

Course: STAA0004B - Research Design

Canvas Site: STAA0004

Duration: 6 weeks: from 12ᵗʰ October 2020 to 29ᵗʰ November 2020

Assumed Knowledge: Basic Statistics (e.g. STAA0001)

Software used: Excel and IBM SPSS Statistics Version 26

Course Description:

Participants will acquire skills and knowledge in observational and experimental studies, designing an experiment, incidence and prevalence statistics, different types of study designs including Cohort and Case-control studies.

Topics will include

- The basic concepts of experimental designs
- Common designs used in health statistics and elsewhere
- Incidence, prevalence and fertility statistics
- Mortality Statistics and Standardisation of rates
- Randomized trials and Cohort studies
- Case control studies

Online quizzes will be available for self-assessment.

Course: STAA0005A - Multiple Linear Regression

Canvas Site: STAA0005

Duration: 2 weeks: from 31ˢᵗ August 2020 to 13ᵗʰ September 2020

Assumed Knowledge: Simple Linear Regression and ANOVA (e.g. STAA0002)

Software used: IBM SPSS Statistics Version 26

Course Description:

In Multiple Regression we will look at simple linear regression and multiple regression using three different strategies (standard regression, stepwise regression and hierarchical regression). Particular attention is paid to report writing, assumption checking, outlier checking and tests for mediation. Make sure that you have access to SPSS and please revise the relevant material for the simple linear regression and ANOVA short course beforehand.

Online quizzes will be available for self-assessment.

Course: STAA0005B - Factor Analysis and MANOVA

Canvas Site: STAA0005

Duration: 6 weeks: from 14ᵗʰ August 2020 to 25ᵗʰ October 2020

Assumed Knowledge: Simple Linear Regression and ANOVA (e.g. STAA0002)

Software used: IBM SPSS Statistics Version 26

Course Description:

Factor Analysis covers exploratory factor analysis (EFA). The various methods for extracting and rotating factors are discussed as are the interpretation of factors and the creation of factor scores and summated.
EFA is a descriptive technique. That is, it is designed to help us understand and explain patterns in the data, without making any formal predictions about what results will look like. However, it is not our data's job to tell us what its underlying structure is and a sound factor analytic study will begin with a great deal of prior thinking about the nature of the concept that we want to understand, appropriate indicators of that concept, appropriate population, and how results of factor analysis will be used. So even before we begin data collection, let alone data analysis, we will have an expectation about what the results might look like. The job of the data is then to show us how well our expectations are reflected in the 'real world'. The results of exploratory factory analysis can then be used inform future hypotheses. These hypotheses are subsequently tested using confirmatory factor analysis (CFA), which is conducted within the structural equation modelling framework (not covered in this subject).

MANOVA examines between subjects, within subjects and mixed multivariate analysis of variance.
Particular attention is paid to assumption checking, the testing of specific contrasts and report writing. Make sure that you have access to SPSS and please revise the relevant material for the ANOVA and Simple Linear Regression short course beforehand.

Make sure that you have access to SPSS and please revise the relevant material for the ANOVA and Simple Linear Regression short course beforehand.

Online quizzes will be available for self-assessment.

Course: STAA0006 – Introduction to SAS

Canvas Site: STAA0006

Duration-6 weeks: from 31ˢᵗ August 2020 to 11ᵗʰ October 2020

Assumed Knowledge: Simple Linear Regression (e.g. STAA0002).and ANOVA

Software used: SAS

Course Description:

This unit is an introduction to one of the leading statistical software packages SAS - a powerful tool for organising and analysing data. Students will use this statistical software to develop their knowledge in data management, data presentation, and statistical analysis.

Students who successfully complete this unit will be able to:

1. Write SAS programs to enter, load, and merge different types of data from multiple sources.
2. Visualise data graphically using SAS graphical procedures.
3. Summarise data and perform descriptive statistical analyses using SAS procedures.
4. Conduct parametric and non-parametric analysis of variance for balanced and unbalanced study designs using SAS and interpret the results.
5. Perform linear regression analysis using SAS and interpret the results.
6. Analyse count data using generalised linear model (GLM) techniques using SAS.

Online quizzes will be available for self-assessment.

Course: STAA0007A - Forecasting

Canvas Site: STAA0005

Duration-6 weeks: from 31ˢᵗ August 2020 to 11ᵗʰ October 2020

Assumed Knowledge: ANOVA and Simple Linear Regression (e.g. STAA0002). Basic knowledge of SPSS package is required.

Software used: EXCEL and SPSS

Course Description:

This unit aims to introduce students to various forecasting methods and their application in business and industry using time series data. The advantages and disadvantages of naïve, moving average, exponential smoothing and decomposition forecasts will be covered.

Participants who successfully complete this Unit should be able to:

- extract time series data from web sites
- compute indices based on time series data
- plot time series and describe their characteristics
- understand the time series and identify the properties like trend, seasonality or cyclic behaviour relevant to the time series under investigation
- use various methods to obtain short-term forecasts
- compare the accuracy of these forecasts using appropriate measures.

Online quizzes will be available for self-assessment.

Course: STAA0007B – Advanced Forecasting with SAS

Canvas Site: STAA0007B

Duration-6 weeks: from 12ᵗʰ October 2020 to 29ᵗʰ November 2020

Assumed Knowledge: Introduction to SAS (e.g. STAA0006) and Forecasting (e.g. STAA0007A)..

Software used: SAS

Course Description:

A range of advanced forecasting methods and their application to planning and decision-making are covered in this unit.

Participants who successfully complete this Unit should be able to:

- Appraise the limitations of regression as a forecasting tool.
- Demonstrate univariate forecasting methods such as ARIMA, ARMA, SARIMA, and identifying appropriate models for a project.
- Identify an appropriate transformation for stationarity.
- Use the autocorrelation function (ACF) and the partial autocorrelation function (PACF) to identify the most appropriate model.
- Assess forecast performance using key measures of forecast accuracy
- Interpret forecast results for a general non-statistical audience.

Online quizzes will be available for self-assessment.

Course: STAA0008 - Survey Sampling

Canvas Site: STAA0008

Duration-11 weeks: from 31st August 2020 to 22nd November 2020

Software used: SAS

Pre-requisite: Introduction to SAS (e.g. STAA0006).

Course Description:

This unit of study gives students an understanding of the main principles and methods of sampling required to collect and analyse social survey data. In particular, it introduces students to key concepts in sampling such as a sampling frame, non-response, variance, survey weights and design effects, and how to select samples using the main probability sampling methods (simple random sampling, systematic sampling, stratified sampling and cluster sampling). Students are shown how to perform estimation techniques for the various sampling methods and how to calculate sample sizes for surveys with different constraints.

After successfully completing this unit, you will be able to:
1. Choose a suitable sampling frame for a study.
2. Compare the basic probability sampling methods including, simple random sampling, systematic sampling, stratified sampling and cluster sampling.
3. Estimate population parameters from complex samples using appropriate software.
4. Evaluate the effect of small populations in the calculation of confidence intervals.
5. Estimate sample sizes for complex designs.
6. Appraise the effects of non-response bias.
7. Construct population weightings for sampled data.

Online quizzes will be available for self-assessment.

Course: STAA0009A – Introductory Structural Equation Modelling with AMOS

Canvas Site: STAA0009A

Duration-5 weeks: from 31st August 2020 to 4th October 2020

Software used: AMOS

Pre-requisite: Multiple Linear Regression (e.g. STAA0005A) and Factor Analysis and MANOVA (e.g. STAA0005B)

Course Description:

This course is designed as an introductory, applied course in the use of Structural Equation Models (SEM) in research. The aim of this subject is to provide students with a broad understanding of structural equation modelling, its underlying theory and potential uses in research, as well as an awareness of its strengths and limitations. It is designed to give students the skills to competently design, assess, and interpret simple research models across multiple discipline areas using structural equation modelling techniques.

- Review Exploratory Factor Analysis (SPSS).
- Introduction to AMOS and Confirmatory Factor Analysis (CFA)
- Confirmatory Factor Analysis (CFA)
- Review Multiple Regression (SPSS) and Path Analysis in AMOS
- Full Structural Models
- Using SEM to assess measures (Congeneric Confirmatory Factor Analysis, Scale Reliability and Validity, Higher Order Confirmatory Factor Analysis)

Online quizzes will be available for self-assessment.

Course: STAA0009B – Advanced Structural Equation Modelling with MPLUS

Canvas Site: STAA0009B

Duration-5 weeks: from 12th October 2020 to 22nd November 2020

Software used: **MPLUS**

Pre-requisite: Multiple Linear Regression (e.g. STAA0005A) and Factor Analysis and MANOVA (e.g. STAA0005B)

Course Description:

This course is designed as an introductory, applied course in the use of Structural Equation Models (SEM) in research. The aim of this subject is to provide students with a broad understanding of structural equation modelling, its underlying theory and potential uses in research, as well as an awareness of its strengths and limitations. It is designed to give students the skills to competently design, assess, and interpret simple research models across multiple discipline areas using structural equation modelling techniques.

- Review Exploratory Factor Analysis (SPSS).
- Introduction to AMOS and Confirmatory Factor Analysis (CFA)
- Confirmatory Factor Analysis (CFA)
- Review Multiple Regression (SPSS) and Path Analysis in AMOS
- Full Structural Models
- Using SEM to assess measures (Congeneric Confirmatory Factor Analysis, Scale Reliability and Validity, Higher Order Confirmatory Factor Analysis)

Online quizzes will be available for self-assessment.

Course: STAA00010 – Introduction to Bayesian Statistics

Canvas Site: STAA00010

Duration: 6 weeks: from 31st August 2020 to 11th October 2020

Software used: **R and R studio**

Pre-requisite: Multiple Linear Regression (e.g. STAA0005A), Introduction to R (e.g. STAA0003A)

Course Description:

This short course introduces the fundamentals of Bayesian statistical modelling. Students will learn the importance of subjective beliefs in Bayesian statistics. Important concepts such as prior distributions, likelihood functions, and posterior distributions will be discussed at length.
Numerical estimation techniques will be introduced. Empirical applications of Bayesian analysis will be performed in an R software environment.

Online quizzes will be available for self-assessment.

Course: STAA0011A –Advanced Data Mining

Canvas Site: STAA00011A

Duration-11 weeks: from 31ˢᵗ August 2020 to 22ⁿᵈ November 2020

Software used: R and R studio

Pre-requisites: Multiple Linear Regression (e.g. STAA0005A) and Factor Analysis and MANOVA (e.g. STAA0005B)

Course Description:

This short course provides and introduction to data mining using SAS Enterprise Miner. In particular it introduces market basket analysis, sequence analysis, link analysis and text analysis before comparing classification methods such as Classification Trees and Logistic Regression, and comparing prediction methods such as Regression and Regression Trees. In addition it describes the use of neural network methods for classification, prediction and segmentation. Furthermore, this short course introduces four statistical methods commonly used for marketing and other purposes. Visualisation is key for these methods. These methods include Multidimensional scaling for displaying patterns of similarity for objects (e.g. similarity of brands in terms of customer perception). Correspondence analysis is used to illustrate the relationships between categorical variables. Conjoint Analysis is used to measure the importance of product attributes underlying product preferences and Cluster Analysis is used to create homogeneous groups of people or objects that can be used for strategic and other purposes.

Online quizzes will be available for self-assessment.