

Liverpool John Moores University

Title: DATA EXPLORATION AND ANALYSIS
Status: Definitive
Code: **4100STATS** (124195)
Version Start Date: 01-08-2021

Owning School/Faculty: Computer Science and Mathematics
Teaching School/Faculty: Computer Science and Mathematics

Team	Leader
Ian Jarman	Y

Academic Level: FHEQ4 **Credit Value:** 20 **Total Delivered Hours:** 57
Total Learning Hours: 200 **Private Study:** 143

Delivery Options

Course typically offered: Semester 2

Component	Contact Hours
Lecture	33
Practical	11
Tutorial	11

Grading Basis: 40 %

Assessment Details

Category	Short Description	Description	Weighting (%)	Exam Duration
Report	AS1	Build a simple relational database to extract dataset for analysis in Minitab	30	
Exam	AS2	Examination	70	2

Aims

To enable the student to carry out an exploratory analysis of a set of data either 'by hand' or using Minitab. This will include building knowledge of a simple relational databases in order to interrogate and filter data for analysis.

To provide the student with the required background knowledge of probability and

random variables so that they can make use of a number of formal statistical models in their analyses.

To enable the student to appreciate the need for, and use of, confidence intervals in a number of commonly occurring data analysis situations.

To enable the student to appreciate the need for, and use of, hypothesis tests in a number of commonly occurring data analysis situations

Learning Outcomes

After completing the module the student should be able to:

- 1 Extract data using simple relational database techniques
- 2 Carry out an exploratory numerical and graphical analysis of a set of data by hand and/or using Minitab
- 3 Demonstrate the ability to calculate and estimate probabilities
- 4 Calculate confidence intervals for parameters of a number of probability models used in data analysis
- 5 Construct and carry out hypothesis tests upon parameters of a number of probability models in data analysis

Learning Outcomes of Assessments

The assessment item list is assessed via the learning outcomes listed:

Data analysis - 2000 words	1	2		
Examination	3	4	5	

Outline Syllabus

Data and tabular display. Graphical displays - pie charts, histograms, stem-and-leaf plots, box plots, scatter diagrams.

Sample summary statistics - mean, median, mode, quartiles, interquartile range, variance, standard deviation

Basic introduction to relational databases to extract appropriate data for exploratory analysis using Minitab

Samples and populations

Probability - definitions, addition rule, multiplication rule, independent events, conditional probability

Random variables - discrete and continuous

Probability distributions - discrete uniform, Bernoulli, Binomial, Poisson, continuous uniform, Normal, Normal probability plots

Sampling distribution of the mean, central limit effect, Normal approximations

Confidence intervals - the mean of a Normal population: one and two sample cases, large and small sample methods, the t-distribution

Hypothesis testing - the mean of a Normal population: one and two sample cases, large and small sample methods, testing equality of variances

Learning Activities

Lecture, tutorials, laboratory sessions, directed study

Notes

This module covers the exploratory analysis of datasets (including basic relational database skills to extract data from appropriate sources), the use of probability to handle uncertainty and develops techniques of hypothesis testing and confidence interval construction