LIVERPOOL JOHN MOORES UNIVERSITY

Data Exploration and Analysis

Module Information

2022.01, Approved

Summary Information

Module Code	4100STATS
Formal Module Title	Data Exploration and Analysis
Owning School	Computer Science and Mathematics
Career	Undergraduate
Credits	20
Academic level	FHEQ Level 4
Grading Schema	40

Teaching Responsibility

LJMU Schools involved in Delivery
Computer Science and Mathematics

Learning Methods

Learning Method Type	Hours
Lecture	33
Practical	11
Tutorial	11

Module Offering(s)

Display Name	Location	Start Month	Duration Number Duration Unit
JAN-CTY	CTY	January	12 Weeks

Aims and Outcomes

Aims	To enable the student to carry out an exploratory analysis of a set of data either 'by hand' or using appropriate software. 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.

After completing the module the student should be able to:

Learning Outcomes

Code	Number	Description
MLO1	1	Extract data using simple relational database techniques.
MLO2	2	Carry out an exploratory numerical and graphical analysis of a set of data by hand and/or using appropriate software.
MLO3	3	Calculate and estimate probabilities and confidence intervals for parameters of a number of probability models used in data analysis.
MLO4	4	Construct and carry out hypothesis tests upon parameters of a number of probability models in data analysis.

Module Content

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 appropriate software.Samples and populations.Probability - definitions, addition rule, multiplication rule, independent events, conditional probability.Random variables - discrete and continuous.Expectation of a random variable.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.
Module Overview	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.
Additional Information	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.

Assessments

Assignment Category	Assessment Name	Weight	Exam/Test Length (hours)	Module Learning Outcome Mapping
Technology	Data analysis	40	0	MLO1, MLO2
Centralised Exam	Examination	60	2	MLO3, MLO4

Module Contacts

Module Leader

Contact Name	Applies to all offerings	Offerings
Ian Jarman	Yes	N/A

Partner Module Team

Contact Name	Applies to all offerings	Offerings