

Liverpool John Moores University

Title: STATISTICAL MODELLING
Status: Definitive
Code: **5123COMP** (121250)
Version Start Date: 01-08-2021

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

| Team | Leader |
|-------------|--------|
| Mark Taylor | Y |

Academic Level: FHEQ5 **Credit Value:** 20 **Total Delivered Hours:** 56.5
Total Learning Hours: 200 **Private Study:** 143.5

Delivery Options

Course typically offered: Semester 1

| Component | Contact Hours |
|-----------|---------------|
| Lecture | 22 |
| Practical | 33 |

Grading Basis: 40 %

Assessment Details

| Category | Short Description | Description | Weighting (%) | Exam Duration |
|----------|-------------------|---------------------|---------------|---------------|
| Report | AS1 | Case Study Analysis | 60 | |
| Exam | AS2 | Examination | 40 | 1.5 |

Aims

To develop a theoretical knowledge of statistical skills to solve data science problems

To develop and display solutions to data science problems by applying statistical theory using appropriate software applications

Learning Outcomes

After completing the module the student should be able to:

- 1 Applying appropriate statistical theory data science problem to derive meaningful solutions.
- 2 Apply appropriate statistical theory and derive meaningful solutions in a suitable programming language

Learning Outcomes of Assessments

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

| | |
|---------------------|---|
| Case Study Analysis | 2 |
| Examination | 1 |

Outline Syllabus

Review summary statistics

Assumption testing for statistical test

Normality

Multivariate normality

Homoscedasticity

etc

Correlation and Covariance

Non-parametric test – Chi Square

T-Tests

One sample T-test

Two sample T-test

Paired Two-sample T-Test

ANOVA

Linear Models

Simple Linear Regression

Multiple Regression

Discussion of Generalized Linear Models

Logistic Regression

Poisson Regression

Model Diagnostics

Residuals – ANOVA – Akaike Information Criteria (AIC)

Cross-Validation

Bootstrap

Nonlinear Models

Nonlinear Least Squares

Generalized Additive Models

Decision trees

Random Forests (Ensemble)

Learning Activities

Lectures will be used to introduce and demonstrate topics, however the key component in the module is the use of accessible practical tasks to reinforce the theoretical aspects of the lecture material which will be reinforced through practical work.

Notes

Although an apparently heavy theoretical treatment of the area, this is intended to be a practical, hands-on exploration of the area.