THE INTERNATIONAL QUALIFICATIONS

AND ASSESSMENT CENTRE (IQAC)

«Approved»

2025 y. «16» 04

Programme	**Level 7 Diploma in Data Science**		
Unit Number/ Unit Title	UNIT 2 ADVANCED STATISTICAL MODELLING AND INFERENCE		
Cohort Code:	L07ASMI-U2		
Unit Level	Level 7		
Total GLH	Total qualification time 200/ Total Guided learning hours 90/ Self-guided learning hours 110		
Credits	20 CATS/ 10 ECTS		
Lecturer			
Start Date		**End Date**	

Unit Aims	This unit deepens students' understanding of advanced statistical techniques used for inference and decision-making in complex data environments. It emphasizes multivariate analysis, Bayesian inference, model selection, and diagnostic techniques essential in predictive modelling and data interpretation.
Differentiation Strategies *(e.g. planned activities or support for individual learners according to their needs)*	The total number of students to be in the lesson is approximately 20. This is a multicultural group of students predominantly between the ages of 24 – 45, with numerous ethnic, gender, and creed background. These are UK academic level 5 students; hence it is assumed that they have practical, theoretical, or technological knowledge and understanding of a subject or field of work to find ways forward in broadly defined, complex contexts. These students must be able to generate information, evaluate, synthesise the use information from a variety of sources. Various approaches to addressing the various identified students needs will be adopted throughout the lesson. Such will include:- 1. Progressive tasks 2. Digital resources 3. Verbal support

	4. Variable outcomes 5. Collaborative learning 6. Ongoing assessment 7. Flexible-pace learning
Equality & Diversity	Variety of teaching techniques will be employed to ensure that the needs of each individual learner are met.
Safeguarding & Prevent	Safeguarding policies and the Prevent duty are strictly observed to ensure the safety, well-being, and inclusivity of all students and staff.
Health & Safety	SIRM H&S policies will be maintained.
Learning Resources	**Teaching and Learning Materials**
	• Gelman, A. et al. (2013). Bayesian Data Analysis. CRC Press. • Montgomery, D. C., & Runger, G. C. (2014). Applied Statistics and Probability for Engineers. Wiley. • Wood, S. (2017). Generalized Additive Models. CRC Press. • Shumway, R. H., & Stoffer, D. S. (2017). Time Series Analysis and Its Applications. Springer.

Learning Outcome	Assessment Criteria
LO1. 1. Apply advanced statistical models to real-world datasets.	**Written Report**: 1.1 Build and interpret multivariate regression models. 1.2 Analyze model assumptions and fit.
LO2. 2. Demonstrate Bayesian reasoning in data analysis.	**Exam**: 2.1 Use Bayes' theorem for probabilistic inference. 2.2 Interpret posterior distributions and credible intervals.
LO3. 3. Evaluate time-series and survival analysis models.	**Portfolio**: 3.1 Apply ARIMA and exponential smoothing methods. 3.2 Conduct survival analysis with censored data.
LO4. 4. Critically assess model diagnostics and selection methods.	**Practical Exercise**: 4.2 Apply AIC/BIC for model comparison. 4.2 Perform residual analysis and hypothesis testing.

No	Learning Outcome / Topic	Learning and Teaching Activities	Which assessment criteria does the session relate to?	Day/month/year/ signature
1.	**Multivariate Linear Regression**	**Multivariate Linear Regression** Model formulation, interpretation of coefficients	LO1: Advanced Statistical Models	
2.	**Model Assumptions & Diagnostics**	**Model Assumptions & Diagnostics** Linearity, homoscedasticity, multicollinearity (VIF)	LO1: Advanced Statistical Models	
3.	**Generalized Linear Models (GLMs)**	**Generalized Linear Models (GLMs)** Logistic regression, Poisson regression	LO1: Advanced Statistical Models	
4.	**Mixed Effects Models**	**Mixed Effects Models** Fixed vs. random effects, hierarchical modeling	LO1: Advanced Statistical Models	
5.	**Nonlinear Regression**	**Nonlinear Regression** Polynomial, spline, and kernel regression	LO1: Advanced Statistical Models	
6.	**Bayesian vs. Frequentist Paradigms**	**Bayesian vs. Frequentist Paradigms Key differences, subjective vs. objective probability**	LO2: Bayesian Inference	
7.	**Bayes' Theorem & Applications**	**Bayes' Theorem & Applications** Disease testing, spam filtering case studies	LO2: Bayesian Inference	
8.	**Half-Term Exam**	**- Review of LO1 topics** **- Practice questions and mock assessment** **- Half-term assessment based on LO1 (theory)**	LO1 LO2	
9.	**Prior Selection**	**Prior Selection** Conjugate priors, non-informative vs. informative priors	LO2: Bayesian Inference	
10.	**Markov Chain Monte Carlo (MCMC)**	**Markov Chain Monte Carlo (MCMC)** Gibbs sampling, Metropolis-Hastings algorithm	LO2: Bayesian Inference	

11.	**Posterior Interpretation**	**Posterior Interpretation** Credible intervals, highest density regions (HDR)	LO2: Bayesian Inference	
12.	**Time-Series Fundamentals**	**Time-Series Fundamentals Stationarity, autocorrelation (ACF/PACF)**	LO3: Time-Series & Survival Analysis	
13.	**ARIMA Models**	**ARIMA Models Identification (p, d, q), SARIMA for seasonality**	LO3: Time-Series & Survival Analysis	
14.	**Final Exam Preparation & Review**	**- Comprehensive review of all learning outcomes** **- Practice questions and revision of key topics**		
15.	**Final Exam**	**- Final-term assessment covering all learning outcomes (theory and practical elements)**		
16.	**Feedback & Reflection**	**- Review of final exam** **- Individual feedback on performance** **- Reflective discussion on key learning points**		
17.	**Exponential Smoothing**	**Exponential Smoothing Holt-Winters, ETS models**	LO3: Time-Series & Survival Analysis	
18.	**Survival Analysis Basics**	**Survival Analysis Basics** Kaplan-Meier estimator, hazard functions	LO3: Time-Series & Survival Analysis	
19.	**Cox Proportional Hazards Model**	**Cox Proportional Hazards Model** Interpretation of coefficients, censoring	LO3: Time-Series & Survival Analysis	
20.	**Bootstrapping**	**Bootstrapping** Nonparametric confidence intervals	LO3: Time-Series & Survival Analysis	

21.	**Information Criteria**	**Information Criteria AIC, BIC, DIC for model comparison**	LO4: Model Evaluation & Selection	
22.	**Cross-Validation Techniques**	**Cross-Validation Techniques k-fold, time-series CV (rolling window)**	LO4: Model Evaluation & Selection	
23.	**Half-Term Exam**	**Project Full statistical analysis report (e.g., marketing ROI)**		
24.	**Residual Analysis**	**Residual Analysis QQ plots, Shapiro-Wilk test, heteroscedasticity tests**	LO4: Model Evaluation & Selection	
25.	**Hypothesis Testing**	**Hypothesis Testing Likelihood ratio test, Wald test**	LO4: Model Evaluation & Selection	
26.	**Bayesian Workflow**	**Bayesian Workflow From prior elicitation to posterior predictive checks**	LO5: Integration & Case Studies	
27.	**Time-Series Forecasting Project**	**Time-Series Forecasting Project Stock prices, COVID-19 case prediction**	LO5: Integration & Case Studies	
28.	**Survival Analysis Application**	**Survival Analysis Application** Clinical trial data (R/Python implementation)	LO5: Integration & Case Studies	
29.	**Final Exam Preparation & Review**	**LO1, LO2, LO3, LO4**	LO1, LO2, LO3, LO4	
30.	**Final Exam**		LO1, LO2, LO3, LO4	