Dublin Business School Assessment Brief

Assessment Details

Unit Title	Statistics for Data Analytics
Unit Code	B9DA101
Unit Leader	Terri Hoare
Level:	9
Assessment Title	Applied Statistics - Modelling
Assessment Number	Summative CA02-1
Assessment Type:	Individual
Assessment Weighting	30%
Issue Date:	Week of 26th March 2018
Hand in Date:	Sunday 22nd April 2018 23:55
Mode of Submission:	On-line

Assessment Task (100 Marks)

Using the Boston data set introduced during LAB work, apply linear regression modelling to predict the per capita crime rate using other variables in the data set. In other words, per capita crime rate is the response and the other variables are predictors.

a) Use descriptive statistics to explore the dataset. [10]

b) For each predictor, fit a simple linear regression model to predict the response. Complete regression analysis. Describe your results. In which of the models is there a statistically significant association between the predictor and the response? Create plots to back up your assertions. [30]

c) Fit a multiple regression model to predict the response using all the predictors. Describe your results. For which predictors can we reject the null hypothesis $H_0 : \beta_j = 0$? [10]

d) How do your results from (b) compare to your results from (c)? Create a plot displaying the univariate regression coefficients from (b) on the x-axis, and the multiple regression coefficients from (c) on the y-axis. That is each predictor is displayed as a single point in the plot. Its coefficient in a simple linear regression model is shown on the x-axis, and its coefficient estimate in the multiple linear regression model is shown on the y-axis. [10]

e) Is there evidence of non-linear association between any of the predictors and the response? To answer this question, for each predictor X, fit a model of the form

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \varepsilon$$

Summarise your findings.

[10]

f) Using the Boston data set, fit classification models in order to predict whether a given suburb has a crime rate above or below the median. Explore both logistic regression and LDA using subsets of the predictors and comment on the use of frequentist and/or Bayesian statistics where appropriate.

[30]

A report should be submitted together with supporting software analyses / R-code via online submission on or before TBC. The grade assessment will be based on the DBS CA grading scheme which has been included in this document.

DBS Grade Assessment Policy (B9DA101)

Module Descriptor	Mark Band	Criteria	Determinator within grade band
A (Outstanding)	80-100	• Displays a thorough and systematic knowledge of module content, solution and handover process and documentation. • Clear grasp of the issues involved, with evidence of innovative and original use of learning resources. • Knowledge beyond module content. • Clear evidence of independence of thought and originality • Methodological rigour • High critical judgement and confident grasp of complex issues	Displaying insight into problem and solution presented.
A (Clear)	70-79	• Methodological rigour • Originality • Critical judgement • Use of additional learning resources	Methodological rigour
B	60-69	• Very good knowledge and understanding of the module content. • Well-argued answer • Some evidence of originality and critical judgement • Sound methodology • Critical judgement and some grasp of complex issues.	Extent of use of additional or non-core learning resources
C	50-59	• Good knowledge and understanding of the module content. • Reasonably well-argued answer • Largely descriptive or narrative in focus • Methodological application is not consistent or thorough	Understanding of the main issues
D	40-49	• Lacking methodological application • Adequately argued • Basic understanding and knowledge • Gaps or inaccuracies but not damaging	Relevance of knowledge displayed
E (Fail)	0-39		Weakness of understanding

General Requirements for Students:

PLEASE READ CAREFULLY

1. It is your responsibility to ensure your file is uploaded correctly.
2. Students are required to retain a copy of each assignment.
3. When an assignment is submitted, it is the student's responsibility to ensure that the file is in the correct format and opens correctly.
4. Students should refer to the assessment regulations in their Course Guide.
5. DBS penalises students who engage in academic impropriety (i.e. plagiarism, Collusion and / or copying). Please refer to the referencing guidelines on Moodle for information on correct referencing.
6. All relevant provisions of the Assessment Regulations must be complied with.
7. Penalties for late submission of assignments are as follows:
    a. 25% penalty for assignments submitted within 5 working days of the deadline.
    b. No marks for assignments submitted more than 5 working days after the deadline.
8. Extensions to assignment submission deadlines will be granted in exceptional circumstances only. The appropriate "Application for Extension" form must be used and supporting documentation (e.g. medical certificate) must be attached. Applications for extensions should be made directly to the Head of Year or Programme Leader in advance of the deadline date.