Masters in Science in Cybersecurity

Research Methods

Project (40%)

Semester 3 2022/2023

Due date: July 16th 2023.

Overview

Repositories of official government statistics and databases are available for Ireland at https://data.gov.ie/, for the United Kingdom from https://data.gov.uk/ and for the United States at https://data.gov.

Using datasets of your choice from one or more of these sites and working in teams of 2 or 3, you are required to undertake two separate pieces of statistical analysis.

1. The first analysis involves the creation and evaluation of a multiple regression model.
2. The second analysis involves the creation and evalaution of a logistic regression model.

You may carry out the analysis using any appropriate software tool of your choice.

Report

Your report should be in PDF format and should contain the following sections for each analysis:

- A description of the sources of the datasets;
- A clear statement of the objectives of the analysis including the null hypothesis and the alternative hypothesis;
- Descriptive statistics for each variable in the dataset;
- A list of the independent/dependent variables in the proposed analysis and their levels of measurement;
- The results of an appropriate regression analysis performed using R, Python or any other software tool of your choice and presented in a manner consistent with the standard of reporting in research articles;
- Diagnostic plots and tests appropriate to the analysis being carried out, accompanied by an interpretation of these tests and plots;
- A clear statement of the decisions and conclusions you have come to as a result of the analysis

All relevant output should be included in the main body of the report.

Note that there is no recommended word count for the report. Instead the emphasis is on the ability to locate suitable datasets, formulate appropriate hypotheses, conduct relevant statistical tests, and to draw conclusions based on your findings. As such, word counts will vary depending on the data used, and interpretation of findings.

Marking

The report will be marked as follows:

- Objectives, null and alternative hypotheses (5% per analysis)
- Descriptive statistics / levels of measurement (5% per analysis)
- Regression Analysis (10% per analysis)
- Diagnostics (10% per analysis)
- Formal report of results (10% per analysis)
- Decision and conclusion (5% per analysis)

The remaining 10% of the overall mark will be allocated to the discussion of the team working aspect of the project described above. This discussion should be suitably detailed.

Submission

The report should be uploaded to the Turnitin link on Moodle by the date shown against the upload link. The standard School of Computing penalties will apply to late submissions

Academic Integrity

Any written work created by others must be properly cited and should be paraphrased or summarised where possible, otherwise it should be included in quotes. Figures not created by you should include an acknowledgment detailing the name(s) of the creator(s).

The use of large language models such as ChatGPT is also **strictly prohibited**.

Students are strongly advised to familiarise themselves with the Guide to Academic Integrity produced by the NCI Library [1].

Note: All submissions will be electronically screened for evidence of academic misconduct, e.g. plagiarism, collusion and misrepresentation. Any submission showing evidence of such misconduct will be referred to the college's academic misconduct committee for disciplinary action.

Your lecturer reserves the right to request a viva presentation with any team should this be deemed necessary for any reason.

Under no circumstances may you upload this assessment or your submission to any internet site, including but not limited to Chegg or CourseHero.

[1] https://libguides.ncirl.ie/academicintegrity