FINAL EXAMINATION
Due date: Midnight, Sunday 29 October 2023
Any extensions to the due date needs to be by special approval from your student office.

1   Brief

The file `compdat.txt` contains data on the number of complaints received, along with some demographic information, for 94 doctors who worked in an emergency service at a hospital.

The following are the variables in the data.

(a) `visits`: the number of patient visits

(b) `complaints`: the number of complaints against the doctor in the previous year

(c) `residency`: is the doctor in residency training (Y = Yes, N = No)

(d) `gender`: gender of the doctor (M = male, F = female)

(e) `revenue`: doctor's hourly income (dollars)

(f) `hours`: total number of hours the doctor worked in a year

Aim of the analysis: To determine the effect of the variables on the number of complaints received.

Notes

(a) Consider appropriate models for the data.

(b) Select your best model and justify your choice.

(c) Note that several models could be equally good. In that case the simpler model is preferred.

(d) You should investigate some meaningful interaction between categorical variables. Do not over-complicate your model.

(e) Give a careful interpretation of your best model.

2   Examination submission

You are required to write a journal style article presenting your analysis. **This is a take home examination. You can use any resource you wish, except consulting with another person. The submission should be entirely your own work. You are not allowed to use any AI tools for this examination. Any suspected breach of this will be reported to the faculty misconduct officer for further action.**
    Submit your article to the LMS link under `Final Exam 2023`.

Your paper will include the following sections.

(a) Title and author information.

(b) Abstract of no more than 500 words.

(c) Introduction of no more than two pages. Here you can support your paper with other information from online searches, relating to the number of complaints received against doctors in hospital emergency departments. As a final paragraph in this section, give an outline of the paper, stating what is contained in each of the sections that follow.

(d) Methodology of no more than a page. Here you can present the statistical modelling without giving mathematical details. You can assume the reader is familiar with standard statistical techniques. You should provide sufficient detail and clarity so that another person can replicate your analysis.

(e) Results, describing the findings of the modelling, in not more than 2 pages. You should use equations to illustrate and assist with your explanations.

(f) Discussion of no more than a page. Discuss your findings with reference to your introduction.

(g) References. As many as you have used. Use any consistent format for references. You may be able to simply copy and paste these.

(h) An appendix that details ALL the `R` code and output used in the analysis. Include only the code that is relevant to your analysis and your report. The code should be clearly annotated so it is clear to follow and understand, and should indicate each stage of the analysis.

3   Marking Scheme

The marks will be allocated as follows.

(a) The Paper: 85 marks.

    (1) Abstract = 10 marks

    (2) Introduction = 20 marks. Particular attention will be given to your relevant literature search.

    (3) Methodology = 10 marks. Clarity of expression is important.

    (4) Results = 25 marks. Clearly describe your findings.

    (5) Discussion = 20 marks. Discuss the findings and any implications.

(b) Appendix A for `R` code: 15 marks. Clearly describe the steps in your modelling. Clearly state what you preferred model is with justification based on verifying model assumptions.

Note: Include only relevant tables and graphs in your report. Any table or graph that is included MUST be discussed in the paper. **Too many or irrelevant tables and graphs will be penalised.**