The University of Western Australia  
SCHOOL OF MATHEMATICS AND STATISTICS  
STAT 1106 Economic and Business Statistics  
Project  
Semester 2, 2007

Due 3 pm Friday 12 October 2007

This assignment is in the form of a report. The report should be written on loose A4 size paper only, not in folders or exercise books. Fill in an official cover sheet (available online), staple it to the front of your report and submit it to your Tutor’s Box by 3 pm Friday 12 October May. Late reports will be penalised 5% of the total marks per day unless they are accompanied by a special consideration approval from your sub dean. Solutions which lack a completed and signed cover sheet, or are not stapled, will not be marked and will be awarded a mark of zero.

Your report must be entirely your own individual work (see University policy on plagiarism). This does not preclude preliminary discussion with colleagues or staff.

This Assignment counts 15% towards your final assessment in STAT1106.

The First National Bank (hereafter called the Bank) is facing a gender discrimination suit. The charge is that it’s female employees are receiving a substantially lower salary than its male employees. The Bank’s employee database is listed in Bank.xls, available online and also in the Maths Lab folder in the Mathematics Computing Laboratory. For each employee, the data set includes the following variables:

- EducLev: education level, a categorical variable with categories 1 (finished high school), 2 (finished some tertiary education), 3 (obtained a bachelor’s degree), 4 (took some postgraduate courses), 5 (obtained a postgraduate degree).
- Job Grade: a categorical variable indicating the current job level, the possible levels being 1 (lowest) to 6 (highest).
- YrHired: year employee was hired.
- YrBorn: year employee was born.
- Gender: a categorical variable with values “Female” and “Male”.
- YrsPrior: number of years of work experience at another bank prior to working at First National.
- PCJob: a categorical yes/no variable indicating whether the employees current job is PC related.
- Salary: current salary in thousands of dollar.

The project is to analyse that data using appropriate statistical techniques to determine if there is gender discrimination at the Bank. The analysis will be conducted in Excel.

Notes

1. Look at the laboratory work that has so far been completed this semester. Use whatever you can from this analysis.

2. You must include some exploratory analysis of the data. This can take the form of table, graphs and/or summary statistics. Select carefully what you include here. Comment on your findings.
3. Carefully describe the steps in your analysis. You should assume that the reader knows statistical methodology, so you do not need to go into the details of the statistical theory. For example, you do not need to explain how to perform a hypothesis test. However, you do need to explain how you arrive at a particular statistical model, or why you chose the method of analysis that you performed.

4. State your conclusions carefully. Explain any surprising results or suggest possible reasons for them if you can. Suggest any improvements to the analysis or any weaknesses in your analysis.

The report should follow the format (but it could differ in some aspects — for example you may name your section differently, or you may not have some sections that are in my example, or you may include other relevant sections) of the sample report available online. It should cover the following aspects.

(a) An executive summary.

(b) Appropriate analysis of the data, checking any model assumptions.

(c) A conclusion that summarises your findings, makes recommendations, discusses any weaknesses of the analysis and suggests any improvements.

1. Some guidance regarding the project will be given in the laboratory classes in the weeks prior to the project due date.

Relevant details of the methodology and any other complementary information (such as complete Excel outputs) should be put in appendices.

Marks will be awarded for presentation and coherence of the report, along with the correctness of analysis and justification for conclusions.

Your report must not be longer than 3,000 words (or 8 typeset pages), not including tables, graphs and the appendices.

A short, coherent and complete report can expect to score a higher mark than a longer report.

You are welcome to discuss any aspects of the project with me or any other STAT1106 staff, or with your classmates, but the analysis and report must be your own work.

©Dr Nazim Khan
Semester 2
2007