Statistics 747 – Assignment 2

Due 10:30am 15th August 2010

The data in ‘enquiry.xls’ from <http://www.stat.auckland.ac.nz/~balemi/>

is for 98 people who were interviewed (randomly selected from a company data base) about aspects of service they encountered when they made and enquiry about their insurance policy. (The company will be referred to as XX for privacy reasons) The question they were asked were:

###### READ

For the next section in the survey I would like you to answer using a rating scale from 1 to 10, where 1 = extremely dissatisfied and 10 = extremely satisfied. Just tell me if a question is not relevant or you don’t know the answer. **CONFIRM RESPONDENT UNDERSTANDS SCALE**

3. Thinking about your dealings overall with XX Insurance how would you rate them in terms of:

|  |  |  |
| --- | --- | --- |
|  | Rating | Reference |
| Their overall quality of service, taking everything into account |  | **Q32A1** |

4b. Using the same scale, where 1 = extremely dissatisfied and 10 = extremely satisfied, how would you rate XX Insurance on the following **(ROTATE AND READ)**:

|  |  |  |
| --- | --- | --- |
|  | Rating | Reference |
| The speed with which the phone was  |  | **Q37A1** |
| Ease of getting through to someone who can help with your enquiry |  | **Q37A2** |
| Listening to and showing an interest in your enquiry |  | **Q37A3** |
| The consultant spoken to was able answer your questions and resolve your enquiry |  | **Q37A4** |
| Following through on promises and keeping you informed  |  | **Q37A11** |
| Dealing with you in a friendly and courteous manner |  | **Q37A12** |
| Providing clear explanations in easy to understand terms |  | **Q37A13** |
| The product knowledge of call centre/XX branch staff  |  | **Q37A14** |
| Treating you as an important customer |  | **Q37A15** |
| Dealing with you in a professional manner |  | **Q37A16** |
| Their speed of response to your requests |  | **Q37A17** |

Along with questions about their overall education and house hold income:

Q33. What was the highest level you completed in your formal education: **READ AND CODE ONE ONLY.**

|  |
| --- |
| Reference |

|  |  |  |
| --- | --- | --- |
| Primary | 1 | Q77 |
| Secondary | 2 |  |
| School Certificate | 3 |  |
| UE/6th Form/Bursary | 4 |  |
| Trade Certificate or Diploma | 5 |  |
| University Degree | 6 |  |
| Post Graduate Degree | 7 |  |
| Refused **(DO NOT READ)** | 8 |  |

 Q34. Which of the following best describes your total household income, before tax? Is it..... **READ AND CODE ONE ONLY.**

|  |
| --- |
| Reference |

|  |  |  |
| --- | --- | --- |
| Under $20,000 | 1 | Q84 |
| $20,000 - $29,999 | 2 |  |
| $30,000 - $39,999 | 3 |  |
| $40,000 - $59,999 | 4 |  |
| $60,000 - $79,999 | 5 |  |
| $80,000 - $99,999 | 6 |  |
| $100,000+ | 7 |  |
| Refused **(DO NOT READ)** | 8 |  |
| Don’t know **(DO NOT READ)** | 9 |  |

Q1 (a)

By inspecting the relationship (if any) between education level and claimed household income levels impute suitable replacement values for missing values (anything code 8 or above) for each of these by imputing them with a suitable representative data point. Justify your answers).

Q1(b)

Use random hot-deck imputation within each income band (which has now been ‘cleaned’ in Q1(a) above) impute the missing values for the satisfaction questions (Q32A1,Q37A1,Q37A2,Q37A3,Q37A4,Q37A11,Q7A12,Q37A13,Q37A14,Q37A15,Q37A16,Q37A17) .

(Note – if any observation has all missing values on

Q37A1,Q37A2,Q37A3,Q37A4,Q37A11,Q7A12,Q37A13,Q37A14,Q37A15,

Q37A16,Q37A17 firstly delete this person from your analysis.)

Provide your program code (in R, Splus or SAS) you used to impute these missing values?

Q1(c)

Assume now that the data is missing completely at random (MCAR). Replace any missing values in the original ‘un-cleaned’ data set by an appropriate value.

(Note – if any observation has all missing values on Q37A1,Q37A2,Q37A3,Q37A4,Q37A11,Q7A12,Q37A13,Q37A14,Q37A15,

Q37A16,Q37A17 firstly delete this person from your analysis.)

Q2

Produce a performance-importance matrix for of each aspect of service for the `cleaned data” from Q1c. Importance, in this case, should be calculated by using simple linear regression or correlations.

As a consequence, indicate how this Insurance company should prioritise improvements to their service.

Q3

Repeat the above analysis, except this time using factor analysis along with multiple logistic regression to calculate the importance of each aspect of service. Produce another performance-importance matrix based on these new values.

Comment and compare these results with what you found in Q1 above.

Note: The usual checks for the validity of these analyses will be reported in a separate technical appendix for each question. The emphasis here is on presentation and interpretation – so make it interesting and presentable. The returned assignment should be given back as a word (or something very similar) document.