

6G7Z1020

Assignment 1

Refer to the work you have done in the lab practical, the relevant examples and the exercises given in the lecture notes. For each answer, include your results in a formatted PDF document that contains your graph, one small paragraph with your comment and a separate section that includes your SAS code.

Questions

1 Melbourne Cup

The Melbourne Cup is the most famous Australian horse race. It is conducted annually on the Flemington Racecourse in Melbourne, Victoria. The file `melb.csv` contains¹ the winner's time in seconds (variable `Time`) for each year (variable `Year`) from 1861 until 2013.

- (a) Plot `Time` as a function of the `Year`. [30]
- (b) Comment on the winner's time as time progresses. Does the plot suggest a linear relation? [20]

2 Sexual Activity

The file `SA.csv` contains the reported number of sexual partners (variable `Partners`) for 1682 males and 1850 females (variable `SEX`) from a US survey² which was conducted interviewing randomly selected households using stratified sampling. Only a subset of the responses is presented and the non-responders have been omitted. Assume that the variable `Partners` is a categorical variable.

- (a) Create a bar graph distinguishing between the Males and the Females. You might as well produce two graphs one for Males and one for Females. Hint: [30]
 - You can generate a grouped barplot using the `/GROUP=` option on `VBAR` statement.
 - You can select the observations that satisfy a condition, e.g. a certain category-value `VALUE`, using the `WHERE=VALUE` statement.
- (b) Using the previous graph, identify on which aspects the sexual activity of the two genders is similar and on which aspects they differ. Justify your response. [20]

¹Adapted from http://en.wikipedia.org/wiki/List_of_Melbourne_Cup_winners

² 1989,1990,1991 General Social Survey