

Exercise #2 Instructions

Crime Analytics (CJUS 6106)

Instructions

For this assignment you will be using the `nlsy97.Rdata` file to run several analyses from the “Interacting, Querying, and Reporting Data” lecture. To load the data file you should use the `load()` function. To make things easier for yourself, be sure to put the data file in the same folder as your `.rmd` file. If you do so, you do not have to specify the entire file path to the data set, just the name of the data set (including its file extension). Be sure to write thorough interpretations where asked and to put all code within code chunks. You should submit both a knitted PDF and your `.rmd` file for this exercise by midnight on the date indicated on Canvas.

Question 1

Use the `summary()` function to provide a summary of the `nlsy97` data frame. Following this, provide an indication of the level of measurement for each variable with a justification for your decision.

Question 2

Choose three outcome variables measured in 1999 to individually match with the following three variables: age at the 1998 interview (`age98`), household structure at the 1998 interview (`hh_struc98`), and number of weeks employed since the 1997 interview (`empweeks_num98`). So, you will have three different pairs, each pair having one of those three variables as its independent variable and one of the variables measured in 1999 as the outcome variable.

Provide the proper data queries available to you (tables, correlations, summaries by nominal/ordinal group, and scatterplots) and provide an interpretation for the relationship between the two variables. You may only use each type of query once.