

Homework 1

POLI 5: Data Analytics for the Social Sciences

Homework 1 is due on Saturday, August 13, at 11:59pm via the [Gradescope](#) tab on Canvas. All submissions must be in PDF format.

If this is your first time using Gradescope, please watch [this video](#) and budget enough time to familiarize yourself with the user interface. You will need to (1) submit the writeup as a PDF document, and (2) tag individual problems according to the instructions.

Part I: Research Questions, Theories, and Hypotheses

Question 1

Assignment Project Exam Help

Formulate a research question that can be empirically tested using one of the datasets in our [Google Drive folder](#).

<https://powcoder.com>
Hint: Review the components of a good empirical research question by watching [this video](#).

Question 2

Add WeChat powcoder

Elaborate a theory to propose an answer to your research question. Remember that we are defining *theory* as “A reasoned and precise speculation about the answer to a research question.” (King, Keohane, and Verba 1994)

Hint: Remember what Marco says in the video from Lesson 1: “A good theory is specific and clear about how the independent variable affects or causes the dependent variable.”

Question 3

State your dependent variable, your independent variable, and pose a hypothesis that can be tested with the dataset of your choice. Remember that we are defining *hypothesis* as “A statement of the relationship that you expect to find between the dependent and the independent variable.” (Powner 2004)

Hint: Review the four components of a good hypothesis in [this video](#). Be sure to pose a hypothesis that is specific, falsifiable, not immediately verifiable, and has directionality.

Part II: Descriptive Analysis using R

Choose two variables from a dataset in our [Google Drive folder](#). You are encouraged to choose variables related to your work for Part I of this homework, but this is entirely optional. You are allowed to choose variables unrelated to Part I or from a different dataset.

Question 1

Perform univariate analyses on each one of your variables, separately. For each variable, please include a graphic (a histogram or a univariate bar plot with frequencies) and descriptive statistics (a distribution table or summary statistics table). Remember that what type of graphic and table to include for each variable will depend on the types of your variables. Be sure to format your graphics appropriately, including titles and axis labels. If you generate a histogram, be sure to set the bar width and the x axis in a way that facilitates interpretation.

Include the code you used in R to generate your plots and tables. Please format your code with **monospaced font** (examples of monospaced font in MS Word include Courier New and Lucida Console.)

Hint: Review the slides for Lessons 3 and 4.

Question 2

Conduct a bivariate analysis with your pair of variables. Include a plot that depicts the relationship between your variables. Remember that what type of plot to include depends on the types of your variables. Be sure to format your plot appropriately, including titles and axis labels.

Include the code you used in R to generate your plot. Please format your code with **monospaced font** (examples of monospaced font in MS Word include Courier New and Lucida Console.)

Hint: Review the slides for Lessons 3 and 4.

Question 3

Write a 4 to 5 sentence paragraph summarizing what your descriptive analysis reveals about the data. How are each of your variables distributed? Are the two variables correlated? Is the correlation in the direction you expected? What are the key takeaways of your descriptive analysis?