CMPSC 301
Data Analytics
Fall 2018

Lab 5: Political Science Exploratory Data Analysis
19th October 2018

Objectives

To learn how to ask exploratory questions while using data from a political science area. To provide the necessary analysis to be able to answer those questions, combining skills, methods and theory from data analytics. To extract meaningful correlations in the data on political cultures in the nature of civil legal representation.

Reading Assignment

Review Chapter 5 and then read chapters 7, 8, and 13 in the course book. Also, please review your notes from the lab discussion of the data which you will utilize for the completion of this lab. Please use the included paper with your lab data entitled, *Funding Poverty Law in an Era of Decentralization*, by Brian M. Harward [1] of Allegheny College. The online version of this article can be found here: `https://www.tandfonline.com/doi/full/10.1080/10875540802350062?scroll=top&needAccess=true`.

GitHub Starter Link: Group work

<div align="center">

`https://classroom.github.com/g/gIXr61Yn`

</div>

To use this link, please follow the steps below.

- Your group leader will click on the link and accept the assignment and prepare a team name. All other members will later click on the link and select their team's name from the list that will appear.

- Once the importing task has completed, click on the created assignment link which will take you to your newly created GitHub repository for this lab.

- Clone this repository (bearing your name) and work on the lab locally.

- As you are working on your lab, you are to commit and push regularly. You can use the following commands to add a single file, you must be in the directory where the file is located (or add the path to the file in the command):

  - `git commit <nameOfFile> -m ``Your notes about commit here''`
  - `git push`

HANDED OUT: 19th OCT 2018

Alternatively, you can use the following commands to add multiple files from your repository:

- `git add -A`
- `git commit -m ''Your notes about commit here''`
- `git push`

Groupwork

You are to work in a group of not more than four people for this lab. Be sure to discuss each of the tasks and proceed after the group has come to a complete agreement. **Each person is to turn in his or her own report and code, however the names of all lab partners should be listed in the submission.**

Political Science Data Analysis

In this lab you will conduct your own exploratory data analysis on a data set provided by Dr. Harward. In particular, you should follow the following steps:

1. Your data set has been included in this lab as well as the article which describes the data of the columns. See Sections **Research Methods** and **Results** of the article for discussion about each of the columns. In these sections, you will also learn of the background to provide some context to the data.

2. One way to open the csv data file is to use the following command.
   `poliSciData <- read.table(file.choose(), header = TRUE, sep = ",")`

3. Spend the necessary time to review the data with your group.

4. This lab asks that you generate five (5) research questions about the included data set. It is expected that you will select questions which may be answered by performing an analysis on the columns of data in the lab.

5. Utilize your analytics and programming skills to write an R program that performs an analysis and provides convincing evidence to answer your research question. Be sure to leave comments in your code to help the reader to understand its flow. Also, be sure to document your results which you will use at the end of the lab for your report document. Note, this document will be used to explain the research questions and detail their conclusions.

6. In research, one is constantly finding new questions which result from the conclusions of others. In the nature of exploration, you may want to add some of your research questions to the lab which come after an initial research question has been resolved. *Remember, you only need five questions and five conclusions with programming solutions.*

Communicating your Results From Programming Solutions

During your lab, you will finalize your solutions and your analysis and write a report with your findings and your reflection. Your final deliverable for this lab will include:

- **Your Political Science Data Analysis Report**: In addition to the R code that provides your solutions, you are to write a concise report (`writing/reportAndReflections.md`) that includes the following details.

    1. **Explain the data behind your work**: A brief explanation of the columns you used from the data for your study. For example, be sure to visit each column you used for analysis to explain what the data is and how it is relevant to your work. Remember to review the article's **Research Methods** and **Results** sections for help learning about the data itself.

    2. **Data and Five (5) Research Questions**: After you have spent sufficient time to review the data, you should be able to ask important questions that can also be addressed using analysis techniques and programming. You will also have these questions selected based on the data that you have in the data table.

       Here, you are to list the questions and a brief justification about why the data column(s) you chose was/were able to answer the question.

    3. **Responses to the Questions**: The responses to your five questions where you discuss the code to perform the analyses.

    1. **Reflection**: Based on Dr. Harward's provided data and article, discuss the processes to make political decisions to fund states' legal services. In addition, discuss how such decisions are affected. Reflect on potential faults in this system (as you understand them) and how data analysis may be able to identify such faults.

Important Details

Lab directory structure: Make sure you have placed your submission materials for this lab into your group submission GitHub repository.

Note: Please remember to include your name, as well as, those of your group members, on everything you submit for the lab.

Required Deliverables

1. In File, `writing/report.md`; Report that includes the data descriptions, exploratory questions and their responses using analysis. Write a short discussion after each question and response to discuss what you have learned from your analysis. At the end of this file, please respond to the reflection question in blue, above.

2. In File, `src/analysis.r` The R source code that you used to answer your questions. Please be sure to list the question and add all necessary code for analysis, following the question. This will help the instructor to follow your analysis and solution.

3. In File: `writing/reportAndReflections.md`, respond to the questions in blue, above.

References

[1] Harward, Brian M. "Funding Poverty Law in an Era of Decentralization." Journal of Poverty 12.4 (2008): 387-410.