

Assignment #2

Complete the following tasks:

- Find a dataset that you find interesting. You can look at a variety of sources for data including:
 - Religion, health, etc.: www.Thearda.com
 - Census, Education, Wealth, Time-Use, etc: www.ipums.org
 - Education: www.nces.ed.gov
 - Mental Health: www.cdc.gov/mentalhealth/data_publications/index.htm
 - Many Other Sources
- Once you have identified a dataset, explain how you accessed it. What links did you click? Was there any registration required? Did you download directly or was there an online system you navigated?
- In R:
 - Import the data
 - Provide a list of variables in the dataset
 - Copy and paste this list in your word document
 - What is the structure of the data? Which variables are character and which are numeric?
 - Copy and paste this list in your word document
 - Describe the data through Central Tendencies:
 - Give me the mean, median, and mode of **five variables**.
 - Describe the data through variation:
 - Choose one variable and give me:
 - The variance, range, and standard deviation.
 - A histogram and either box-plot or stem-leaf plot

Assignment #2

What to turn in

- Publish your homework to GitHub under your user account, the appropriate format (check syllabus for instructions). This means turn in a separate document from your code that answers the questions above.
- Upload the code you used to GitHub.
- Provide me the links of these two files via e-mail no later than Thursday, September 20th at 6:00pm.