CADS - Health Group Meetings Overview

* **2/2/2019 - Introduction to R**

Goals: introduce ourselves, brainstorm ideas for the final project, and work with data types and structures

R file = “HGroup\_2:2:2019.Rmd” - introduction to R variable assignment, types of data (numerical, character, boolean, factors), structures (vectors, matrices, data frames), data exploration, and one variable linear models

**Preparation for next meeting**: please take a look at this description of [linear models](http://www.sthda.com/english/articles/40-regression-analysis/167-simple-linear-regression-in-r/).

If you have time, these websites on [multilinear regression](http://www.sthda.com/english/articles/40-regression-analysis/168-multiple-linear-regression-in-r/) and [interaction effects](http://www.sthda.com/english/articles/40-regression-analysis/164-interaction-effect-in-multiple-regression-essentials/) are helpful. We will be applying regression concepts to data for the next meeting, and possibly taking a look at Tidyverse ([ggplot2](http://www.sthda.com/english/wiki/be-awesome-in-ggplot2-a-practical-guide-to-be-highly-effective-r-software-and-data-visualization)) packages if we have enough time. See you then!

* **2/9/2019 - Project, Linear Models and Intro to R Packages**

Goals: decide on project idea, interpret linear model output, introduce plotting packages and visualization with R Shiny

R file = “HGroup\_2-9-2019.Rmd” - interpret parameters of linear models from previous meeting and try out different models that will fit the data. Applying interaction effects to clustered data. Using ggplot2 and intro to R Shiny

**Preparation for next meeting:** today we decided on an idea for our semester project suggested by Jenna and Jane: app visualization of vaccinated population across the United States using CDC survey data. Please look through this [webpage](https://www.cdc.gov/vaccines/imz-managers/nis/datasets.html) for data/survey design and choose some variables that might be interesting to look at. Also, check out this link on creating [Shiny](http://rstudio.github.io/shiny/tutorial/#hello-shiny) apps. Thanks!

* **2/16/2019 - Exploratory Look at CDC Dataset**

Goals: observe distribution of data by creating plots, understand more about the data we will be working with for upcoming weeks, and troubleshooting download concerns

R file = “HGroup\_2-16-2019.Rmd” - graphing variables of interest, formatting individuals by state in preparation for input into R Shiny

**Preparation for next meeting:** during today’s meeting, we looked at the distribution of certain covariates and noticed that we needed to address for missing variables, categories that required additional formatting, and other areas to address. We have entered the data cleaning stage, so take a glance at this user’s guide to [dplyr](https://rpubs.com/bradleyboehmke/data_wrangling) for tips on data wrangling. Also, once we are done formatting the data, we will be creating a map in R Shiny. For now, please take a look at how to incorporate a map into Shiny [here](https://shiny.rstudio.com/tutorial/written-tutorial/lesson5/). Also, please let me know if you have any questions or concerns!

* **3/2/2019 - CDC Dataset and Applications in Shiny**

Goals: visualize dataset in R Shiny, learn about identifying errors and debugging code, and plan our next steps for the semester project

R file = “HGroup\_3-2-2019.Rmd” - data cleaning, creating helper functions to be used in UI - Server and running a prototype of our app

**Preparation for next meeting:** for today’s meeting, we applied our cleaned data to an R Shiny app, but noticed that there were some bugs in our code. Before our next meeting, please try to work on fixing the bugs related to legend intervals, state location mismatch, or anything else you can find. Methods for addressing location mismatch include incorporating coordinates [here](https://quantdev.ssri.psu.edu/tutorials/intro-geospatial-data-and-maps-r) and FIPS [here](https://mgimond.github.io/ES218/Week12a.html) . We also talked about building upon our app to be supported in Leaflet for the upcoming weeks. If you have time, please read through this description of [Leaflet](https://rstudio.github.io/leaflet/).

* **3/30/2019 - Introduction to Leaflet**

Goals: understand coordinate/FIPS assignment, introduce ourselves to Leaflet and prepare for our final project presentations

R file = “HGroup\_3\_30\_19.Rmd” - visualization of vaccination data with Leaflet

**Preparation for next meeting:** it’s all coming together now! For our next meeting, we will integrate R Shiny and Leaflet components to get an interactive map of health determinant variables. For the upcoming weeks, we will be finalizing our product and adding various components to better explain the data (charts, statistical models, etc.). Please read about visualizing Leaflet with Shiny [here](https://rstudio.github.io/leaflet/shiny.html).

**Useful Resources:**

[Introduction to R](https://www.statmethods.net/r-tutorial/index.html) - side tabs include information on importing data, functions, and more

[Shiny app examples](https://shiny.rstudio.com/gallery/) - motivation for final product

[Leaflet](https://rstudio.github.io/leaflet/) - an intro to incorporating our Shiny app into Leaflet