**Agenda 4/30/2020: Descriptive Statistics**

* *Introduction* 
  + *Stress-free learning space to refine our knowledge in R and Python* 
    - *Topics in Rotation:*
      * *Data Cleaning/Processing (e.g. reading, formatting variables, renaming, subsetting, merging data)*
      * *Visualization (e.g. creating tables, plotting methods, formatting plots, dashboards)*
      * *Reporting (e.g. formatting output in pdf, csv, best practices in reporting in DOH, automating reports)*
  + Week 2 Agenda:
    - Review Types of Variables
      * R: numerical, character, logical; matrix, data frame; factor (dplyr glimpse)
      * Python: everything is an object with an associated type; scalars (dtypes)
    - Discrete Variables: Frequency Tables
      * R: tables, prop.tables, crosstabs
      * Python: value\_count, pd.crosstab
    - IR Variables: Summary Stats
      * R: summary; summarytools; gmodel
      * Python: describe, quantile
    - Future last step: Tie together with SMH data / use SMH
  + Potential Future Topics/Structure
    - Reading/Writing Data – 4/17/2020
      * CSV, Excel (large, small)
      * JSON
      * Via link
    - Looking at Data
      * Basic Descriptive Statistics (Discrete/IR) – 4/30/20
    - Data Types
      * Functions/limitations of each data type
      * Dates/Times
      * Converting between data types
    - Dplyr?
    - Data Cleaning
      * Missing data
      * Duplicates
      * Renaming & Subsetting (grep vs in)
      * Creating new variables/data frames
    - Data Wrangling
      * Joining
      * Concatenating
      * Reshaping
    - Visualizing, Reporting, etc.
* Walkthrough exercise with group in R and Python
* Thoughts/Discussion