Code

* [Show All Code](#gjdgxs)
* [Hide All Code](#gjdgxs)
* [Download Rmd](#gjdgxs)

# R Notebook

Count() verb: learn to aggregate data, take many observations and summarize them in one. This is a commom data science strategy for making datasets manageable and interpretable. It’s a very useful type of aggregation when analyzing a dataset.

counties %>%  
 count(state, sort=TRUE)

Add weight: We can add the argument wt, which stands for “weight”, equals population.

counties %>%  
 count(state, wt=population, sort = TRUE)

This means that the column after the “state” will be weighted by the population. In the result, instead of seeing the number of counties in each state, it shows the total population.

========================================================================================================= Counting by region

# Use count to find the number of counties in each region  
counties\_selected %>%  
 count(region, sort = TRUE)

A tibble: 4 x 2

region n 1 South 1420 2 North Central 1054 3 West 447 4 Northeast 217 ========================================================================================================= Counting citizens by state

# Find number of counties per state, weighted by citizens  
counties\_selected %>%  
 count(state, wt=citizens, sort = TRUE)

A tibble: 50 x 2

state n 1 California 24280349 2 Texas 16864864 3 Florida 13933052 4 New York 13531404 5 Pennsylvania 9710416 6 Illinois 8979999 7 Ohio 8709050 8 Michigan 7380136 9 North Carolina 7107998 10 Georgia 6978660 # … with 40 more rows ========================================================================================================= Mutating and counting: “What are the US states where the most people walk to work?”

counties\_selected %>%  
 # Add population\_walk containing the total number of people who walk to work   
 mutate(population\_walk = population \* walk / 100) %>%  
 # Count weighted by the new column  
 count(state, wt = population\_walk, sort = TRUE)

A tibble: 50 x 2

state n 1 New York 1237938. 2 California 1017964. 3 Pennsylvania 505397. 4 Texas 430783. 5 Illinois 400346. 6 Massachusetts 316765. 7 Florida 284723. 8 New Jersey 273047. 9 Ohio 266911. 10 Washington 239764. # … with 40 more rows =========================================================================================================

