Manuscript for Data Analysis Exercise

Vijay Panthayi

#Summary/Abstract This is a manuscript for the data analysis exercise

#Introduction

##General Background Information The NCHS, CDC, and National Vital Statistics System keep a running collection on the mortality data of the United States by year/state.

##Description of data and data source The data set I used for the data analysis exercise contains the death count/rates for the top 10 causes of death in the United States. The data can be found [here](https://data.cdc.gov/NCHS/NCHS-Leading-Causes-of-Death-United-States/bi63-dtpu). The data is split up by year, then the 113 Cause of Death, then the cause of death, then by which state the numbers are from. The information needed to understand the data are: this includes annual data from 1999-2017, the adjusted death rates are per 100,000 population, the adjusted death rates are based on the 2000 U.S. standard population, and the 113 Cause of Death is an official list of causes of deaths (there are currently 115, but the name 113 Cause of Death stuck).

##Data manipulation methods For this exercise, I focused on only the state of Georgia and the “Cause Name” of Suicide. I also deleted the “113 Cause Name” column to avoid any possible confusion with the “Cause Name” variable.

#References