# PS6 Data Science

William Townsend

Department of Economics, The University of Oklahoma, Norman

March 22, 2022

1 PS6a

As a whole for this assignment, I collected homicide rate data in Mexico at the state level. I scraped the data from wikipedia and the specific site is in my R script. The data assesses homicide rates at each individual state level on an annual basis from 2006 to 2020. The image below (PS6a) shows a linear upward trend of homicides in Mexico at the state level. Each dot is an observation representing the homicide rate per 100,000 residents in each state. The vertical axis represents homicide rates per 100,000 residents while the horizontal axis represents the years 2006 to 2020. The blue line indicates that there has been a gradual increase in homicide rates in Mexico from 2006 to 2020. In conclusion, this image shows us that Mexico is facing an overall increase in homicide rates
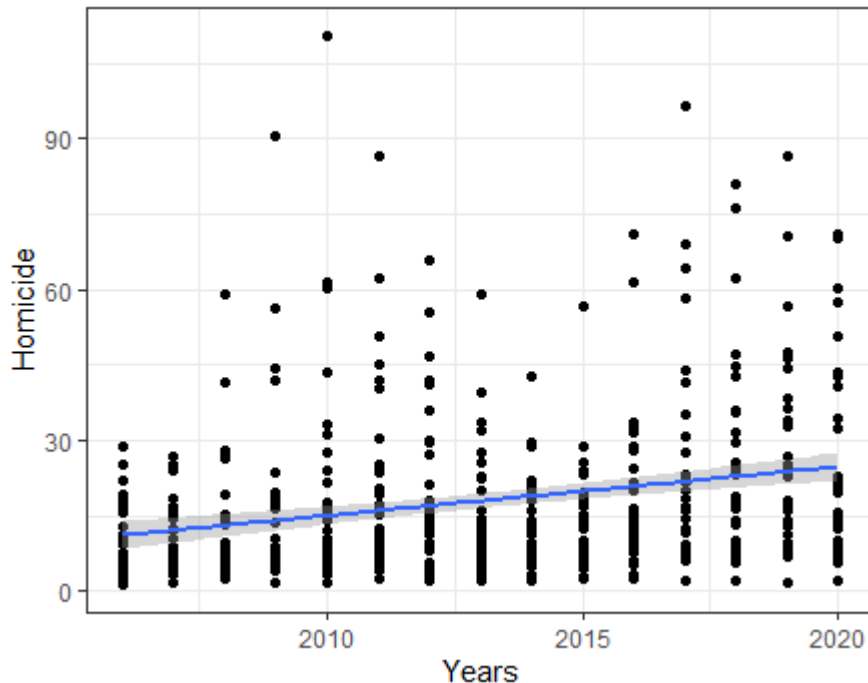
2   PS6b

Image PS6b (below) shows the distribution of the homicide rates per 100,000 residents over the year 2006 to 2020 at the state level. As we should expect, more areas have lower crime than vice versa. We see a right tail distribution, which shows that there is a smaller portion of Mexico that is facing comparatively more severe high homicide rates.
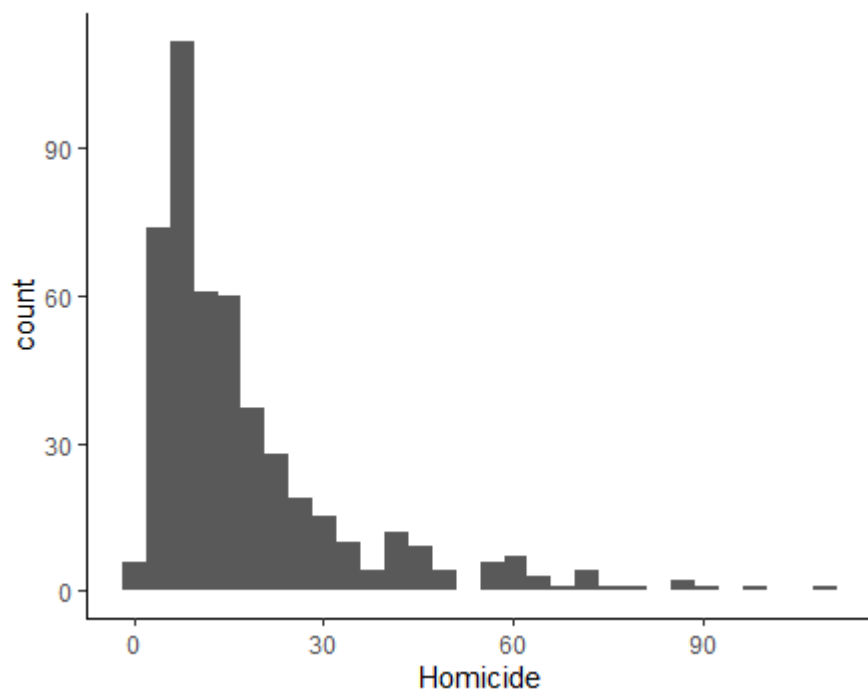
3   PS6c

The image in this section represents the distribution of homicide rates once I treat homicides as a log function. Thus, this image includes the same data as PS6b. The only difference is I logged homicides and added a value of 1 to insure my log function had no errors. In this image, we now see an image displaying something more similar to a normal distribution. Both PS6b and PS6c allow us to see what type of distribution our data displays. Such information allows us to begin assessing the data with concepts and theories that relate to normal distribution.