RISI Data Analyst Candidate Exercise

Drew Rosebush

7/28/2021

Short analysis

During the cleaning process I identified that there were inconsistence levels of information on the upper and lower bounds of the age groups by ACP region typology. To treat each group equally, I dropped ages below 15 and above 79. I also imputed “Unreliable” crude\_rates using population and death estimates provided for the corresponding group.

Suicides are on the rise (Figure 1). Across all age groups and all ACP region typologies. However, some groups are more vulnerable than others. In particular, Teens (15-19) over the past 20 years have experienced almost a 40% increase in the suicide rates (7.68 to 10.69), with the largest increases coming in the last 5 years (Figure 4). When examining ACP regions over the whole 20 years, I noticed that the regions with the largest raw count of suicides (Figure 2) are not the regions with the highest suicides rates (Figure 3). The 3 of the top 6 regions (Big Cities, Urban Suburbs, and Exurbs) in terms of raw suicide counts (top 2 quintiles), are in the lowest 2 quintiles in terms of suicide rates. The inverse seems to hold too. 3 of the top 6 regions (Native American Lands, LDS Enclaves, and Military Posts) in terms of suicide rates are in the lowest 2 quintiles in raw suicide counts. The regions experiencing the highest rates of suicides, along with Teens, deserve further examination as they seem to be the most vulnerable populations.

My steps in further this analysis would be an examination of the groups emerging in the data with perhaps a cluster analysis or spatial analysis using ACP regions.

Figure 1

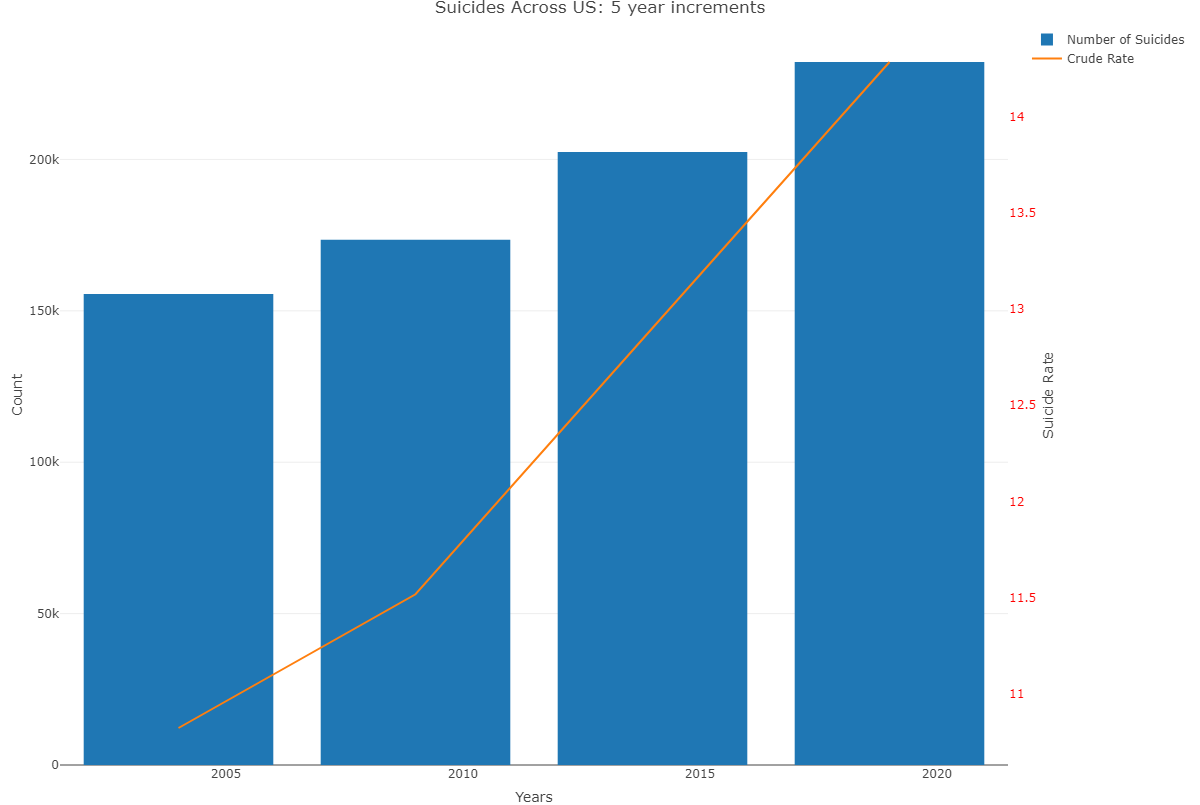


Figure 2

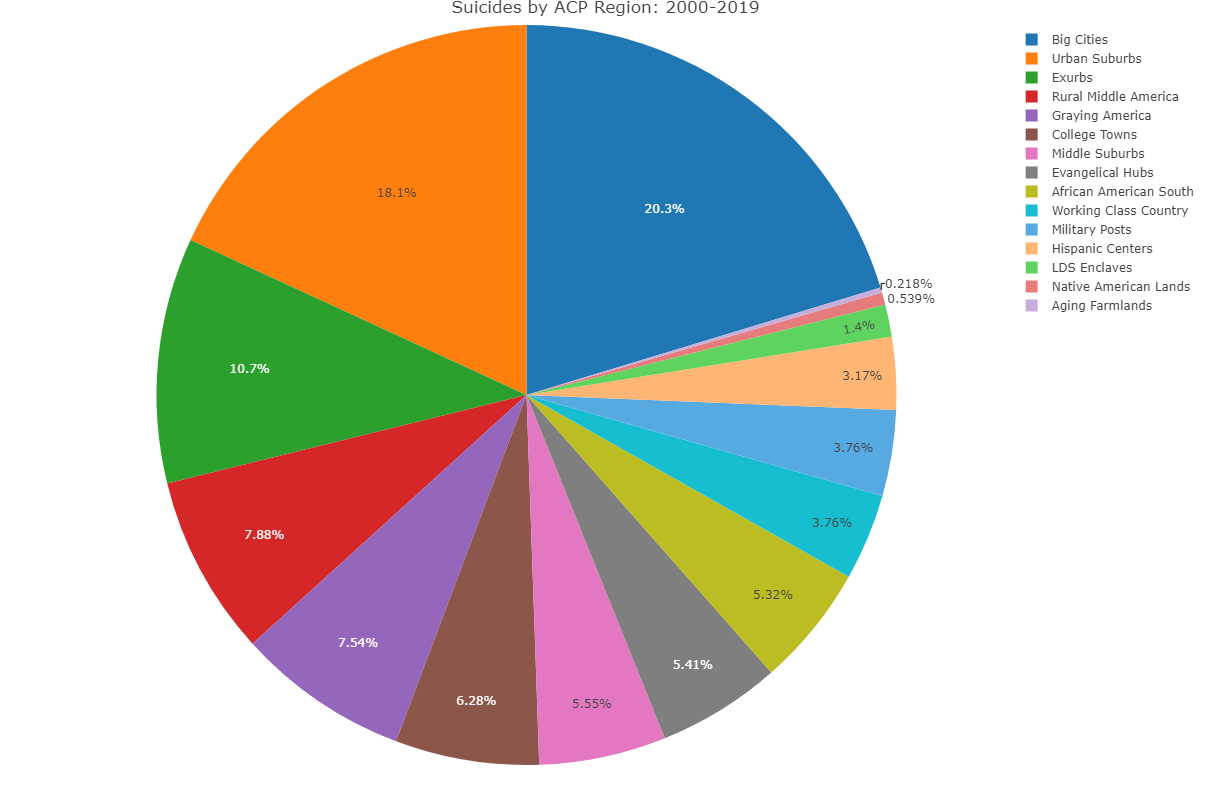


Figure 3



Figure 4

