Carolina Data Challenge 2018: US Health data
Group: 404   Members: Weifang Liu, Minxin Lu and Wancen Mu

Our group focuses on health data in the US to visualize and explore regional and time differences of the prevalence of common diseases, such as Alzheimer's disease, heart disease, cancer, and high blood pressure. The first dataset, which includes health data for 500 cities in the US in year 2014 and 2015, gives us the prevalence of 28 conditions or diseases at track level. We transfer the data values from track level to city level in order to better visualize them using maps.

At city level, we produce an interactive bubble map for adults with high blood pressure in 2015. A bubble map is a map on which we add markers that can be bubble with a size and color relative to the percentage of people who have high blood pressure. We can interact with this plot to zoom in and see the value for each city. We could easily find that high rates of high blood pressure persist in US Southeast, almost 75 quantiles are in southeast, while regional differences could be exist in dietary intake of potassium, magnesium, and protein, which have been implicated as factors in hypertension, as suggested by some articles. In addition, environmental toxins such as lead and cadmium may contribute to geographic differences in hypertension levels in some rural areas.

For the second dataset which includes data of the leading causes of death in the US from 1999 to 2016 at state level, we first plot the trend lines of the mortalities of these eleven causes. Then we focus on Alzheimer's disease, heart disease and cancer and construct three choropleth maps of, which are thematic maps that have shaded areas in proportion to the measurement of the statistical variable, in this case the number of deaths per year or the age-adjusted death rate. We find both heart disease and cancer decrease obviously, which are decreasing from 0.26% in 1999 to 0.196% in 2016, with annual decreasing rate 1.6% and decreasing from 0.198% in 1999 to 0.185% in 2016, with annual decreasing rate 0.4%, respecivitely. From the trend our plot shows, if current trends continue, cancer will become the leading cause of death in the future. When we trying to figure out the reason, we think it may be attributed to further reductions in the major risk factors — total cholesterol levels, high blood pressure, and smoking — and increased physical activity
In addition, an interesting condition is that the death rate of Alzheimer's disease is increasing. The reason are partly due to the growing number of older adults in the United States, as Alzheimer's disease most commonly affects adults ages 65 years and older, as well as due to an increase among doctors, coroners and medical examiners specifically reporting Alzheimer's disease as a cause of death.

Sources:
https://www.livescience.com/59261-alzheimers-deaths-increase.html
https://www.cdc.gov/pcd/issues/2016/16_0211.htm
https://www.webmd.com/hypertension-high-blood-pressure/news/20000106/high-blood-pressure-more-prevalent-in-south#1