Geographic Distribution of Businesses in Los Angeles

Yunbin Peng

In this report I use the **Listing of Active Businesses** and the **Population by Council Districts** datasets from *Los Angeles City Data Portal (data.lacity.org)* to answer the following questions. 1. Is there any pattern of geographic distribution of active businesses in Los Angeles? 2. Which council district has the lowest population-to-business ratio?

Active Businesses dataset lists all active businesses currently registered with the Office of Finance in Los Angeles City and surrounding counties. From this dataset, I use the *location* variable which lists the location of a business in (longitude, latitude) format. **Population by Council District** gives the population for each of 15 council districts in Los Angeles, I use the *value* variable which gives population from 2010 Census in each district.

To produce heat map, I extract longitude and latitude for each business using string manipulation method, longitude and latitude are rounded off to 3 decimal places. This effectively divides LA into small blocks of 0.001 longitude by 0.001 latitude, all businesses within the same block are "moved" to the center of the block. I use the combination of longitude and latitude as the grouping variable and count number of active businesses in each block. I plot these blocks as points using *ggmap* package where the color of each point depends on the log of the number of active businesses.

For the population to business ratio, I use the dplyr package to summarize the number of active businesses in each council district from the *Active Business* dataset. Then the count data is merged with the **Population by Council Districts** dataset. The population to business ratio for each district is computed by dividing the population by the number of active businesses, I arrange the ratio in ascending order and use *ggplot2* package to produce a bar plot.

In conclusion, most active businesses are concentrated in the downtown and along major road networks (for example Highway 101 in the north of downtown). There is also a great disparity of distribution of businesses among districts. For example, the 5th District has the lowest population to business ratio around 5, while the 15th District has the ratio of 22.5, almost 4 times as large.
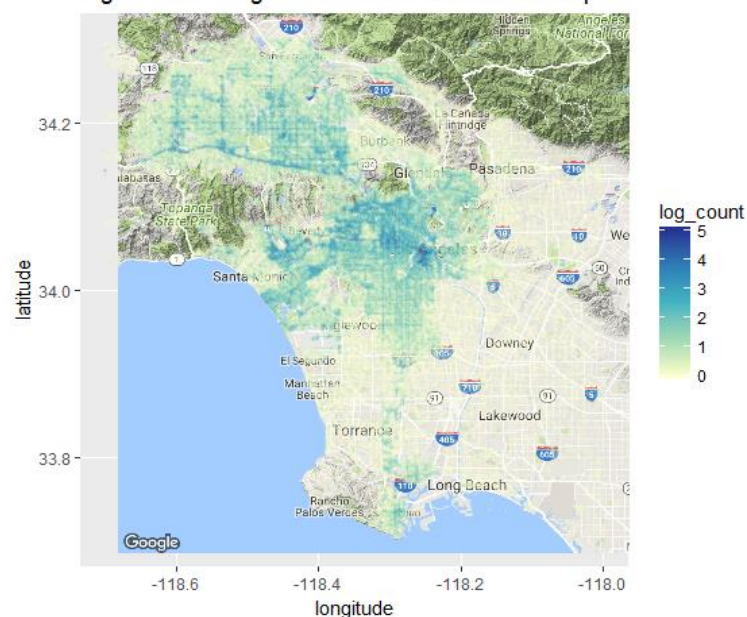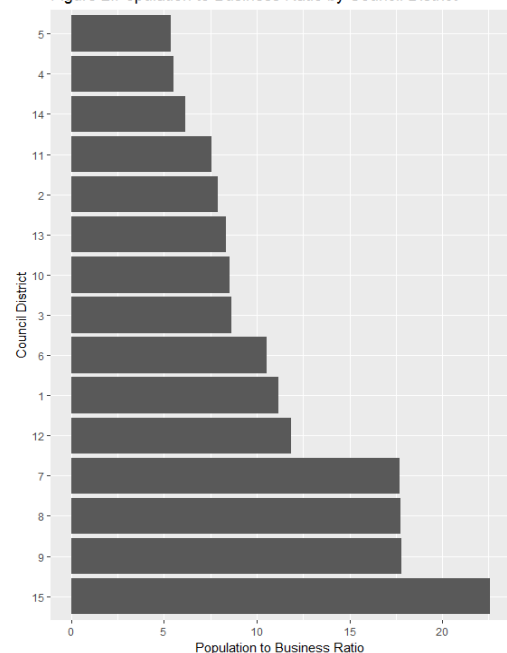

Figure 1:Los Angeles Active Business Heat Map

Figure 2:Population to Business Ratio by Council District

Appendix

Figure 1:Los Angeles Active Business Heat Map