IE6600-Sec05-Group17_Hackathon

Arjun Janardhan, Ajay Parthiban Senthilvel, Chaitanya mandava

10/12/2021

Introduction

Homelessness in the USA is one of the significant problems associated with the nation's current-day development. The data we got from the United States census bureau has helped us visualize in this hackathon. Apart from this, the information we use in this dataset has been taken from the national alliance to end homelessness, which allowed us to analyze the different types of homelessness and relate homelessness to the chronic diseases people are suffering from. The Homeless research Institute has been trying to reduce homelessness over the past few years.

Problem Statement

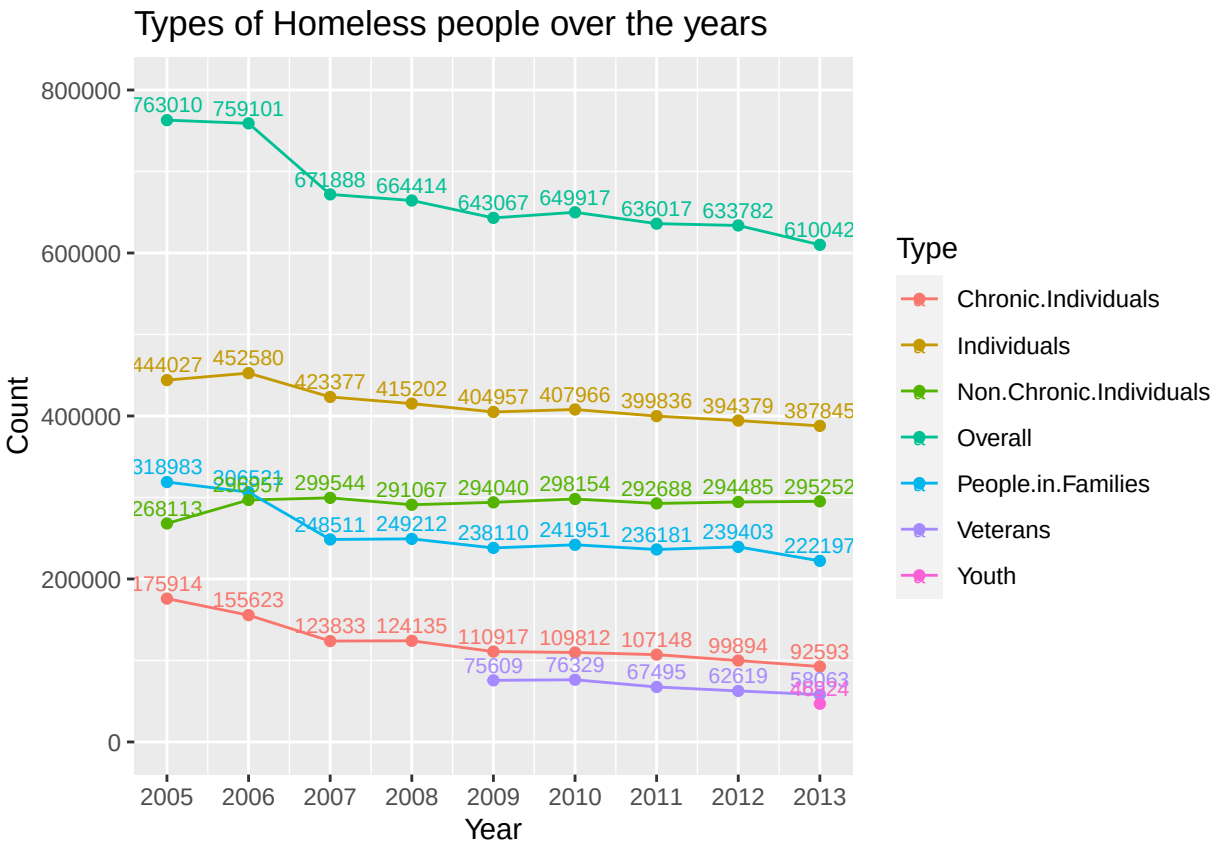
Working on these datasets based on 7-8 years old helped us read the scenario of the dataset. We have seen all the different possibilities of homelessness and the reasons for homelessness. Most chronically diseased people are homeless; followed by this, the dataset significantly impacts individuals rather than the families. In addition to this, major cities have more homeless people like New York, Massachusetts, North Dakota, Oregon, Nevada, and California. Initially, we established the trends in the dataset by following the data through the years 2005-2013. Supplementing this, we found out the different categories of homeless people in the year 2013. We split the data into two categories based on population to establish a relationship to look into the change over last year's homeless people count. Furthermore, the homeless population also had non-chronic who were lost, which helped us find the difference between chronic and non-chronic people who were homeless. This data was until 2013, but as the trends keep on updating, the data keeps on updating, helping us analyze the decrement of homeless over the years.

Q1: How has the distribution of homeless people been over the years, and what are the different types of the homeless population?

```
## Warning: Removed 12 row(s) containing missing values (geom_path).
```

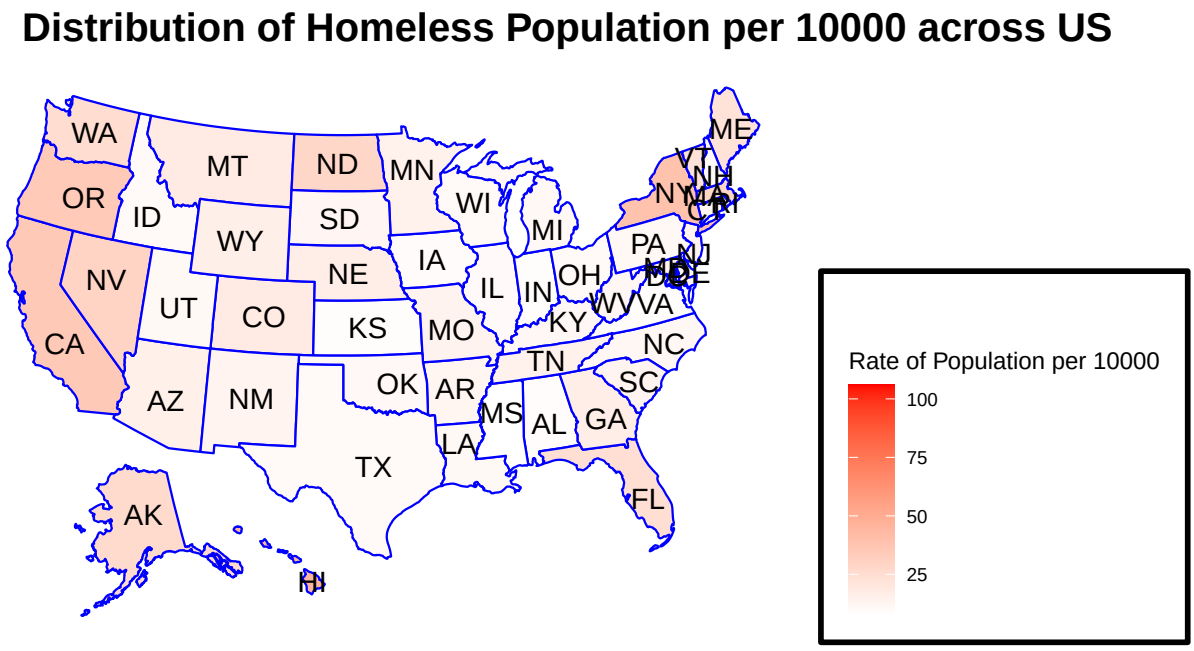
```
## Warning: Removed 12 rows containing missing values (geom_point).
```

```
## Warning: Removed 12 rows containing missing values (geom_text).
```

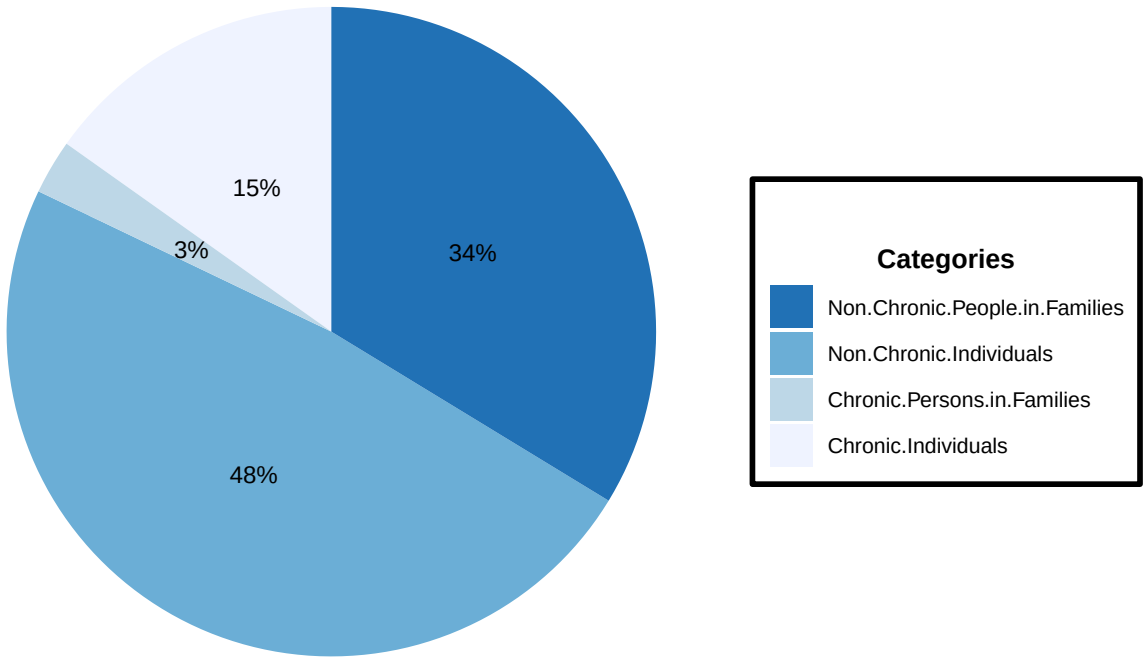


Output We decided to initially plot a line chart that helped us analyze the count of homeless in the United States from this data. Moreover, we could see a decrease in homeless people count over the past few years. We could also see what kind of population saw the most decrement in the number of homeless people.

Q2: How is the density of the homeless population distributed across the United States in the year 2013? How has this been categorized across the United States?

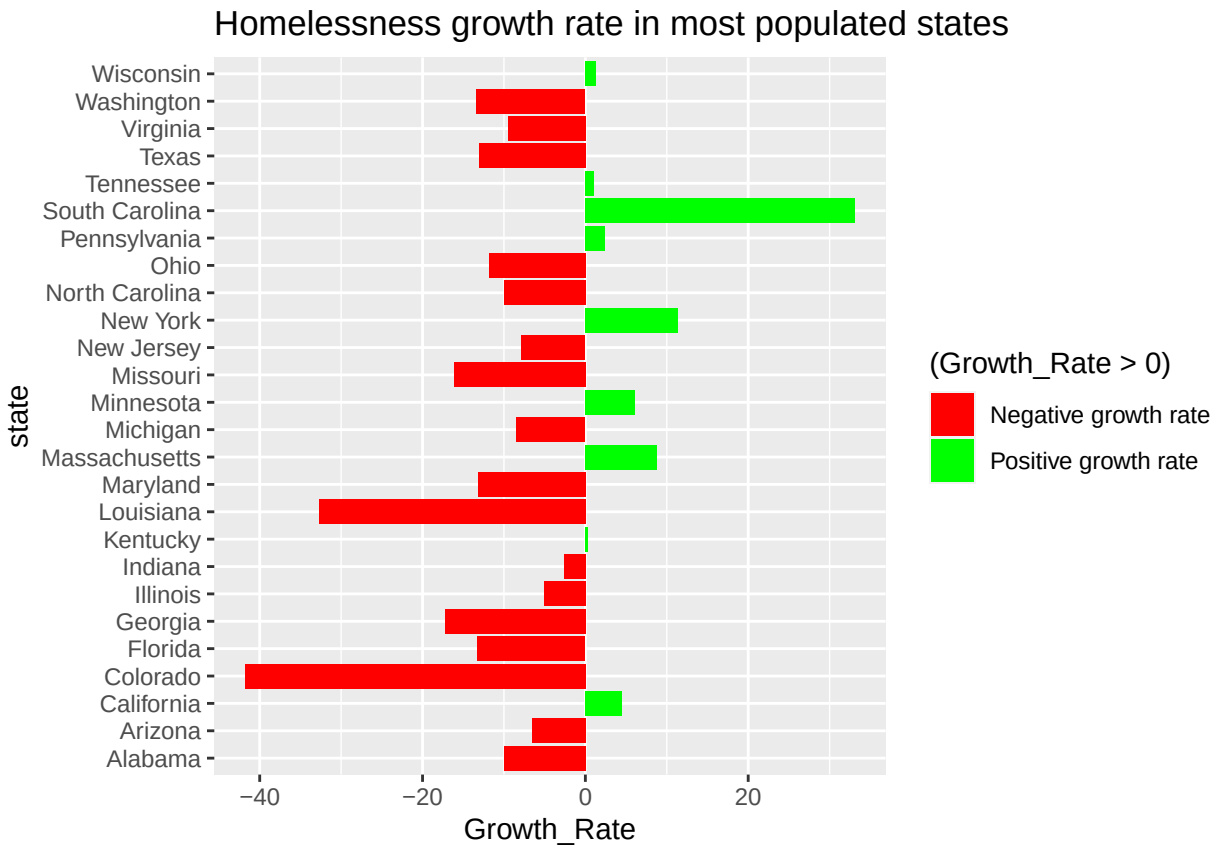


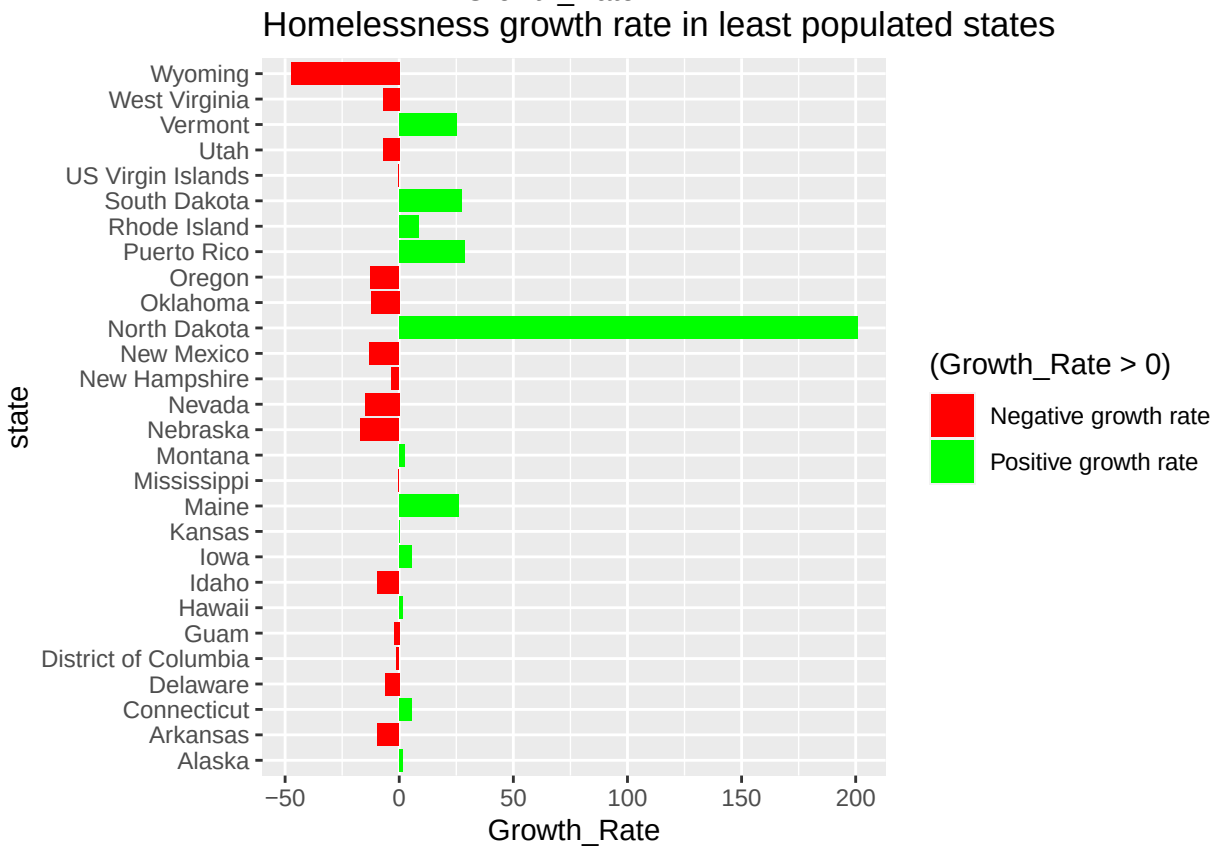
```
## Warning: Ignoring unknown parameters: face

## Warning: Use of ‘df_stats_2013$Category_Percent‘ is discouraged. Use
## ‘Category_Percent‘ instead.
```



Output For this analysis, we plot the distribution of homeless people across the United States, using a map of the United States. Furthermore, this data helped us categorize the homeless population in the United States in 2013 primarily as chronic and non-chronic. Through this, we understand how the various types of homeless people are distributed across the United States.

Q3: How can we analyze the correlation between population and change in homelessness? Does the population influence the change in homelessness?

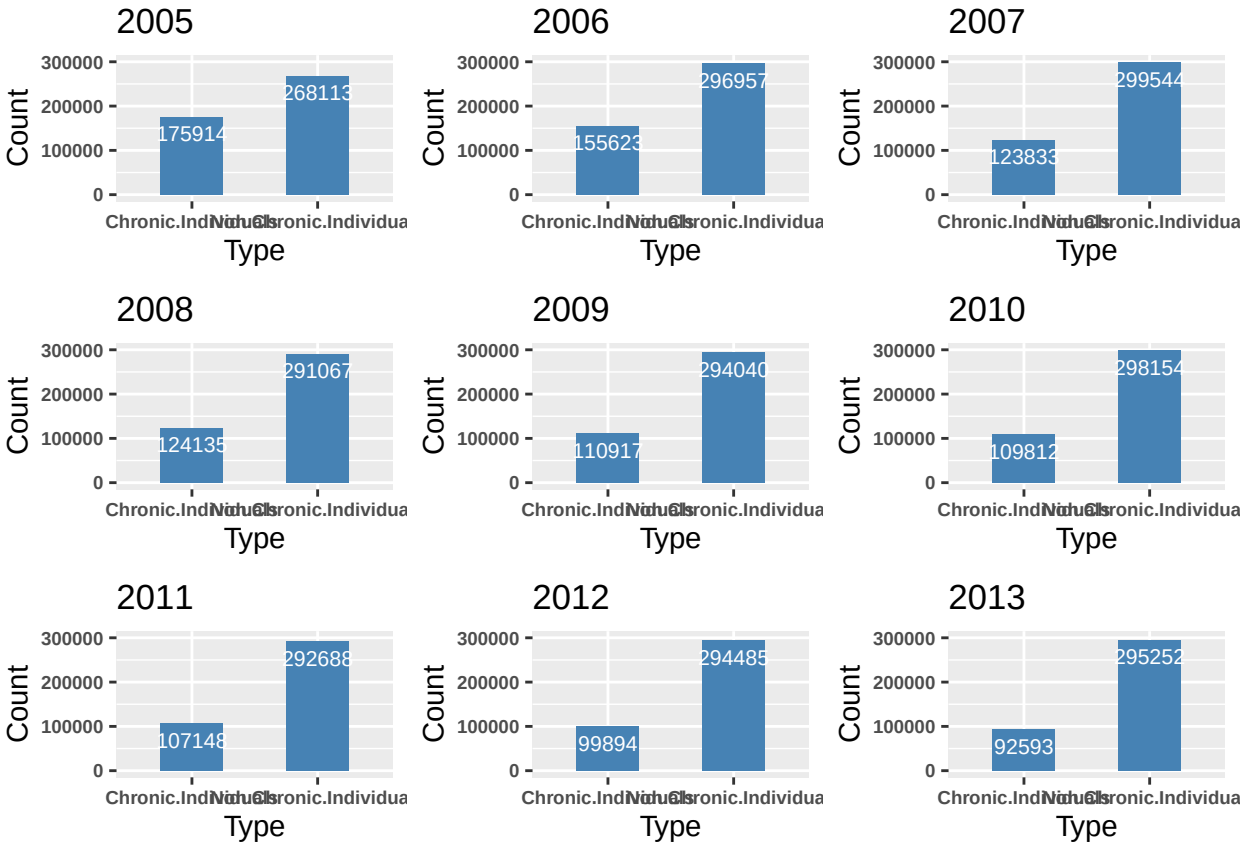


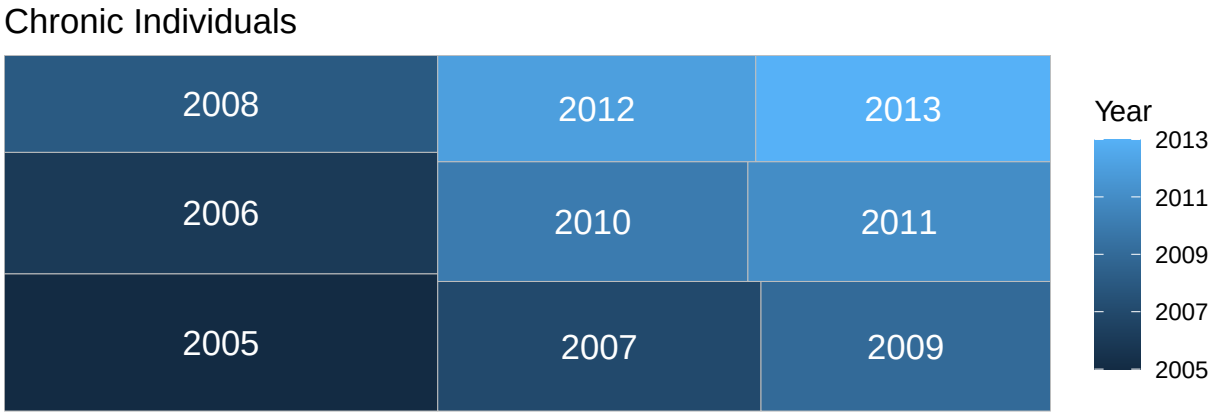


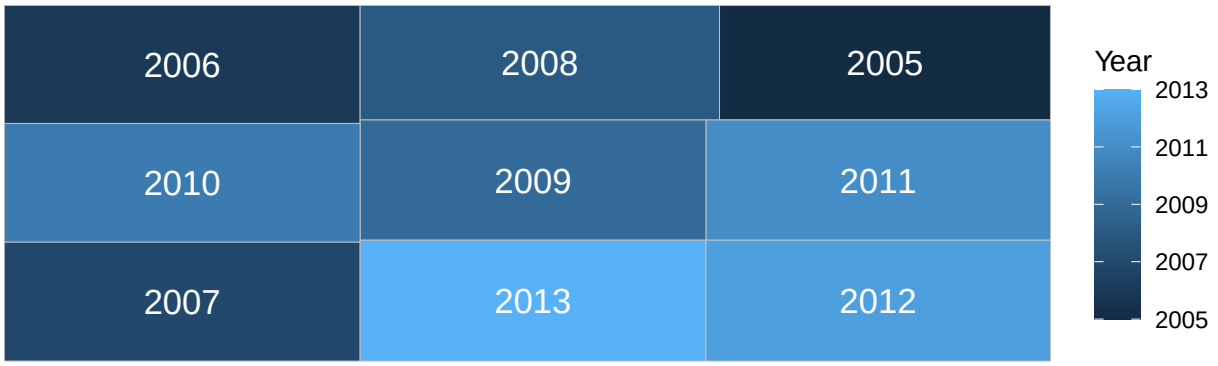
Output Looking at the concentration of population of the various states in the United States, we categorized

the data as above 4 million and below 4 million, and later we have analyzed the trends using a bar graph as the negative and positive scale of the percentage change in the homeless population over the year 2012-2013.

Q4: Only the non-chronic people tend to be more homeless than the chronic. Justify the answer.







Output As we have analyzed the data, we got some relations between homeless and chronically diseased people. Passing through the years, more non-chronic people are homeless than chronic people. There might be several reasons for this, but we could exactly show the difference between chronically diseased people and non-chronically diseased people.

Conclusion Working with extensive data opens a gate to explore a lot of visualizations. The data we have worked on in this hackathon has more usage in the government and many non-profitable organizations, which helps improve the situations in the country addressing this issue. Initially, looking at the dataset, we started to reach out how the population has been changing over the years and the categories in the weightage of homeless people. Followed by this, we also got the distribution of homeless people across the country and what percentage of the homeless are families and individuals and who are chronic and non-chronic. Until now, we only saw the population distribution across the country. Further, we expanded the categories in this distribution and counted the number of chronic and non-chronic people in the total number of homeless people using the bar graphs. In addition to this, we plotted a treemap graph to show the density of homeless people over the years. Analyzing such kinds of data gives us a base to explore through many visualizations and supporting any non-analyst how to understand the data easily.

References:

1. https://www.rdocumentation.org/packages/ggplot2/versions/3.3.5
2. https://jtr13.github.io/cc19/different-ways-of-plotting-u-s-map-in-r.html#using-usmap-package