**Interactive Homeless Operations Planner and Tracker**

**for Los Angeles City and County**

DSO545 Final Project – Fall 2107

Group 14

Mohammad Ganji, Modan Wang, Muhammad Musthofa, Naiyuan Xiao, Wei Tang, Yufei Wang

**Introduction**

The City of Los Angeles asked our team to build an interactive dashboard that shows the predictive modeling of different homelessness intervention strategies using geospatial analysis and regression models. Our dashboard also includes descriptive statistics about crimes and 311 calls.

The goal of this project is to provide insightful geographic, demographic, and time indicators of homelessness in The City of LA. All the reports in our dashboard are integrated into an online mapping application that will enable Mayor’s office staff to evaluate and track the effects of existing and upcoming homelessness interventions. Also, it will allow Mayor’s office to analyze data in a more systematic way and facilitate the rapid deployment of resources and services.

This report begins with an introduction of different datasets which are used for the time, social, and geospatial analysis. Each analysis is followed by the generated insights and the report concludes actionable recommendations for the City of LA.

**Part I: Data Sources**

1. 311 Homeless Encampment Requests

* Data from August 2015 – November 2017
* Contains information about homeless encampment requests from different channels or source (mobile phone, web portal, etc.) and the response time

1. Crime Homeless Victim

* Data from January 2015 – November 2017
* Contains information about crime homeless victims, location, time, and modus operandi.

1. Homeless Count by Census Tract

* Data from 2015 – 2017
* Contains information about the number of homeless people by location based on census tract.

1. Homeless Shelters and Services

* Contains information about number of homeless shelters and the service type.

1. Geographic Info

* Contains information about population density and area density in the given period.

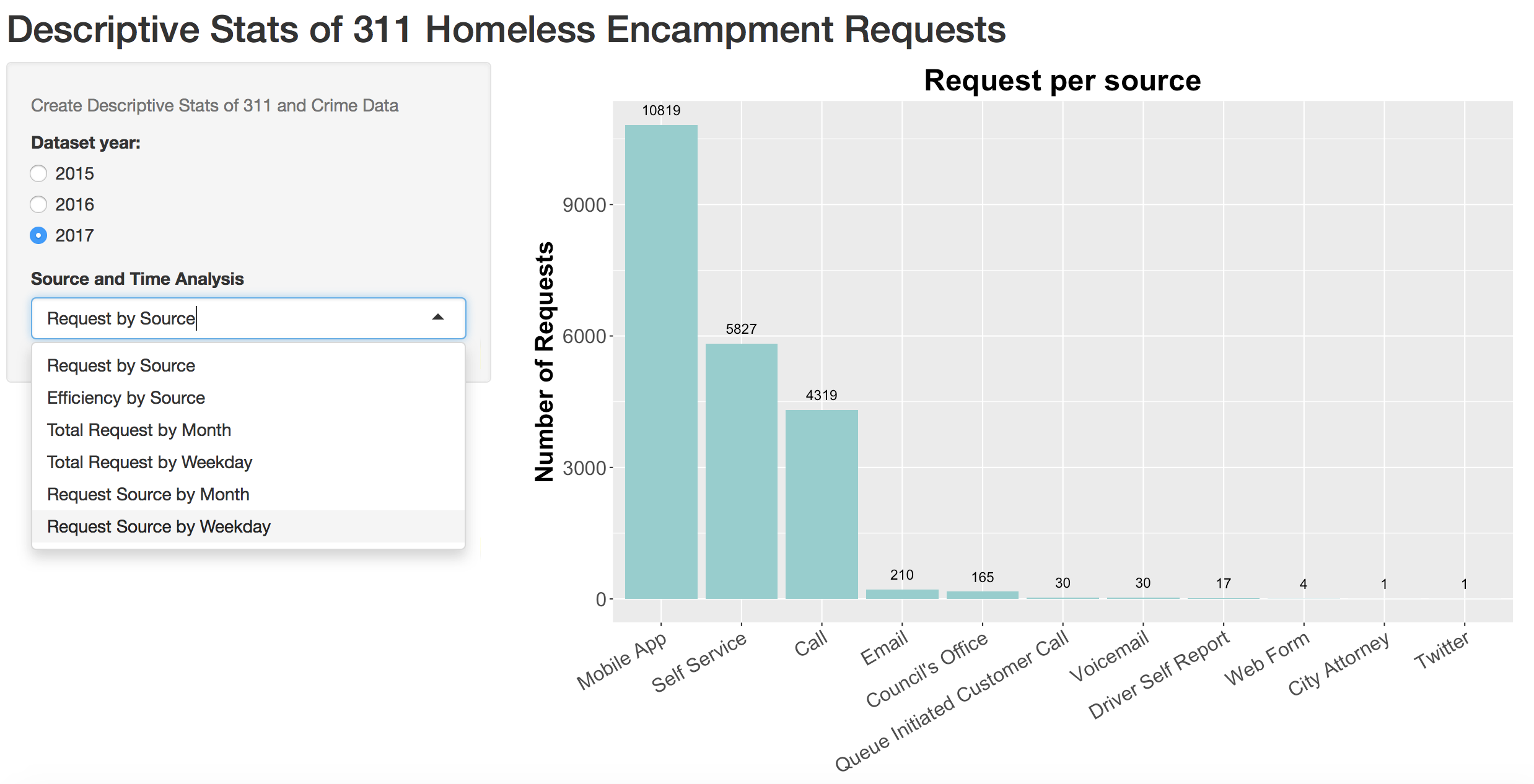
All these datasets are mapped and analyzed based on census tract as one unique single ID. Also, we added council district as another location ID in some datasets such as 311 homeless requests and crimes.

**Part II: Analysis and Insights**

* **Time Based Analysis**

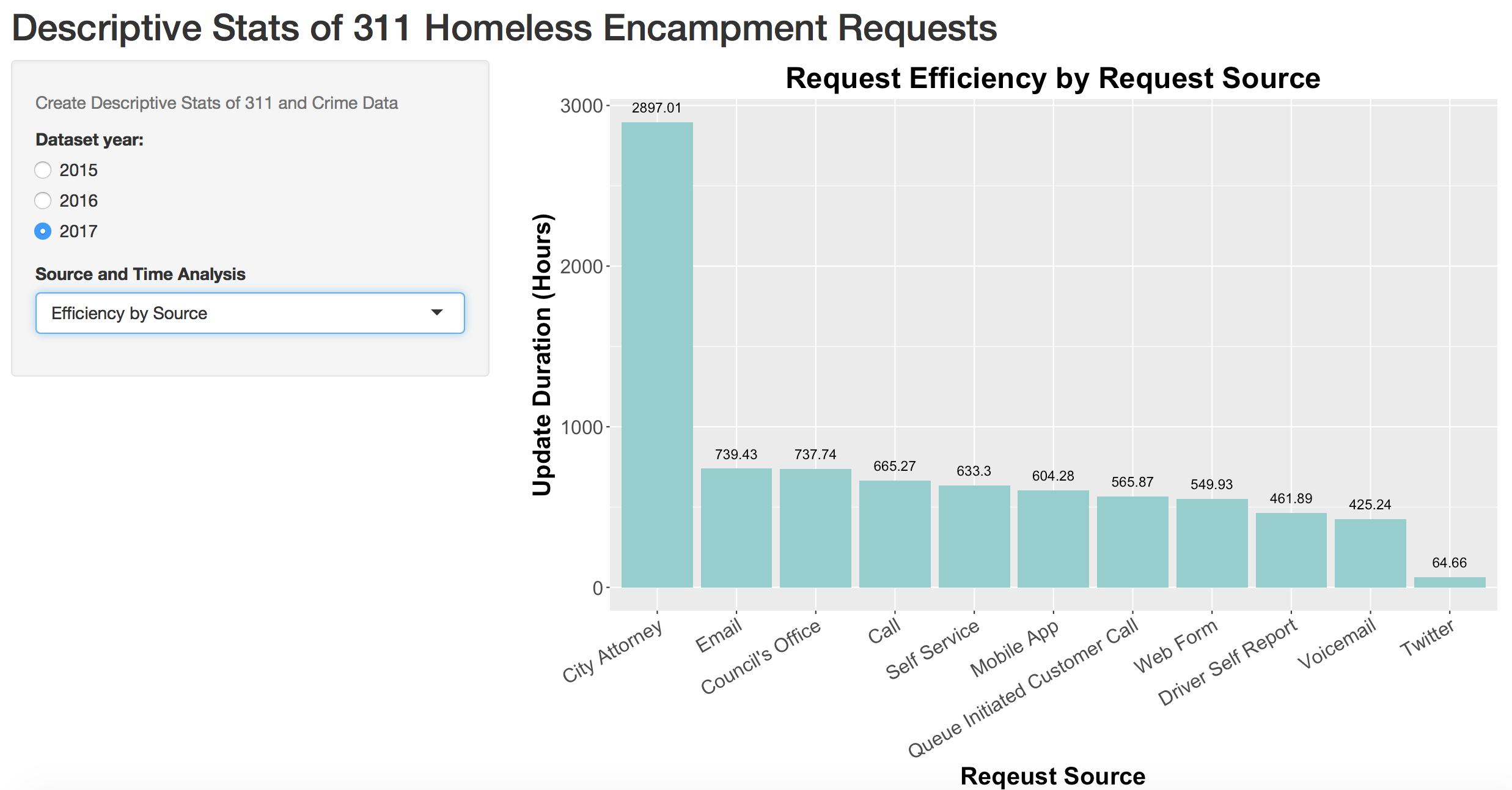
Datasets: 311 homeless encampment requests and crime homeless victim

1. **Time based analysis on 311 homeless encampment requests**
2. Distribution of Requests by Source



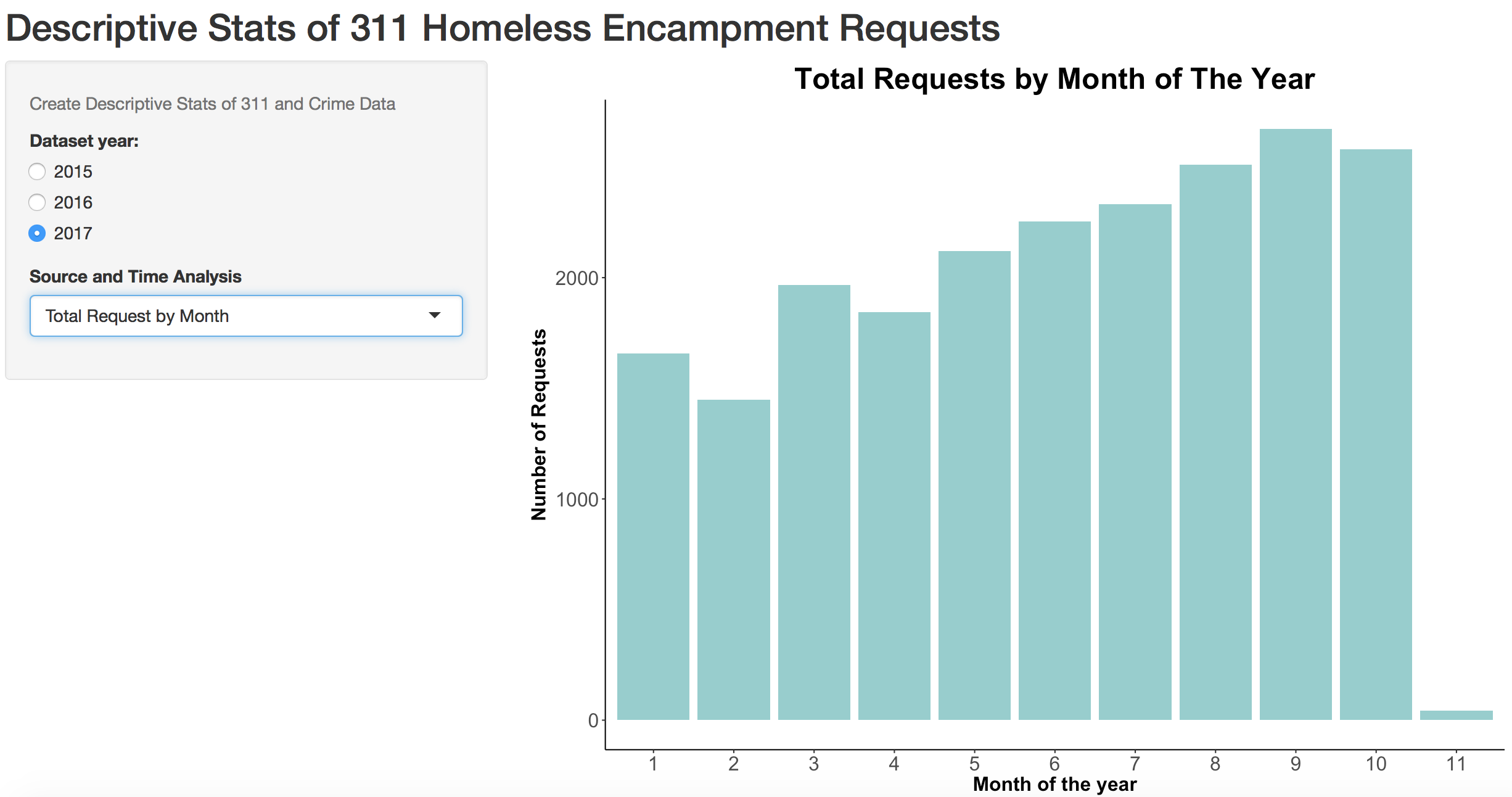
* In this dashboard, we analyzed descriptive statistics of the number homeless encampment requests per source. The dashboard has three dataset from 2015 - 2017. The most popular channel or source people used to call 311 is mobile app. As the more smart-phones are affordable, mobile apps become a powerful tool to call 311. For these reasons, to anticipate the peak concurrent users in any given period, the City of LA should maintain the reliability of mobile app.
* Other channels or sources which have higher active users are self service and call. Self service means people access the web portal of LA 311 to request and call means they have a phone call directly to the 311 office station. The use of social media is relatively small such as twitter, we hope that the City of LA could encourage people who see homelessness to report it via social media. It is important to filter the redundancy of report and also the credibility of report.

1. Distribution of Service Efficiency by Source



* In this graph, we analyzed the homeless request efficiency by source or channels. The request efficiency means how fast the City of LA update any homeless request to the system from the time they received the request (update date - received date). We excluded the response date (the time of homeless requests solved) because almost 60% of these datasets (from 2015-2016) has NA values.
* The graph shows the higher update duration, the response from a particular channel or source should be improved. Twitter has the lowest update duration which is the most efficient among other sources. However, as we can see from the previous graph, the usage of twitter is still low. Although mobile app is the most popular source to use, the efficiency is still worst than web form or voice mail.

1. Total Request by Month



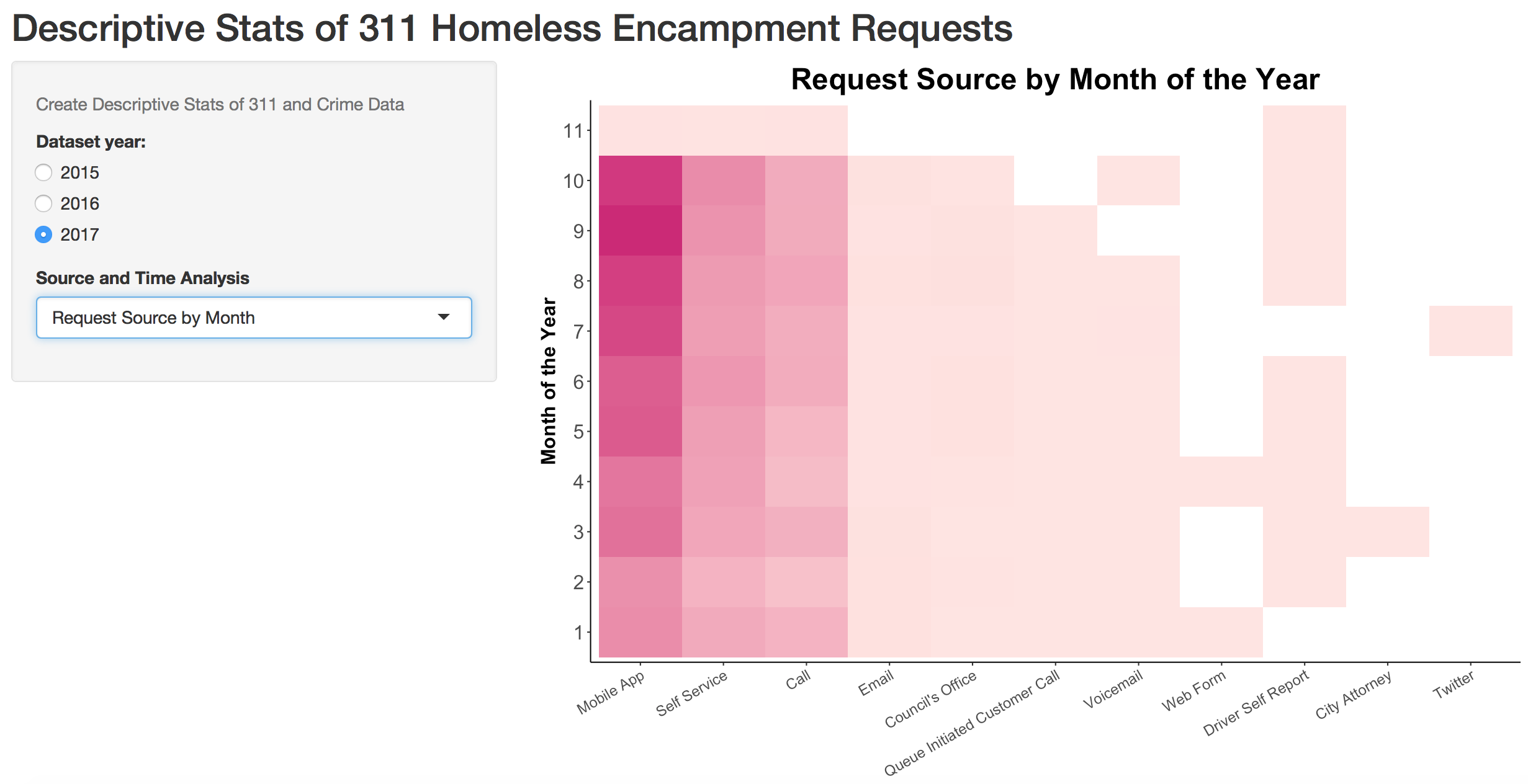
* Total request by month of the year (2017) shows the number of homeless encampments requests gets higher in summer. We analyzed this phenomena happened due to temperature. As temperature becomes hot, more people go out, see homelessness around and report it to the City of LA. This trend not only occurred in 2017 but also in 2015 and 2016.

1. Total Request by Weekday



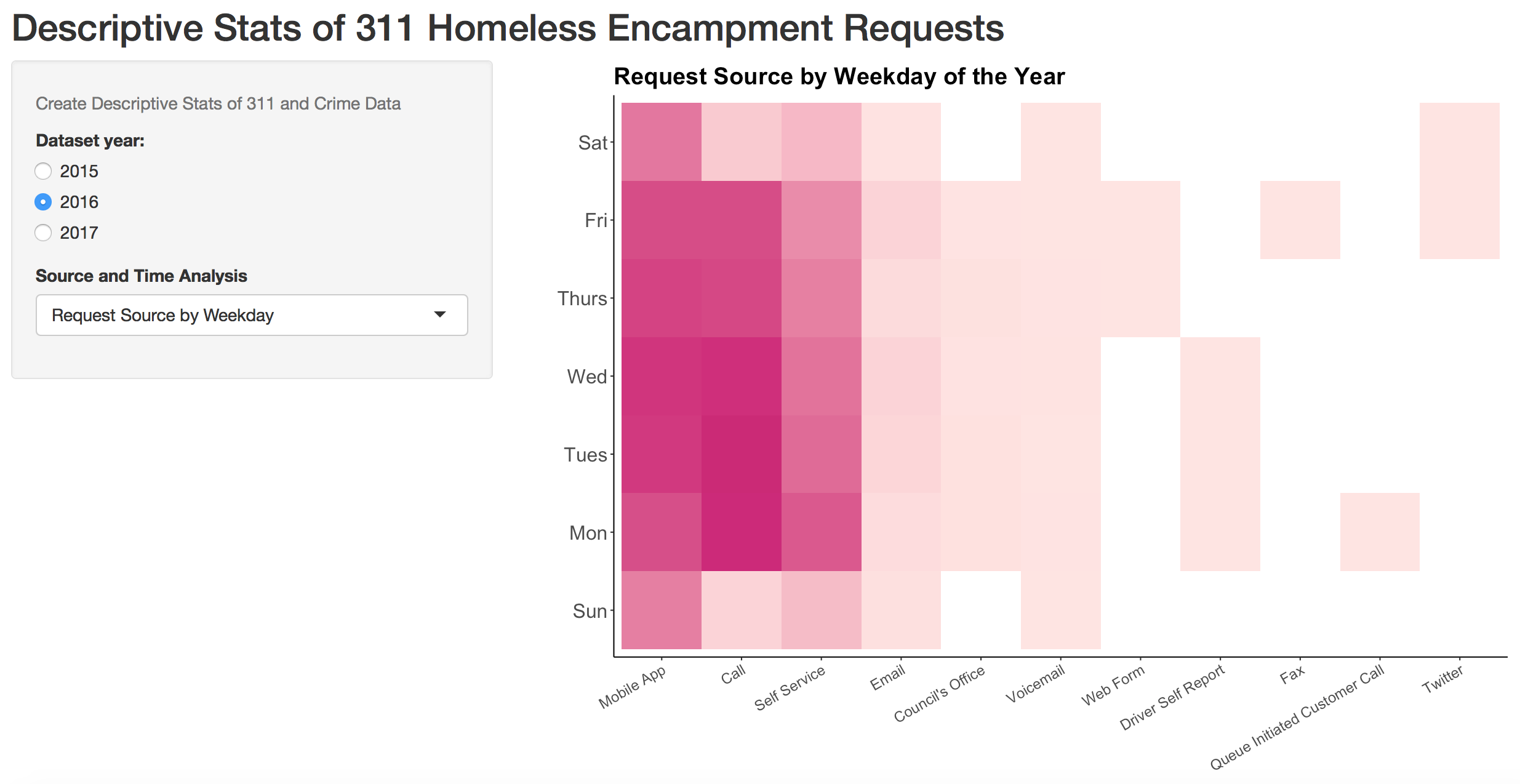
* The City of LA received more requests during weekday than weekend. We analyzed this phenomena happens because of the human activities. For example, during weekday, employees go to office or restaurants, see homeless people who ask for help and report it.

1. Distribution of Request with Different Sources or Channel by Month



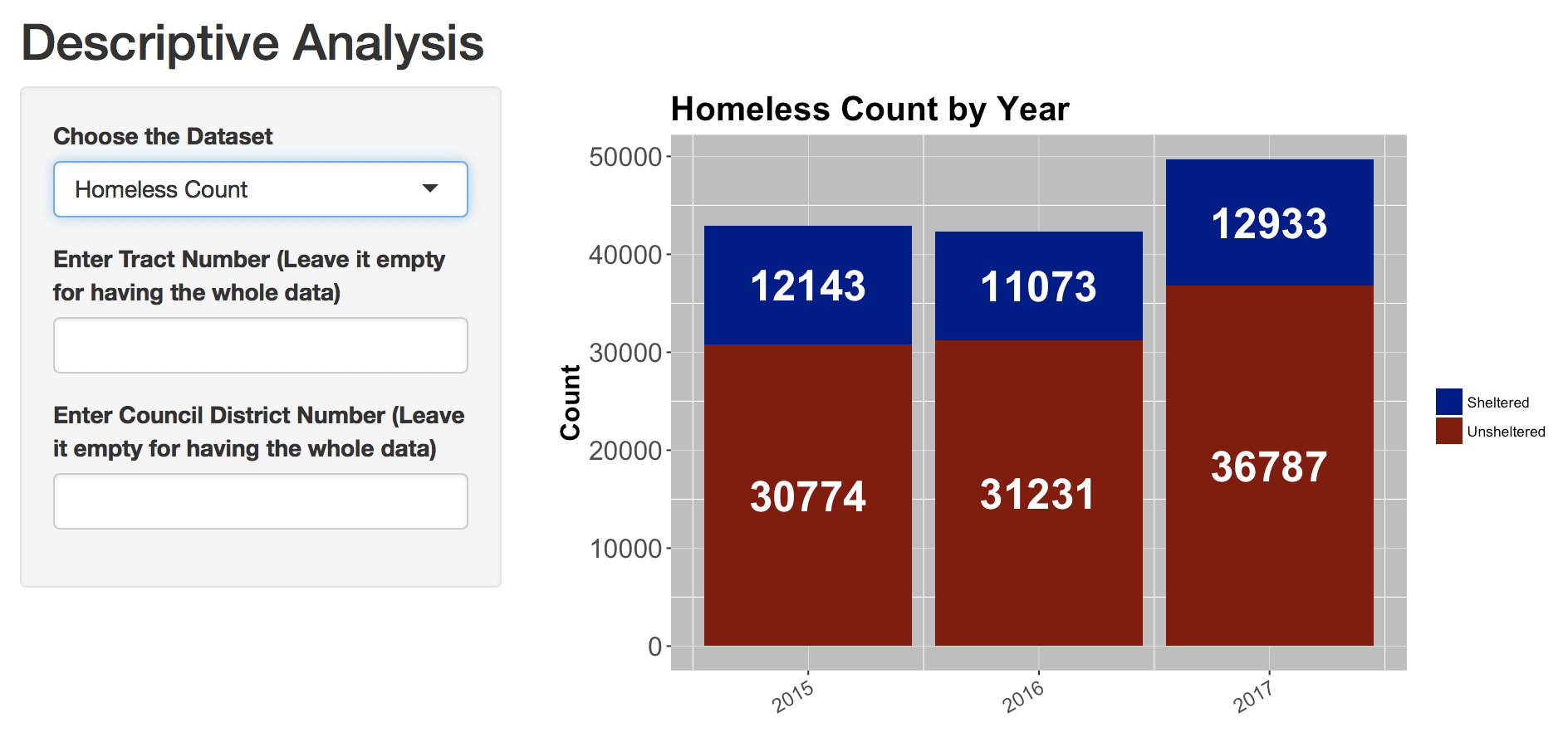
* If the City of LA wants to break down the number of request by a particular source and by month of the year, we have provided this interactive dashboard to drill down any source or channel which people used to report homelessness.
* Mobile app, service call and call are still the more popular source to use to report homelessness. Also, the request from these sources gets higher in July, August, September and October.

1. Distribution of Request with Different Sources or Channel by Weekday



* Using this dashboard, the City of LA could break down the distribution of request by a particular source in a particular day. From 2016 datasets, requests from mobile app, call and self-service were still common and these requests happened during weekday.

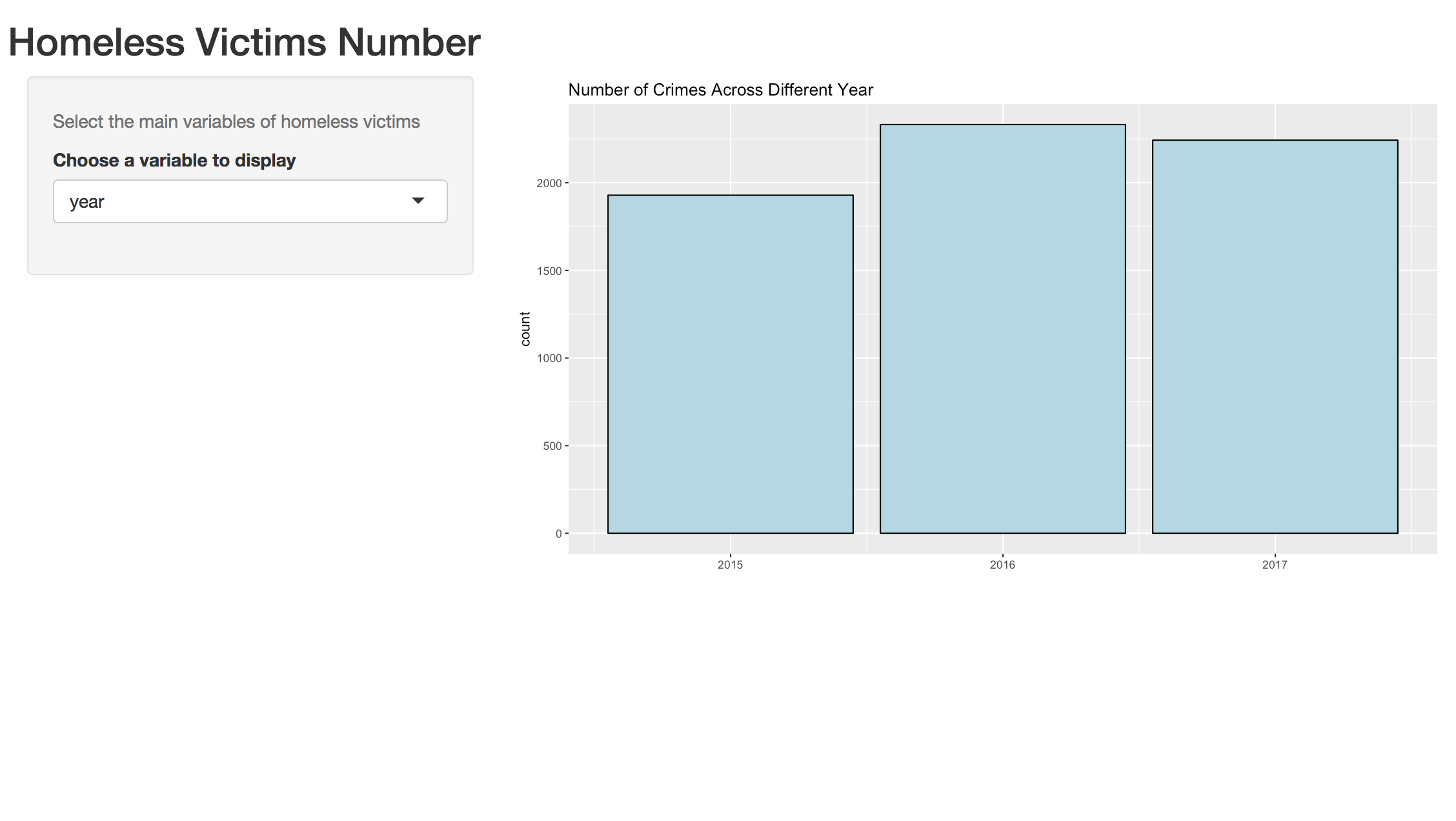
1. **Time based analysis on homeless count**



We also analyzed descriptive statistics of the number homeless population. We used three datasets from 2015 – 2017 and created the graph above. From this bar chart, we can find that the homeless count increases significantly from 2016 to 2017, mainly due to the increase in unsheltered people.

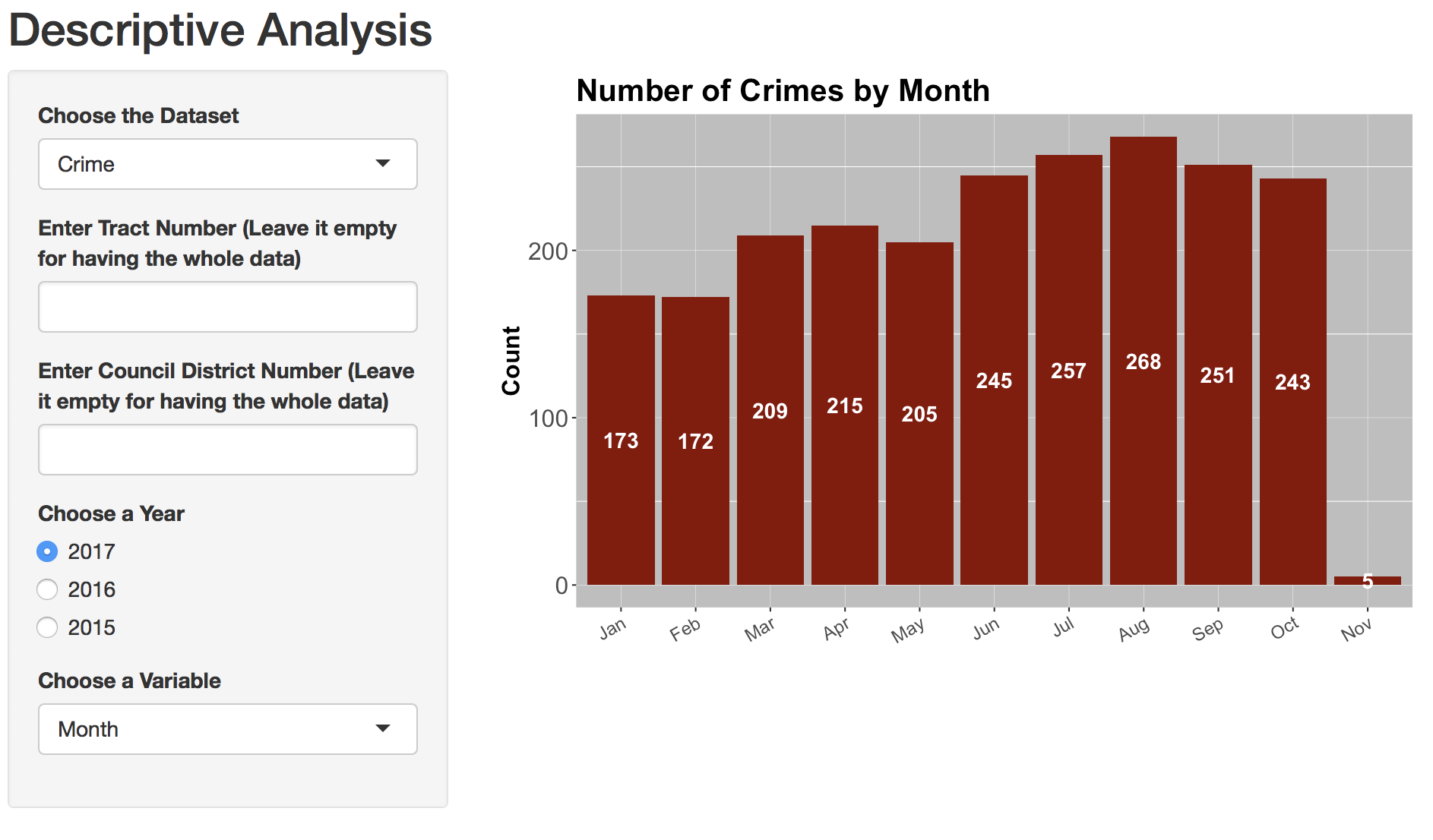
1. **Time based analysis on homeless victims(crime)**
2. Number of Homeless Victims (Crime)Across Years

This bar chart shows the number of crime across years. From this chart we can find that homeless victims increased significantly from 2015 to 2016, but decreased from 2016 to 2017.



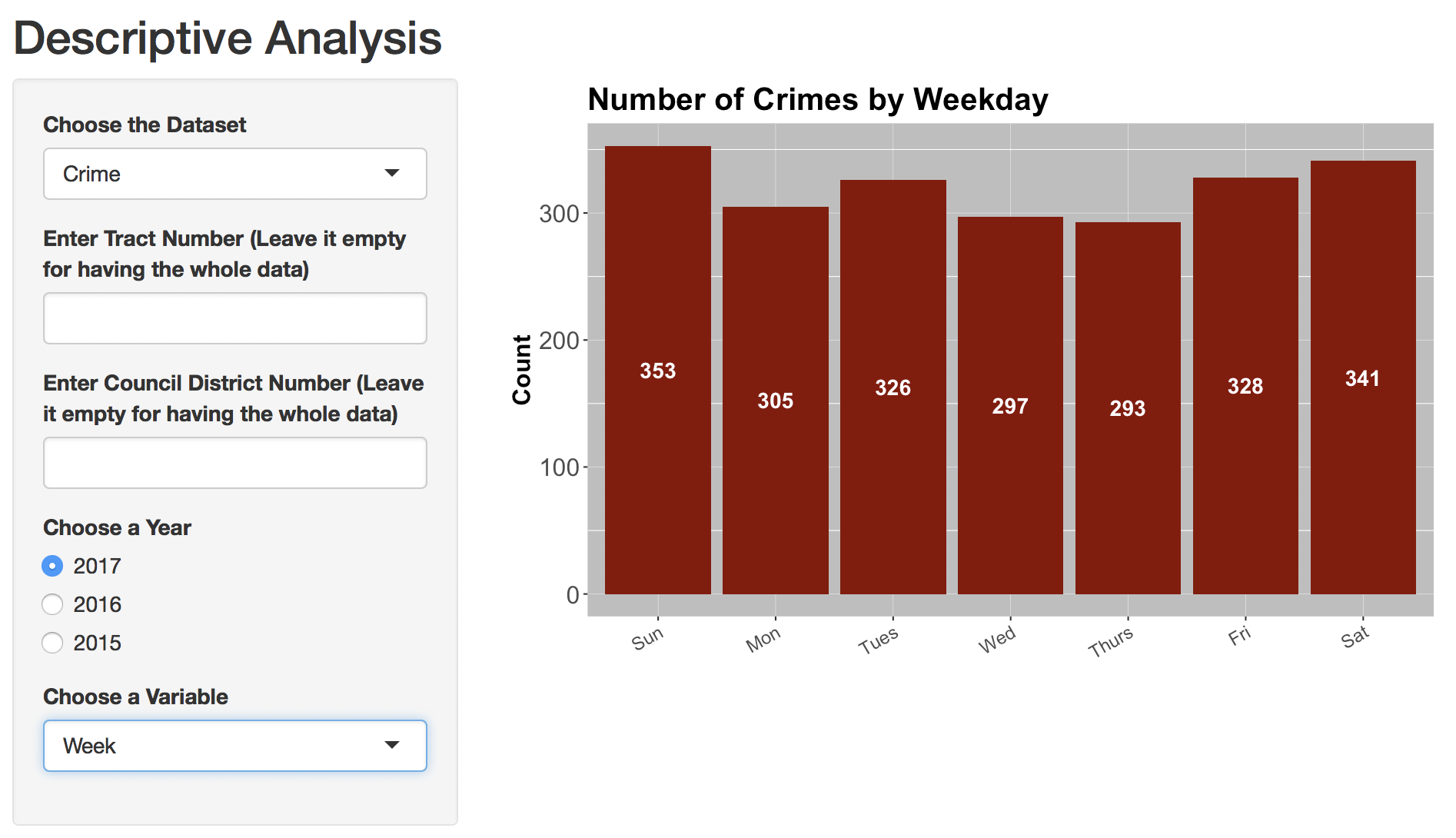
1. Number of Homeless People(Crime)Across Different Month

This tables shows the number of crime across different month. From this table, we can find that July is the month with most homeless victims. November and December are the months with least homeless victims.



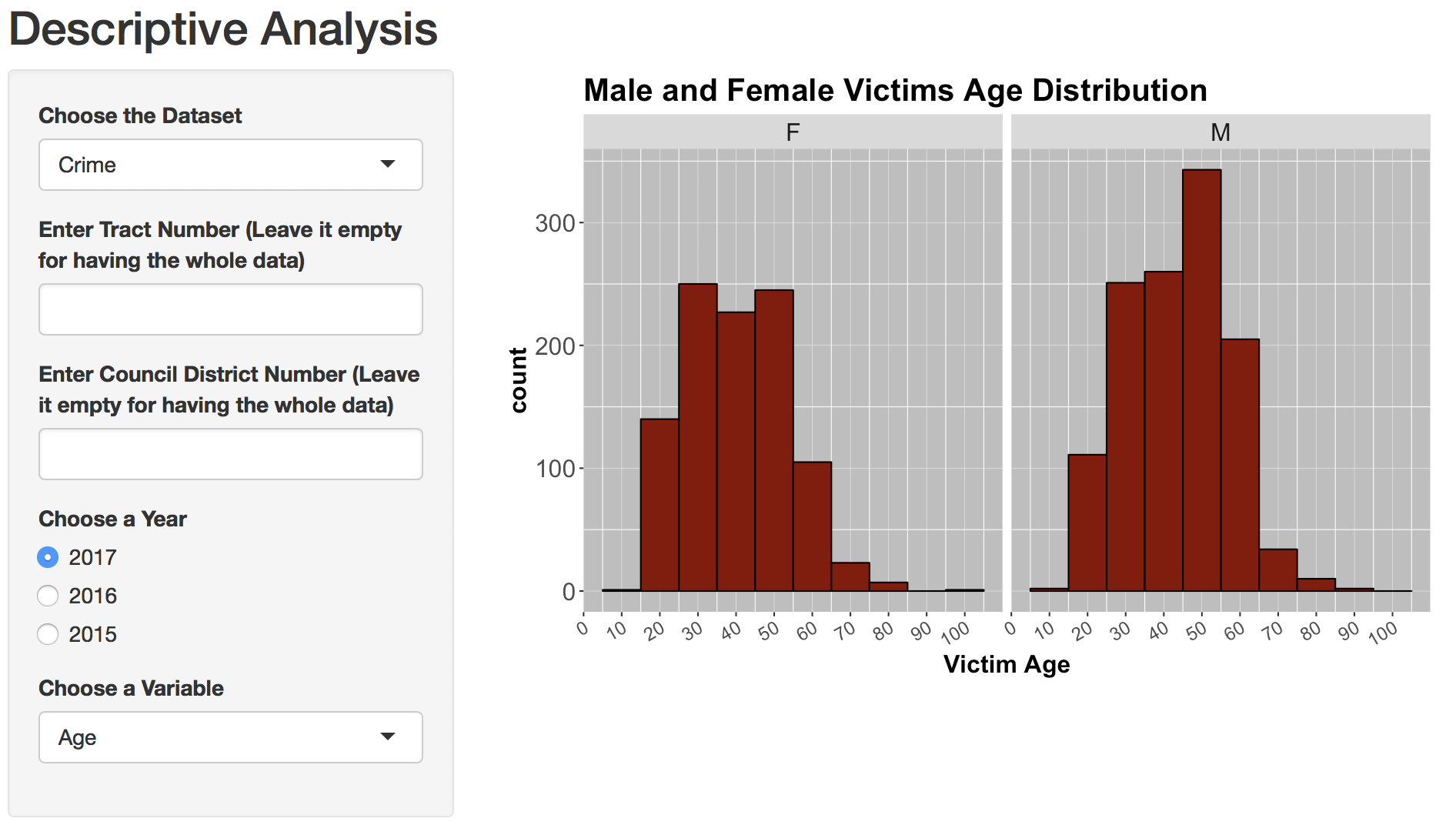
1. Number of Homeless People (Crime) Across Weekdays

This table shows the number of crimes across weekdays. During the week, most crimes involved homeless victims happened on weekend.



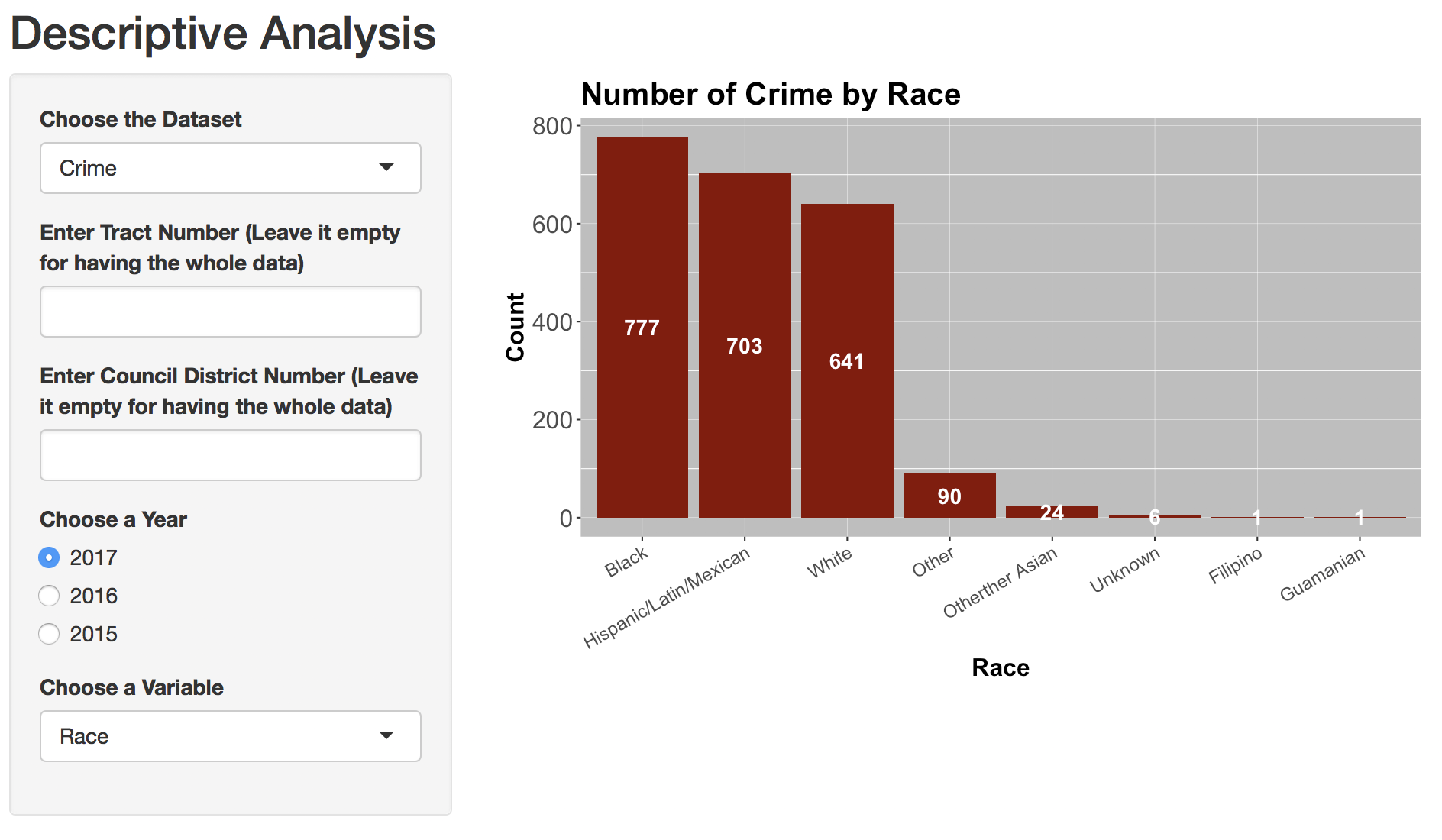
1. Number of (Homeless People) Crime across Age

* This chart shows the number of victim across the age. The one in the left shows the number of female victims in different age group, and the one in the right one shows the number of male victims in different age group.
* From these two bar charts, we can find that most of Female victims are around 30-60 years old. Distribution of male victims are relatively skew compared to the distribution of female victims. Most Male victims are around 40-60 years old.



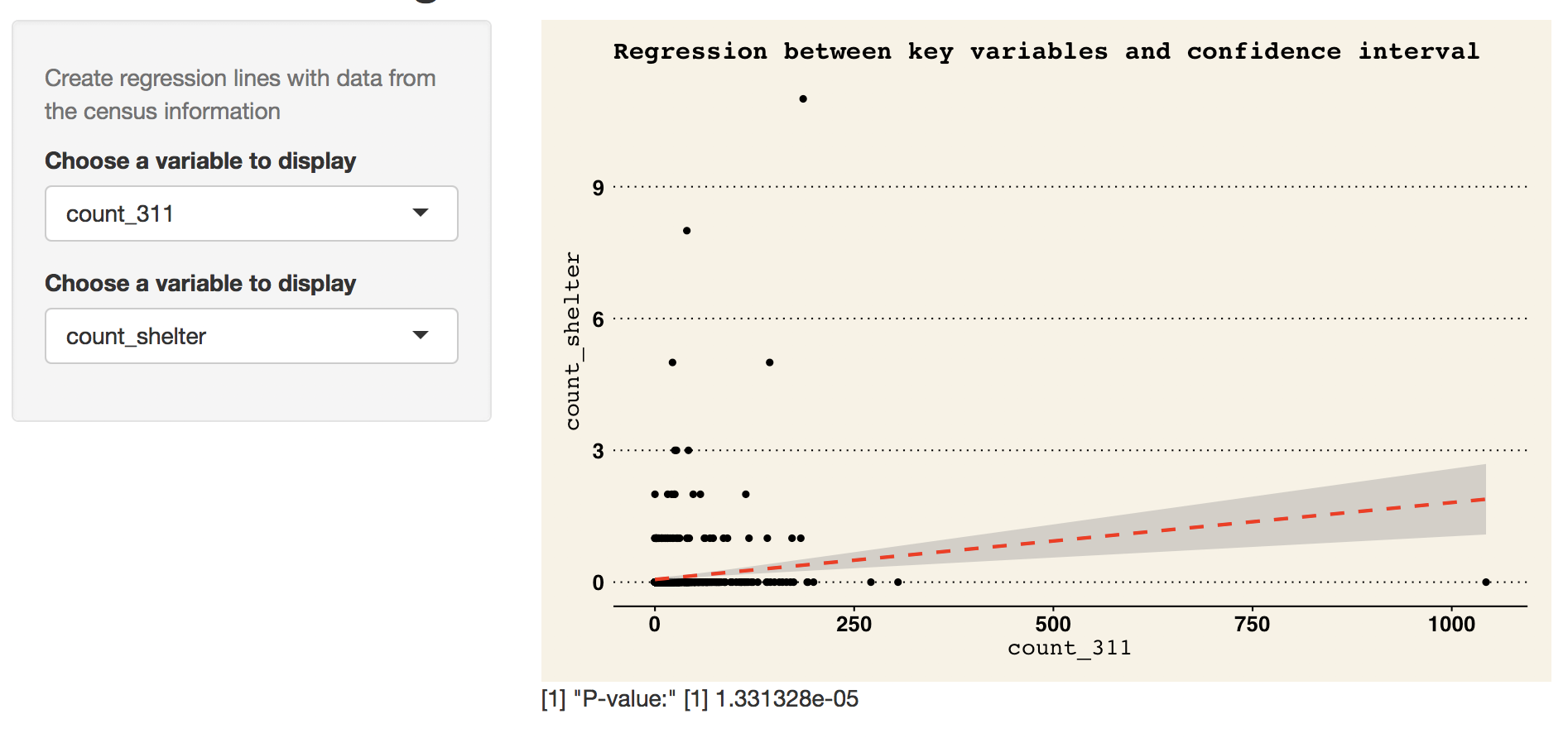
1. Number of Homeless People(Crime) across different Descents

This bar chart shows the number of crime across the descent. From the table, we can find that the Black, Hispanic/Latin/Mexican and White are most likely to be homeless victims.

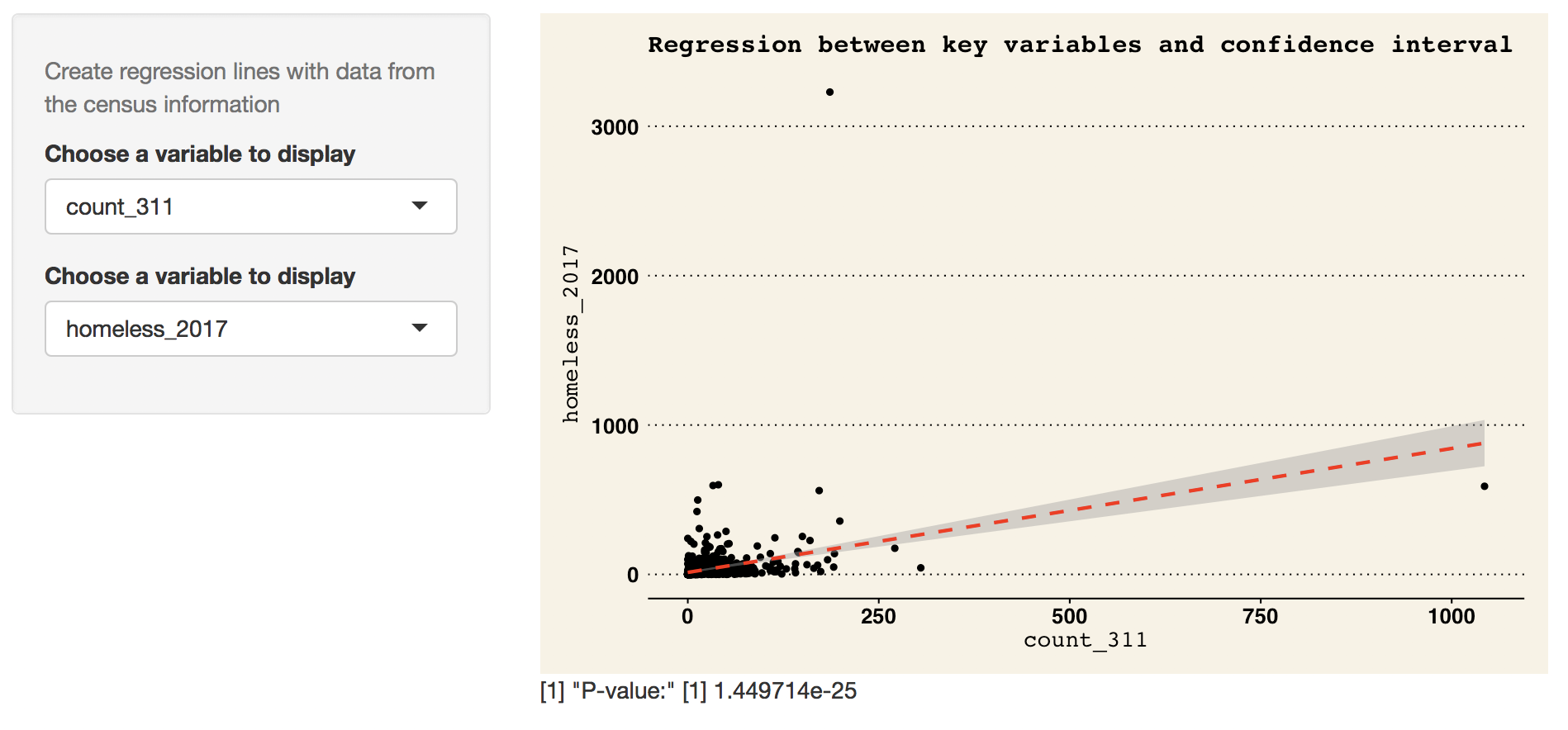


* **Predictive Modeling Analysis**

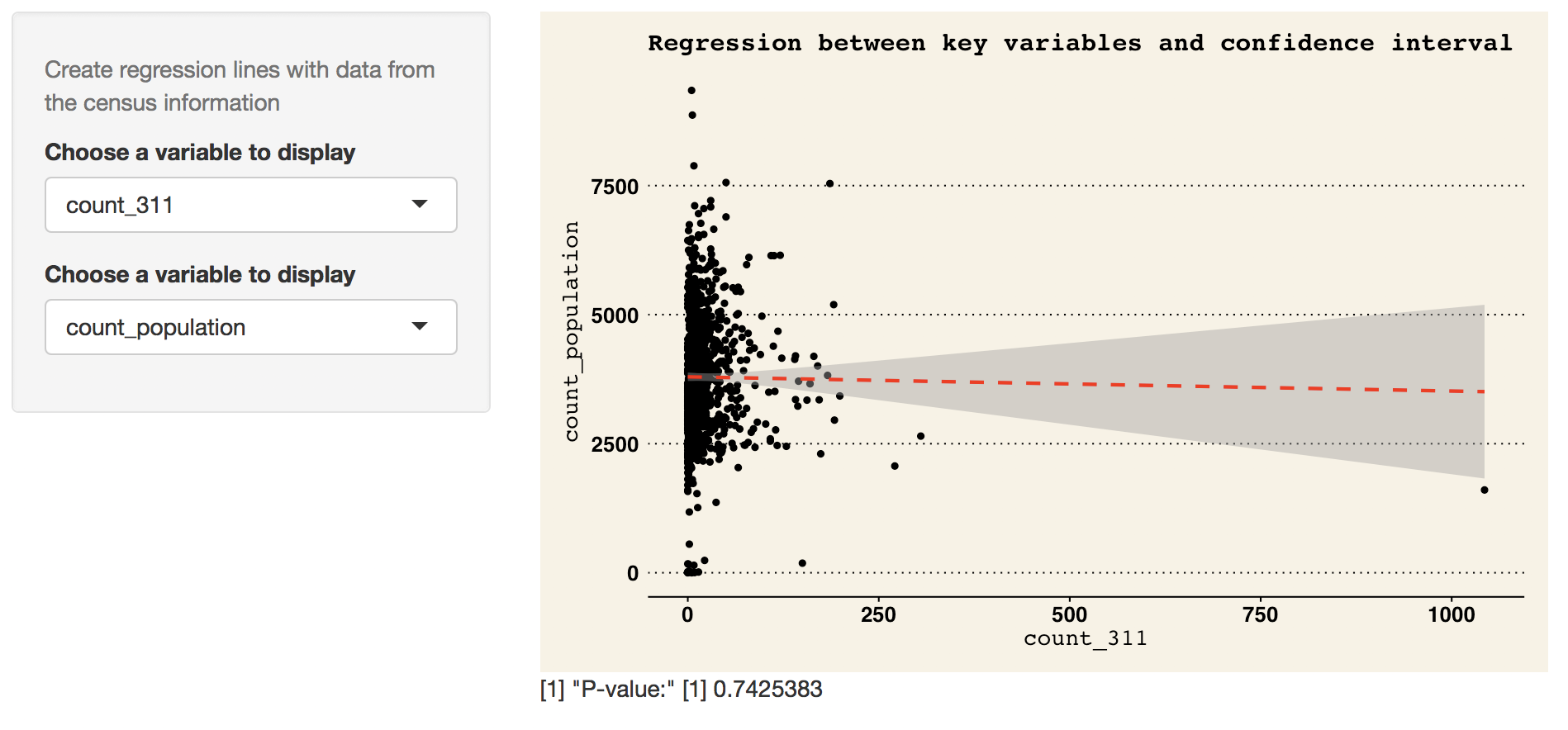
1. Shelter and 311\_calls

****

1. 311 and Homeless



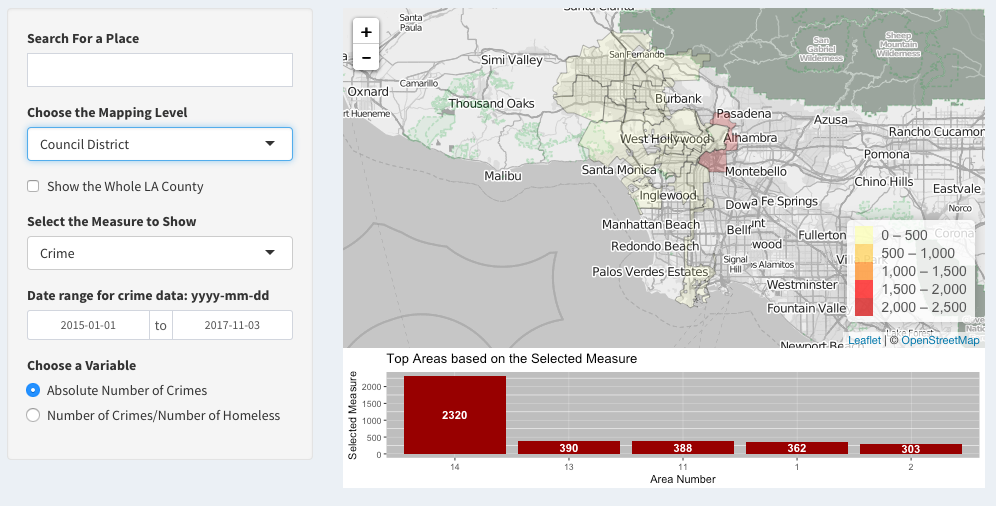
1. 311 and Population



This dashboard shows us the linear regression results between two key variables which are chosen from 311 calls number, population, crime number, shelter number and homeless people number. The regression results tell us that all the combinations of 2 variables have significantly positive relationship, except the combination of 311 calls number and population. The grey areas show a 95% confidence interval for the regression line. But what we can get is just a positive relationship, rather than a causal relationship. For example, we cannot conclude that the more shelters, the more 311 calls. Rather, it’s because shelters are built in the area where there is more homeless people’s demand. To some degree, is shows the current distribution of shelters is reasonable.

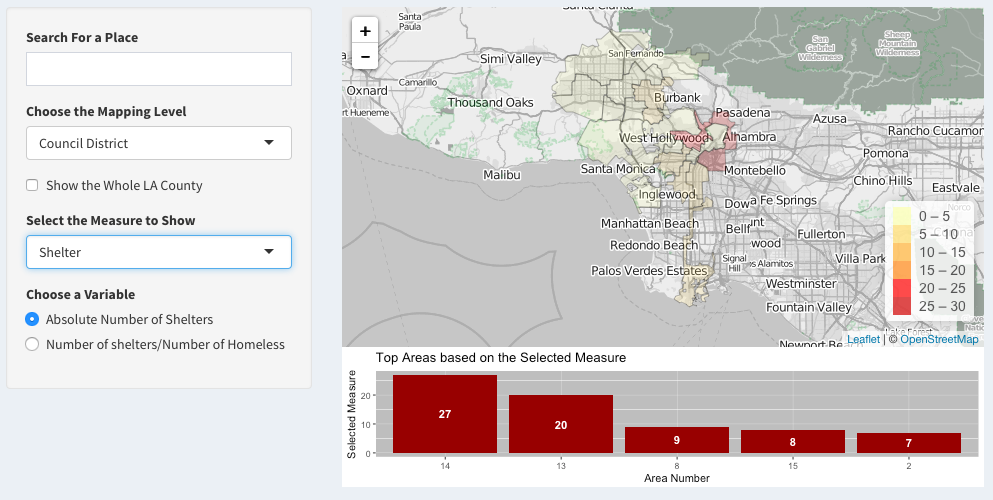
* **Geographical Analysis**

1. Crime Heat Map based on Council District



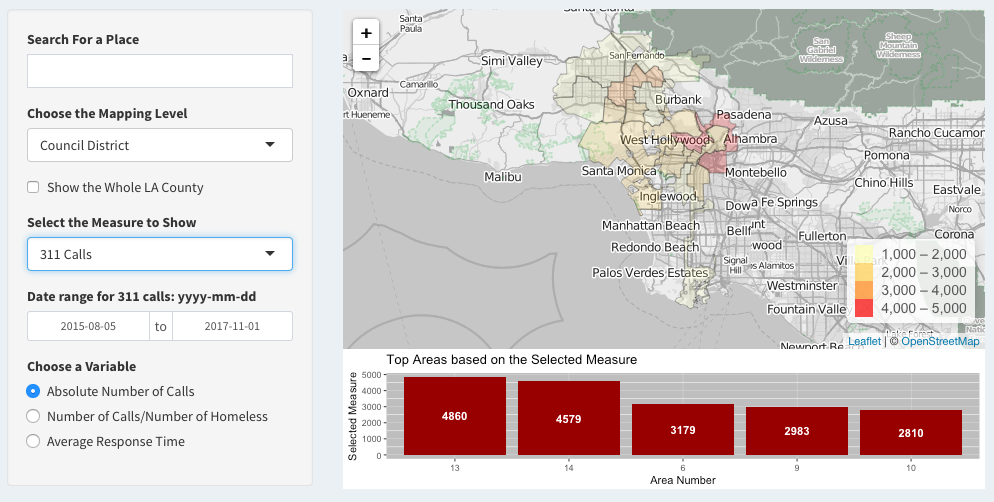
This heat map is to show the number of crime in different council districts. Users can choose the specific place to show, choose council district or census tract, choose data range of data and choose different variables to show. From the bar chart below the map, user can get information about which area has the highest count of crime.

1. Shelter Distribution Map based on Council District



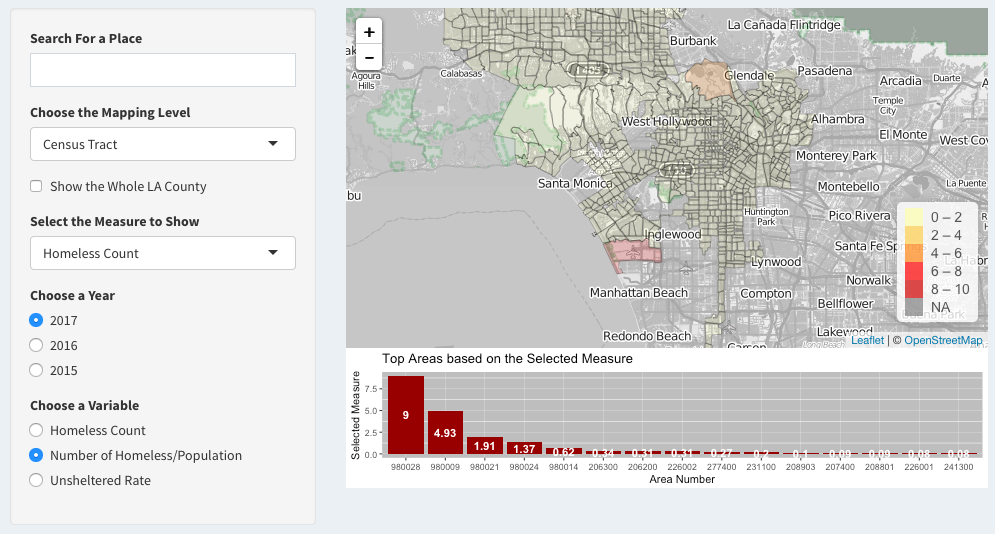
This heat map is to show the number of shelter in different council districts. Users can choose the specific place to show, choose council district or census tract and choose different variables to show. From the bar chart below the map, user can get information about which area has the highest amount of shelter.

1. 311 Calls Map based on Council District



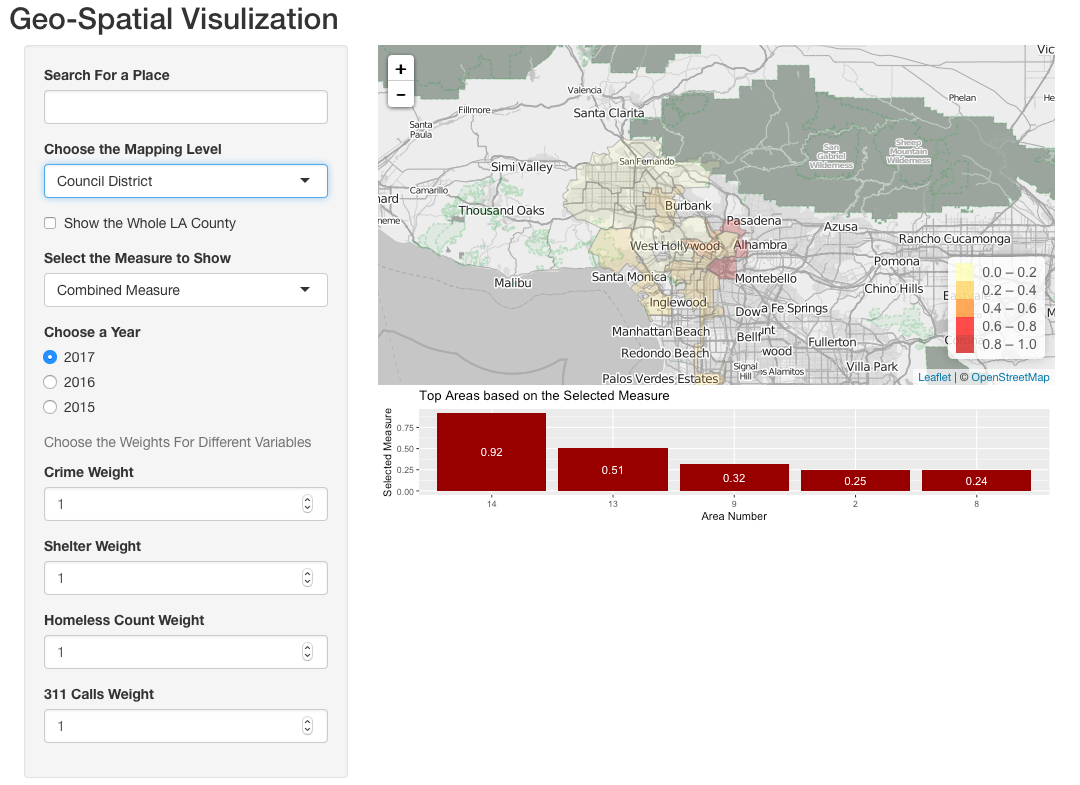
This heat map is to show the number of 311 calls in different council districts. Users can choose the specific place to show, choose council district or census tract and choose different variables to show. From the bar chart below the map, user can get information about which area has the highest amount of 311 calls.

1. Homeless Distribution Map based on Council District



This heat map is to show the count of homeless in different census tracts. Users can choose the specific place to show, choose council district or census tract and choose different variables to show. From the bar chart below the map, user can get information about which area has the highest amount of homeless.

1. Area Risk Measure



* Choose the mapping level, select the combined measure to show a map about the risk index of different area based on council district or census tract. User can manually enter the risk weights for those four variables (crime, shelter, homeless, 311 calls), then get the area with highest risk index. We also have unique features such as “Search for a Place” to help users get the geospatial analysis data from a certain location such as USC.
* The dashboard offers several adjustable measurements based on certain parameters to users. For example, crime measurements help users navigate certain locations which have crimes (higher or lower). The main goal of this dashboard is to help the City of LA analyzing potential locations of homeless and already risked locations of homeless.

**Part III: Conclusion**

With the help of our geospatial dashboard, the user can decide which areas are of the highest priority based on the risk measure. the risk measure is consisted of four variables and the relative importance of this variables can be chosen by the user in the UI. these variables are: percentage of total crimes occurring in that area, (1-percentage of total shelters located in that area), percentage of 311 calls made in that area and percentage of total homeless people that are in that area. By defining these variables, all of them are normalized and so they can be added together in a meaningful way. By choosing the weights, the user can see the distribution of the measure on the map and also the top 15 census tracts or top 5 council districts based on that measure. As an actionable and tangible solution to increase the living quality of homeless people, the government should build more shelters in the areas with the highest risk index. Another solution that can lead to higher living quality of homeless people like lower rate of crime, is allocating resources more efficiently. With the help of our descriptive dashboard, some useful information about density of crimes and requests across time (day of week, hour of the day, month) and efficiency of 311 system in updating requests can be extracted. The dashboard also gives the possibility to filter the results for the desired census tract or council district. In addition to build more shelters in riskiest areas, the government can allocate more resources in the periods when there is highest number of crimes or highest number of requests.

We encourage further research to be conducted to determine other important aspects so that we could reduce or prevent homelessness in LA City and County.