Geographic Information Science III - Lab 3

Mark Baker

Loading Necessary Libraries

```
library(jsonlite)
library(ggmap)
```

```
## Loading required package: ggplot2
```

```
## Google's Terms of Service: https://cloud.google.com/maps-platform/terms/.
```

```
## Please cite ggmap if you use it! See citation("ggmap") for details.
```

Bike Share Data

For this lab, I will be using live station data from the San Francisco bike share scheme, which was presented in *Singleton et al* lab, "Linking R to the Web."

```
bikes <-fromJSON(txt="http://feeds.bayareabikeshare.com/stations/stations.json")
## Accessing dataframe
bikes_SF <- data.frame(bikes[2])
head(bikes_SF)
```

```
##   stationBeanList.id                  stationBeanList.stationName
## 1                  3 Powell St BART Station (Market St at 4th St)
## 2                  4                  Cyril Magnin St at Ellis St
## 3                  5 Powell St BART Station (Market St at 5th St)
## 4                  6                The Embarcadero at Sansome St
## 5                  7                          Frank H Ogawa Plaza
## 6                  8                The Embarcadero at Vallejo St
##   stationBeanList.availableDocks stationBeanList.totalDocks
## 1                              7                         34
## 2                             15                         34
## 3                             11                         35
## 4                             14                         22
## 5                             23                         35
## 6                             17                         22
##   stationBeanList.latitude stationBeanList.longitude
## 1                 37.78638                -122.4049
## 2                 37.78588                -122.4089
## 3                 37.78390                -122.4084
## 4                 37.80477                -122.4032
## 5                 37.80456                -122.2717
## 6                 37.79995                -122.3985
##   stationBeanList.statusValue stationBeanList.statusKey
## 1                  In Service                         1
## 2                  In Service                         1
```

```
## 3                  In Service                          1
## 4                  In Service                          1
## 5                  In Service                          1
## 6                  In Service                          1
##   stationBeanList.availableBikes                    stationBeanList.stAddress1
## 1                             27 Powell St BART Station (Market St at 4th St)
## 2                             19                  Cyril Magnin St at Ellis St
## 3                             24 Powell St BART Station (Market St at 5th St)
## 4                              8                The Embarcadero at Sansome St
## 5                             12                          Frank H Ogawa Plaza
## 6                              5                The Embarcadero at Vallejo St
##   stationBeanList.stAddress2 stationBeanList.city stationBeanList.postalCode
## 1
## 2
## 3
## 4
## 5
## 6
##   stationBeanList.location stationBeanList.altitude stationBeanList.testStation
## 1                                                                         FALSE
## 2                                                                         FALSE
## 3                                                                         FALSE
## 4                                                                         FALSE
## 5                                                                         FALSE
## 6                                                                         FALSE
##   stationBeanList.lastCommunicationTime stationBeanList.landMark
## 1                2020-04-26 08:02:06 PM
## 2                2020-04-26 07:59:55 PM
## 3                2020-04-26 07:59:55 PM
## 4                2020-04-26 08:02:20 PM
## 5                2020-04-26 07:59:54 PM
## 6                2020-04-26 07:59:54 PM
```

In this dataset, we have access to the following parameters: station name, available docks, total docks, latitude, longitude, number of available bikes, and the stations address.

Data Cleaning

In this portion of the lab, I will an entries from the data that do not have Latitude or Longitude Values. Additionally, I drop variables that will not be examined in this analysis, and will generate binary variables to signal locations where there are either fewer than 5 bikes available or fewer than 5 dock left open.

```
bikes_SF <- bikes_SF[!is.na(bikes_SF$stationBeanList.longitude),]
bikes_SF <- bikes_SF[!is.na(bikes_SF$stationBeanList.latitude),]
bikes_SF_clean <-subset(bikes_SF, select = -c(stationBeanList.stAddress2,
                                              stationBeanList.statusKey,
                                              stationBeanList.statusValue,
                                              stationBeanList.testStation,
                                              stationBeanList.lastCommunicationTime,
                                              stationBeanList.landMark))
# Stations with more than 5 bikes available.
bikes_SF_clean$fewerthan5ab <- ifelse(bikes_SF_clean$stationBeanList.availableBikes < 5, 1, 0)
# Stations with more than 5 available docs
bikes_SF_clean$lessthan5ad <- ifelse(bikes_SF_clean$stationBeanList.availableDocks < 5, 1, 0)
```

```
summary(bikes_SF_clean)
```

```
##  stationBeanList.id stationBeanList.stationName stationBeanList.availableDocks
##  Min.   :  3.0      Length:454                  Min.   : 0.00
##  1st Qu.:132.2      Class :character            1st Qu.: 6.25
##  Median :253.5      Mode  :character            Median :10.00
##  Mean   :262.2                                  Mean   :11.01
##  3rd Qu.:397.8                                  3rd Qu.:15.00
##  Max.   :525.0                                  Max.   :37.00
##  stationBeanList.totalDocks stationBeanList.latitude stationBeanList.longitude
##  Min.   : 0.00              Min.   :37.31            Min.   :-122.5
##  1st Qu.:16.00              1st Qu.:37.75            1st Qu.:-122.4
##  Median :19.50              Median :37.78            Median :-122.4
##  Mean   :20.21              Mean   :37.71            Mean   :-122.3
##  3rd Qu.:23.00              3rd Qu.:37.80            3rd Qu.:-122.3
##  Max.   :38.00              Max.   :37.88            Max.   :-121.9
##  stationBeanList.availableBikes stationBeanList.stAddress1 stationBeanList.city
##  Min.   : 0.000                 Length:454                 Length:454
##  1st Qu.: 5.000                 Class :character           Class :character
##  Median : 8.000                 Mode  :character           Mode  :character
##  Mean   : 8.762
##  3rd Qu.:12.000
##  Max.   :31.000
##  stationBeanList.postalCode stationBeanList.location stationBeanList.altitude
##  Length:454                 Length:454               Length:454
##  Class :character           Class :character         Class :character
##  Mode  :character           Mode  :character         Mode  :character
##
##
##
##   fewerthan5ab     lessthan5ad
##  Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.0000   Median :0.0000
##  Mean   :0.2379   Mean   :0.1762
##  3rd Qu.:0.0000   3rd Qu.:0.0000
##  Max.   :1.0000   Max.   :1.0000
```

Generating Base Map using ggmap

```
## Generating base map
register_google(key = "AIzaSyDixf2h9hBKOIhXOHMZm3DF3BK18e8Tcz0") #Note API regenerated after lab was su
map <- get_map(c(-122.4, 37.75), zoom= 12,maptype = "roadmap")
```

```
## Source : https://maps.googleapis.com/maps/api/staticmap?center=37.75,-122.4&zoom=12&size=640x640&scal
```

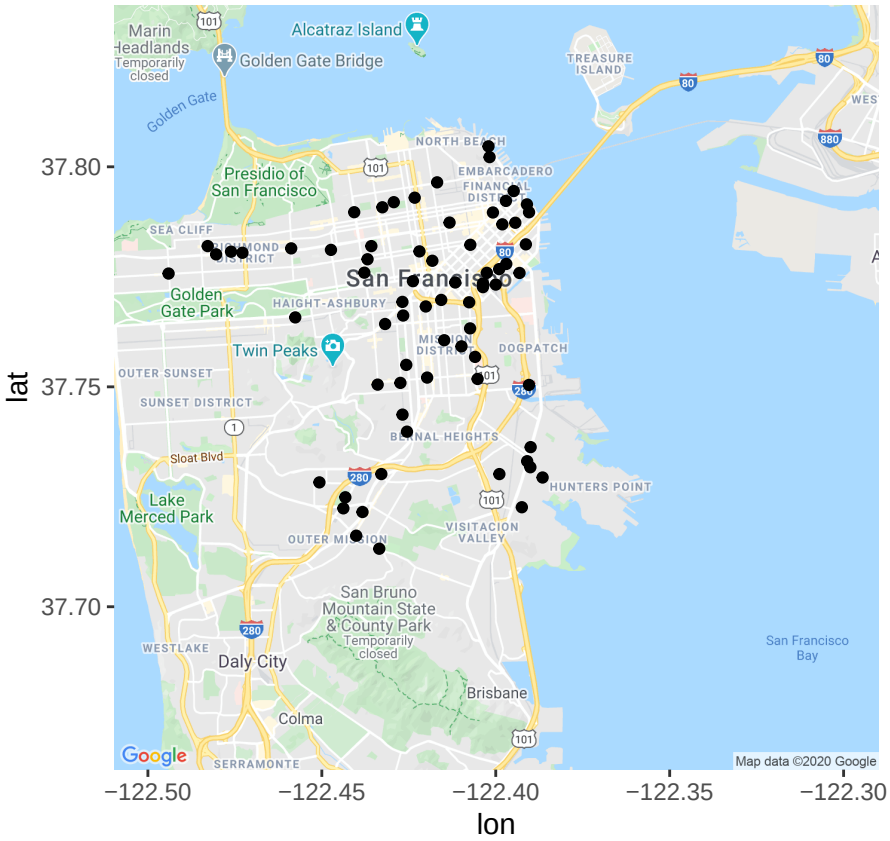
```
P <- ggmap(map)
```

Locations with Fewer Than 5 Bikes Available

In this section, I will highlight the areas that have fewer than 5 bikes, emphasizing stations that are more likely to not have available bikes at all times.

```
bikes_SF_5_bikes <- bikes_SF_clean[bikes_SF_clean$fewerthan5ab == 1,]
P + geom_point(data=bikes_SF_5_bikes, aes(x=stationBeanList.longitude, y=stationBeanList.latitude),size=
```

```
## Warning: Removed 38 rows containing missing values (geom_point).
```



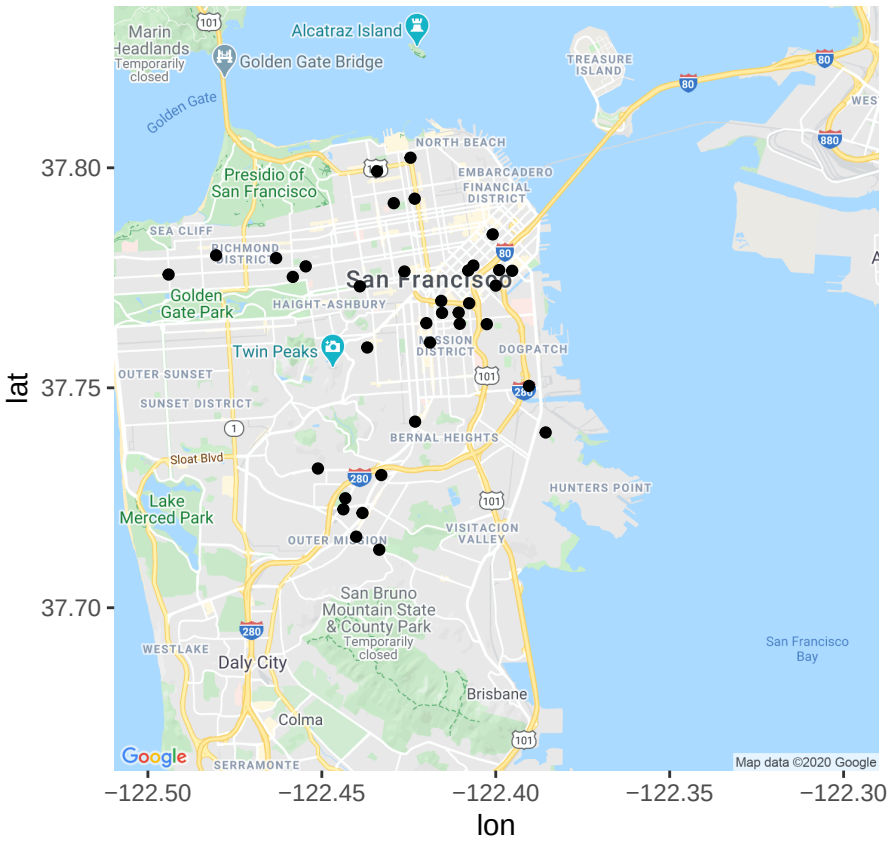
Based on the resulting output, it appears that individuals near the Financial District are more likely to encounter stations with fewer than 5 bikes. This observation means that during peak usage hours, it may be likely that users in this area are going to experience some difficulty in initially locating available bikes.

Locations with Fewer Than 5 Docks Available

In this section, I will highlight the areas that have fewer than 5 docks, emphasizing stations that are more likely to not have available docks at all times.

```
bikes_SF_5_docks <- bikes_SF_clean[bikes_SF_clean$lessthan5ad == 1,]
P + geom_point(data=bikes_SF_5_docks, aes(x=stationBeanList.longitude, y=stationBeanList.latitude),size=
```

```
## Warning: Removed 44 rows containing missing values (geom_point).
```



Based on this output, it appears that users near the Mission District may have difficulties dropping off their bikes. Additionally, in San Fransico, there appear to be fewer stations that have fewer than 5 spaces avaible, signalling that users most likely do as frequently run into problems with drop offs compared to finding available.

Trouble Shooting Log:

During this Lab, I encountered difficulties setting up the **ggmaps** library. This resulted from the new requirement that an API key be utilized. After creating an API Key, I did not encounter any other major problems.