ALY 6070:

COMMUNICATION AND VISUALIZATION FOR DATA ANALYTICS

Assignment 3:

Airbnb New York Analysis (2011 - 2019)
Choropleth map using ShinyR

Submitted to:
Prof. Fatemeh Ahmadi Abkenari

Submitted by:

Abhilash Dikshit

Milan Prajapati

Shamim Sherafati

Smit Parmar

College of Professional Studies,
Northeastern University
Vancouver, Canada

Introduction:

Airbnb has been facilitating unique and tailored travel experiences for both guests and hosts since 2008. The present dataset provides comprehensive information on the metrics and listing activity specific to New York City in 2019.

Dataset:

This dataset contains information about Airbnb host listings, neighbourhoods, price, minimum nights, review counts/rate, longitude, latitude, and the availability throughout the year.

Limitations:

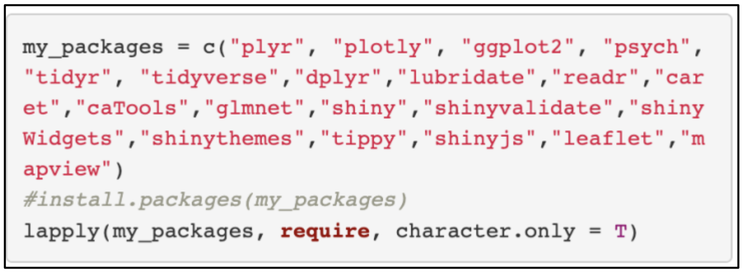
No data on customer stay duration, timeline, review score for each listing and no. of tourist attractions nearby.

Acknowledgements:

Dataset taken from Kaggle. This is a **public dataset of Airbnb**, and the original source can be found in their website.

Approach:

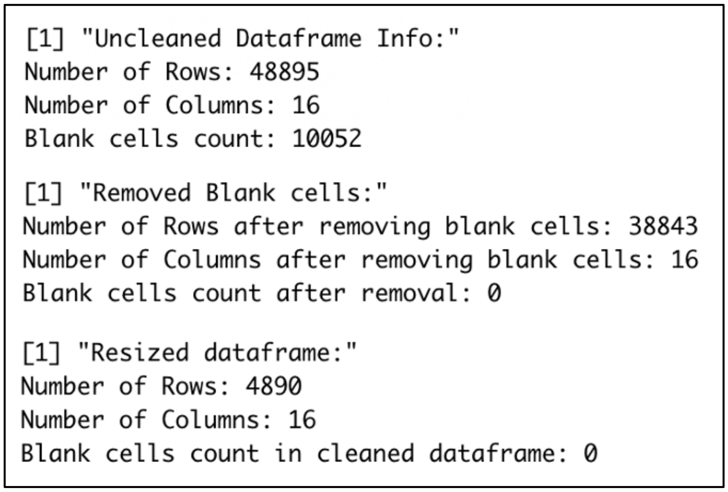
First, let's install and load the necessary packages - shiny, leaflet, and dplyr.



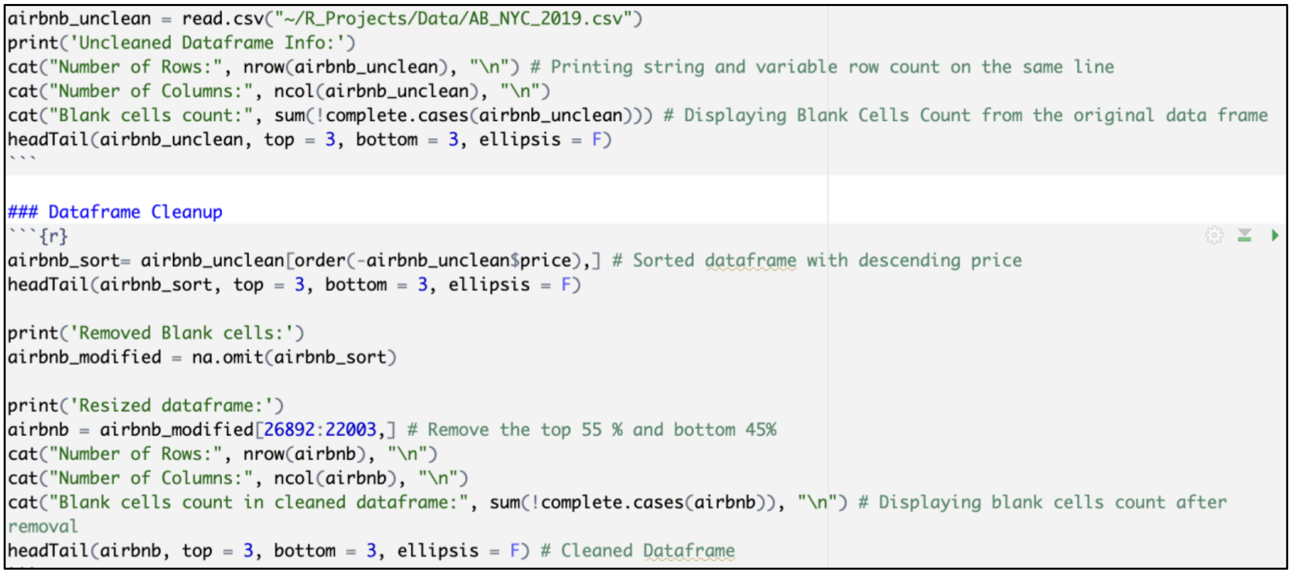
```r
my_packages = c("plyr", "plotly", "ggplot2", "psych", "tidyr", "tidyverse","dplyr","lubridate","readr","caret","caTools","glmnet","shiny","shinyvalidate","shinyWidgets","shinythemes","tippy","shinyjs","leaflet","mapview")
#install.packages(my_packages)
lapply(my_packages, require, character.only = T)
```

Dataframe cleanup:

- Sorted as per price.
- Removed blank cells.
- Resized the data frame to rows< 4890 and attributes> 16 by removimg the top 55 % and bottom 45%.



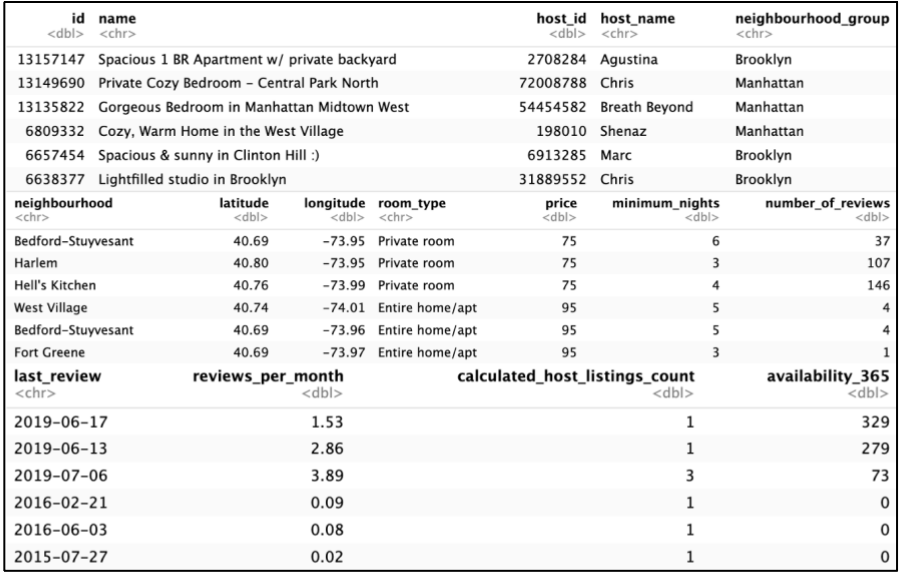
```
[1] "Uncleaned Dataframe Info:"
Number of Rows: 48895
Number of Columns: 16
Blank cells count: 10052

[1] "Removed Blank cells:"
Number of Rows after removing blank cells: 38843
Number of Columns after removing blank cells: 16
Blank cells count after removal: 0

[1] "Resized dataframe:"
Number of Rows: 4890
Number of Columns: 16
Blank cells count in cleaned dataframe: 0
```



```
airbnb_unclean = read.csv("~/R_Projects/Data/AB_NYC_2019.csv")
print('Uncleaned Dataframe Info:')
cat("Number of Rows:", nrow(airbnb_unclean), "\n") # Printing string and variable row count on the same line
cat("Number of Columns:", ncol(airbnb_unclean), "\n")
cat("Blank cells count:", sum(!complete.cases(airbnb_unclean))) # Displaying Blank Cells Count from the original data frame
headTail(airbnb_unclean, top = 3, bottom = 3, ellipsis = F)
```

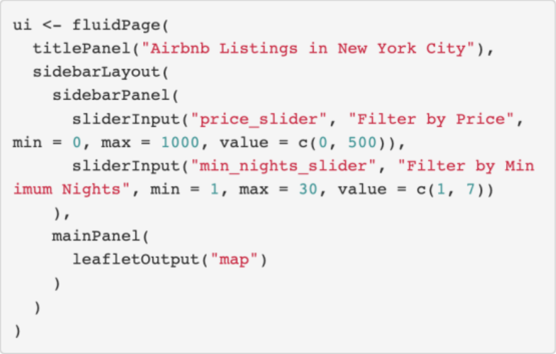
~~~
### Dataframe Cleanup
```{r}
airbnb_sort= airbnb_unclean[order(-airbnb_unclean$price),] # Sorted dataframe with descending price
headTail(airbnb_sort, top = 3, bottom = 3, ellipsis = F)

print('Removed Blank cells:')
airbnb_modified = na.omit(airbnb_sort)

print('Resized dataframe:')
airbnb = airbnb_modified[26892:22003,] # Remove the top 55 % and bottom 45%
cat("Number of Rows:", nrow(airbnb), "\n")
cat("Number of Columns:", ncol(airbnb), "\n")
cat("Blank cells count in cleaned dataframe:", sum(!complete.cases(airbnb)), "\n") # Displaying blank cells count after removal
headTail(airbnb, top = 3, bottom = 3, ellipsis = F) # Cleaned Dataframe
~~~

Final Dataset Output:



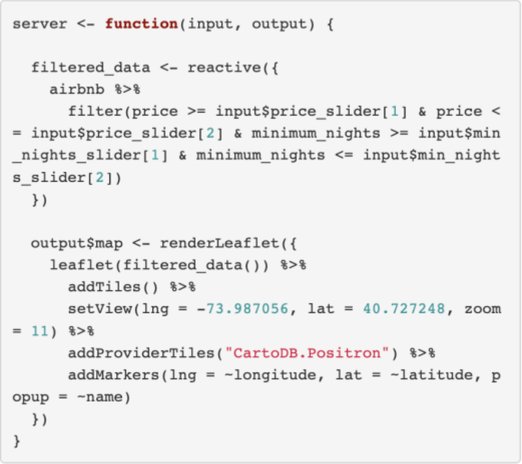
id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	minimum_nights	number_of_reviews	last_review	reviews_per_month	calculated_host_listings_count	availability_365
<dbl>	<chr>	<dbl>	<chr>	<chr>	<chr>	<dbl>	<dbl>	<chr>	<dbl>	<dbl>	<dbl>	<chr>	<dbl>	<dbl>	<dbl>
13157147	Spacious 1 BR Apartment w/ private backyard	2708284	Agustina	Brooklyn	Bedford-Stuyvesant	40.69	-73.95	Private room	75	6	37	2019-06-17	1.53	1	329
13149690	Private Cozy Bedroom - Central Park North	72008788	Chris	Manhattan	Harlem	40.80	-73.95	Private room	75	3	107	2019-06-13	2.86	1	279
13135822	Gorgeous Bedroom in Manhattan Midtown West	54454582	Breath Beyond	Manhattan	Hell's Kitchen	40.76	-73.99	Private room	75	4	146	2019-07-06	3.89	3	73
6809332	Cozy, Warm Home in the West Village	198010	Shenaz	Manhattan	West Village	40.74	-74.01	Entire home/apt	95	5	4	2016-02-21	0.09	1	0
6657454	Spacious & sunny in Clinton Hill :)	6913285	Marc	Brooklyn	Bedford-Stuyvesant	40.69	-73.96	Entire home/apt	95	5	4	2016-06-03	0.08	1	0
6638377	Lightfilled studio in Brooklyn	31889552	Chris	Brooklyn	Fort Greene	40.69	-73.97	Entire home/apt	95	3	1	2015-07-27	0.02	1	0

Now, we can create our Shiny app UI and server functions. In the UI function, we will create a leafletOutput object to display the choropleth map, and sliders to allow the user to filter the data based on the price and minimum_nights columns.



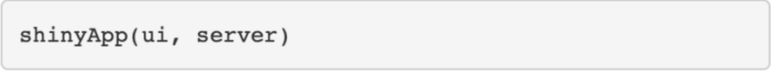
```
ui <- fluidPage(
  titlePanel("Airbnb Listings in New York City"),
  sidebarLayout(
    sidebarPanel(
      sliderInput("price_slider", "Filter by Price", min = 0, max = 1000, value = c(0, 500)),
      sliderInput("min_nights_slider", "Filter by Minimum Nights", min = 1, max = 30, value = c(1, 7))
    ),
    mainPanel(
      leafletOutput("map")
    )
  )
)
```

In the server function, we will filter the Airbnb based on the user's selected price and minimum nights values and create a leaflet object with a choropleth layer to display the data on a map.



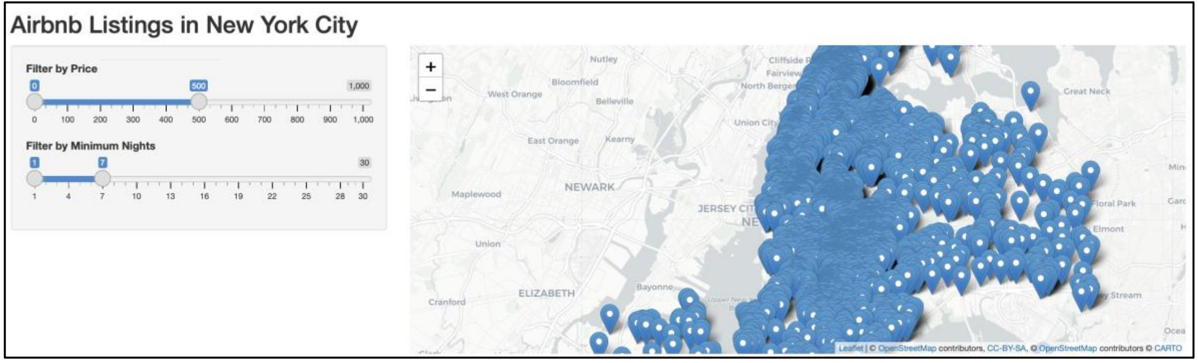
```
server <- function(input, output) {

  filtered_data <- reactive({
    airbnb %>%
      filter(price >= input$price_slider[1] & price <
= input$price_slider[2] & minimum_nights >= input$min
_nights_slider[1] & minimum_nights <= input$min_night
s_slider[2])
  })

  output$map <- renderLeaflet({
    leaflet(filtered_data()) %>%
      addTiles() %>%
      setView(lng = -73.987056, lat = 40.727248, zoom
= 11) %>%
      addProviderTiles("CartoDB.Positron") %>%
      addMarkers(lng = ~longitude, lat = ~latitude, p
opup = ~name)
  })
}
```

Finally, we can run the Shiny app using the shinyApp function. This will open a Shiny app in web browser, where we can use the sliders to filter the Airbnb data based on price and minimum nights, and view the filtered data on a choropleth map of New York City.



```
shinyApp(ui, server)
```

Output:



Airbnb Listings in New York City

Filter by Price: 0 – 500 (0 to 1,000)

Filter by Minimum Nights: 1 – 7 (1 to 30)

Conclusion:

The Airbnb listings can be filtered depending on the user's chosen price and minimum nights values, and it will build a leaflet object with a choropleth layer to display the information on a map. This gives a better clarity to customers to book there Airbnb's based on some of the factors listed below:

- Good hospitality service and corporate discounts.
- Breakfast, parking, airport pickup/drop facility.
- Better transport and infrastructure.

Reference:

1. Chang, W., Cheng, J., Allaire, J., Xie, Y., & McPherson, J. (2020). shiny: Web Application Framework for R. R package version 1.6.0. Retrieved from https://CRAN.R-project.org/package=shiny
2. RStudio. (n.d.). Leaflet for R. Retrieved March 19, 2023, from https://rstudio.github.io/leaflet/