R Notebook

This is an R Markdown (http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Cmd+Shift+Enter*.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
```

```
transport_data <- read.csv("/Users/vaishnavishankardevadig/Downloads/CTA_-_Ridership_
-__L__Station_Entries_-_Monthly_Day-Type_Averages___Totals-2.csv")
distinct_data = subset(transport_data, !duplicated(stationame))
distinct_data$stationame
```

```
##   [1] "Howard"                  "Jarvis"
##   [3] "Morse"                   "Loyola"
##   [5] "Granville"               "Thorndale"
##   [7] "Bryn Mawr"               "Berwyn"
##   [9] "Argyle"                  "Lawrence"
##  [11] "Wilson"                  "Sheridan"
##  [13] "Addison-North Main"      "Belmont-North Main"
##  [15] "Fullerton"               "North/Clybourn"
##  [17] "Clark/Division"          "Chicago/State"
##  [19] "Grand/State"             "Lake/State"
##  [21] "Washington/State"        "Monroe/State"
##  [23] "Jackson/State"           "Harrison"
##  [25] "Roosevelt"               "Cermak-Chinatown"
```

```
##  [27] "Sox-35th-Dan Ryan"        "47th-Dan Ryan"
##  [29] "Garfield-Dan Ryan"        "63rd-Dan Ryan"
##  [31] "69th"                     "79th"
##  [33] "87th"                     "95th/Dan Ryan"
##  [35] "Linden"                   "Central-Evanston"
##  [37] "Noyes"                    "Foster"
##  [39] "Davis"                    "Dempster"
##  [41] "Main"                     "South Boulevard"
##  [43] "Skokie"                   "O'Hare Airport"
##  [45] "Rosemont"                 "Cumberland"
##  [47] "Harlem-O'Hare"            "Jefferson Park"
##  [49] "Montrose-O'Hare"          "Irving Park-O'Hare"
##  [51] "Addison-O'Hare"           "Belmont-O'Hare"
##  [53] "Logan Square"             "California/Milwaukee"
##  [55] "Western/Milwaukee"        "Damen/Milwaukee"
##  [57] "Division/Milwaukee"       "Chicago/Milwaukee"
##  [59] "Grand/Milwaukee"          "Washington/Dearborn"
##  [61] "Monroe/Dearborn"          "Jackson/Dearborn"
##  [63] "LaSalle"                  "Clinton-Forest Park"
##  [65] "UIC-Halsted"              "Racine"
##  [67] "Medical Center"           "Western-Forest Park"
##  [69] "Kedzie-Homan-Forest Park" "Pulaski-Forest Park"
##  [71] "Cicero-Forest Park"       "Austin-Forest Park"
##  [73] "Oak Park-Forest Park"     "Harlem-Forest Park"
##  [75] "Forest Park"              "Polk"
##  [77] "18th"                     "Damen-Cermak"
##  [79] "Western-Cermak"           "California-Cermak"
##  [81] "Kedzie-Cermak"            "Central Park"
##  [83] "Pulaski-Cermak"           "Kostner"
##  [85] "Cicero-Cermak"            "54th/Cermak"
##  [87] "Harlem-Lake"              "Oak Park-Lake"
##  [89] "Ridgeland"                "Austin-Lake"
##  [91] "Central-Lake"             "Laramie"
##  [93] "Cicero-Lake"              "Pulaski-Lake"
##  [95] "Kedzie-Lake"              "California-Lake"
##  [97] "Ashland-Lake"             "Clinton-Lake"
##  [99] "35-Bronzeville-IIT"       "Indiana"
## [101] "43rd"                     "47th-South Elevated"
## [103] "51st"                     "Garfield-South Elevated"
## [105] "King Drive"               "East 63rd-Cottage Grove"
## [107] "Halsted/63rd"             "Ashland/63rd"
## [109] "Kimball"                  "Kedzie-Brown"
## [111] "Francisco"                "Rockwell"
## [113] "Western-Brown"            "Damen-Brown"
## [115] "Montrose-Brown"           "Irving Park-Brown"
## [117] "Addison-Brown"            "Paulina"
## [119] "Southport"                "Wellington"
```

```
## [121] "Diversey"                 "Armitage"
## [123] "Sedgwick"                 "Chicago/Franklin"
## [125] "Merchandise Mart"         "Midway Airport"
## [127] "Pulaski-Orange"           "Kedzie-Midway"
## [129] "Western-Orange"           "35th/Archer"
## [131] "Ashland-Orange"           "Halsted-Orange"
## [133] "Washington/Wells"         "Quincy/Wells"
## [135] "LaSalle/Van Buren"        "Library"
## [137] "Adams/Wabash"             "Madison/Wabash"
## [139] "Randolph/Wabash"          "State/Lake"
## [141] "Clark/Lake"               "Conservatory"
## [143] "Homan"                    "Dempster-Skokie"
## [145] "Oakton-Skokie"            "Morgan-Lake"
## [147] "Cermak-McCormick Place"   "Washington/Wabash"
```

```
summary(transport_data)
```

```
##    station_id     stationame        month_beginning    avg_weekday_rides
##  Min.   :40010   Length:37622       Length:37622       Length:37622
##  1st Qu.:40370   Class :character   Class :character   Class :character
##  Median :40760   Mode  :character   Mode  :character   Mode  :character
##  Mean   :40766
##  3rd Qu.:41150
##  Max.   :41700
##  avg_saturday_rides avg_Sunday_holiday_rides  monthtotal
##  Length:37622       Length:37622             Length:37622
##  Class :character   Class :character         Class :character
##  Mode  :character   Mode  :character         Mode  :character
##
##
##
```

```
df= data.frame(transport_data)
df$avg_weekday_rides <- gsub("[^0-9]", "", df$avg_weekday_rides)
df$avg_weekday_rides <- as.numeric(df$avg_weekday_rides)
df$avg_saturday_rides <- gsub("[^0-9]", "", df$avg_saturday_rides)
df$avg_saturday_rides <- as.numeric(df$avg_saturday_rides)
df$avg_Sunday_holiday_rides <- gsub("[^0-9]", "", df$avg_Sunday_holiday_rides)
df$avg_Sunday_holiday_rides <- as.numeric(df$avg_Sunday_holiday_rides)
df_means <- df %>% group_by(stationame) %>% summarize(avg_weekday_rides = mean(avg_we
ekday_rides),
                                                      avg_saturday_rides = mean(avg_satur
day_rides),
                                                      avg_Sunday_holiday_rides = mean(avg
_Sunday_holiday_rides))
# Sort data frame by avg_weekday_rides in descending order
sum_by_station <- df_means[order(df_means$avg_weekday_rides, decreasing = TRUE),]

# Subset data frame to top 10 stations
sum_by_station <- head(sum_by_station, n = 30)
sum_by_station$stationame <- factor(sum_by_station$stationame, levels = sum_by_statio
n$stationame[order(sum_by_station$avg_weekday_rides)])
colors <- rainbow(nrow(sum_by_station))
# Create horizontal bar graph with ggplot2
ggplot(sum_by_station, aes(x = avg_weekday_rides, y = stationame, fill=stationame)) +
  geom_bar(stat = "identity") +
  labs(x = "Average Weekday Rides", y = "Station Name") +
  ggtitle("Top 30 Stations by Average Weekday Rides")
```

```
sum_by_station <- df_means[order(df_means$avg_saturday_rides, decreasing = TRUE),]

# Subset data frame to top 10 stations
sum_by_station <- head(sum_by_station, n = 30)
sum_by_station$stationame <- factor(sum_by_station$stationame, levels = sum_by_statio
n$stationame[order(sum_by_station$avg_saturday_rides)])

# Create horizontal bar graph with ggplot2
ggplot(sum_by_station, aes(x = avg_saturday_rides, y = stationame, fill = stationame)
) +
  geom_bar(stat = "identity") +
  labs(x = "Average Saturday Rides", y = "Station Name") +
  ggtitle("Top 30 Stations by Average Saturday Rides")
```

```
sum_by_station <- df_means[order(df_means$avg_Sunday_holiday_rides, decreasing = TRUE
),]

# Subset data frame to top 10 stations
sum_by_station <- head(sum_by_station, n = 30)
sum_by_station$stationame <- factor(sum_by_station$stationame, levels = sum_by_statio
n$stationame[order(sum_by_station$avg_Sunday_holiday_rides)])

# Create horizontal bar graph with ggplot2
ggplot(sum_by_station, aes(x = avg_Sunday_holiday_rides, y = stationame, fill = stati
oname)) +
  geom_bar(stat = "identity") +
  labs(x = "Average Sunday and Holiday Rides", y = "Station Name") +
  ggtitle("Top 30 Stations by Average Sunday and Holiday Rides")
```

Top 30 Stations by Average Sunday and Holiday Rides
Station Name
O'Hare Airport
Chicago/State
Grand/State
Belmont-North Main
Lake/State
Roosevelt
Addison-North Main
Fullerton
95th/Dan Ryan
Clark/Division
Clark/Lake
State/Lake
79th
Midway Airport
Washington/Dearborn
Jackson/State
North/Clybourn
Howard
Cermak-Chinatown
Damen/Milwaukee
Monroe/State
Wilson
Loyola
69th
Washington/Wabash
Sheridan
Sox-35th-Dan Ryan
Adams/Wabash
Logan Square
Bryn Mawr
0e+00
2e+05
4e+05
6e+05
Average Sunday and Holiday Rides
stationame
Bryn Mawr
Logan Square
Adams/Wabash
Sox-35th-Dan Ryan
Sheridan
Washington/Wabash
69th
Loyola
Wilson
Monroe/State
Damen/Milwaukee
Cermak-Chinatown
Howard
North/Clybourn
Jackson/State
Washington/Dearborn
Midway Airport
79th
State/Lake
Clark/Lake
Clark/Division
95th/Dan Ryan
Fullerton
Addison-North Main
Roosevelt
Lake/State
Belmont-North Main
Grand/State
Chicago/State
O'Hare Airport

R Notebook

This is an R Markdown (http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Cmd+Shift+Enter*.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
transport_data <- read.csv("/Users/vaishnavishankardevadig/Downloads/CTA_-_Ridership_
-_Bus_Routes_-_Monthly_Day-Type_Averages___Totals-2.csv")
distinct_data = subset(transport_data, !duplicated(routename))
distinct_data$routename
```

```
##   [1] "Indiana/Hyde Park"                "Hyde Park Express"
##   [3] "King Drive"                       "Cottage Grove"
##   [5] "Jackson Park Express"             "Harrison"
##   [7] "Halsted"                          "South Halsted"
##   [9] "Ashland"                          "Museum of S & I"
##  [11] "Lincoln/Sedgwick"                 "Roosevelt"
##  [13] "Jeffery Express"                  "Westchester"
##  [15] "16th/18th"                        "United Center Express"
##  [17] "Madison"                          "Cermak"
##  [19] "Cermak Express"                   "Clark"
##  [21] "Wentworth"                        "West Cermak"
##  [23] "South Deering"                    "Stony Island"
##  [25] "State"                            "South Chicago"
##  [27] "Mag Mile Express"                 "South Michigan"
```

```
##  [29] "35th"                                "Broadway"
##  [31] "Sedgwick"                            "Pershing"
##  [33] "43rd"                                "Wallace-Racine"
##  [35] "47th"                                "South Damen"
##  [37] "Western"                             "South Western"
##  [39] "North Western"                       "Western Express"
##  [41] "Damen"                               "51st"
##  [43] "Kedzie/California"                   "South Kedzie"
##  [45] "Pulaski"                             "South Pulaski"
##  [47] "Cicero"                              "North Cicero/Skokie Blvd."
##  [49] "South Cicero"                        "Garfield"
##  [51] "55th/Narragansett"                   "Milwaukee"
##  [53] "North Milwaukee"                     "Laramie"
##  [55] "59th/61st"                           "Blue Island/26th"
##  [57] "Archer"                              "Archer/Harlem"
##  [59] "63rd"                                "West 63rd"
##  [61] "Foster-Canfield"                     "Grand"
##  [63] "Chicago"                             "67th-69th-71st"
##  [65] "Northwest Highway"                   "Cumberland/East River"
##  [67] "Division"                            "71st/South Shore"
##  [69] "North"                               "Armitage"
##  [71] "Fullerton"                           "74th-75th"
##  [73] "Diversey"                            "Belmont"
##  [75] "Montrose"                            "79th"
##  [77] "Irving Park"                         "Lawrence"
##  [79] "West Lawrence"                       "Kimball-Homan"
##  [81] "Peterson"                            "Central"
##  [83] "North Central"                       "Narragansett/Ridgeland"
##  [85] "87th"                                "Higgins"
##  [87] "Harlem"                              "North Harlem"
##  [89] "Austin"                              "Foster"
##  [91] "California/Dodge"                    "South California"
##  [93] "93rd-95th"                           "West 95th"
##  [95] "Lunt"                                "Skokie"
##  [97] "Jeffery Manor Express"               "West 103rd"
##  [99] "East 103rd"                          "Halsted/95th"
## [101] "Pullman/111th/115th"                 "Vincennes/111th"
## [103] "Michigan/119th"                      "Ogilvie/Wacker Express"
## [105] "Union/Wacker Express"                "Illinois Center/Ogilvie Express"
## [107] "Illinois Center/Union Express"       "Water Tower Express"
## [109] "Jackson"                             "Madison/Roosevelt Circulator"
## [111] "West Loop/South Loop"                "Clarendon/LaSalle Express"
## [113] "Sheridan/LaSalle Express"            "Wilson/Michigan Express"
## [115] "Inner Drive/Michigan Express"        "Outer Drive Express"
## [117] "Sheridan"                            "Addison"
## [119] "Devon"                               "LaSalle"
## [121] "Streeterville/Taylor"                "69th-UPS Express"
```

```
## [123] "U. of Chicago/Midway"              "U. of Chicago/Hyde Park"
## [125] "U. of Chicago/Kenwood"             "U. of Chicago/Lakeview Express"
## [127] "Central/Ridge"                     "Main/Emerson"
## [129] "Ridge/Grant"                       "Dodge"
## [131] "Soldier Field Express"             "Navy Pier"
## [133] "Wrigley Field Express"             "Museum Campus"
## [135] "UIC-Pilsen Express"                "Irving Park Express"
## [137] "Avon Express"                      "Garfield Express"
## [139] "69th Bus Pre-Paid Area"            "South Shore Night Bus"
## [141] "Main Shuttle"                      "Chicago/Golf"
## [143] "Evanston Circulator"               "Jeffery Local"
## [145] "South Shore Express"               "Stockton/LaSalle Express"
## [147] "Stockton/Michigan Express"         "Marine/Michigan Express"
## [149] "Shuttle/Special Event Route"       "Clarendon/Michigan Express"
## [151] "Chinatown/Pilsen Shuttle"          "Chicago Manufacturing Campus"
## [153] "Stony Island Express"              "South Pulaski Limited"
## [155] "King Drive Express"                "Cottage Grove Express"
## [157] "U. of Chicago Hospitals Express"   "Special Event Route"
## [159] "Ashland Express"                   "Washington/Madison Express"
## [161] "Ogden/Taylor"                      "Cicero Express"
## [163] "U. of Chicago/Garfield Stations"   "55th/Austin"
## [165] "Goose Island Express"              "West 65th"
## [167] "Touhy Supplement"                  "Central/Sherman"
## [169] "31st/35th"                         "Bronzeville/Union Station"
## [171] "Lincoln"                           "111th/King Drive"
## [173] "Ogilvie/Streeterville Express"     "Union/Streeterville Express"
## [175] "Jeffery Jump"                      "Pullman/115th"
## [177] "Cermak-Roosevelt Express"          "Dan Ryan OWL Shuttle"
## [179] "Dan Ryan Local Shuttle"            "69th-Garfield Express Shuttle"
## [181] "79th-Garfield Express Shuttle"     "87th-Garfield Express Shuttle"
## [183] "95th-Garfield Express Shuttle"     "Pershing Shuttle"
## [185] "Pullman Shuttle"                   "31st"
## [187] "95th"                              "ROAD CALL"
## [189] "Special Dest Signs"                "Kedzie"
## [191] "California"                        "O'Hare Shuttle"
## [193] "Inner Lake Shore/Michigan Express" "Outer DuSable Lake Shore Express"
```

```
summary(transport_data)
```

```
##     route            routename         month_beginning    Avg_Weekday_Rides
##  Length:35966       Length:35966       Length:35966       Min.   :    0
##  Class :character   Class :character   Class :character   1st Qu.: 1239
##  Mode  :character   Mode  :character   Mode  :character   Median : 3732
##                                                           Mean   : 6246
##                                                           3rd Qu.: 9569
##                                                           Max.   :37787
##  Avg_Saturday_Rides Avg_Sunday_Holiday_Rides   MonthTotal
##  Min.   :    0      Min.   :    0            Min.   :      1
##  1st Qu.:    0      1st Qu.:    0            1st Qu.:  28338
##  Median : 1796      Median : 1122            Median :  92612
##  Mean   : 3961      Mean   : 2771            Mean   : 163017
##  3rd Qu.: 6106      3rd Qu.: 4357            3rd Qu.: 248763
##  Max.   :30645      Max.   :24111            Max.   :1058879
```

```
df= data.frame(transport_data)
df
```

route	routename	month_beginning	Avg_Weekday_Rides	Avg_Saturday_I
<chr>	<chr>	<chr>	<dbl>	
1	Indiana/Hyde Park	1/1/01	6982.6	
2	Hyde Park Express	1/1/01	1000.0	
3	King Drive	1/1/01	21406.5	13
4	Cottage Grove	1/1/01	22432.2	17
6	Jackson Park Express	1/1/01	18443.0	13
7	Harrison	1/1/01	5504.4	
8	Halsted	1/1/01	19582.2	12
8A	South Halsted	1/1/01	3196.5	3
9	Ashland	1/1/01	29265.4	22
10	Museum of S & I	1/1/01	0.0	

1-10 of 10,000 rows | 1-5 of 7 columns Previous **1** 2 3 4 5 6 … 1000 Next

```
df_means <- df %>% group_by(routename) %>% summarize(Avg_Weekday_Rides = mean(Avg_Wee
kday_Rides),
                                                     Avg_Saturday_Rides = mean(Avg_Satur
day_Rides),
                                                     Avg_Sunday_Holiday_Rides = mean(Avg
_Sunday_Holiday_Rides))
sum_by_route <- df_means[order(df_means$Avg_Weekday_Rides, decreasing = TRUE),]

# Subset data frame to top 10 stations
sum_by_route <- head(sum_by_route, n = 30)
sum_by_route$routename <- factor(sum_by_route$routename, levels = sum_by_route$routen
ame[order(sum_by_route$Avg_Weekday_Rides)])
colors <- rainbow(nrow(sum_by_route))
# Create horizontal bar graph with ggplot2
ggplot(sum_by_route, aes(x = Avg_Weekday_Rides, y = routename, fill=routename)) +
  geom_bar(stat = "identity") +
  labs(x = "Average Weekday Rides", y = "Route Name") +
  ggtitle("Top 30 Routes by Average Weekday Rides")
```

```
sum_by_route <- df_means[order(df_means$Avg_Saturday_Rides, decreasing = TRUE),]

# Subset data frame to top 10 stations
sum_by_route <- head(sum_by_route, n = 30)
sum_by_route$routename <- factor(sum_by_route$routename, levels = sum_by_route$routen
ame[order(sum_by_route$Avg_Saturday_Rides)])
colors <- rainbow(nrow(sum_by_route))
# Create horizontal bar graph with ggplot2
ggplot(sum_by_route, aes(x = Avg_Saturday_Rides, y = routename, fill=routename)) +
  geom_bar(stat = "identity") +
  labs(x = "Average Saturday Rides", y = "Route Name") +
  ggtitle("Top 30 Routes by Average Saturday Rides")
```

```
sum_by_route <- df_means[order(df_means$Avg_Sunday_Holiday_Rides, decreasing = TRUE),
]

# Subset data frame to top 10 stations
sum_by_route <- head(sum_by_route, n = 30)
sum_by_route$routename <- factor(sum_by_route$routename, levels = sum_by_route$routen
ame[order(sum_by_route$Avg_Sunday_Holiday_Rides)])
colors <- rainbow(nrow(sum_by_route))
# Create horizontal bar graph with ggplot2
ggplot(sum_by_route, aes(x = Avg_Sunday_Holiday_Rides, y = routename, fill=routename)
) +
  geom_bar(stat = "identity") +
  labs(x = "Average Sunday and Holiday Rides", y = "Route Name") +
  ggtitle("Top 30 Routes by Average Sunday and Holiday Rides")
```

```
library(ggplot2)

data <- read.csv("C:\\Users\\hp\\OneDrive\\Documents\\Semester 2\\Data Prep\\Project\\CTA_-_Ride
rship_-__L__Station_Entries_-_Monthly_Day-Type_Averages___Totals.csv", header=TRUE)

ggplot(data, aes(x=monthtotal)) +
  geom_line(aes(y = Weekday_Rides ), color = "darkred",linetype="twodash") +
  geom_line(aes(y = Saturday_Rides), color="steelblue", linetype="twodash")+
  geom_line(aes(y = Sunday_Holiday_Rides), color="darkgreen", linetype="twodash")+
  ggtitle("Comparing L-Train Rides based on days")+
  xlab("Monthly Rides")+ylab("Rides") +
  scale_color_manual(name = "Days",
                     values = c("darkred", "steelblue", "darkgreen"),
                     labels = c("Weekday", "Saturday", "Sunday/Holiday"))
```

```
data<-read.csv("C:\\Users\\hp\\OneDrive\\Documents\\Semester 2\\Data Prep\\Project\\CTA_-_Riders
hip_-_Bus_Routes_-_Monthly_Day-Type_Averages___Totals.csv")

ggplot(data, aes(x=monthtotal)) +
  geom_line(aes(y = Weekday_Rides ), color = "darkred",linetype="twodash") +
  geom_line(aes(y = Saturday_Rides), color="steelblue", linetype="twodash")+
  geom_line(aes(y = Sunday_Holiday_Rides), color="darkgreen", linetype="twodash")+
  ggtitle("Comparing Bus Rides based on days")+
  xlab("Monthly Rides")+ylab("Rides") +
  scale_color_manual(name = "Days",
                     values = c("darkred", "steelblue", "darkgreen"),
                     labels = c("Weekday", "Saturday", "Sunday/Holiday"))
```

```
crime_data <- read.csv('C:/Users/siddh/OneDrive - hawk.iit.edu/Desktop/DPA/Proj_siddhi/CTA_Crim
e.csv', header=TRUE, stringsAsFactors=FALSE)
ridership_data <- read.csv("C:/Users/siddh/OneDrive - hawk.iit.edu/Desktop/DPA/Proj_siddhi/CTA_-
_Ridership_-_Daily_Boarding_Totals.csv", header=TRUE, stringsAsFactors=FALSE)
```

```
crime_data$Date <- as.Date(crime_data$Date, "%m/%d/%Y %I:%M:%S %p")
crime_data$day_of_week <- weekdays(crime_data$Date)
daily_crime_count <- aggregate(crime_data$ID, by=list(crime_data$Date), FUN=length)
colnames(daily_crime_count) <- c("Date", "Crime_Count")
```

```
library(ggplot2)
ggplot(daily_crime_count, aes(x=Date, y=Crime_Count)) +
  geom_line() +
  labs(x="Date", y="Daily Crime Count", title="Daily Crime Count in Chicago Transit Authority")
```

dataPrepWeather

2023-03-28

```
weatherData = read.csv(file.path("/Users/pranitkotkar/Downloads/DPA_Proj/Q4/weatherData.
csv"))
summary(weatherData)
```

```
##      Date             TAVG              TMAX             TMIN       
##  Length:33332       Mode:logical   Min.   :-11.00   Min.   :-25.00  
##  Class :character   NA's:33332     1st Qu.: 41.00   1st Qu.: 28.00  
##  Mode  :character                  Median : 60.00   Median : 42.00  
##                                    Mean   : 58.93   Mean   : 41.81  
##                                    3rd Qu.: 78.00   3rd Qu.: 58.00  
##                                    Max.   :107.00   Max.   : 84.00  
##                                    NA's   :31       NA's   :41      
##       PRCP             SNOW              SNWD       
##  Min.   :0.0000   Min.   : 0.0000   Min.   : 0.0000  
##  1st Qu.:0.0000   1st Qu.: 0.0000   1st Qu.: 0.0000  
##  Median :0.0000   Median : 0.0000   Median : 0.0000  
##  Mean   :0.1014   Mean   : 0.1119   Mean   : 0.5116  
##  3rd Qu.:0.0400   3rd Qu.: 0.0000   3rd Qu.: 0.0000  
##  Max.   :6.1600   Max.   :17.6000   Max.   :29.0000  
##  NA's   :40       NA's   :137       NA's   :145      
```

```
weatherData = cbind(weatherData[1],weatherData[3],weatherData[4],weatherData[5],weatherD
ata[6])
```

```
# convert the date column to a Date object
weatherData$Date <- as.Date(weatherData$Date, format = "%m/%d/%Y")

# format the date column to the desired format
weatherData$Date <- format(weatherData$Date, "%Y-%m-%d")
```

```
weatherData$Date <- as.Date(weatherData$Date)
summary(weatherData)
```

```
##       Date                  TMAX             TMIN             PRCP
##  Min.   :0000-01-01   Min.   :-11.00   Min.   :-25.00   Min.   :0.0000
##  1st Qu.:0022-10-24   1st Qu.: 41.00   1st Qu.: 28.00   1st Qu.:0.0000
##  Median :0054-03-15   Median : 60.00   Median : 42.00   Median :0.0000
##  Mean   :0051-11-07   Mean   : 58.93   Mean   : 41.81   Mean   :0.1014
##  3rd Qu.:0077-01-06   3rd Qu.: 78.00   3rd Qu.: 58.00   3rd Qu.:0.0400
##  Max.   :0099-12-31   Max.   :107.00   Max.   : 84.00   Max.   :6.1600
##                       NA's   :31       NA's   :41       NA's   :40
##       SNOW
##  Min.   : 0.0000
##  1st Qu.: 0.0000
##  Median : 0.0000
##  Mean   : 0.1119
##  3rd Qu.: 0.0000
##  Max.   :17.6000
##  NA's   :137
```

```
newData <- subset(weatherData, Date >= "0001-01-01" & Date <= "0022-06-30")
```

```
newData$Date <- gsub("^\\d{2}", "20", newData$Date)
summary(newData)
```

```
##      Date                 TMAX             TMIN             PRCP
##  Length:7851        Min.   : -9.00   Min.   :-22.00   Min.   :0.0000
##  Class :character   1st Qu.: 43.00   1st Qu.: 30.00   1st Qu.:0.0000
##  Mode  :character   Median : 62.00   Median : 44.00   Median :0.0000
##                     Mean   : 60.26   Mean   : 44.17   Mean   :0.1144
##                     3rd Qu.: 79.00   3rd Qu.: 61.00   3rd Qu.:0.0500
##                     Max.   :105.00   Max.   : 83.00   Max.   :5.1100
##                     NA's   :1        NA's   :3
##       SNOW
##  Min.   : 0.0000
##  1st Qu.: 0.0000
##  Median : 0.0000
##  Mean   : 0.1127
##  3rd Qu.: 0.0000
##  Max.   :17.2000
##  NA's   :123
```

```
newData$Date <- as.Date(newData$Date)
summary(newData)
```

```
##       Date                 TMAX             TMIN             PRCP
##  Min.   :2001-01-01   Min.   : -9.00   Min.   :-22.00   Min.   :0.0000
##  1st Qu.:2006-05-17   1st Qu.: 43.00   1st Qu.: 30.00   1st Qu.:0.0000
##  Median :2011-10-01   Median : 62.00   Median : 44.00   Median :0.0000
##  Mean   :2011-10-01   Mean   : 60.26   Mean   : 44.17   Mean   :0.1144
##  3rd Qu.:2017-02-13   3rd Qu.: 79.00   3rd Qu.: 61.00   3rd Qu.:0.0500
##  Max.   :2022-06-30   Max.   :105.00   Max.   : 83.00   Max.   :5.1100
##                       NA's   :1        NA's   :3
##       SNOW
##  Min.   : 0.0000
##  1st Qu.: 0.0000
##  Median : 0.0000
##  Mean   : 0.1127
##  3rd Qu.: 0.0000
##  Max.   :17.2000
##  NA's   :123
```

```
which(is.na(newData$TMAX))
```

```
## [1] 2259
```

```
newData <- newData[-2259, ]
summary(newData)
```

```
##       Date                 TMAX             TMIN             PRCP
##  Min.   :2001-01-01   Min.   : -9.00   Min.   :-22.00   Min.   :0.0000
##  1st Qu.:2006-05-17   1st Qu.: 43.00   1st Qu.: 30.00   1st Qu.:0.0000
##  Median :2011-10-01   Median : 62.00   Median : 44.00   Median :0.0000
##  Mean   :2011-10-01   Mean   : 60.26   Mean   : 44.17   Mean   :0.1144
##  3rd Qu.:2017-02-13   3rd Qu.: 79.00   3rd Qu.: 61.00   3rd Qu.:0.0500
##  Max.   :2022-06-30   Max.   :105.00   Max.   : 83.00   Max.   :5.1100
##                                        NA's   :3
##       SNOW
##  Min.   : 0.0000
##  1st Qu.: 0.0000
##  Median : 0.0000
##  Mean   : 0.1128
##  3rd Qu.: 0.0000
##  Max.   :17.2000
##  NA's   :123
```

```
which(is.na(newData$TMIN))
```

```
## [1]  299 2228 2562
```

```
newData <- newData[-299,]
summary(newData)
```

```
##       Date                 TMAX             TMIN             PRCP
##  Min.   :2001-01-01   Min.   : -9.00   Min.   :-22.00   Min.   :0.0000
##  1st Qu.:2006-05-18   1st Qu.: 43.00   1st Qu.: 30.00   1st Qu.:0.0000
##  Median :2011-10-02   Median : 62.00   Median : 44.00   Median :0.0000
##  Mean   :2011-10-01   Mean   : 60.26   Mean   : 44.17   Mean   :0.1144
##  3rd Qu.:2017-02-14   3rd Qu.: 79.00   3rd Qu.: 61.00   3rd Qu.:0.0500
##  Max.   :2022-06-30   Max.   :105.00   Max.   : 83.00   Max.   :5.1100
##                                        NA's   :2
##       SNOW
##  Min.   : 0.0000
##  1st Qu.: 0.0000
##  Median : 0.0000
##  Mean   : 0.1128
##  3rd Qu.: 0.0000
##  Max.   :17.2000
##  NA's   :123
```

```
which(is.na(newData$TMIN))
```

```
## [1] 2227 2561
```

```
newData <- newData[-2227,]
summary(newData)
```

```
##       Date                 TMAX             TMIN             PRCP
##  Min.   :2001-01-01   Min.   : -9.00   Min.   :-22.00   Min.   :0.0000
##  1st Qu.:2006-05-17   1st Qu.: 43.00   1st Qu.: 30.00   1st Qu.:0.0000
##  Median :2011-10-02   Median : 62.00   Median : 44.00   Median :0.0000
##  Mean   :2011-10-01   Mean   : 60.26   Mean   : 44.17   Mean   :0.1144
##  3rd Qu.:2017-02-14   3rd Qu.: 79.00   3rd Qu.: 61.00   3rd Qu.:0.0500
##  Max.   :2022-06-30   Max.   :105.00   Max.   : 83.00   Max.   :5.1100
##                                        NA's   :1
##       SNOW
##  Min.   : 0.0000
##  1st Qu.: 0.0000
##  Median : 0.0000
##  Mean   : 0.1124
##  3rd Qu.: 0.0000
##  Max.   :17.2000
##  NA's   :123
```

```
which(is.na(newData$TMIN))
```

```
## [1] 2560
```

```
newData <- newData[-2560,]
summary(newData)
```

```
##       Date                 TMAX             TMIN             PRCP
##  Min.   :2001-01-01   Min.   : -9.00   Min.   :-22.00   Min.   :0.0000
##  1st Qu.:2006-05-17   1st Qu.: 43.00   1st Qu.: 30.00   1st Qu.:0.0000
##  Median :2011-10-03   Median : 62.00   Median : 44.00   Median :0.0000
##  Mean   :2011-10-02   Mean   : 60.26   Mean   : 44.17   Mean   :0.1144
##  3rd Qu.:2017-02-14   3rd Qu.: 79.00   3rd Qu.: 61.00   3rd Qu.:0.0500
##  Max.   :2022-06-30   Max.   :105.00   Max.   : 83.00   Max.   :5.1100
##
##       SNOW
##  Min.   : 0.0000
##  1st Qu.: 0.0000
##  Median : 0.0000
##  Mean   : 0.1125
##  3rd Qu.: 0.0000
##  Max.   :17.2000
##  NA's   :123
```

```
newData$SNOW <- ifelse(is.na(newData$SNOW), 0, newData$SNOW)
summary(newData)
```

```
##       Date                 TMAX             TMIN             PRCP
##  Min.   :2001-01-01   Min.   : -9.00   Min.   :-22.00   Min.   :0.0000
##  1st Qu.:2006-05-17   1st Qu.: 43.00   1st Qu.: 30.00   1st Qu.:0.0000
##  Median :2011-10-03   Median : 62.00   Median : 44.00   Median :0.0000
##  Mean   :2011-10-02   Mean   : 60.26   Mean   : 44.17   Mean   :0.1144
##  3rd Qu.:2017-02-14   3rd Qu.: 79.00   3rd Qu.: 61.00   3rd Qu.:0.0500
##  Max.   :2022-06-30   Max.   :105.00   Max.   : 83.00   Max.   :5.1100
##       SNOW
##  Min.   : 0.0000
##  1st Qu.: 0.0000
##  Median : 0.0000
##  Mean   : 0.1107
##  3rd Qu.: 0.0000
##  Max.   :17.2000
```

precpt

2023-03-28

```
weatherData = read.csv(file.path("/Users/pranitkotkar/Downloads/DPA_Proj/weather_ride_me
rge.csv"))
summary(weatherData)
```

```
##      Date           station_name         station_id    number_of_trips
##  Length:37336       Length:37336       Min.   :40010   Min.   :     0
##  Class :character   Class :character   1st Qu.:40370   1st Qu.: 33289
##  Mode  :character   Mode  :character   Median :40760   Median : 67315
##                                        Mean   :40766   Mean   : 93358
##                                        3rd Qu.:41150   3rd Qu.:123272
##                                        Max.   :41700   Max.   :670496
##
##     Avg_temp          SNOW            PRCP            SNWD
##  Min.   :-2.00   Min.   :0.000   Min.   :0.000   Min.   : 0.000
##  1st Qu.:36.00   1st Qu.:0.000   1st Qu.:0.000   1st Qu.: 0.000
##  Median :52.00   Median :0.000   Median :0.000   Median : 0.000
##  Mean   :52.19   Mean   :0.005   Mean   :0.083   Mean   : 0.145
##  3rd Qu.:69.50   3rd Qu.:0.000   3rd Qu.:0.030   3rd Qu.: 0.000
##  Max.   :89.00   Max.   :1.200   Max.   :1.290   Max.   :17.000
##  NA's   :3290    NA's   :3290    NA's   :3290    NA's   :3290
```

```
weatherData = cbind(weatherData[4],weatherData[7])
summary(weatherData)
```

```
##  number_of_trips       PRCP
##  Min.   :     0   Min.   :0.000
##  1st Qu.: 33289   1st Qu.:0.000
##  Median : 67315   Median :0.000
##  Mean   : 93358   Mean   :0.083
##  3rd Qu.:123272   3rd Qu.:0.030
##  Max.   :670496   Max.   :1.290
##                   NA's   :3290
```

```
mean_value <- mean(weatherData$PRCP, na.rm = TRUE)
weatherData$PRCP <- ifelse(is.na(weatherData$PRCP), mean_value, weatherData$PRCP)
summary(weatherData)
```

```
##  number_of_trips       PRCP
##  Min.   :     0   Min.   :0.0000
##  1st Qu.: 33289   1st Qu.:0.0000
##  Median : 67315   Median :0.0000
##  Mean   : 93358   Mean   :0.0828
##  3rd Qu.:123272   3rd Qu.:0.0828
##  Max.   :670496   Max.   :1.2900
```

```
par(mfrow = c(2, 1))
plot(x=weatherData$PRCP,y=weatherData$PRCP,xlab = "PRCP" ,sub = "PRCP vs no. of trip",yl
ab =  "Number of trips")
```

snow

2023-03-28

```
weatherData = read.csv(file.path("/Users/pranitkotkar/Downloads/DPA_Proj/weather_ride_me
rge.csv"))
summary(weatherData)
```

```
##      Date           station_name         station_id    number_of_trips
##  Length:37336       Length:37336       Min.   :40010   Min.   :     0
##  Class :character   Class :character   1st Qu.:40370   1st Qu.: 33289
##  Mode  :character   Mode  :character   Median :40760   Median : 67315
##                                        Mean   :40766   Mean   : 93358
##                                        3rd Qu.:41150   3rd Qu.:123272
##                                        Max.   :41700   Max.   :670496
##
##     Avg_temp          SNOW            PRCP            SNWD
##  Min.   :-2.00   Min.   :0.000   Min.   :0.000   Min.   : 0.000
##  1st Qu.:36.00   1st Qu.:0.000   1st Qu.:0.000   1st Qu.: 0.000
##  Median :52.00   Median :0.000   Median :0.000   Median : 0.000
##  Mean   :52.19   Mean   :0.005   Mean   :0.083   Mean   : 0.145
##  3rd Qu.:69.50   3rd Qu.:0.000   3rd Qu.:0.030   3rd Qu.: 0.000
##  Max.   :89.00   Max.   :1.200   Max.   :1.290   Max.   :17.000
##  NA's   :3290    NA's   :3290    NA's   :3290    NA's   :3290
```

```
weatherData = cbind(weatherData[1],weatherData[3],weatherData[4],weatherData[6])
summary(weatherData)
```

```
##      Date             station_id    number_of_trips       SNOW
##  Length:37336       Min.   :40010   Min.   :     0   Min.   :0.000
##  Class :character   1st Qu.:40370   1st Qu.: 33289   1st Qu.:0.000
##  Mode  :character   Median :40760   Median : 67315   Median :0.000
##                     Mean   :40766   Mean   : 93358   Mean   :0.005
##                     3rd Qu.:41150   3rd Qu.:123272   3rd Qu.:0.000
##                     Max.   :41700   Max.   :670496   Max.   :1.200
##                                                      NA's   :3290
```

```
weatherData$SNOW <- ifelse(is.na(weatherData$SNOW), 0, weatherData$SNOW)
summary(weatherData)
```

```
##      Date             station_id    number_of_trips       SNOW
##  Length:37336       Min.   :40010   Min.   :     0   Min.   :0.000000
##  Class :character   1st Qu.:40370   1st Qu.: 33289   1st Qu.:0.000000
##  Mode  :character   Median :40760   Median : 67315   Median :0.000000
##                     Mean   :40766   Mean   : 93358   Mean   :0.004944
##                     3rd Qu.:41150   3rd Qu.:123272   3rd Qu.:0.000000
##                     Max.   :41700   Max.   :670496   Max.   :1.200000
```

```
weatherData$Date <- gsub("^\\d{2}", "20", weatherData$Date)
weatherData$Date <- as.Date(weatherData$Date)
weatherData$Date <- gsub("^\\d{2}", "20", weatherData$Date)
weatherData$Date <- as.Date(weatherData$Date)
summary(weatherData)
```

```
##       Date              station_id    number_of_trips       SNOW
##  Min.   :2001-01-20   Min.   :40010   Min.   :     0   Min.   :0.000000
##  1st Qu.:2003-01-20   1st Qu.:40370   1st Qu.: 33289   1st Qu.:0.000000
##  Median :2006-01-20   Median :40760   Median : 67315   Median :0.000000
##  Mean   :2008-09-19   Mean   :40766   Mean   : 93358   Mean   :0.004944
##  3rd Qu.:2009-01-20   3rd Qu.:41150   3rd Qu.:123272   3rd Qu.:0.000000
##  Max.   :2020-01-20   Max.   :41700   Max.   :670496   Max.   :1.200000
```

```
par(mfrow = c(2, 1))
plot(x=weatherData$SNOW,y=weatherData$number_of_trips,xlab = "SNOW" ,sub = "SNOW vs no.
of trip",ylab =  "Number of trips")
```

Number of trips
0e+00
0.0
0.2
0.4
0.6
0.8
1.0
1.2
SNOW
SNOW vs no. of trip

temp

2023-03-28

```
weatherData = read.csv(file.path("/Users/pranitkotkar/Downloads/DPA_Proj/weather_ride_me
rge.csv"))
summary(weatherData)
```

```
##      Date           station_name         station_id    number_of_trips
##  Length:37336       Length:37336       Min.   :40010   Min.   :     0
##  Class :character   Class :character   1st Qu.:40370   1st Qu.: 33289
##  Mode  :character   Mode  :character   Median :40760   Median : 67315
##                                        Mean   :40766   Mean   : 93358
##                                        3rd Qu.:41150   3rd Qu.:123272
##                                        Max.   :41700   Max.   :670496
##
##     Avg_temp          SNOW            PRCP            SNWD
##  Min.   :-2.00   Min.   :0.000   Min.   :0.000   Min.   : 0.000
##  1st Qu.:36.00   1st Qu.:0.000   1st Qu.:0.000   1st Qu.: 0.000
##  Median :52.00   Median :0.000   Median :0.000   Median : 0.000
##  Mean   :52.19   Mean   :0.005   Mean   :0.083   Mean   : 0.145
##  3rd Qu.:69.50   3rd Qu.:0.000   3rd Qu.:0.030   3rd Qu.: 0.000
##  Max.   :89.00   Max.   :1.200   Max.   :1.290   Max.   :17.000
##  NA's   :3290    NA's   :3290    NA's   :3290    NA's   :3290
```

```
weatherData = cbind(weatherData[4],weatherData[5])
summary(weatherData)
```

```
##  number_of_trips     Avg_temp
##  Min.   :     0   Min.   :-2.00
##  1st Qu.: 33289   1st Qu.:36.00
##  Median : 67315   Median :52.00
##  Mean   : 93358   Mean   :52.19
##  3rd Qu.:123272   3rd Qu.:69.50
##  Max.   :670496   Max.   :89.00
##                   NA's   :3290
```

```
mean_value <- mean(weatherData$Avg_temp, na.rm = TRUE)
weatherData$Avg_temp <- ifelse(is.na(weatherData$Avg_temp), mean_value, weatherData$Avg_
temp)
summary(weatherData)
```

```
##  number_of_trips     Avg_temp
##  Min.   :     0   Min.   :-2.00
##  1st Qu.: 33289   1st Qu.:38.00
##  Median : 67315   Median :52.19
##  Mean   : 93358   Mean   :52.19
##  3rd Qu.:123272   3rd Qu.:68.50
##  Max.   :670496   Max.   :89.00
```

```
par(mfrow = c(2, 1))
plot(x=weatherData$Avg_temp,y=weatherData$Avg_temp,xlab = "Avg_temp" ,sub = "Avg_temp vs
no. of trip",ylab =  "Number of trips")
```

model

2023-03-28

```
weatherData = read.csv(file.path("/Users/pranitkotkar/Downloads/DPA_Proj/weather_ride_me
rge.csv"))
summary(weatherData)
```

```
##      Date           station_name         station_id    number_of_trips
##  Length:37336       Length:37336       Min.   :40010   Min.   :     0
##  Class :character   Class :character   1st Qu.:40370   1st Qu.: 33289
##  Mode  :character   Mode  :character   Median :40760   Median : 67315
##                                        Mean   :40766   Mean   : 93358
##                                        3rd Qu.:41150   3rd Qu.:123272
##                                        Max.   :41700   Max.   :670496
##
##     Avg_temp          SNOW            PRCP            SNWD
##  Min.   :-2.00   Min.   :0.000   Min.   :0.000   Min.   : 0.000
##  1st Qu.:36.00   1st Qu.:0.000   1st Qu.:0.000   1st Qu.: 0.000
##  Median :52.00   Median :0.000   Median :0.000   Median : 0.000
##  Mean   :52.19   Mean   :0.005   Mean   :0.083   Mean   : 0.145
##  3rd Qu.:69.50   3rd Qu.:0.000   3rd Qu.:0.030   3rd Qu.: 0.000
##  Max.   :89.00   Max.   :1.200   Max.   :1.290   Max.   :17.000
##  NA's   :3290    NA's   :3290    NA's   :3290    NA's   :3290
```

```
weatherData = cbind( weatherData[4], weatherData[5], weatherData[6], weatherData[7])
summary(weatherData)
```

```
##  number_of_trips     Avg_temp          SNOW            PRCP
##  Min.   :     0   Min.   :-2.00   Min.   :0.000   Min.   :0.000
##  1st Qu.: 33289   1st Qu.:36.00   1st Qu.:0.000   1st Qu.:0.000
##  Median : 67315   Median :52.00   Median :0.000   Median :0.000
##  Mean   : 93358   Mean   :52.19   Mean   :0.005   Mean   :0.083
##  3rd Qu.:123272   3rd Qu.:69.50   3rd Qu.:0.000   3rd Qu.:0.030
##  Max.   :670496   Max.   :89.00   Max.   :1.200   Max.   :1.290
##                   NA's   :3290    NA's   :3290    NA's   :3290
```

```
mean_value <- mean(weatherData$Avg_temp, na.rm = TRUE)
weatherData$Avg_temp <- ifelse(is.na(weatherData$Avg_temp), mean_value, weatherData$Avg_
temp)
summary(weatherData)
```

```
##  number_of_trips     Avg_temp          SNOW            PRCP
##  Min.   :     0   Min.   :-2.00   Min.   :0.000   Min.   :0.000
##  1st Qu.: 33289   1st Qu.:38.00   1st Qu.:0.000   1st Qu.:0.000
##  Median : 67315   Median :52.19   Median :0.000   Median :0.000
##  Mean   : 93358   Mean   :52.19   Mean   :0.005   Mean   :0.083
##  3rd Qu.:123272   3rd Qu.:68.50   3rd Qu.:0.000   3rd Qu.:0.030
##  Max.   :670496   Max.   :89.00   Max.   :1.200   Max.   :1.290
##                                   NA's   :3290    NA's   :3290
```

```
weatherData$SNOW <- ifelse(is.na(weatherData$SNOW), 0, weatherData$SNOW)
summary(weatherData)
```

```
##  number_of_trips     Avg_temp          SNOW               PRCP
##  Min.   :     0   Min.   :-2.00   Min.   :0.000000   Min.   :0.000
##  1st Qu.: 33289   1st Qu.:38.00   1st Qu.:0.000000   1st Qu.:0.000
##  Median : 67315   Median :52.19   Median :0.000000   Median :0.000
##  Mean   : 93358   Mean   :52.19   Mean   :0.004944   Mean   :0.083
##  3rd Qu.:123272   3rd Qu.:68.50   3rd Qu.:0.000000   3rd Qu.:0.030
##  Max.   :670496   Max.   :89.00   Max.   :1.200000   Max.   :1.290
##                                                      NA's   :3290
```

```
mean_value <- mean(weatherData$PRCP, na.rm = TRUE)
weatherData$PRCP <- ifelse(is.na(weatherData$PRCP), mean_value, weatherData$PRCP)
summary(weatherData)
```

```
##  number_of_trips     Avg_temp          SNOW               PRCP
##  Min.   :     0   Min.   :-2.00   Min.   :0.000000   Min.   :0.0000
##  1st Qu.: 33289   1st Qu.:38.00   1st Qu.:0.000000   1st Qu.:0.0000
##  Median : 67315   Median :52.19   Median :0.000000   Median :0.0000
##  Mean   : 93358   Mean   :52.19   Mean   :0.004944   Mean   :0.0828
##  3rd Qu.:123272   3rd Qu.:68.50   3rd Qu.:0.000000   3rd Qu.:0.0828
##  Max.   :670496   Max.   :89.00   Max.   :1.200000   Max.   :1.2900
```

```
# create the multiple linear regression model
model <- lm(weatherData$number_of_trips ~ weatherData$Avg_temp + weatherData$PRCP + weat
herData$SNOW)

# view the summary output of the model
summary(model)
```

```
##
## Call:
## lm(formula = weatherData$number_of_trips ~ weatherData$Avg_temp +
##     weatherData$PRCP + weatherData$SNOW)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -99711 -59953 -25931  30094 575150
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          82902.35    1317.50  62.924   <2e-16 ***
## weatherData$Avg_temp   197.52      23.58   8.376   <2e-16 ***
## weatherData$PRCP      1915.02    2226.50   0.860     0.39
## weatherData$SNOW     -2239.39    6028.46  -0.371     0.71
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 85990 on 37332 degrees of freedom
## Multiple R-squared:  0.001946,   Adjusted R-squared:  0.001865
## F-statistic: 24.26 on 3 and 37332 DF,  p-value: 1.125e-15
```

```
length(weatherData$SNOW)
```

```
## [1] 37336
```

```
length(weatherData$PRCP)
```

```
## [1] 37336
```

```
length(weatherData$Avg_temp)
```

```
## [1] 37336
```

```
length(weatherData$number_of_trips)
```

```
## [1] 37336
```

```
# generate predicted values for number_of_trips based on Avg_temp, precipitation, and sn
ow
pred <- predict(model, newdata = weatherData[, c("Avg_temp", "PRCP", "SNOW")])
library(plot3D)

# plot the 3D scatter plot
scatter3D(weatherData$Avg_temp, weatherData$PRCP, weatherData$SNOW,
          colvar = weatherData$number_of_trips, pch = 16, cex = 0.5,
          xlab = "Average daily temperature", ylab = "Precipitation", zlab = "Snow")
```

lasso

2023-03-28

```
weatherData = read.csv(file.path("/Users/pranitkotkar/Downloads/DPA_Proj/weather_ride_me
rge.csv"))
summary(weatherData)
```

```
##      Date           station_name         station_id    number_of_trips
##  Length:37336       Length:37336       Min.   :40010   Min.   :     0
##  Class :character   Class :character   1st Qu.:40370   1st Qu.: 33289
##  Mode  :character   Mode  :character   Median :40760   Median : 67315
##                                        Mean   :40766   Mean   : 93358
##                                        3rd Qu.:41150   3rd Qu.:123272
##                                        Max.   :41700   Max.   :670496
##
##     Avg_temp          SNOW            PRCP            SNWD
##  Min.   :-2.00   Min.   :0.000   Min.   :0.000   Min.   : 0.000
##  1st Qu.:36.00   1st Qu.:0.000   1st Qu.:0.000   1st Qu.: 0.000
##  Median :52.00   Median :0.000   Median :0.000   Median : 0.000
##  Mean   :52.19   Mean   :0.005   Mean   :0.083   Mean   : 0.145
##  3rd Qu.:69.50   3rd Qu.:0.000   3rd Qu.:0.030   3rd Qu.: 0.000
##  Max.   :89.00   Max.   :1.200   Max.   :1.290   Max.   :17.000
##  NA's   :3290    NA's   :3290    NA's   :3290    NA's   :3290
```

```
weatherData$Date <- gsub("^\\d{2}", "20", weatherData$Date)
weatherData$Date <- as.Date(weatherData$Date)
weatherData$Date <- gsub("^\\d{2}", "20", weatherData$Date)
weatherData$Date <- as.Date(weatherData$Date)
summary(weatherData)
```

```
##       Date            station_name         station_id    number_of_trips
##  Min.   :2001-01-20   Length:37336       Min.   :40010   Min.   :     0
##  1st Qu.:2003-01-20   Class :character   1st Qu.:40370   1st Qu.: 33289
##  Median :2006-01-20   Mode  :character   Median :40760   Median : 67315
##  Mean   :2008-09-19                      Mean   :40766   Mean   : 93358
##  3rd Qu.:2009-01-20                      3rd Qu.:41150   3rd Qu.:123272
##  Max.   :2020-01-20                      Max.   :41700   Max.   :670496
##
##     Avg_temp          SNOW            PRCP            SNWD
##  Min.   :-2.00   Min.   :0.000   Min.   :0.000   Min.   : 0.000
##  1st Qu.:36.00   1st Qu.:0.000   1st Qu.:0.000   1st Qu.: 0.000
##  Median :52.00   Median :0.000   Median :0.000   Median : 0.000
##  Mean   :52.19   Mean   :0.005   Mean   :0.083   Mean   : 0.145
##  3rd Qu.:69.50   3rd Qu.:0.000   3rd Qu.:0.030   3rd Qu.: 0.000
##  Max.   :89.00   Max.   :1.200   Max.   :1.290   Max.   :17.000
##  NA's   :3290    NA's   :3290    NA's   :3290    NA's   :3290
```

```
mean_value <- mean(weatherData$Avg_temp, na.rm = TRUE)
weatherData$Avg_temp <- ifelse(is.na(weatherData$Avg_temp), mean_value, weatherData$Avg_
temp)
summary(weatherData)
```

```
##       Date            station_name         station_id    number_of_trips
##  Min.   :2001-01-20   Length:37336       Min.   :40010   Min.   :     0
##  1st Qu.:2003-01-20   Class :character   1st Qu.:40370   1st Qu.: 33289
##  Median :2006-01-20   Mode  :character   Median :40760   Median : 67315
##  Mean   :2008-09-19                      Mean   :40766   Mean   : 93358
##  3rd Qu.:2009-01-20                      3rd Qu.:41150   3rd Qu.:123272
##  Max.   :2020-01-20                      Max.   :41700   Max.   :670496
##
##     Avg_temp          SNOW            PRCP            SNWD
##  Min.   :-2.00   Min.   :0.000   Min.   :0.000   Min.   : 0.000
##  1st Qu.:38.00   1st Qu.:0.000   1st Qu.:0.000   1st Qu.: 0.000
##  Median :52.19   Median :0.000   Median :0.000   Median : 0.000
##  Mean   :52.19   Mean   :0.005   Mean   :0.083   Mean   : 0.145
##  3rd Qu.:68.50   3rd Qu.:0.000   3rd Qu.:0.030   3rd Qu.: 0.000
##  Max.   :89.00   Max.   :1.200   Max.   :1.290   Max.   :17.000
##                  NA's   :3290    NA's   :3290    NA's   :3290
```

```
weatherData$SNOW <- ifelse(is.na(weatherData$SNOW), 0, weatherData$SNOW)
summary(weatherData)
```

```
##       Date            station_name         station_id    number_of_trips
##  Min.   :2001-01-20   Length:37336       Min.   :40010   Min.   :     0
##  1st Qu.:2003-01-20   Class :character   1st Qu.:40370   1st Qu.: 33289
##  Median :2006-01-20   Mode  :character   Median :40760   Median : 67315
##  Mean   :2008-09-19                      Mean   :40766   Mean   : 93358
##  3rd Qu.:2009-01-20                      3rd Qu.:41150   3rd Qu.:123272
##  Max.   :2020-01-20                      Max.   :41700   Max.   :670496
##
##     Avg_temp          SNOW               PRCP            SNWD
##  Min.   :-2.00   Min.   :0.000000   Min.   :0.000   Min.   : 0.000
##  1st Qu.:38.00   1st Qu.:0.000000   1st Qu.:0.000   1st Qu.: 0.000
##  Median :52.19   Median :0.000000   Median :0.000   Median : 0.000
##  Mean   :52.19   Mean   :0.004944   Mean   :0.083   Mean   : 0.145
##  3rd Qu.:68.50   3rd Qu.:0.000000   3rd Qu.:0.030   3rd Qu.: 0.000
##  Max.   :89.00   Max.   :1.200000   Max.   :1.290   Max.   :17.000
##                                     NA's   :3290    NA's   :3290
```

```
mean_value <- mean(weatherData$PRCP, na.rm = TRUE)
weatherData$PRCP <- ifelse(is.na(weatherData$PRCP), mean_value, weatherData$PRCP)
summary(weatherData)
```

```
##       Date            station_name         station_id    number_of_trips
##  Min.   :2001-01-20   Length:37336       Min.   :40010   Min.   :     0
##  1st Qu.:2003-01-20   Class :character   1st Qu.:40370   1st Qu.: 33289
##  Median :2006-01-20   Mode  :character   Median :40760   Median : 67315
##  Mean   :2008-09-19                      Mean   :40766   Mean   : 93358
##  3rd Qu.:2009-01-20                      3rd Qu.:41150   3rd Qu.:123272
##  Max.   :2020-01-20                      Max.   :41700   Max.   :670496
##
##     Avg_temp          SNOW               PRCP             SNWD
##  Min.   :-2.00   Min.   :0.000000   Min.   :0.0000   Min.   : 0.000
##  1st Qu.:38.00   1st Qu.:0.000000   1st Qu.:0.0000   1st Qu.: 0.000
##  Median :52.19   Median :0.000000   Median :0.0000   Median : 0.000
##  Mean   :52.19   Mean   :0.004944   Mean   :0.0828   Mean   : 0.145
##  3rd Qu.:68.50   3rd Qu.:0.000000   3rd Qu.:0.0828   3rd Qu.: 0.000
##  Max.   :89.00   Max.   :1.200000   Max.   :1.2900   Max.   :17.000
##                                                      NA's   :3290
```

```
weatherData = cbind(weatherData[1],weatherData[3],weatherData[4],weatherData[5],weatherD
ata[6], weatherData[7])
summary(weatherData)
```

```
##       Date              station_id    number_of_trips     Avg_temp
##  Min.   :2001-01-20   Min.   :40010   Min.   :     0   Min.   :-2.00
##  1st Qu.:2003-01-20   1st Qu.:40370   1st Qu.: 33289   1st Qu.:38.00
##  Median :2006-01-20   Median :40760   Median : 67315   Median :52.19
##  Mean   :2008-09-19   Mean   :40766   Mean   : 93358   Mean   :52.19
##  3rd Qu.:2009-01-20   3rd Qu.:41150   3rd Qu.:123272   3rd Qu.:68.50
##  Max.   :2020-01-20   Max.   :41700   Max.   :670496   Max.   :89.00
##       SNOW               PRCP
##  Min.   :0.000000   Min.   :0.0000
##  1st Qu.:0.000000   1st Qu.:0.0000
##  Median :0.000000   Median :0.0000
##  Mean   :0.004944   Mean   :0.0828
##  3rd Qu.:0.000000   3rd Qu.:0.0828
##  Max.   :1.200000   Max.   :1.2900
```

```
library(mice)
```

```
##
## Attaching package: 'mice'
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
## The following objects are masked from 'package:base':
##
##     cbind, rbind
```

```
library(glmnet)
```

```
## Loading required package: Matrix
```

```
## Loaded glmnet 4.1-6
```

```
# create a simulated dataset with missing values
set.seed(123)
n <- 100
date <- sample(seq(as.Date('2001-01-20'), as.Date('2020/12/31'), by="day"), n, replace=T
RUE)
station_id <- sample(1:10, n, replace=TRUE)
num_trips <- rpois(n, 5)
avg_trip <- rnorm(n, 10, 2)
snow <- rpois(n, 3)
precipitation <- rnorm(n, 0, 1)
y <- rnorm(n, 50, 10)
data <- data.frame(date, station_id, num_trips, avg_trip, snow, precipitation, y)

# randomly insert some missing values
data[sample(1:n, n/10), "precipitation"] <- NA

# impute missing values using mice
imp <- mice(data, m = 5, maxit = 50)
```

```
##
##  iter imp variable
##   1   1  precipitation
##   1   2  precipitation
##   1   3  precipitation
##   1   4  precipitation
##   1   5  precipitation
##   2   1  precipitation
##   2   2  precipitation
##   2   3  precipitation
##   2   4  precipitation
##   2   5  precipitation
##   3   1  precipitation
##   3   2  precipitation
##   3   3  precipitation
##   3   4  precipitation
##   3   5  precipitation
##   4   1  precipitation
##   4   2  precipitation
##   4   3  precipitation
##   4   4  precipitation
##   4   5  precipitation
##   5   1  precipitation
##   5   2  precipitation
##   5   3  precipitation
##   5   4  precipitation
##   5   5  precipitation
##   6   1  precipitation
##   6   2  precipitation
##   6   3  precipitation
##   6   4  precipitation
##   6   5  precipitation
##   7   1  precipitation
##   7   2  precipitation
##   7   3  precipitation
##   7   4  precipitation
##   7   5  precipitation
##   8   1  precipitation
##   8   2  precipitation
##   8   3  precipitation
##   8   4  precipitation
##   8   5  precipitation
##   9   1  precipitation
##   9   2  precipitation
##   9   3  precipitation
##   9   4  precipitation
##   9   5  precipitation
##   10   1  precipitation
##   10   2  precipitation
##   10   3  precipitation
##   10   4  precipitation
##   10   5  precipitation
```

```
##   11   1  precipitation
##   11   2  precipitation
##   11   3  precipitation
##   11   4  precipitation
##   11   5  precipitation
##   12   1  precipitation
##   12   2  precipitation
##   12   3  precipitation
##   12   4  precipitation
##   12   5  precipitation
##   13   1  precipitation
##   13   2  precipitation
##   13   3  precipitation
##   13   4  precipitation
##   13   5  precipitation
##   14   1  precipitation
##   14   2  precipitation
##   14   3  precipitation
##   14   4  precipitation
##   14   5  precipitation
##   15   1  precipitation
##   15   2  precipitation
##   15   3  precipitation
##   15   4  precipitation
##   15   5  precipitation
##   16   1  precipitation
##   16   2  precipitation
##   16   3  precipitation
##   16   4  precipitation
##   16   5  precipitation
##   17   1  precipitation
##   17   2  precipitation
##   17   3  precipitation
##   17   4  precipitation
##   17   5  precipitation
##   18   1  precipitation
##   18   2  precipitation
##   18   3  precipitation
##   18   4  precipitation
##   18   5  precipitation
##   19   1  precipitation
##   19   2  precipitation
##   19   3  precipitation
##   19   4  precipitation
##   19   5  precipitation
##   20   1  precipitation
##   20   2  precipitation
##   20   3  precipitation
##   20   4  precipitation
##   20   5  precipitation
##   21   1  precipitation
##   21   2  precipitation
```

```
##   21   3  precipitation
##   21   4  precipitation
##   21   5  precipitation
##   22   1  precipitation
##   22   2  precipitation
##   22   3  precipitation
##   22   4  precipitation
##   22   5  precipitation
##   23   1  precipitation
##   23   2  precipitation
##   23   3  precipitation
##   23   4  precipitation
##   23   5  precipitation
##   24   1  precipitation
##   24   2  precipitation
##   24   3  precipitation
##   24   4  precipitation
##   24   5  precipitation
##   25   1  precipitation
##   25   2  precipitation
##   25   3  precipitation
##   25   4  precipitation
##   25   5  precipitation
##   26   1  precipitation
##   26   2  precipitation
##   26   3  precipitation
##   26   4  precipitation
##   26   5  precipitation
##   27   1  precipitation
##   27   2  precipitation
##   27   3  precipitation
##   27   4  precipitation
##   27   5  precipitation
##   28   1  precipitation
##   28   2  precipitation
##   28   3  precipitation
##   28   4  precipitation
##   28   5  precipitation
##   29   1  precipitation
##   29   2  precipitation
##   29   3  precipitation
##   29   4  precipitation
##   29   5  precipitation
##   30   1  precipitation
##   30   2  precipitation
##   30   3  precipitation
##   30   4  precipitation
##   30   5  precipitation
##   31   1  precipitation
##   31   2  precipitation
##   31   3  precipitation
##   31   4  precipitation
```

```
##   31   5  precipitation
##   32   1  precipitation
##   32   2  precipitation
##   32   3  precipitation
##   32   4  precipitation
##   32   5  precipitation
##   33   1  precipitation
##   33   2  precipitation
##   33   3  precipitation
##   33   4  precipitation
##   33   5  precipitation
##   34   1  precipitation
##   34   2  precipitation
##   34   3  precipitation
##   34   4  precipitation
##   34   5  precipitation
##   35   1  precipitation
##   35   2  precipitation
##   35   3  precipitation
##   35   4  precipitation
##   35   5  precipitation
##   36   1  precipitation
##   36   2  precipitation
##   36   3  precipitation
##   36   4  precipitation
##   36   5  precipitation
##   37   1  precipitation
##   37   2  precipitation
##   37   3  precipitation
##   37   4  precipitation
##   37   5  precipitation
##   38   1  precipitation
##   38   2  precipitation
##   38   3  precipitation
##   38   4  precipitation
##   38   5  precipitation
##   39   1  precipitation
##   39   2  precipitation
##   39   3  precipitation
##   39   4  precipitation
##   39   5  precipitation
##   40   1  precipitation
##   40   2  precipitation
##   40   3  precipitation
##   40   4  precipitation
##   40   5  precipitation
##   41   1  precipitation
##   41   2  precipitation
##   41   3  precipitation
##   41   4  precipitation
##   41   5  precipitation
##   42   1  precipitation
```

```
##   42   2  precipitation
##   42   3  precipitation
##   42   4  precipitation
##   42   5  precipitation
##   43   1  precipitation
##   43   2  precipitation
##   43   3  precipitation
##   43   4  precipitation
##   43   5  precipitation
##   44   1  precipitation
##   44   2  precipitation
##   44   3  precipitation
##   44   4  precipitation
##   44   5  precipitation
##   45   1  precipitation
##   45   2  precipitation
##   45   3  precipitation
##   45   4  precipitation
##   45   5  precipitation
##   46   1  precipitation
##   46   2  precipitation
##   46   3  precipitation
##   46   4  precipitation
##   46   5  precipitation
##   47   1  precipitation
##   47   2  precipitation
##   47   3  precipitation
##   47   4  precipitation
##   47   5  precipitation
##   48   1  precipitation
##   48   2  precipitation
##   48   3  precipitation
##   48   4  precipitation
##   48   5  precipitation
##   49   1  precipitation
##   49   2  precipitation
##   49   3  precipitation
##   49   4  precipitation
##   49   5  precipitation
##   50   1  precipitation
##   50   2  precipitation
##   50   3  precipitation
##   50   4  precipitation
##   50   5  precipitation
```

```
# extract the completed data
data_imputed <- complete(imp)

# split the data into training and testing sets
train_idx <- sample(1:n, n/2)
train_data <- data_imputed[train_idx,]
test_data <- data_imputed[-train_idx,]

# fit a Lasso regression model using glmnet
x_train <- model.matrix(y ~ ., data = train_data)[,-1]
y_train <- train_data$y
lasso_fit <- glmnet(x_train, y_train, alpha = 1)

# use cross-validation to choose the best lambda value
cv_fit <- cv.glmnet(x_train, y_train, alpha = 1)

# plot the cross-validation results
plot(cv_fit)
```

```
# use the selected lambda value to make predictions on the test set
x_test <- model.matrix(y ~ ., data = test_data)[,-1]
y_test <- test_data$y
lasso_pred <- predict(lasso_fit, s = cv_fit$lambda.min, newx = x_test)

# calculate the mean squared error on the test set
mse <- mean((lasso_pred - y_test)^2)


# calculate residual standard error
rss <- sum((lasso_pred - y_test)^2)
n <- length(y_test)
p <- ncol(x_test)
rse <- sqrt(rss/(n-p))

# calculate multiple R-squared
rsq <- 1 - rss/sum((y_test-mean(y_test))^2)

# calculate adjusted R-squared
adj_rsq <- 1-(rss/(n-p))/((n-1)/(n-p-1))


# calculate F statistic
f_stat <- (sum((lasso_pred-mean(y_test))^2)/p) / (rss/(n-p))
```

Ridership patterns

April 30, 2023

[ ]:

```

```

[ ]:

```
pip install kneed
```

[1]:

```
# Import necessary libraries
import pandas as pd
import numpy as np
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt
from kneed import KneeLocator

# Read the dataset from CSV file
bus_data = pd.read_csv("/Users/vaishnavishankardevadig/Downloads/
↪CTA_-_Ridership_-_Bus_Routes_-_Monthly_Day-Type_Averages___Totals-2.csv")

# Select the relevant columns for clustering
bus_cluster_data = bus_data.loc[:, ["Avg_Weekday_Rides", "Avg_Saturday_Rides",␣
↪"Avg_Sunday_Holiday_Rides"]]

# Scale the data
scaler = StandardScaler()
bus_cluster_data_scaled = scaler.fit_transform(bus_cluster_data)

# Determine the optimal number of clusters using the elbow method
sse = []
for k in range(1, 11):
    kmeans = KMeans(n_clusters=k, random_state=0)
    kmeans.fit(bus_cluster_data_scaled)
    sse.append(kmeans.inertia_)

kl = KneeLocator(range(1,11), sse, curve="convex", direction="decreasing")
optimal_k = kl.elbow

# Apply k-means clustering with the optimal number of clusters
kmeans_model = KMeans(n_clusters=optimal_k, random_state=0)
kmeans_model.fit(bus_cluster_data_scaled)
```

```
bus_data["Cluster"] = kmeans_model.labels_

# Visualize the clusters
fig, ax = plt.subplots(figsize=(10,6))
ax.scatter(bus_cluster_data_scaled[:, 0], bus_cluster_data_scaled[:, 1],␣
 ↪c=kmeans_model.labels_, cmap="rainbow")
ax.set_title("Clusters based on similar ridership patterns")
plt.show()

# Print the results
print(bus_data)
```

```
      route                        routename month_beginning  \
0         1                Indiana/Hyde Park          1/1/01
1         2                Hyde Park Express          1/1/01
2         3                       King Drive          1/1/01
3         4                    Cottage Grove          1/1/01
4         6             Jackson Park Express          1/1/01
…       …                                 …               …
35961   172            U. of Chicago/Kenwood         12/1/22
35962   192  U. of Chicago Hospitals Express         12/1/22
35963   201                     Central/Ridge         12/1/22
35964   206              Evanston Circulator         12/1/22
35965  1001      Shuttle/Special Event Route         12/1/22
```

```
       Avg_Weekday_Rides  Avg_Saturday_Rides  Avg_Sunday_Holiday_Rides  \
0                 6982.6                 0.0                       0.0
1                 1000.0                 0.0                       0.0
2                21406.5             13210.7                    8725.3
3                22432.2             17994.0                   10662.2
4                18443.0             13088.2                    7165.6
…                    …                   …                         …
35961              811.8               242.7                     211.2
35962              317.1                 0.0                       0.0
35963             1149.6               657.8                       0.0
35964              227.1                 0.0                       0.0
35965               67.3               151.5                     152.2

       MonthTotal  Cluster
0          153617        0
1           22001        0
2          567413        2
3          618796        2
4          493926        2
…           …       …
35961       19318        0
35962        6660        0
35963       27431        0
35964        4770        0
35965        2932        0

[35966 rows x 8 columns]
```

[2]:
```
for cluster_label, group in bus_data.groupby("Cluster"):
    print(f"Cluster {cluster_label}:")
    print(group.head())
```

```
Cluster 0:
  route          routename month_beginning  Avg_Weekday_Rides  \
0     1  Indiana/Hyde Park          1/1/01             6982.6
1     2  Hyde Park Express          1/1/01             1000.0
5     7           Harrison          1/1/01             5504.4
7    8A      South Halsted          1/1/01             3196.5
9    10    Museum of S & I          1/1/01                0.0

   Avg_Saturday_Rides  Avg_Sunday_Holiday_Rides  MonthTotal  Cluster
0                 0.0                       0.0      153617        0
1                 0.0                       0.0       22001        0
5                 0.0                       0.0      121097        0
7              3006.6                    1336.2       89030        0
9               562.6                     372.9        4115        0
Cluster 1:
   route       routename month_beginning  Avg_Weekday_Rides  \
```

```
11    12       Roosevelt          1/1/01            10763.5
17    21          Cermak          1/1/01             7423.7
23    28    Stony Island          1/1/01            11964.9
27    34  South Michigan          1/1/01             7911.6
34    47            47th          1/1/01             8069.8

    Avg_Saturday_Rides  Avg_Sunday_Holiday_Rides  MonthTotal  Cluster
11              6950.6                    4691.0      288055        1
17              7747.2                    4781.0      218216        1
23             11269.6                    5587.8      336246        1
27              5744.2                    3580.5      214933        1
34              6422.7                    3490.7      220680        1
Cluster 2:
  route             routename month_beginning  Avg_Weekday_Rides  \
2     3            King Drive          1/1/01            21406.5
3     4         Cottage Grove          1/1/01            22432.2
4     6  Jackson Park Express          1/1/01            18443.0
6     8               Halsted          1/1/01            19582.2
8     9               Ashland          1/1/01            29265.4

   Avg_Saturday_Rides  Avg_Sunday_Holiday_Rides  MonthTotal  Cluster
2             13210.7                    8725.3      567413        2
3             17994.0                   10662.2      618796        2
4             13088.2                    7165.6      493926        2
6             12420.0                    8280.8      521892        2
8             22621.7                   15336.1      811006        2
```

[3]:
```python
from sklearn.metrics import silhouette_score, calinski_harabasz_score,␣
↪davies_bouldin_score

# Silhouette score
silhouette_avg = silhouette_score(bus_cluster_data_scaled, kmeans_model.labels_)
print("Silhouette score:", silhouette_avg)

# Calinski-Harabasz score
ch_score = calinski_harabasz_score(bus_cluster_data_scaled, kmeans_model.
↪labels_)
print("Calinski-Harabasz score:", ch_score)

# Davies-Bouldin score
db_score = davies_bouldin_score(bus_cluster_data_scaled, kmeans_model.labels_)
print("Davies-Bouldin score:", db_score)
```

```
Silhouette score: 0.6333605652308159
Calinski-Harabasz score: 104757.73820429464
Davies-Bouldin score: 0.5665584420999883
```

[ ]:
```python
#This is just a test Example
```

[6]:
```python
print("Number of clusters:", optimal_k)
print(bus_data[["route", "Cluster"]])
```

```
Number of clusters: 3
      route  Cluster
0         1        0
1         2        0
2         3        2
3         4        2
4         6        2
…       …        …
35961   172        0
35962   192        0
35963   201        0
35964   206        0
35965  1001        0

[35966 rows x 2 columns]
```

[ ]:
```python

```

R Notebook

This is an R Markdown (http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Cmd+Shift+Enter*.

```
library(tidyverse)
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
## ✔ dplyr     1.1.0     ✔ readr     2.1.4
## ✔ forcats   1.0.0     ✔ stringr   1.5.0
## ✔ ggplot2   3.4.2     ✔ tibble    3.1.8
## ✔ lubridate 1.9.2     ✔ tidyr     1.3.0
## ✔ purrr     1.0.1
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the ]8;;http://conflicted.r-lib.org/conflicted package]8;; to force all
conflicts to become errors
```

```
transit_data <- read.csv("/Users/vaishnavishankardevadig/Downloads/CTA_-_Ridership_-_
Bus_Routes_-_Monthly_Day-Type_Averages___Totals-2.csv")
transit_data <- transit_data %>% filter(!is.na(Avg_Weekday_Rides) & !is.na(Avg_Saturd
ay_Rides) & !is.na(Avg_Sunday_Holiday_Rides))
create_bar_graph <- function(route_name) {
  route_data <- transit_data %>% filter(routename == route_name)
  route_data_long <- route_data %>% pivot_longer(cols = c("Avg_Weekday_Rides", "Avg_S
aturday_Rides", "Avg_Sunday_Holiday_Rides"), names_to = "day_type", values_to = "avg_
ridership")
  ggplot(route_data_long, aes(x = day_type, y = avg_ridership, fill = day_type)) +
    geom_col() +
    labs(x = "Day Type", y = "Average Ridership", fill = "Day Type", title = paste("A
verage Ridership for", route_name))
}

unique_stations <- unique(transit_data$routename)
for (station in unique_stations) {
  graph <- create_bar_graph(station)
  print(graph)
}
```

Average Ridership for Indiana/Hyde Park
Average Ridership
4e+05
2e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Hyde Park Express
Average Ridership
6e+05
4e+05
2e+05
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Average Ridership for King Drive
5e+06
4e+06
3e+06
2e+06
1e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Cottage Grove
Average Ridership
4e+06
2e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Jackson Park Express
Average Ridership
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Harrison
1500000
1000000
500000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Halsted
5e+06
4e+06
3e+06
2e+06
1e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for South Halsted
8e+05
6e+05
4e+05
2e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Ashland
6e+06
4e+06
2e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Museum of S & I
200000
150000
100000
50000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Lincoln/Sedgwick
Average Ridership
6e+05
4e+05
2e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Roosevelt
3e+06
2e+06
1e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Jeffery Express
Average Ridership
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Westchester
Average Ridership
60000
40000
20000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 16th/18th
Average Ridership
6e+05
4e+05
2e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for United Center Express
Average Ridership
60000
40000
20000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Madison
5e+06
4e+06
3e+06
2e+06
1e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Cermak
Average Ridership
2000000
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Cermak Express
Average Ridership
60000
40000
20000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Clark
5e+06
4e+06
3e+06
2e+06
1e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Wentworth
8e+05
6e+05
4e+05
2e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for West Cermak
30000
20000
10000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for South Deering
Average Ridership
150000
100000
50000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Stony Island
Average Ridership
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for State
Average Ridership
3e+06
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for South Chicago
8e+05
6e+05
4e+05
2e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Mag Mile Express
Average Ridership
75000
50000
25000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for South Michigan
Average Ridership
1e+06
5e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 35th
8e+05
6e+05
4e+05
2e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Broadway
Average Ridership
3e+06
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Sedgwick
Average Ridership
3e+05
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Pershing
5e+05
4e+05
3e+05
2e+05
1e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 43rd
4e+05
3e+05
2e+05
1e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Wallace-Racine
Average Ridership
1250000
1000000
750000
500000
250000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 47th
Average Ridership
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for South Damen
Average Ridership
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Western
5e+06
4e+06
3e+06
2e+06
1e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for South Western
Average Ridership
75000
50000
25000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for North Western
Average Ridership
1e+06
5e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Western Express
Average Ridership
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Damen
Average Ridership
2000000
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 51st
Average Ridership
4e+05
2e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Kedzie/California
Average Ridership
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for South Kedzie
Average Ridership
1200000
800000
400000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Pulaski
5e+06
4e+06
3e+06
2e+06
1e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for South Pulaski
Average Ridership
2000000
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Cicero
Average Ridership
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for North Cicero/Skokie Blvd.
250000
200000
150000
100000
50000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for South Cicero
Average Ridership
750000
500000
250000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Garfield
Average Ridership
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 55th/Narragansett
Average Ridership
150000
100000
50000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Milwaukee
Average Ridership
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for North Milwaukee
Average Ridership
90000
60000
30000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Laramie
Average Ridership
6e+05
4e+05
2e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 59th/61st
Average Ridership
750000
500000
250000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Blue Island/26th
3e+06
2e+06
1e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Archer
3e+06
2e+06
1e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Archer/Harlem
3e+05
2e+05
1e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 63rd
Average Ridership
4e+06
3e+06
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for West 63rd
Average Ridership
4e+05
3e+05
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Foster-Canfield
Average Ridership
20000
10000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Grand
Average Ridership
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Chicago
Average Ridership
4e+06
2e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 67th-69th-71st
Average Ridership
3e+06
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Northwest Highway
Average Ridership
3e+05
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Cumberland/East River
Average Ridership
60000
40000
20000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Division
2500000
2000000
1500000
1000000
500000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 71st/South Shore
Average Ridership
2000000
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for North
Average Ridership
3e+06
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Armitage
Average Ridership
1e+06
5e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Fullerton
3e+06
2e+06
1e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 74th-75th
Average Ridership
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Diversey
3e+06
2e+06
1e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Belmont
Average Ridership
5e+06
4e+06
3e+06
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Montrose
Average Ridership
2000000
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 79th
Average Ridership
6e+06
4e+06
2e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Irving Park
3e+06
2e+06
1e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Lawrence
Average Ridership
3e+06
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for West Lawrence
4e+05
3e+05
2e+05
1e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Kimball-Homan
Average Ridership
4e+06
3e+06
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Peterson
1000000
750000
500000
250000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Central
Average Ridership
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for North Central
Average Ridership
200000
150000
100000
50000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Narragansett/Ridgeland
Average Ridership
4e+05
2e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 87th
Average Ridership
3e+06
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Higgins
Average Ridership
3e+05
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Harlem
1250000
1000000
750000
500000
250000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for North Harlem
60000
40000
20000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Austin
2000000
1500000
1000000
500000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Foster
Average Ridership
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for California/Dodge
8e+05
6e+05
4e+05
2e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for South California
Average Ridership
2000000
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 93rd-95th
Average Ridership
750000
500000
250000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for West 95th
8e+05
6e+05
4e+05
2e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Lunt
Average Ridership
200000
150000
100000
50000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Skokie
8e+05
6e+05
4e+05
2e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Jeffery Manor Express
Average Ridership
200000
150000
100000
50000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for West 103rd
8e+05
6e+05
4e+05
2e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for East 103rd
Average Ridership
4e+05
3e+05
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Halsted/95th
5e+05
4e+05
3e+05
2e+05
1e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Pullman/111th/115th
Average Ridership
750000
500000
250000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Vincennes/111th
Average Ridership
6e+05
4e+05
2e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Michigan/119th
Average Ridership
1e+06
5e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Ogilvie/Wacker Express
200000
150000
100000
50000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Union/Wacker Express
Average Ridership
200000
150000
100000
50000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Illinois Center/Ogilvie Express
100000
75000
50000
25000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Illinois Center/Union Express
Average Ridership
100000
75000
50000
25000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Water Tower Express
Average Ridership
4e+05
3e+05
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Jackson
Average Ridership
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Madison/Roosevelt Circulator
Average Ridership
60000
40000
20000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for West Loop/South Loop
Average Ridership
1e+05
5e+04
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Clarendon/LaSalle Express
8e+05
6e+05
4e+05
2e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Sheridan/LaSalle Express
5e+05
4e+05
3e+05
2e+05
1e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Wilson/Michigan Express
1000000
750000
500000
250000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Inner Drive/Michigan Express
Average Ridership
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Outer Drive Express
3e+06
2e+06
1e+06
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Sheridan
Average Ridership
4e+06
3e+06
2e+06
1e+06
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Addison
Average Ridership
2500000
2000000
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Devon
Average Ridership
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for LaSalle
2000000
1500000
1000000
500000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Streeterville/Taylor
Average Ridership
9e+05
6e+05
3e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 69th-UPS Express
Average Ridership
60000
40000
20000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for U. of Chicago/Midway
60000
40000
20000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for U. of Chicago/Hyde Park
3e+05
2e+05
1e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for U. of Chicago/Kenwood
Average Ridership
3e+05
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for U. of Chicago/Lakeview Express
Average Ridership
15000
10000
5000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Central/Ridge
Average Ridership
4e+05
3e+05
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Main/Emerson
Average Ridership
10000
5000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Ridge/Grant
Average Ridership
40000
30000
20000
10000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Dodge
Average Ridership
40000
30000
20000
10000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Soldier Field Express
100000
75000
50000
25000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Navy Pier
Average Ridership
3e+05
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Wrigley Field Express
Average Ridership
75000
50000
25000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Museum Campus
Average Ridership
90000
60000
30000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for UIC-Pilsen Express
8000
6000
4000
2000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Irving Park Express
3e+05
2e+05
1e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Avon Express
Average Ridership
30000
20000
10000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Garfield Express
Average Ridership
200000
150000
100000
50000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 69th Bus Pre-Paid Area
Average Ridership
5000
4000
3000
2000
1000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for South Shore Night Bus
Average Ridership
75000
50000
25000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Main Shuttle
Average Ridership
9000
6000
3000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Chicago/Golf
Average Ridership
150000
100000
50000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Evanston Circulator
Average Ridership
150000
100000
50000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Jeffery Local
Average Ridership
1500000
1000000
500000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for South Shore Express
Average Ridership
6e+05
4e+05
2e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Stockton/LaSalle Express
Average Ridership
4e+05
2e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Stockton/Michigan Express
3e+05
2e+05
1e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Marine/Michigan Express
125000
100000
75000
50000
25000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Shuttle/Special Event Route
Average Ridership
6e+05
4e+05
2e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Clarendon/Michigan Express
Average Ridership
4e+05
3e+05
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Chinatown/Pilsen Shuttle
Average Ridership
4000
2000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Chicago Manufacturing Campus
Average Ridership
2000
1000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Stony Island Express
Average Ridership
3e+05
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for South Pulaski Limited
Average Ridership
30000
20000
10000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for King Drive Express
Average Ridership
90000
60000
30000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Cottage Grove Express
Average Ridership
120000
80000
40000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for U. of Chicago Hospitals Express
Average Ridership
1e+05
5e+04
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Special Event Route
Average Ridership
40000
20000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Ashland Express
1000000
750000
500000
250000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Washington/Madison Express
100000
75000
50000
25000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Ogden/Taylor
Average Ridership
60000
40000
20000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Cicero Express
250000
200000
150000
100000
50000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for U. of Chicago/Garfield Stations
Average Ridership
10000
5000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 55th/Austin
Average Ridership
40000
30000
20000
10000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Goose Island Express
Average Ridership
40000
30000
20000
10000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for West 65th
Average Ridership
20000
15000
10000
5000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Touhy Supplement
Average Ridership
150
100
50
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Central/Sherman
2500
2000
1500
1000
500
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 31st/35th
Average Ridership
4e+05
2e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Bronzeville/Union Station
Average Ridership
150000
100000
50000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Lincoln
Average Ridership
150000
100000
50000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 111th/King Drive
Average Ridership
3e+05
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Ogilvie/Streeterville Express
Average Ridership
80000
60000
40000
20000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Union/Streeterville Express
120000
90000
60000
30000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Jeffery Jump
1200000
900000
600000
300000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Pullman/115th
Average Ridership
3e+05
2e+05
1e+05
0e+00
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Cermak-Roosevelt Express
20000
15000
10000
5000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Dan Ryan OWL Shuttle
Average Ridership
5000
4000
3000
2000
1000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Dan Ryan Local Shuttle
Average Ridership
40000
30000
20000
10000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 69th-Garfield Express Shuttle
Average Ridership
10000
7500
5000
2500
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 79th-Garfield Express Shuttle
Average Ridership
20000
15000
10000
5000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 87th-Garfield Express Shuttle
Average Ridership
15000
10000
5000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 95th-Garfield Express Shuttle
Average Ridership
60000
40000
20000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Pershing Shuttle
400
300
200
100
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Pullman Shuttle
20000
15000
10000
5000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 31st
30000
20000
10000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for 95th
3e+05
2e+05
1e+05
0e+00
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for ROAD CALL
Average Ridership
2.0
1.5
1.0
0.5
0.0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Special Dest Signs
Average Ridership
3
2
1
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Kedzie
150000
100000
50000
0
Average Ridership
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for California
Average Ridership
150000
100000
50000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for O'Hare Shuttle
Average Ridership
8
6
4
2
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

Average Ridership for Inner Lake Shore/Michigan Express
Average Ridership
90000
60000
30000
0
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides
Day Type
Day Type
Avg_Saturday_Rides
Avg_Sunday_Holiday_Rides
Avg_Weekday_Rides

```
summary(transit_data)
```

```
##     route            routename         month_beginning    Avg_Weekday_Rides
##  Length:35966       Length:35966       Length:35966       Min.   :    0
##  Class :character   Class :character   Class :character   1st Qu.: 1239
##  Mode  :character   Mode  :character   Mode  :character   Median : 3732
##                                                           Mean   : 6246
##                                                           3rd Qu.: 9569
##                                                           Max.   :37787
##  Avg_Saturday_Rides Avg_Sunday_Holiday_Rides   MonthTotal
##  Min.   :    0      Min.   :    0            Min.   :      1
##  1st Qu.:    0      1st Qu.:    0            1st Qu.:  28338
##  Median : 1796      Median : 1122            Median :  92612
##  Mean   : 3961      Mean   : 2771            Mean   : 163017
##  3rd Qu.: 6106      3rd Qu.: 4357            3rd Qu.: 248763
##  Max.   :30645      Max.   :24111            Max.   :1058879
```