Uber data plots

Daniel Oyeleye

2023-10-24

## R Markdown

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)

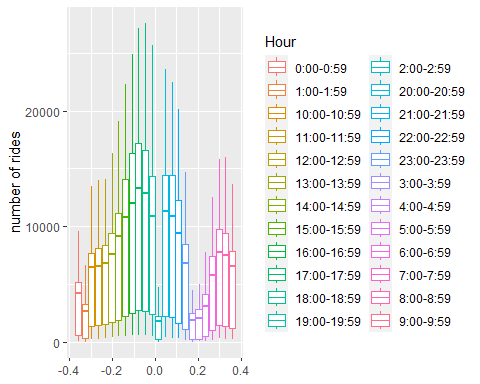
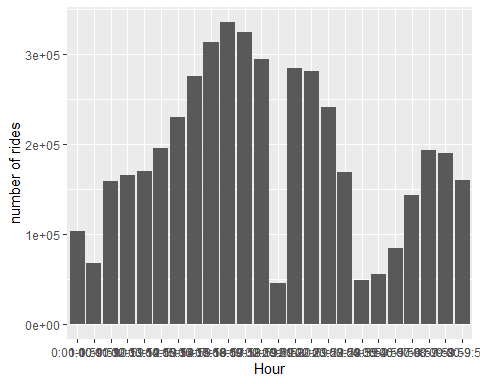
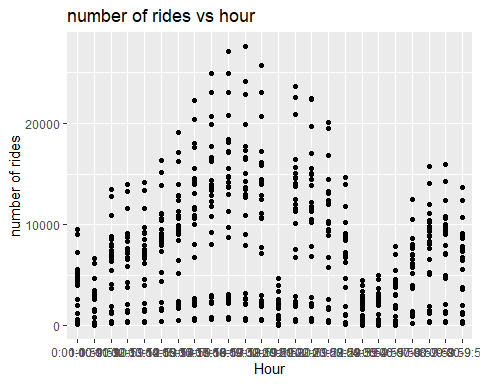
## Including Plots

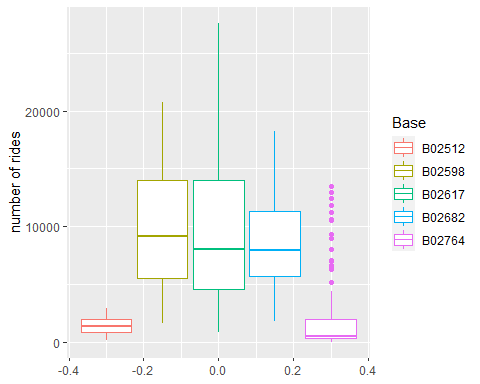
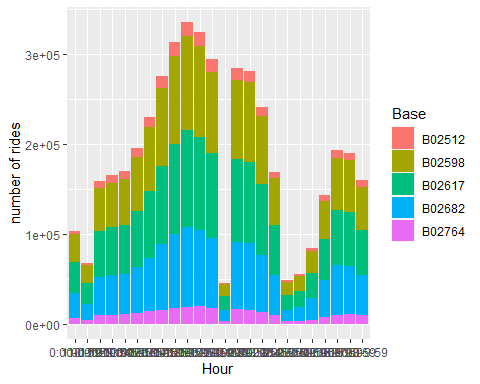
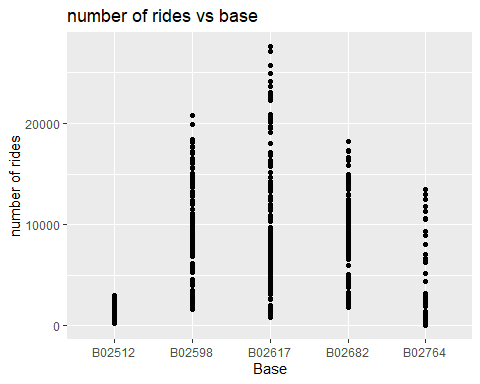
You can also embed plots, for example:

Returns a data frame with the variables total number of rides for each hour of the day, base and month

rides\_per\_hour\_base\_month <- function(df){  
 return (df %>% group\_by(Hour, Base, Month) %>% dplyr:: summarise(num\_rides = n()))  
}

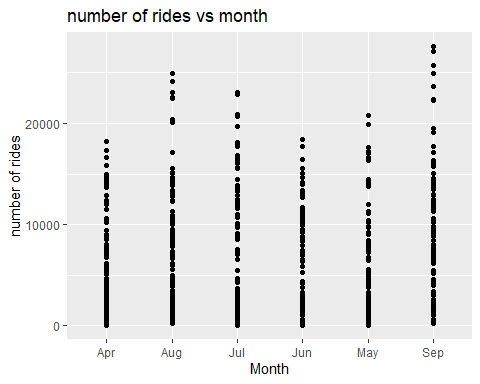
april\_sep14 <- read.csv("C:\\Users\\danie\\Documents\\projects\\data projects\\uber data\\uber\_april\_sep14.csv")  
rides\_hbm <- rides\_per\_hour\_base\_month(april\_sep14)

Scatter plot, box plot and histogram of number of rides per hour of day 

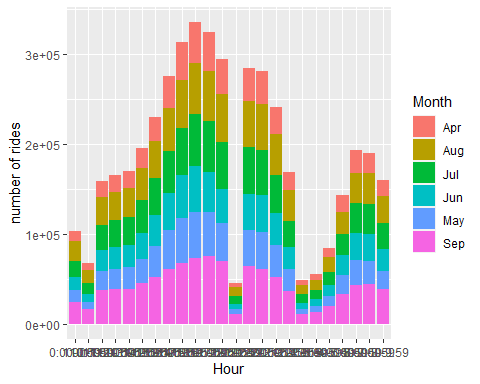
Histogram, scatter plot and box plot of number of rides per base 

Histogram, scatter plot and box plot of number of rides per month

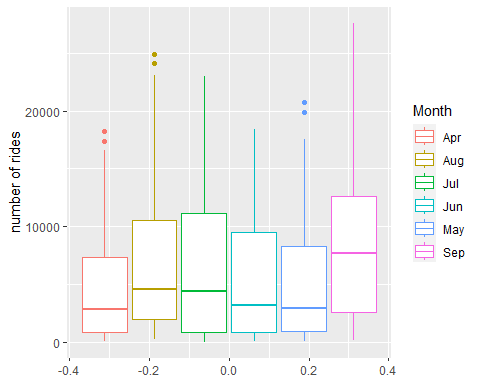
ggplot(data=rides\_hbm) +  
 geom\_point(mapping = aes(x=Month, y = num\_rides)) + labs(title = "number of rides vs month", y="number of rides")

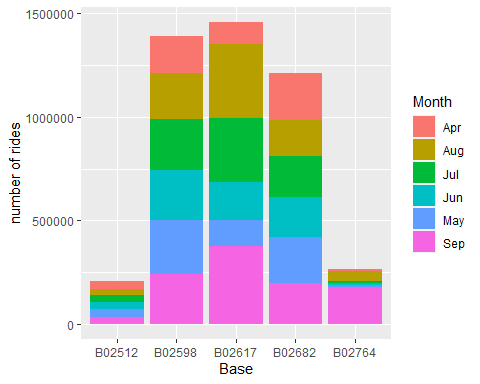


ggplot(data=april\_sep14) +  
 geom\_bar(mapping = aes(x=Hour, fill=Month)) + labs(y = "number of rides")

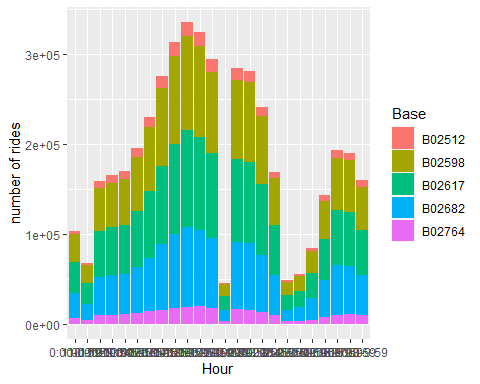


ggplot(data=rides\_hbm) +  
 geom\_boxplot(mapping = aes(y=num\_rides, color=Month)) + labs(y = "number of rides")



Histogram of number of rides per base and month 

Histogram of number of rides per hour of day, base and month 

Histogram of number of rides per hour of day and base 

Histogram of number of rides per hour of day and month 