PSTAT 274 Project

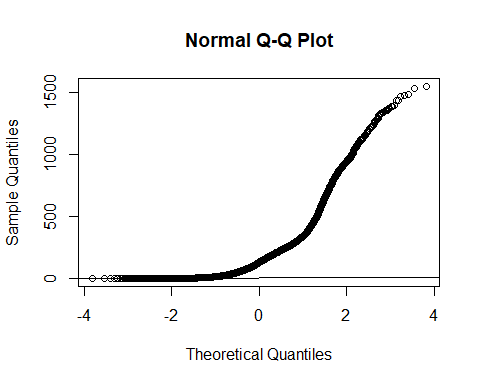
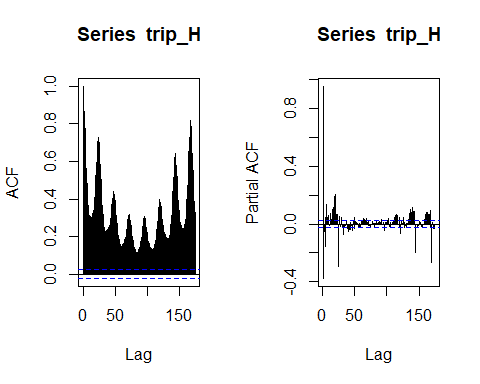
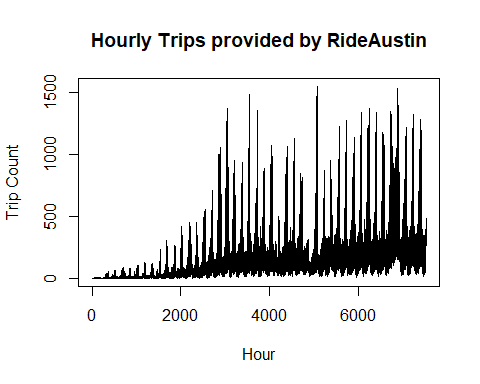
Carlos Baez

November 20, 2018

#Generate CSV to Explore with Kepler.GL  
longrides <- filter(Rides\_DataA, distance\_travelled >= 50000) %>% dplyr::select(start\_location\_lat, start\_location\_long, end\_location\_lat, end\_location\_long, distance\_travelled)  
write.csv(longrides, "longrides.csv")  
  
#Filter out data with more than 100,000 m in distance\_travelled  
Trip\_Data <-filter(Rides\_DataA, distance\_travelled <= 100000) %>% dplyr::select(start\_location\_lat, start\_location\_long, end\_location\_lat, end\_location\_long, distance\_travelled, started\_on, completed\_on)  
rm(Rides\_DataA)

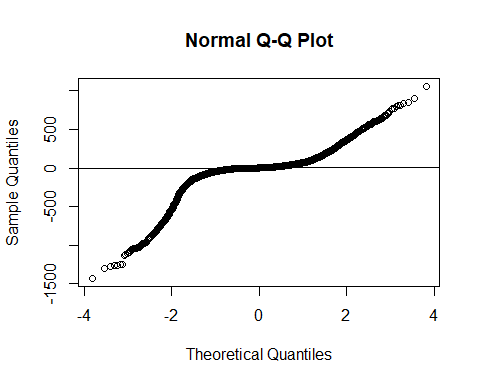
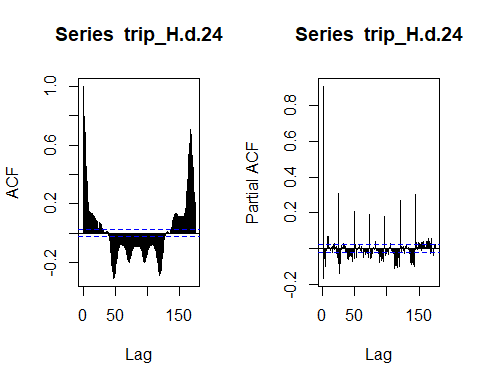
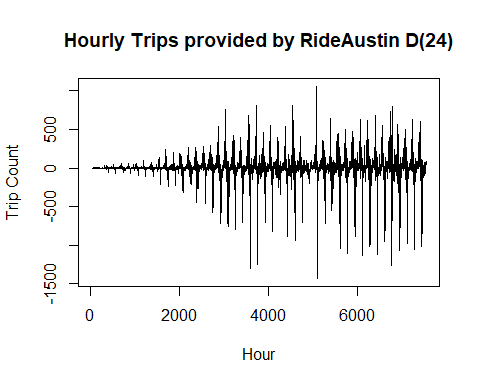
#Hourly   
Trip\_hourly <- Trip\_Data %>% mutate(day = as.numeric(date(started\_on)), wday = wday(started\_on), hour = hour(started\_on)) %>% group\_by(day, wday, hour) %>% summarize(count = n())  
  
#Get Hours where count is zero  
full\_dates <- cbind(rep(16956:17269,each=24),seq(0,23))  
full\_dates <-as.data.frame(full\_dates)  
full\_dates <- full\_dates %>% mutate(wday = wday(as\_date(V1)))  
Trip\_hourly <- full\_join(Trip\_hourly,full\_dates, by = c("day" = "V1", "wday" ="wday", "hour"= "V2"))  
Trip\_hourly <- arrange(Trip\_hourly, day, hour) %>% mutate(count = ifelse(is.na(count), 0, count))  
write.csv(Trip\_hourly,"trip\_hourly.csv")  
rm(full\_dates)  
  
#Make time series  
trip\_H <- ts(Trip\_hourly)[,4]

# Exploring Data

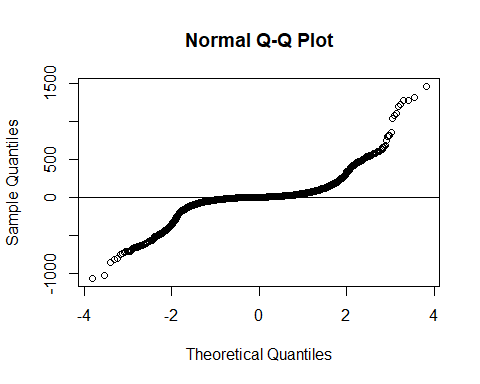
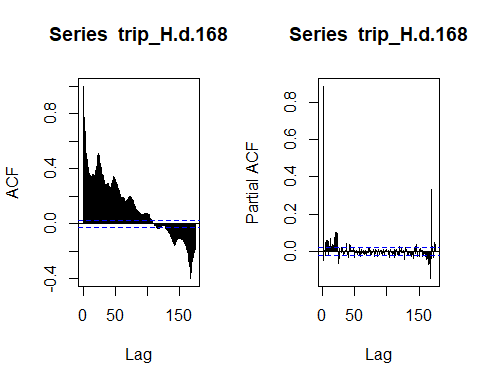
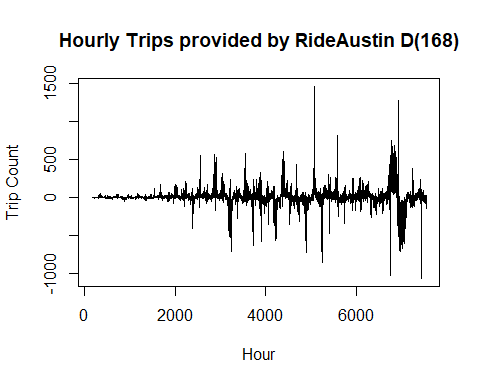


## Differencing

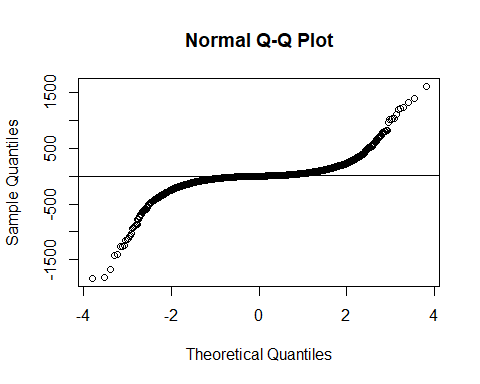
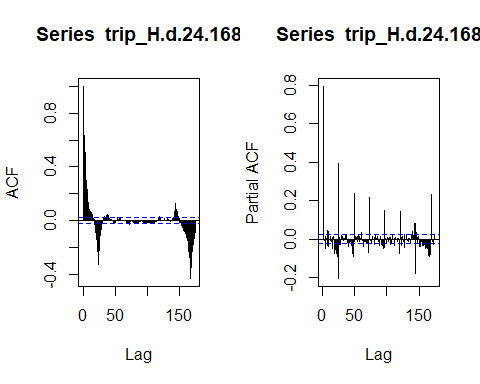
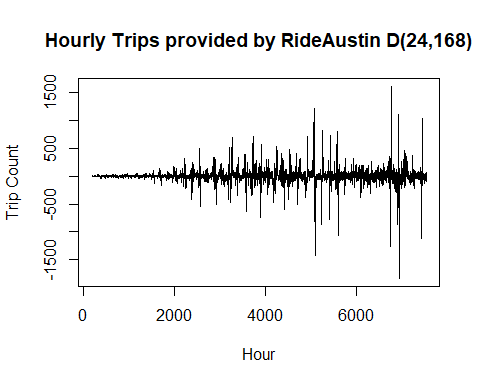
### 24 Hour Differencing



### 168 Hour Differencing

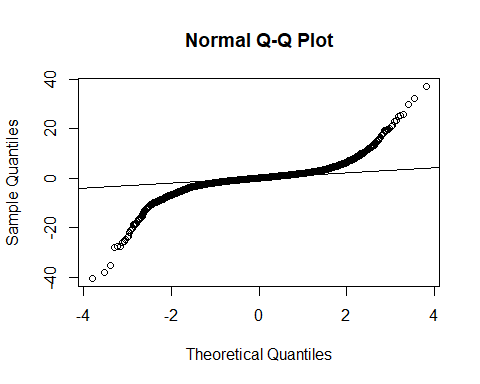
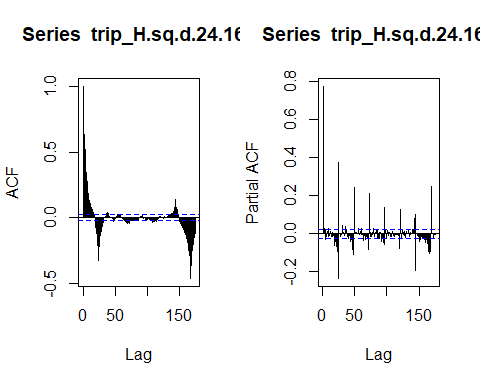
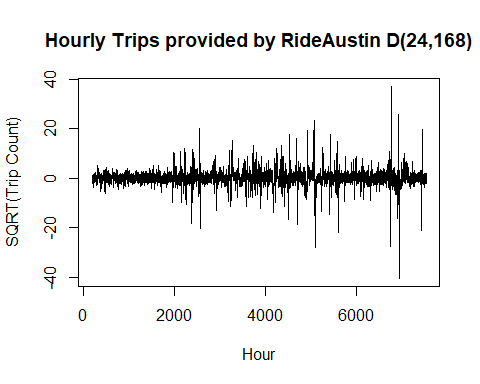
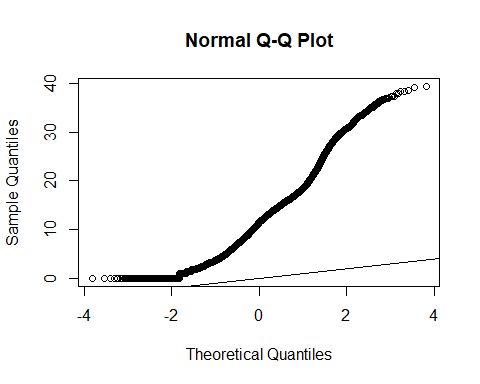
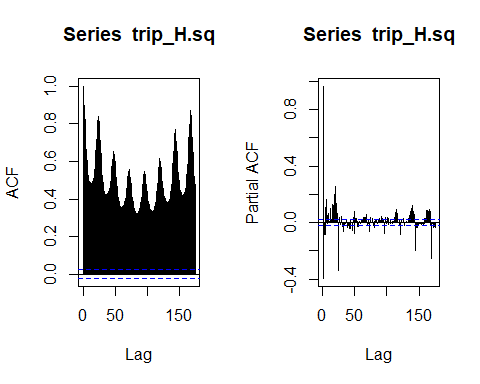
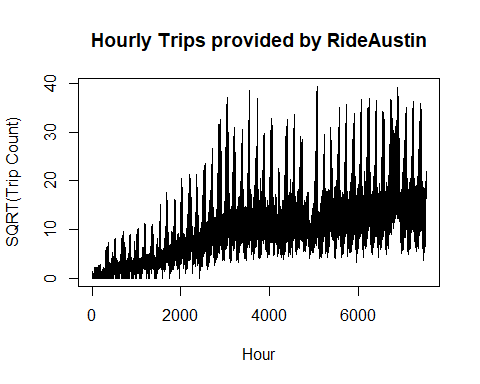


### 24 and 168 hour differencing



## Transformations

### Square Root Transofrmation

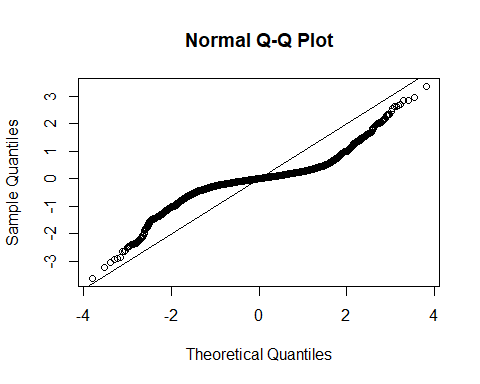
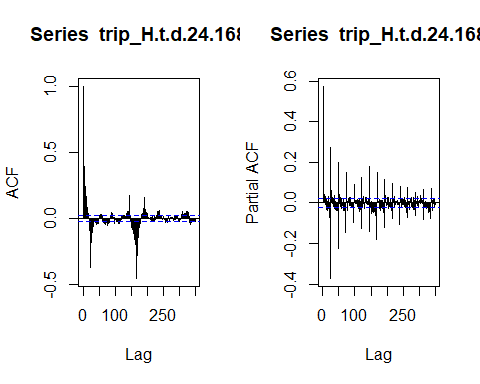
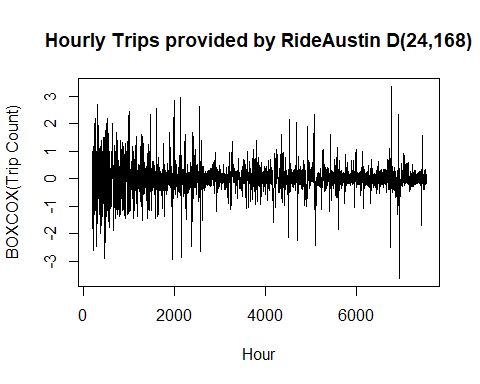
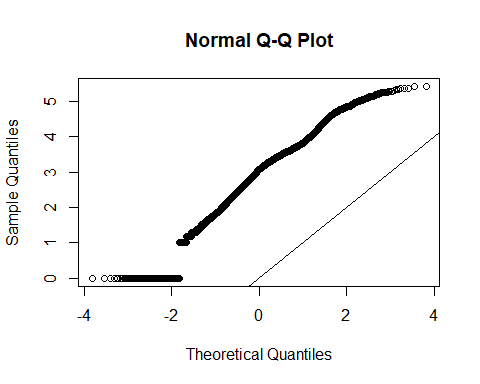
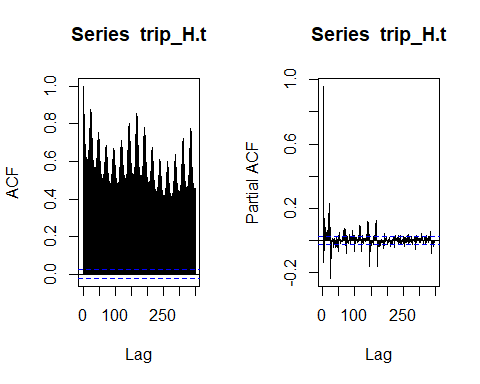
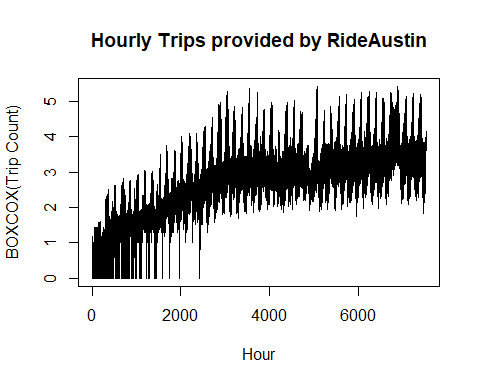


Before differencing:

After Differencing:

### Box-Cox Transformation (lambda = 0.23)

## yjPower Transformation to Normality   
## Est Power Rounded Pwr Wald Lwr Bnd Wald Upr Bnd  
## Y1 0.2314 0.23 0.2181 0.2447  
##   
## Likelihood ratio test that transformation parameter is equal to 0  
## LRT df pval  
## LR test, lambda = (0) 1243.934 1 < 2.22e-16



Before Differencing:

After Differencing: