Stat 508 - Final Project

```
suppressMessages(library(lubridate))
suppressMessages(library(caret))
suppressMessages(library(corrplot))
suppressMessages(library(sugrrants))
suppressMessages(library(dplyr))
```

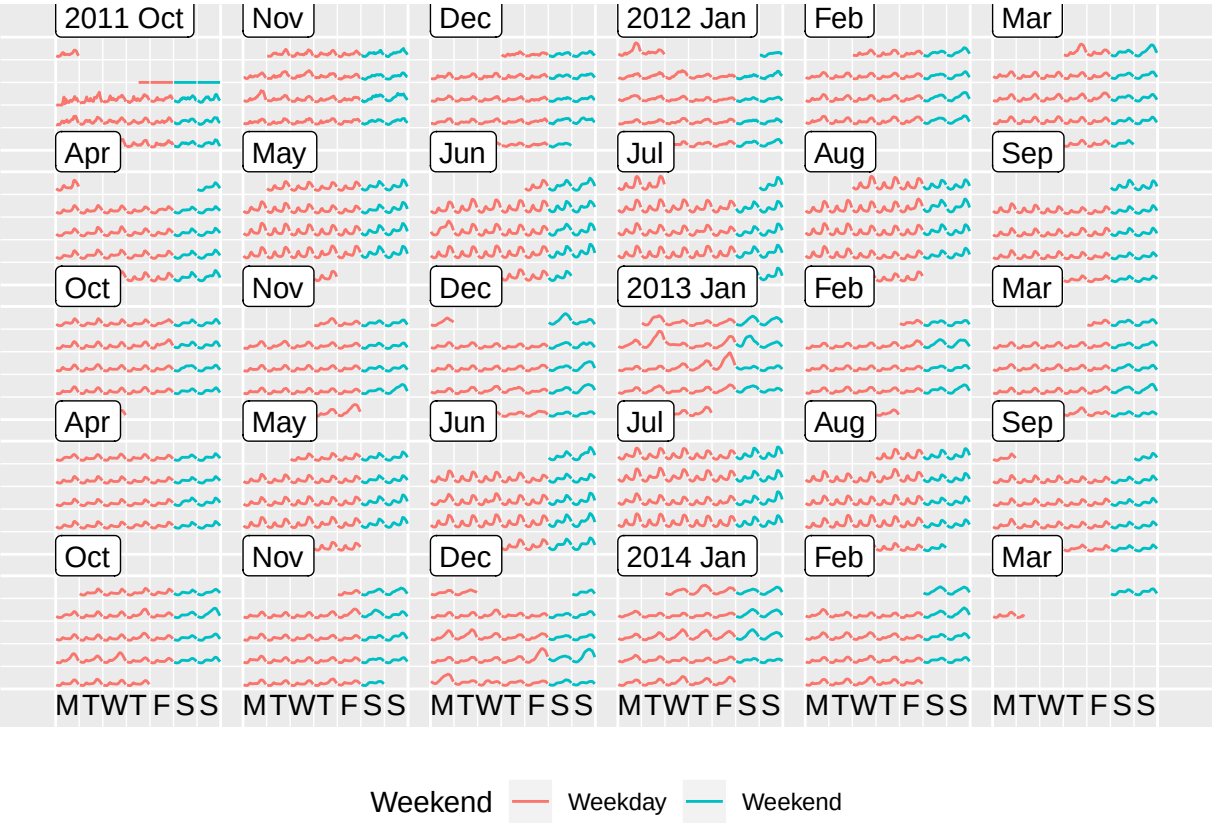
Introduction

Data

```
EnergyDataAggregate <- read.csv("EnergyDataAggregate.csv",stringsAsFactors = FALSE)
EnergyDataAggregate$DATE <- as.Date(EnergyDataAggregate$DATE,format="%m/%d/%Y")
```

Temporal Trend

```
p <- EnergyDataAggregate %>%  mutate(Weekend = if_else(DAYNAME %in% c("Saturday", "Sunday"), "Weekend",
  frame_calendar(x = HOUR, y = GENERAL_SUPPLY_KWH, date = DATE) %>%
  ggplot(aes(x = .HOUR, y = .GENERAL_SUPPLY_KWH, group = DATE, colour = Weekend)) +
  geom_line() +
  theme(legend.position = "bottom")
prettify(p)
```



Analysis

```
EnergyDataAggregate$HighEnergyUse <- ifelse(EnergyDataAggregate$GENERAL_SUPPLY_KWH>0.5,1,0)

set.seed(1)

#training and test set
trainIndex <- createDataPartition(EnergyDataAggregate$GENERAL_SUPPLY_KWH, p = .5, list = FALSE, times =
energyData.train=EnergyDataAggregate[trainIndex,]
energyData.test=EnergyDataAggregate[-trainIndex,]
```