

# weather.Rmd

Collin Fabian

10/27/2021

Libraries

Load and clean data as needed

```
df <- read.csv('/Users/fabian_coll/Desktop/Pure-Solutions/pure_dash/data/weather/rain_rev_21.csv')
df$DayOfWeekNum <- as.factor(df$DayOfWeekNum)
df$DoneDateFormatted <- as.Date(df$DoneDateFormatted, format= "%Y-%m-%d")
summary(df)
```

```
## DoneDateFormatted    Precipitation      AllRev      TickRev
## Min.      :2021-03-01  Min.      :0.0000  Min.      :    0  Min.      :    0.0
## 1st Qu.:2021-04-29  1st Qu.:0.0000  1st Qu.: 1276  1st Qu.:   364.5
## Median :2021-06-28  Median :0.0000  Median :21645  Median :13763.5
## Mean    :2021-06-28  Mean    :0.1943  Mean    :16733  Mean    :10372.4
## 3rd Qu.:2021-08-27  3rd Qu.:0.1025  3rd Qu.:28080  3rd Qu.:17228.8
## Max.    :2021-10-26  Max.    :4.7500  Max.    :38711  Max.    :25145.0
```

```
##
## TurfRev      DayOfWeekNum  DayOfWeekName
## Min.      :    0  0:35      Friday   :34
## 1st Qu.:    0  1:35      Monday    :35
## Median : 6208  2:34      Saturday :34
## Mean    : 5693  3:34      Sunday    :34
## 3rd Qu.: 9534  4:34      Thursday  :34
## Max.    :23034  5:34      Tuesday   :35
##                      6:34      Wednesday:34
```

```
## [1] "Monday"    "Tuesday"   "Wednesday" "Thursday"  "Friday"    "Saturday"
## [7] "Sunday"
```

Filter data into separate dataframes based on program start-date. Also relevel DaysOfWeekName column so that groups are graphed in order by day.

```
df.turf <- df %>%
  filter(DoneDateFormatted >= "2021-03-31")

df.turf$DayOfWeekName <- factor(
  df.turf$DayOfWeekName,
  levels = c('Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday', 'Saturday', 'Sunday')
)

df.tick <- df %>%
```

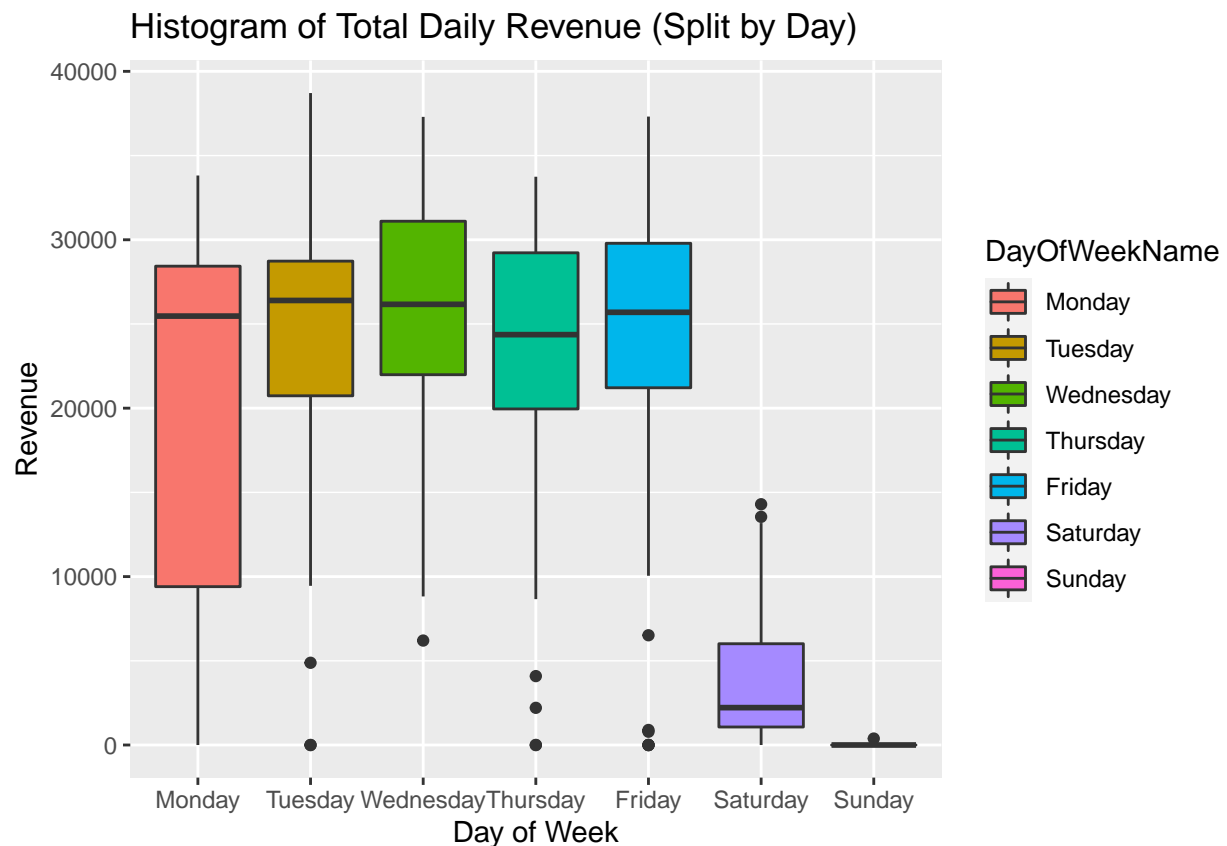
```

filter(DoneDateFormatted >= "2021-03-01")

df.tick$DayOfWeekName <- factor(
  df.tick$DayOfWeekName,
  levels = c('Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday', 'Saturday', 'Sunday')
)

hist.all <- ggplot(df, aes(x=DayOfWeekName, y=AllRev, group=DayOfWeekName)) +
  geom_boxplot(aes(fill=DayOfWeekName)) +
  labs(x='Day of Week', y='Revenue', title = 'Histogram of Total Daily Revenue (Split by Day)')
hist.all

```



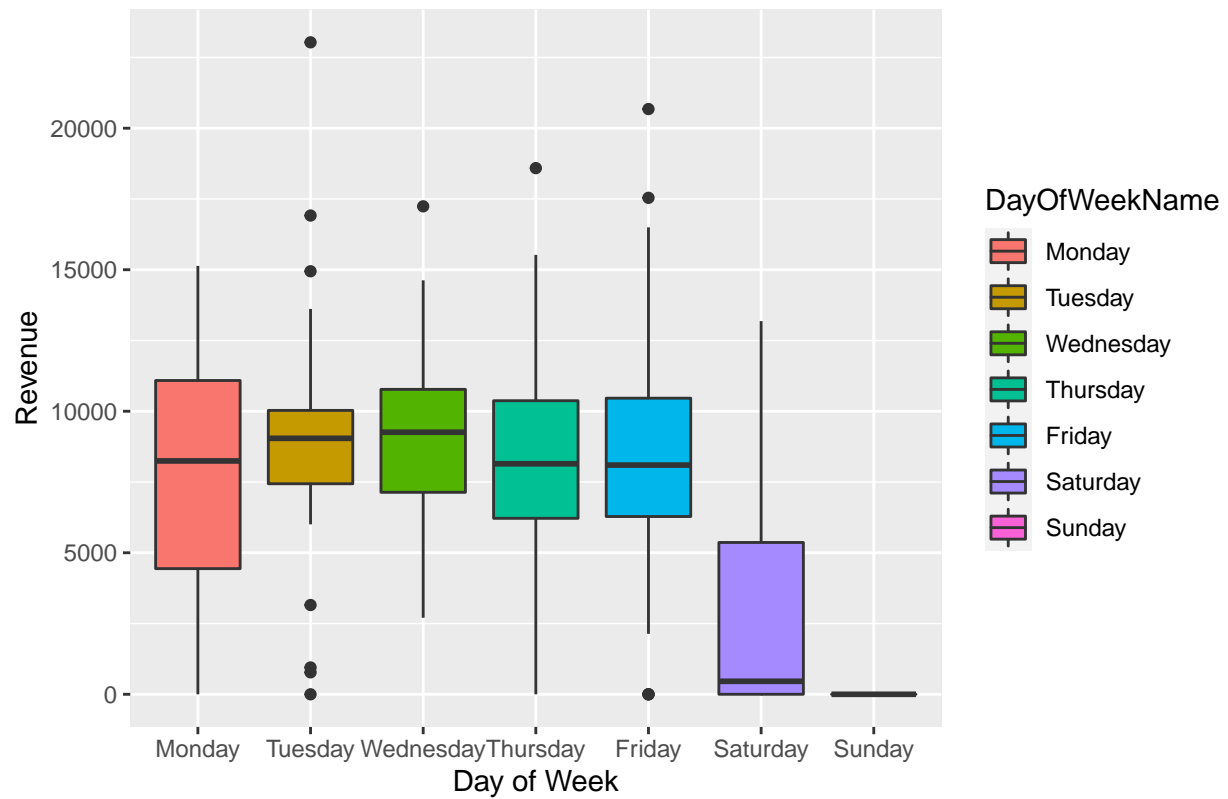
Relevel DaysOfWeekName column (as to be graphed in order)

```

hist.turf <- ggplot(df.turf, aes(x=DayOfWeekName, y=TurfRev, group=DayOfWeekName)) +
  geom_boxplot(aes(fill=DayOfWeekName)) +
  labs(x='Day of Week', y='Revenue', title = 'Histogram of Turf Daily Revenue (Split by Day)')
hist.turf

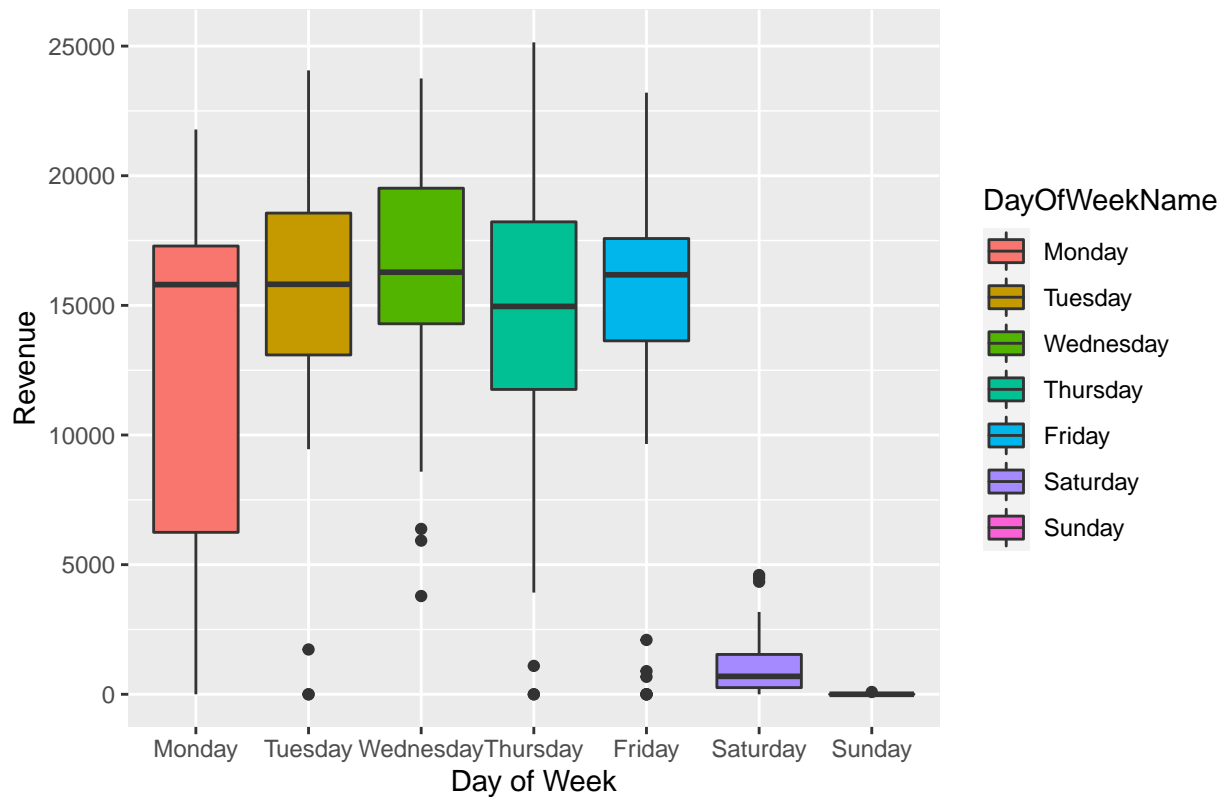
```

Histogram of Turf Daily Revenue (Split by Day)



```
hist.tick <- ggplot(df, aes(x=DayOfWeekName, y=TickRev, group=DayOfWeekName)) +  
  geom_boxplot(aes(fill=DayOfWeekName)) +  
  labs(x='Day of Week', y='Revenue', title = 'Histogram of Tick Daily Revenue (Split by Day)')  
hist.tick
```

### Histogram of Tick Daily Revenue (Split by Day)



```
for (day in levels(df$DayOfWeekName)) {
  temp.df <- df %>% filter(DayOfWeekName == day)
  print(day)
  print(summary(temp.df$AllRev))
}
```

```
## [1] "Monday"
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0   9403   25468   19934   28432   33814
## [1] "Tuesday"
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0  20735   26396   23740   28730   38711
## [1] "Wednesday"
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  6202  21988   26166   25160   31099   37297
## [1] "Thursday"
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0  19954   24366   22026   29224   33743
## [1] "Friday"
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0  21212   25692   21859   29789   37315
## [1] "Saturday"
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0   1069   2220   4104   6012   14296
## [1] "Sunday"
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

##	0.00	0.00	0.00	11.26	0.00	383.00
----	------	------	------	-------	------	--------