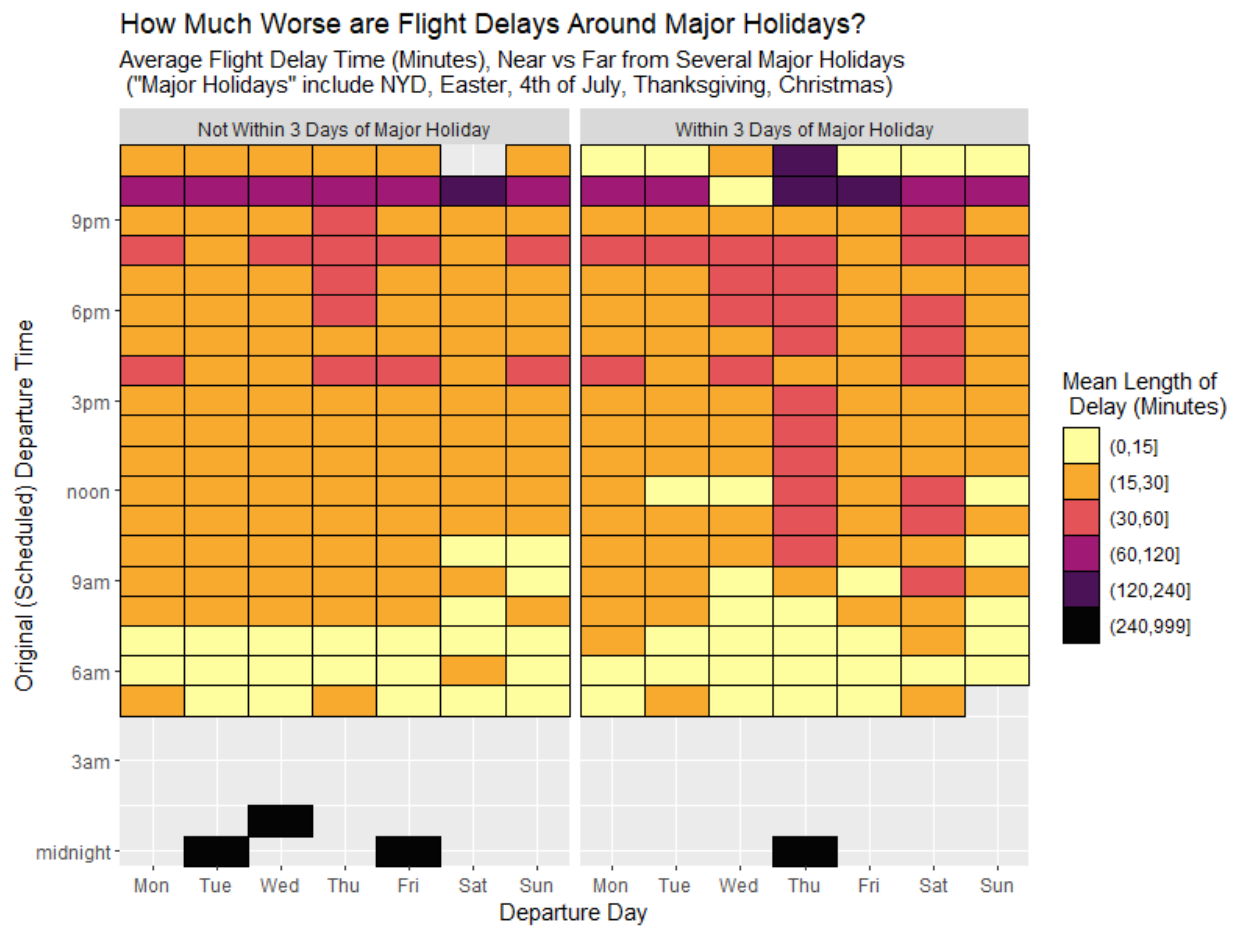


Alt Delay Plot

Amy Ly, Miles Moran, Abraham Mendoza, Brandon Booth, Connor Crane

4/24/2022

Final Result:



Code: (V.6) Heat Map Facet by Proximity to Holidays (within 3 days vs not)

Note: code has been set to `eval=FALSE` in case you want to run it yourself! Additionally, we've split it across 3 pages by task / goal of the chunk.

```
DepTimeToHour <- function(DepTime, DepDelay) {
  DepTime[DepTime<DepDelay] <- DepTime[DepTime<DepDelay] + 2400
  return( trunc((DepTime-DepDelay)/100) )
}

hflights <- hflights:hflights

#####
### 1. Filter out NAs ###
### 2. Combine Year/Month/DayofMonth variables into a single "Date" variable ###
### 3. Turn the DayOfWeek variable into a nice factor ###
### 4. Calculate the hour in which each departure was *supposed* to leave ###
#####

DayOfWeek.lbls <- c("Mon", "Tue", "Wed", "Thu", "Fri", "Sat", "Sun")

hflights_df <-
  hflights %>%
  drop_na(Year, Month, DayofMonth, DayOfWeek, DepTime) %>%
  as_tibble() %>%
  unite("date", c(Year, Month, DayofMonth), remove = FALSE) %>%
  mutate(date = ymd(date, tz="US/Central"),
         DayOfWeek = factor(DayOfWeek, labels = DayOfWeek.lbls),
         DepHourOg = DepTimeToHour(DepTime, DepDelay)
  )
```

```
#####
### Add a new variable for those close to holidays #####
#####

July4th <- ymd("2011-07-04", tz="US/Central")
EasterD <- ymd("2011-04-24", tz="US/Central")
ThnksGv <- ymd("2011-11-24", tz="US/Central")
XmasDay <- ymd("2011-12-25", tz="US/Central")
NYD2011 <- ymd("2011-01-01", tz="US/Central")
NYD2012 <- ymd("2012-01-01", tz="US/Central")

hflights_df <-
  hflights_df %>%
  mutate(close_to_holiday =
    (abs(difftime(date, July4th, units="days")) <= 3) |
    (abs(difftime(date, EasterD, units="days")) <= 3) |
    (abs(difftime(date, ThnksGv, units="days")) <= 3) |
    (abs(difftime(date, XmasDay, units="days")) <= 3) |
    (abs(difftime(date, NYD2011, units="days")) <= 3) |
    (abs(difftime(date, NYD2012, units="days")) <= 3),
    close_to_holiday = factor(close_to_holiday,
                              labels = c("Not Within 3 Days of Major Holiday",
                                           "Within 3 Days of Major Holiday")))
```

```
#####
### Calculate the Mean DepDelay for each Hour in each Day of the Week #####
#####

hour_labs <- c("midnight", "3am", "6am", "9am", "noon", "3pm", "6pm", "9pm", "midnight")
facet_labs <- c("Normal Days", "Within 3 Days of a Major Holiday")
delay_intervals <- c(0,15,30,60,120,240,999)

hflights_df_summary <-
  hflights_df                                %>%
  filter(DepDelay > 0)                        %>%
  group_by(close_to_holiday, DayOfWeek, DepHourOg) %>%
  summarize(meanDDelay = mean(DepDelay),
            meanDDelay_binned = cut(meanDDelay, breaks=delay_intervals)
            )                                %>%
  ungroup()

ggplot(hflights_df_summary) +
  geom_tile(aes(DayOfWeek, DepHourOg, fill = meanDDelay_binned),
            colour = "black",
            width = 1) +
  facet_wrap(.~close_to_holiday) +
  ggtitle("How Much Worse are Flight Delays Around Major Holidays?",
          subtitle = "Average Flight Delay Time (Minutes), Near vs Far from Several Major Holidays \n")
  scale_y_continuous("Original (Scheduled) Departure Time",
                    breaks = seq(0, 24, 3),
                    labels = hour_labs,
                    expand = c(0, 0)
                    ) +
  scale_x_discrete("Departure Day", expand=c(0,0)) +
  scale_fill_discrete_sequential(name="Mean Length of \n Delay (Minutes)",
                                palette = "Inferno",
                                guide = "legend"
                                )

```