

main

nwfried

2025-02-03

Load libraries.

```
library(sf)
```

```
## Linking to GEOS 3.13.0, GDAL 3.10.1, PROJ 9.5.1; sf_use_s2() is TRUE
```

```
library(tigris)
```

```
## To enable caching of data, set `options(tigris_use_cache = TRUE)`  
## in your R script or .Rprofile.
```

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --  
## v dplyr      1.1.4      v readr      2.1.5  
## v forcats    1.0.0      v stringr    1.5.1  
## v ggplot2    3.5.1      v tibble     3.2.1  
## v lubridate  1.9.3      v tidyr      1.3.1  
## v purrr      1.0.2
```

```
## -- Conflicts ----- tidyverse_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()     masks stats::lag()  
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(jsonlite)
```

```
##  
## Attaching package: 'jsonlite'  
##  
## The following object is masked from 'package:purrr':  
##  
##   flatten
```

```
library(knitr)
```

```
options(tigris_use_cache = TRUE)
```

```
filepath <- "~/Documents/congestion_pricing/mapbox_drive_times/data"
```

Initialise zipcodes for congestion zone as list.

```
cong_zips <-c("10036", "10038", "10280", "10282")  
for (i in 10001:10022) {  
  cong_zips <- append(cong_zips, as.character(i))  
}
```

Pull GIS info for NYC.

```
newyork <- tigris::counties(state = "NY", class = "sf") %>%
  st_transform(crs = "WGS84")
```

## Retrieving data for the year 2022

```
manhattan <- newyork %>% filter(NAME == "New York") %>%
  st_transform(crs = "+proj=longlat +datum=WGS84")
brooklyn <- newyork %>% filter(NAME == "Kings") %>%
  st_transform(crs = "WGS84")
bronx <- newyork %>% filter(NAME == "Bronx") %>%
  st_transform(crs = "WGS84")
staten_island <- newyork %>% filter(NAME == "Richmond") %>%
  st_transform(crs = "WGS84")
queens <- newyork %>% filter(NAME == "Queens") %>%
  st_transform(crs = "WGS84") #probably a better way to just pull in WGS84 coords lol
nyc <- newyork %>% filter(NAME == "New York" | NAME == "Kings" | NAME == "Bronx" | NAME == "Richmond" |
manhattan_planar <- manhattan %>% st_transform(32618)
```

Pull GIS info for New Jersey.

```
nj_counties <- counties(state = "NJ", cb = TRUE) %>%
  filter(NAME %in% c("Hudson", "Bergen", "Essex", "Union")) %>% st_transform(crs = "WGS84")
```

## Retrieving data for the year 2022

Pull GIS info for congestion zip codes from list.

```
cong_zone <- tigris::zctas(year = "2020", class = "sf")
cong_zone <- cong_zone %>% filter(ZCTA5CE20 %in% cong_zips) %>% st_transform(crs = "WGS84")
```

Define function to intersect sf file with relevant GIS objects from above (e.g. congestion zone, NYC).

```
analyse_sf <- function(sf){
  sf <- sf %>% mutate(
    cong = sapply(st_intersects(geometry, cong_zone), function(x) length(x) > 0),
    manhattan = sapply(st_intersects(geometry, manhattan), function(x) length(x) > 0),
    brooklyn = sapply(st_intersects(geometry, brooklyn), function(x) length(x) > 0),
    queens = sapply(st_intersects(geometry, queens), function(x) length(x) > 0),
    bronx = sapply(st_intersects(geometry, bronx), function(x) length(x) > 0),
    staten_island = sapply(st_intersects(geometry, staten_island), function(x) length(x) > 0),
    nyc = sapply(st_intersects(geometry, nyc), function(x) length(x) > 0))%>%
    st_transform(32618) %>% #st_crosses requires reprojection to 32618 (planar) coordinate system
    mutate(
      crosses_manhattan = lengths(st_crosses(geometry, manhattan_planar)) > 0
    )
  return(sf)
}
```

Define function to intersect sf file with only the congestion zone (as above, just simplified).

```
analyse_cong <- function(sf) {
  sf <- sf %>% mutate(
    cong = sapply(st_intersects(geometry, cong_zone), function(x) length(x) > 0)
  )
}
```

Read in Jan 15 data (first dataset with coordinates), select only columns that correspond to relevant geometry (lat/long + coordinates), and apply analyse\_sf function from above.

```

jan15_less <- read_delim(file.path(filepath, "mapbox_output_2025-01-15_HH08.csv")) %>% select(orig_long,
  filter(coordinates != '{}') %>%
  mutate(
    geometry = list(
      st_linestring(
        as.matrix(fromJSON(coordinates))
      )
    )
  ) %>%
  st_as_sf(crs = 4326)

```

```

## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## * `` -> `...1`

```

```

jan15_less <- analyse_sf(jan15_less)
jan15_less <- jan15_less %>% st_drop_geometry()

```

Define functions to sort dataset by relevant categories (in congestion zone, outside of NYC) and compute the mean and median duration of selected trips.

```

cong_info <- function(df) {
  mean <- df %>%
    ungroup() %>%
    filter(cong == TRUE) %>%
    summarise(mean_duration = mean(duration_min), median = median(duration_min), n = n())
  return(mean)
}

trip_info <- function(df) {
  mean <- df %>%
    ungroup() %>%
    summarise(mean_duration = mean(duration_min), median = median(duration_min), n = n())
  return(mean)
}

control_info <- function(df) {
  mean <- df %>% ungroup() %>% filter(nyc == FALSE) %>%
    summarise(mean_duration = mean(duration_min), median = median(duration_min), n = n())
  return(mean)
}

```

Define function to extract date/hour from formatting of mapbox filenames.

```

extract_date_hour <- function(file_name) {
  match <- str_match(file_name, "mapbox_output_(\\d{4}-\\d{2}-\\d{2})_HH(\\d{2})")
  list(
    date = match[2], # Extracted date (YYYY-MM-DD)
    hour = match[3] # Extracted hour (HH)
  )
}

```

Define function to initialise an empty dataframe for results, then iterate over all files within a specified folder and then perform the following: 1. Extract date/hour from filename. 2. Merge file with categorical data from Jan 15. 3. Compute mean/median of various selections of trips using functions defined above. 4. Store this result based on date/hour of filename in initialised dataframe.

```
# Read all CSV files in a folder
process_csv_files <- function(folder_path) {
  # List all CSV files in the folder
  csv_files <- list.files(path = folder_path, pattern = "\\\\.csv$", full.names = TRUE)
  # Empty dataframe to store results
  results_df <- data.frame(
    date = character(),
    hour = character(),
    cong = numeric(),
    trip = numeric(),
    control = numeric(),
    stringsAsFactors = FALSE
  )

  # Loop over all files
  for (file in csv_files) {
    extracted <- extract_date_hour(basename(file))
    date_hour <- paste0(extracted$date, "_", extracted$hour) # Combine date and hour for unique naming

    # Read the CSV file
    data <- read.csv(file)
    # Apply the functions
    sf_data <- left_join(data, jan15_less, by = c("orig_long", "orig_lat", "dest_long", "dest_lat"), re
    congestion_result <- cong_info(sf_data)
    trip_result <- trip_info(sf_data)
    control_result <- control_info(sf_data)
    # Store the result in results dataframe
    results_df <- rbind(
      results_df,
      data.frame(
        date = extracted$date,
        hour = extracted$hour,
        cong = congestion_result,
        trip = trip_result,
        control = control_result,
        stringsAsFactors = FALSE
      )
    )
  }

  return(results_df)
}
```

Apply this function to folder with mapbox data.

```
results <- process_csv_files(filepath)
```

Define function to iterate over all files within a folder, select only first 20 trips commuting to 10001 and merge duration info for these trips into a results dataset.

```

process_locations <- function(folder_path) {
  # List all CSV files in the folder
  csv_files <- list.files(path = folder_path, pattern = "\\*.csv$", full.names = TRUE)
  results_df <- read_csv(file.path(folder_path, "mapbox_output_2024-12-08_HH08.csv")) %>% filter(zip_commute > 0)
  for (file in csv_files) {
    date_hour <- str_extract(file, "\\d{4}-\\d{2}-\\d{2}_HH\\d{2}")
    data <- read_csv(file)
    # Remove slicing eventually (just to get around memory limit)
    data <- data %>% filter(zip_commute == 10001 & duration_min != 0) %>% select(orig_long, orig_lat, dest_long, dest_lat)
    results_df <- left_join(results_df, data, by = c("orig_long", "orig_lat", "dest_long", "dest_lat"))
  }
  return(results_df)
}

```

```
location_results <- process_locations(folder_path)
```

```

## New names:
## Rows: 1737 Columns: 19
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, de...
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 1737 Columns: 19
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, de...
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 19
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, de...
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 19
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, de...
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 19
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, de...

```

```
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
```







```

## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp

```

```
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,  
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp  
## i Use `spec()` to retrieve the full column specification for this data. i  
## Specify the column types or set `show_col_types = FALSE` to quiet this message.  
## New names:  
## Rows: 2131 Columns: 20  
## -- Column specification  
## ----- Delimiter: "," chr  
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,  
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp  
## i Use `spec()` to retrieve the full column specification for this data. i  
## Specify the column types or set `show_col_types = FALSE` to quiet this message.  
## New names:  
## Rows: 2131 Columns: 20  
## -- Column specification  
## ----- Delimiter: "," chr  
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,  
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp  
## i Use `spec()` to retrieve the full column specification for this data. i  
## Specify the column types or set `show_col_types = FALSE` to quiet this message.  
## New names:  
## Rows: 2131 Columns: 20  
## -- Column specification  
## ----- Delimiter: "," chr  
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,  
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp  
## i Use `spec()` to retrieve the full column specification for this data. i  
## Specify the column types or set `show_col_types = FALSE` to quiet this message.  
## New names:  
## Rows: 2131 Columns: 20  
## -- Column specification  
## ----- Delimiter: "," chr  
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,  
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp  
## i Use `spec()` to retrieve the full column specification for this data. i  
## Specify the column types or set `show_col_types = FALSE` to quiet this message.  
## New names:  
## Rows: 2131 Columns: 20  
## -- Column specification  
## ----- Delimiter: "," chr  
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,  
## zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp  
## i Use `spec()` to retrieve the full column specification for this data. i  
## Specify the column types or set `show_col_types = FALSE` to quiet this message.  
## New names:  
## Rows: 2131 Columns: 20  
## -- Column specification
```

```

## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.

```

```

## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp

```

```

## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (4): weight_name, waypoints, code, uuid dbl (15): ...1, X, inputzip,
## zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dttm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr

```

```
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,  
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp  
## i Use `spec()` to retrieve the full column specification for this data. i  
## Specify the column types or set `show_col_types = FALSE` to quiet this message.  
## New names:  
## Rows: 2131 Columns: 22  
## -- Column specification  
## ----- Delimiter: "," chr  
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,  
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp  
## i Use `spec()` to retrieve the full column specification for this data. i  
## Specify the column types or set `show_col_types = FALSE` to quiet this message.  
## New names:  
## Rows: 2131 Columns: 22  
## -- Column specification  
## ----- Delimiter: "," chr  
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,  
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp  
## i Use `spec()` to retrieve the full column specification for this data. i  
## Specify the column types or set `show_col_types = FALSE` to quiet this message.  
## New names:  
## Rows: 2131 Columns: 22  
## -- Column specification  
## ----- Delimiter: "," chr  
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,  
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp  
## i Use `spec()` to retrieve the full column specification for this data. i  
## Specify the column types or set `show_col_types = FALSE` to quiet this message.  
## New names:  
## Rows: 2131 Columns: 22  
## -- Column specification  
## ----- Delimiter: "," chr  
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,  
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp  
## i Use `spec()` to retrieve the full column specification for this data. i  
## Specify the column types or set `show_col_types = FALSE` to quiet this message.  
## New names:  
## Rows: 2131 Columns: 22  
## -- Column specification  
## ----- Delimiter: "," chr  
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,  
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp  
## i Use `spec()` to retrieve the full column specification for this data. i  
## Specify the column types or set `show_col_types = FALSE` to quiet this message.  
## New names:  
## Rows: 2131 Columns: 22  
## -- Column specification
```

```
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
```



```

## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp

```

```
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## New names:
## Rows: 2131 Columns: 22
## -- Column specification
## ----- Delimiter: "," chr
## (6): weight_name, coordinates, type, waypoints, code, uuid dbl (15): ...1, X,
## inputzip, zip_commute, orig_long, orig_lat, dest_long, d... dtm (1): timestamp
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## * `` -> `...1`
```

Clean result from 10001 trips, filter dates by before and after congestion zone and then calculate respective means. Plot this on NYC/NJ shapefile based on origin.

```
location_results <- location_results %>% select(where(~ !any(is.na(.))), -duration_min, -zip_commute)
cong_cutoff <- "2025-01-04_HH08"
cong_cols<- location_results %>%
  select(matches("\\d{4}-\\d{2}-\\d{2}_HH\\d{2}")) %>% # Select date columns
  select(which(names(.) > cong_cutoff)) # Keep only columns after the cutoff
```

```

before_cong_cols<- location_results %>%
  select(matches("\\d{4}-\\d{2}-\\d{2}_HH\\d{2}")) %>% # Select date columns
  select(which(names(.) <= cong_cutoff)) # Keep only columns before the cutoff
# Compute row-wise average for filtered columns
location_results <- location_results %>%
  mutate(avg_after_cong= rowMeans(cong_cols, na.rm = TRUE), avg_before_cong = rowMeans(before_cong_cols

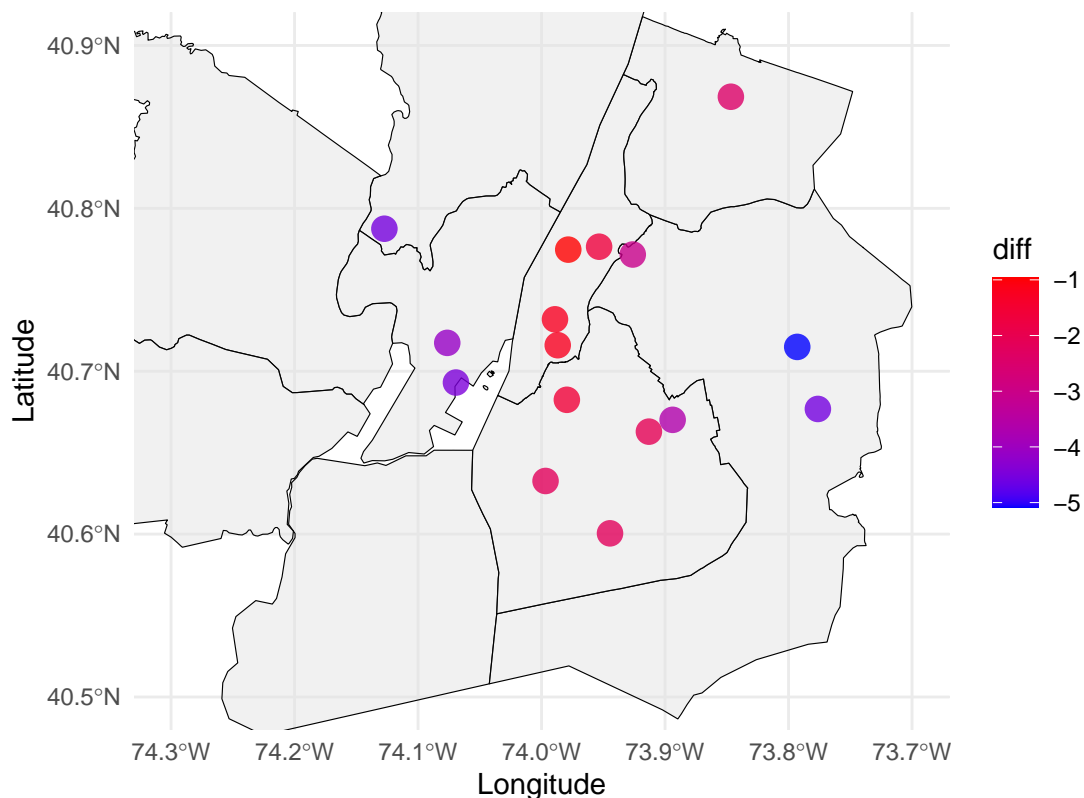
location_results <- location_results %>% mutate(diff = .data[["avg_after_cong"]] - .data[["avg_before_c

# Convert df to an sf object (spatial format)
location_sf <- st_as_sf(location_results, coords = c("orig_long", "orig_lat"), crs = 4326)

nyc_nj_map <- bind_rows(nyc, nj_counties)
# Load NYC shapefile or use OpenStreetMap background (alternative)
ggplot() +
  geom_sf(data = nyc_nj_map, fill = "gray90", color = "black", alpha = 0.5) +
  geom_sf(data = location_sf, aes(color = diff), size = 4, alpha = 0.8) +
  scale_color_gradient(low = "blue", high = "red") +
  coord_sf(xlim = c(-74.3, -73.7), ylim = c(40.5, 40.9)) + # NYC bounding box
  labs(title = "NYC Locations Colored by Diff Value", x = "Longitude", y = "Latitude") +
  theme_minimal()

```

NYC Locations Colored by Diff Value



Define function to determine day of the week:

```

is_weekday_or_weekend <- function(date) {
  day <- weekdays(as.Date(date))
  ifelse(day %in% c("Saturday", "Sunday"), "Weekend", "Weekday")
}

```

```
}
```

```
results <- results %>% mutate(day = is_weekday_or_weekend(as.POSIXct(date, format = "%Y-%m-%d")))
```

Filter results by weekday/weekend/hour.

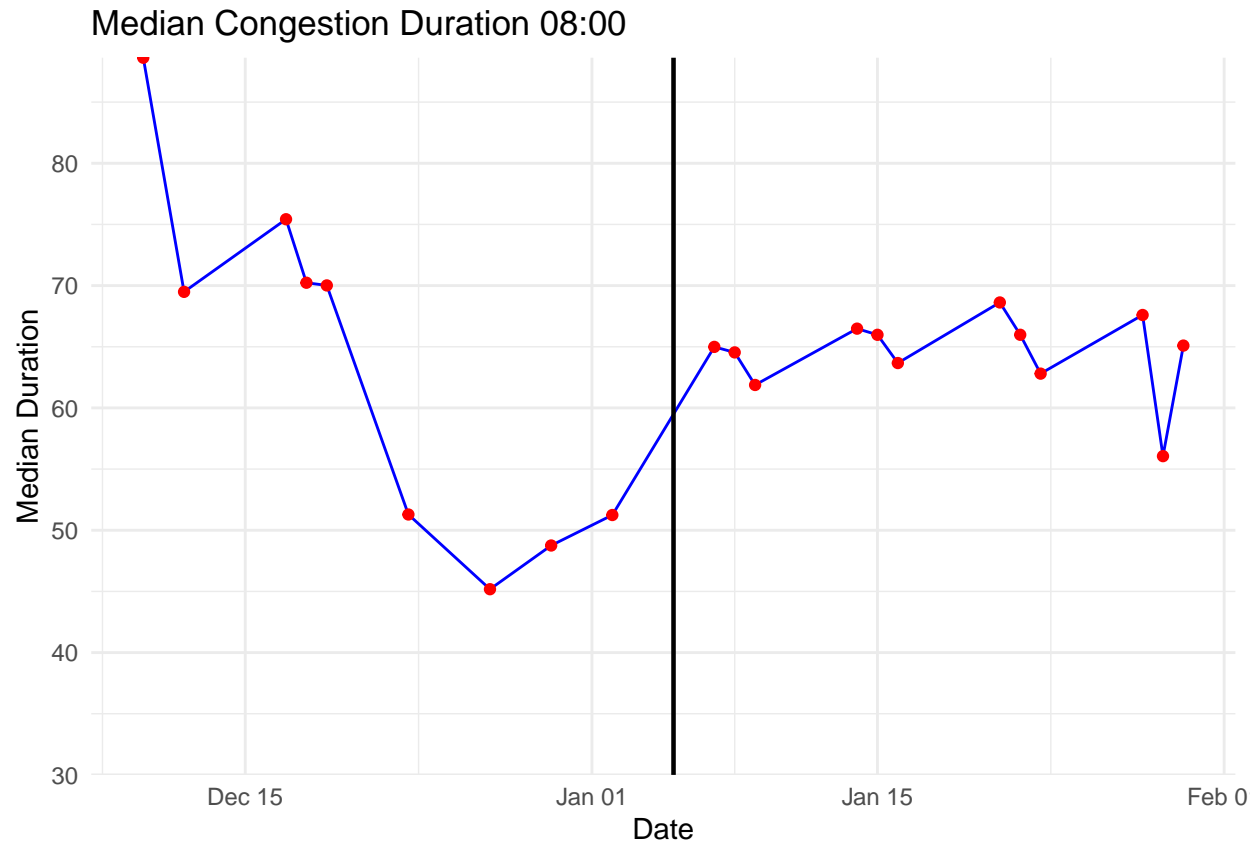
```
results_wd <- results %>% filter(day == "Weekday")
results_we <- results %>% filter(day == "Weekend")
results_wd_08 <- results_wd %>% filter(hour == '08') %>% select(-hour)
results_wd_12 <- results_wd %>% filter(hour == '12') %>% select(-hour)
results_wd_17 <- results_wd %>% filter(hour == '17') %>% select(-hour)

results_we_08 <- results_we %>% filter(hour == '08') %>% select(-hour)
results_we_12 <- results_we %>% filter(hour == '12') %>% select(-hour)
results_we_17 <- results_we %>% filter(hour == '17') %>% select(-hour)
```

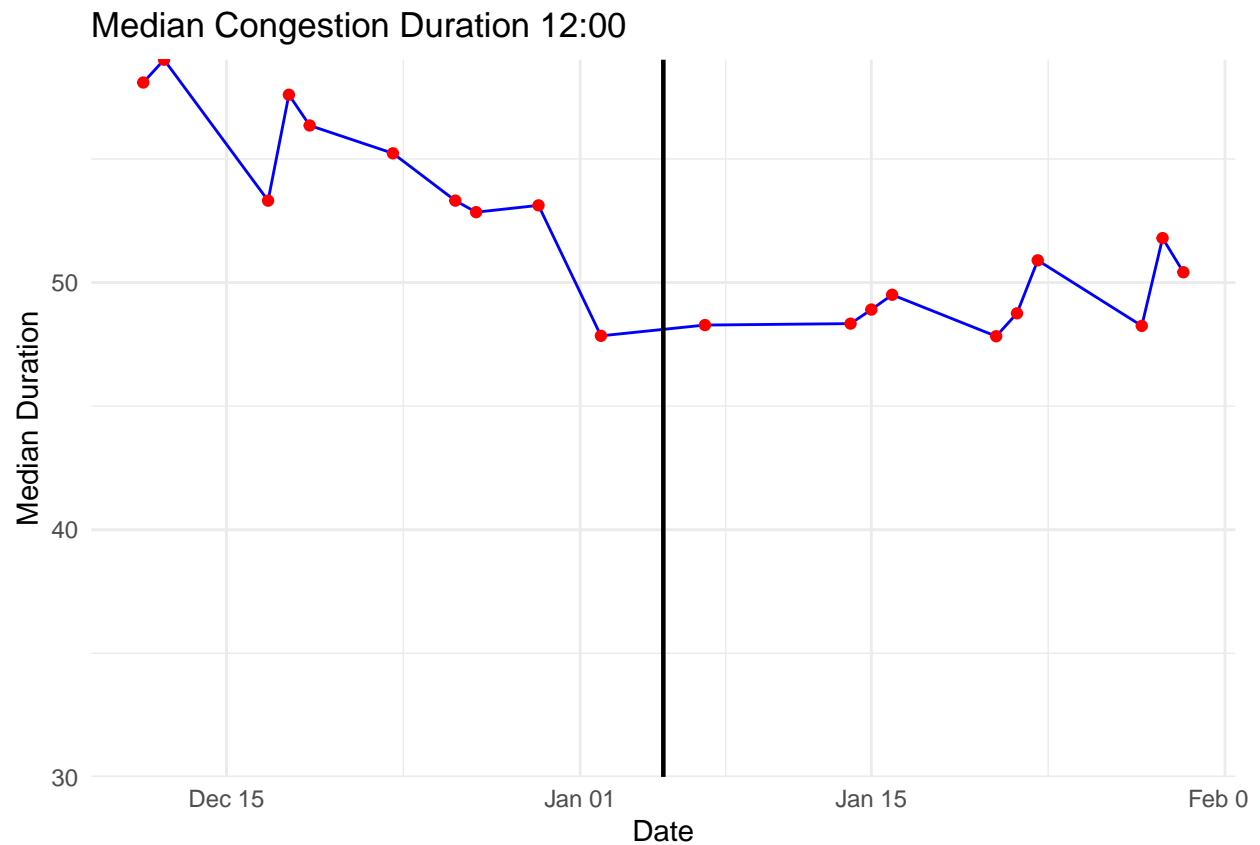
Plot median weekday congestion at 08:00, 12:00, 17:00:

```
ggplot(results_wd_08, aes(x = as.Date(date), y = cong.median)) +
  geom_line(color = "blue") +
  geom_point(color = "red") +
  geom_vline(xintercept = as.Date("2025-01-05"), linetype = "solid", color = "black", size = 0.8) +
  labs(
    x = "Date",
    y = "Median Duration",
    title = "Median Congestion Duration 08:00"
  ) + scale_y_continuous(expand = c(0, 0), limits = c(30, NA)) +
  theme_minimal()
```

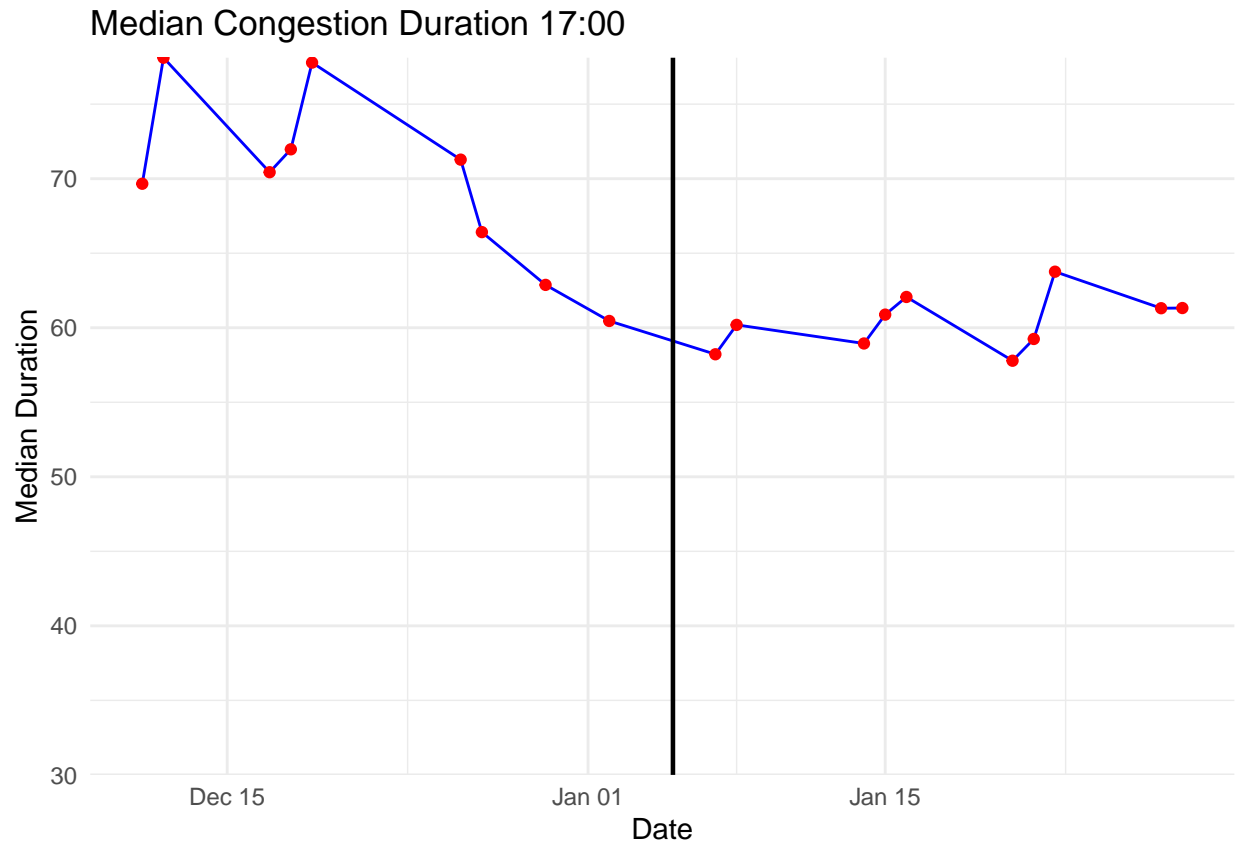
```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```



```
ggplot(results_wd_12, aes(x = as.Date(date), y = cong.median)) +
  geom_line(color = "blue") +
  geom_point(color = "red") +
  geom_vline(xintercept = as.Date("2025-01-05"), linetype = "solid", color = "black", size = 0.8) +
  labs(
    x = "Date",
    y = "Median Duration",
    title = "Median Congestion Duration 12:00"
  ) + scale_y_continuous(expand = c(0, 0), limits = c(30, NA)) + # Ensure the y-axis starts at 100
  theme_minimal()
```



```
ggplot(results_wd_17, aes(x = as.Date(date), y = cong.median)) +
  geom_line(color = "blue") +
  geom_point(color = "red") +
  geom_vline(xintercept = as.Date("2025-01-05"), linetype = "solid", color = "black", size = 0.8) +
  labs(
    x = "Date",
    y = "Median Duration",
    title = "Median Congestion Duration 17:00"
  ) + scale_y_continuous(expand = c(0, 0), limits = c(30, NA)) + # Ensure the y-axis starts at 100
  theme_minimal()
```

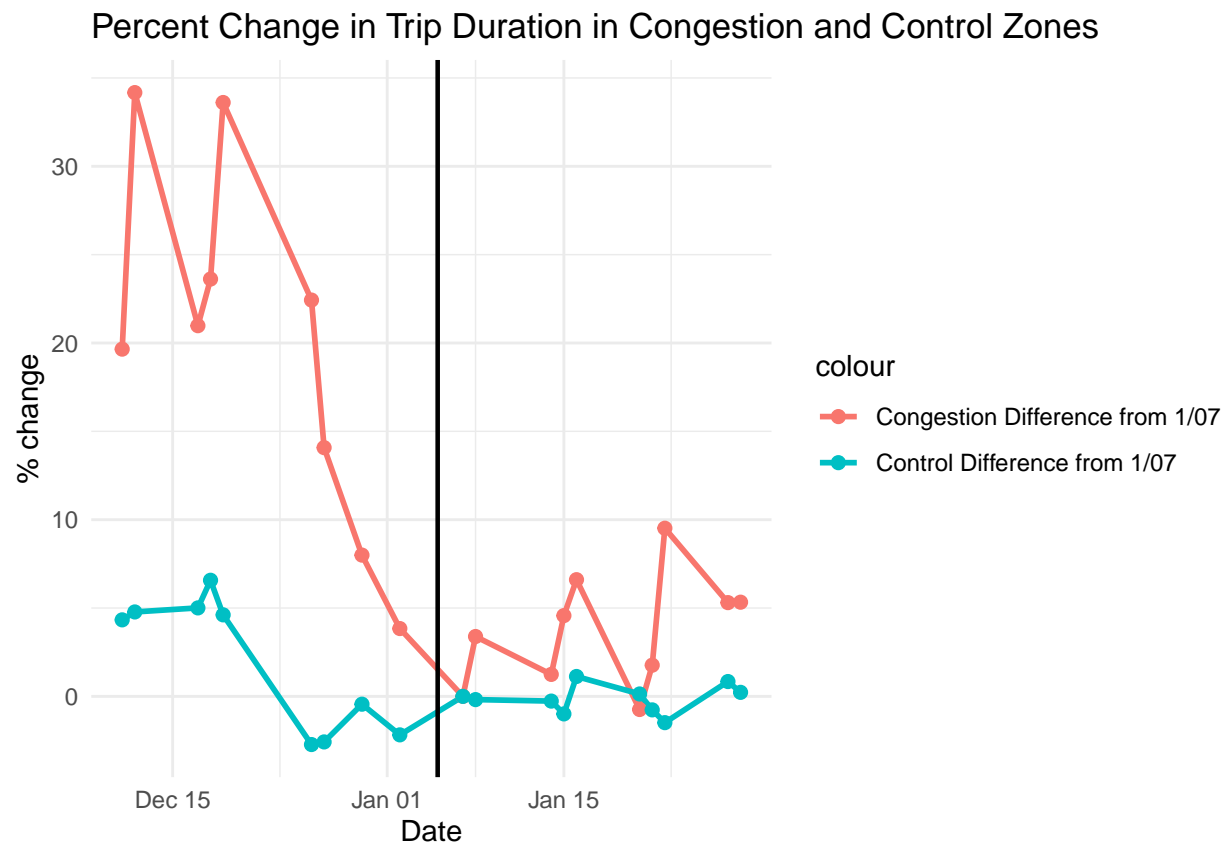


Define function to compute the percent change in median trip time from 01/07 for congestion and control trips. Plot this for the same datasets as above.

```
plot_pct_change <- function(df) {
  cong_ref_value <- df %>%
    filter(date == "2025-01-07") %>%
    pull(cong.median)
  control_ref_value <- df %>%
    filter(date == "2025-01-07") %>%
    pull(control.median)
  # Compute percent change
  df <- df %>%
    mutate(pct_change_cong = (cong.median - cong_ref_value) / cong_ref_value * 100, pct_change_control = (control.median - control_ref_value) / control_ref_value * 100)
  #Plot
  ggplot(df, aes(x = as.Date(date))) +
    geom_line(aes(y = pct_change_cong, color = "Congestion Difference from 1/07"), size = 1) +
    geom_line(aes(y = pct_change_control, color = "Control Difference from 1/07"), size = 1) +
    geom_point(aes(y = pct_change_cong, color = "Congestion Difference from 1/07"), size = 2) +
    geom_point(aes(y = pct_change_control, color = "Control Difference from 1/07"), size = 2) +
    geom_vline(xintercept = as.Date("2025-01-05"), linetype = "solid", color = "black", size = 0.8) +
    labs(
      x = "Date",
      y = "% change",
      title = "Percent Change in Trip Duration in Congestion and Control Zones"
    ) +
    theme_minimal()
}
```

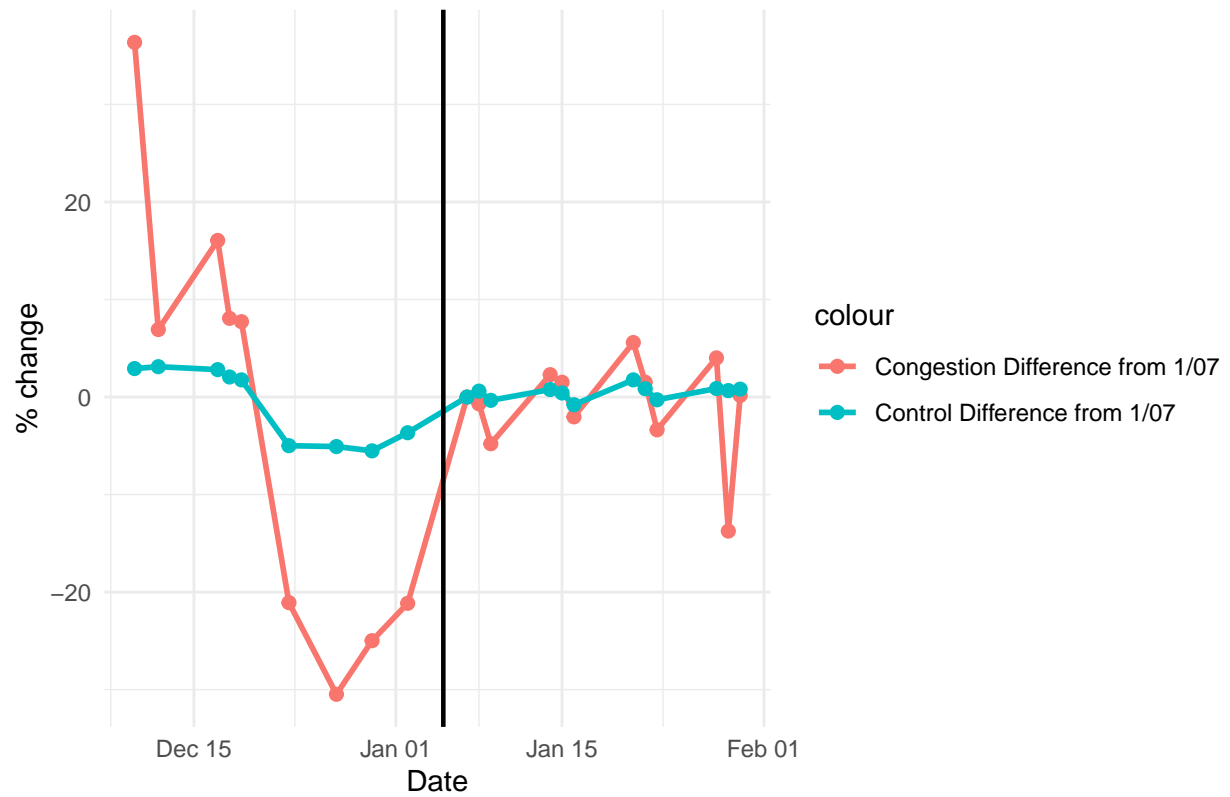


```
plot_pct_change(results_wd_17)
```



```
plot_pct_change(results_wd_08)
```

Percent Change in Trip Duration in Congestion and Control Zones



```
plot_pct_change(results_wd_12)
```

