

Advanced Data Visualizations

West - Week 1

AUTHOR
Jeff Goseland

PUBLISHED
November 9, 2025

Week 1 Homework

Load Libraries

```
#loads needed libraries
suppressMessages({
  library(tidyverse)
  library(RColorBrewer)
  library(janitor)
  library(scales)
})
```

Open Files

```
filename <- '../data/311_Phone_Call_Log_Mod.csv'

calls_raw <- read_csv(filename) %>% clean_names() %>% drop_na()
```

```
Rows: 61075 Columns: 7
— Column specification —————
Delimiter: ","
chr  (3): Work_Order_Type, Called_About, Department
dbl  (2): FID, duration_Seconds
dtm  (1): Call_Date
time (1): Duration_Minutes_
```

- Use ``spec()`` to retrieve the full column specification for this data.
- Specify the column types or set ``show_col_types = FALSE`` to quiet this message.

```
#glimpses dataset
glimpse(calls_raw)
```

```
Rows: 5,695
Columns: 7
$ call_date      <dtm> 2017-03-24 18:22:00, 2017-03-24 14:02:00, 2017-03-24...
$ work_order_type <chr> "General Service Request", "General Service Request",...
$ called_about    <chr> "PKMS - General Concern", "SWGC - General Concern", "...
$ department      <chr> "Parks - Maintenance", "Solid Waste", "Solid Waste", ...
$ duration_minutes <time> 04:41:00, 02:10:00, 02:13:00, 02:36:00, 01:47:00, 01...
$ fid             <dbl> 7005, 7006, 7103, 7111, 7113, 7143, 7152, 7157, 7164,...
$ duration_seconds <dbl> 281, 130, 133, 156, 107, 61, 66, 217, 80, 76, 242, 15...
```

```
head(calls_raw)
```

	call_date	work_order_type
	<dtm>	<chr>
	2017-03-24 18:22:00	General Service Request
	2017-03-24 14:02:00	General Service Request
	2017-03-24 19:15:00	General Service Request
	2017-03-23 22:15:00	General Service Request
	2017-03-24 14:50:00	General Service Request
	2017-03-24 16:32:00	General Service Request

6 rows | 1-2 of 7 columns

Part B

Given a dataset of and using tools you already know in R (i.e., ggplot) create a chart/data visualization that applies strategies presented this week. When creating this graph, keep in mind the examples you found in Part

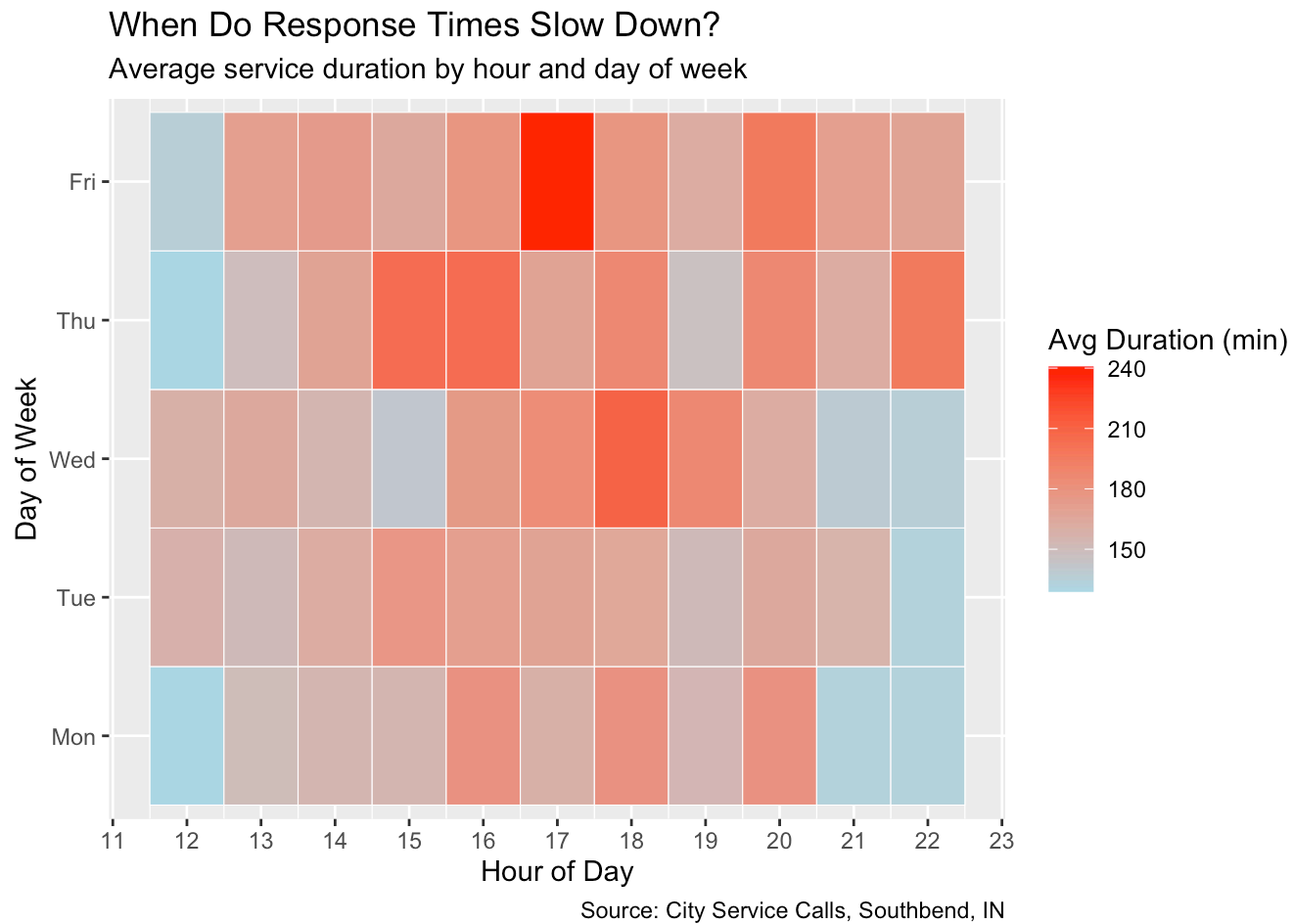
A.

1. Choose variables you would like to visualize. Use what you learned in Storytelling or the to choose a chart type. You could show how long calls from a particular category last on average, or the total call volume per year, or anything else you think would be interesting! Create a visualization in R. Provide both the visualization and the code used to generate it. 5 pts

```
# let's make a heat map, they are fun.
calls_raw <- calls_raw %>%
  mutate(call_date = as.POSIXct(call_date, tz = "UTC"),
         hour = hour(call_date),
         wday = wday(call_date, label = TRUE, abbr = TRUE))

calls_heat <- calls_raw %>%
  group_by(wday, hour) %>%
  summarise(avg_duration = mean(duration_minutes, na.rm = TRUE)/60,

ggplot(calls_heat,
  aes(x = hour, y = wday, fill = avg_duration)) +
  geom_tile(color = "white") +
  scale_x_continuous(breaks = 0:30) +
  scale_fill_gradient(low = "lightblue", high = "red", name = "Avg Duration (min)") +
  labs(title = "When Do Response Times Slow Down?",
       subtitle = "Average service duration by hour and day of week",
       caption = "Source: City Service Calls, Southbend, IN",
       x = "Hour of Day",
       y = "Day of Week")
```



2. Submit a few sentences discussing your visualization, explaining what you did, why you did it, and what you learned about the dataset by visualizing it. Submit as a PDF (or your RMD file) including your analysis

Unfortunately, I went straight to an operational mindset and I wanted to find out when the calls take the longest to reconcile. Thus, I created a heatmap that looked at the data by day of week and time of day. The hope was to see visually some outliers. This type of visualization could help with scheduling or (by deep-diving more) finding what calls are creating those long durations.

The heatmap shows us that the longest calls happen in the middle of the day and seem to be worse Wednesday - Friday, with Thursday and Friday having the worst average call duration (long)