

Data Viz: The Importance of Iteration

Birth Patterns in the US over a Single Year

```
# Load tidyverse (which contains ggplot2)
library(tidyverse)

# Load ggrepel to help with labeling points
library(ggrepel)

# Load package that will graph emojis
library(emoGG)

# Load library that has dataset of interest
library(mosaicData)

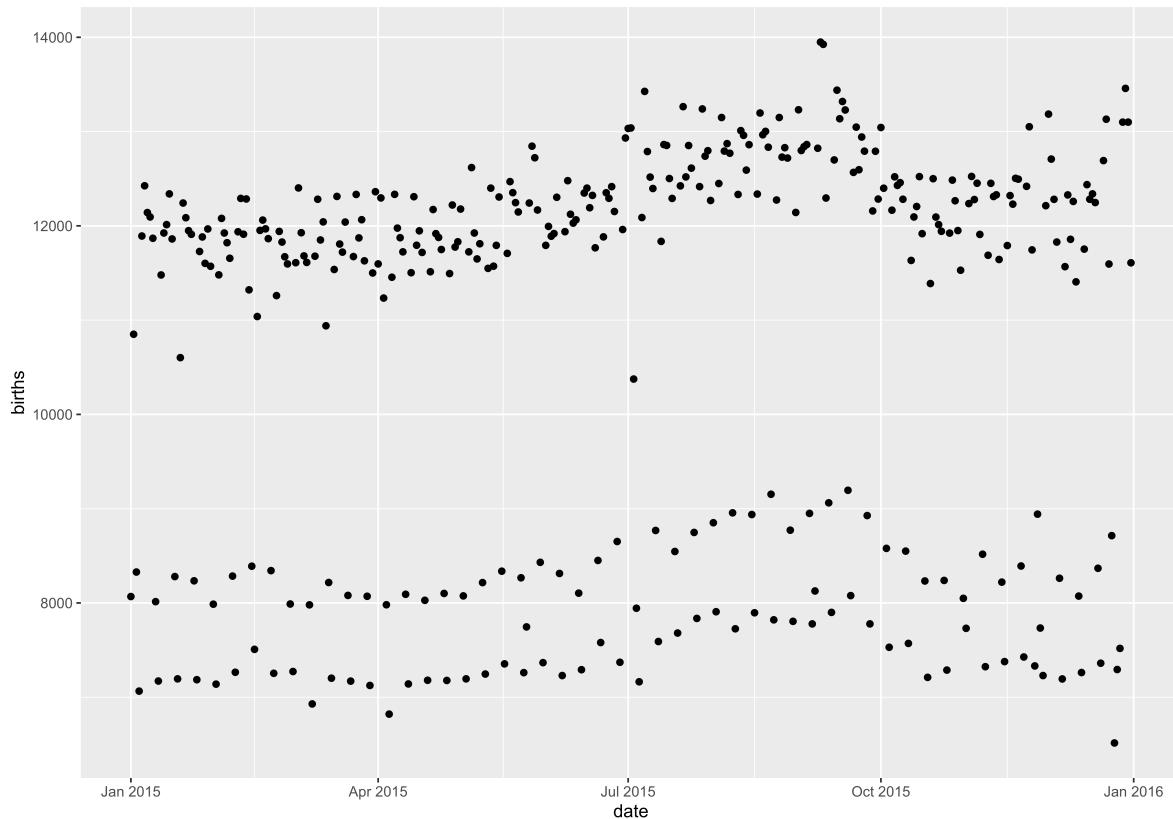
# Grab data
data(Births2015)

# Look at the data
glimpse(Births2015)
```

```
Rows: 365  
Columns: 8  
# date      <date> 2015-01-01, 2015-01-02, 2015-01-03, 2015-01-04, 2015-01-~  
# births    <dbl> 8068, 10850, 8328, 7065, 11892, 12425, 12141, 12094, 1186~  
# wday      <ord> Thu, Fri, Sat, Sun, Mon, Tue, Wed, Thu, Fri, Sat, Sun, Mo~  
# year      <dbl> 2015, 2015, 2015, 2015, 2015, 2015, 2015, 2015, 2015, 201~  
# month     <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ~  
# day_of_year <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17~  
# day_of_month <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17~  
# day_of_week  <dbl> 5, 6, 7, 1, 2, 3, 4, 5, 6, 7, 1, 2, 3, 4, 5, 6, 7, 1, 2, ~
```

Version 1: Graph with the Defaults

```
ggplot(data = Births2015,  
       mapping = aes(x = date, y = births)) +  
       geom_point()
```



What stories do you see in this data visualization?

How could we make the story clearer?

Iterations

Edits to make:

- Don't keep line.
- Add color.

- Change colors.
- Move legend to fix aspect ratio.
- Consider different breaks.
- Size of points.
- Highlight the holidays: labels, arrows, text, emojis
- Font size

Final Versions

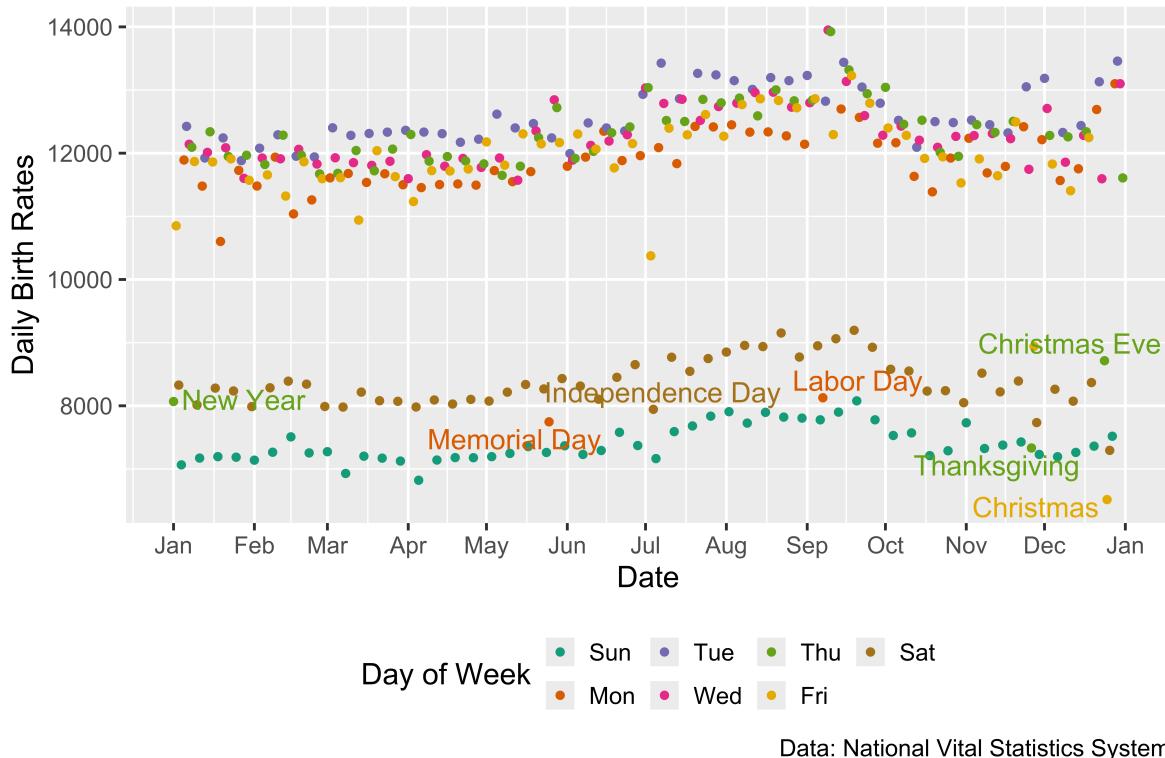
```

holidays <-
  data.frame(date = ymd("2015-01-01", "2015-05-25", "2015-07-04",
                        "2015-12-25", "2015-11-26", "2015-12-24",
                        "2015-09-07"),
             occasion = c("New Year", "Memorial Day",
                           "Independence Day", "Christmas",
                           "Thanksgiving", "Christmas Eve",
                           "Labor Day"),
             emoji = c("1f389", "1f396", "1f386", "1f384",
                       "1f983", "1f381", "1f477"))
holidays <- left_join(holidays, Births2015)

ggplot(data = Births2015,
       mapping = aes(x = date, y = births,
                     color = wday)) +
  geom_point(size = 2) +
  scale_x_date(date_labels = "%b",
               date_breaks = "1 month") +
  scale_color_brewer(type = "qual", palette = 2) +
  labs(x = "Date",
       y = "Daily Birth Rates",
       title = "2015 Trend of Births in the US",
       caption = "Data: National Vital Statistics System",
       color = "Day of Week") +
  geom_text_repel(data = holidays,
                 mapping = aes(label = occasion), size = 6,
                 show.legend = FALSE) +
  theme_gray(base_size = 18) +
  theme(legend.position = "bottom")

```

2015 Trend of Births in the US



```
# Another way of signifying holidays
# Create a story label
label_data <- data.frame(date = ymd("2015-01-01"),
                           births = max(Births2015$births),
                           label = "The frequency of births on holidays \nfollows weekend \ntrends.")

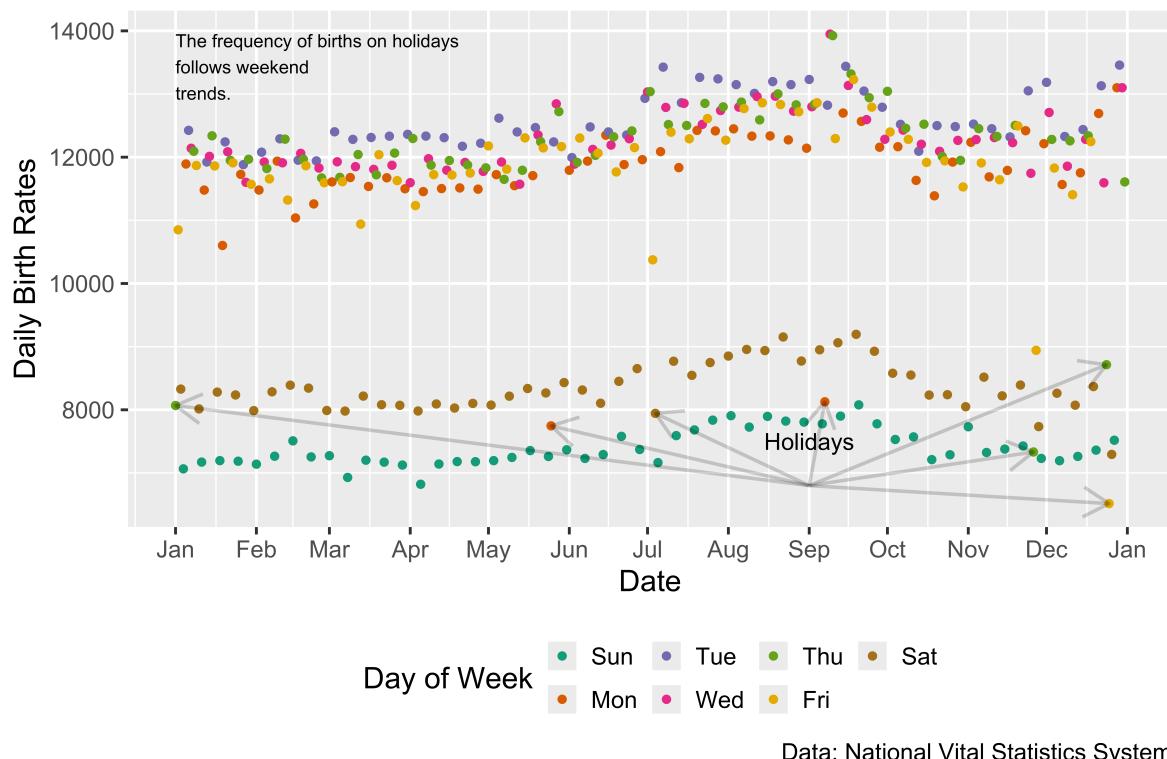
ggplot(data = Births2015,
       mapping = aes(x = date, y = births,
                     color = wday)) +
  geom_point(size = 2) +
  scale_x_date(date_labels = "%b",
               date_breaks = "1 month") +
  scale_color_brewer(type = "qual", palette = 2) +
  labs(x = "Date",
       y = "Daily Birth Rates",
       title = "2015 Trend of Births in the US",
       caption = "Data: National Vital Statistics System",
       color = "Day of Week") +
```

```

annotate("segment", colour = "black",
        x = as_date("2015-09-01"),
        xend = holidays$date,
        y = 6800, yend = holidays$births,
        size = 1, alpha = 0.2, arrow = arrow())+
annotate("text", x = as_date("2015-09-01"),
        y = 7500, label = "Holidays",
        color="black", size=5) +
geom_text(mapping = aes(label = label),
        data = label_data,
        color = "black", vjust = "top",
        hjust = "left") +
theme_gray(base_size = 18) +
theme(legend.position = "bottom")

```

2015 Trend of Births in the US



```

# Yet another option
ggplot(data = Births2015,
        mapping = aes(x = date, y = births,

```

```
          color = wday)) +
geom_point(size = 2) +
scale_x_date(date_labels = "%b",
             date_breaks = "1 month") +
scale_color_brewer(type = "qual", palette = 2) +
labs(x = "Date",
     y = "Daily Birth Rates",
     title = "2015 Trend of Births in the US",
     caption = "Data: National Vital Statistics System",
     color = "Day of Week") +
geom_text(mapping = aes(label = label),
          data = label_data,
          color = "black", vjust = "top",
          hjust = "left") +
geom_emoji(data = holidays,
           mapping = aes(emoji = emoji,
                         x = date,
                         y = births),
           inherit.aes = FALSE) +
theme_gray(base_size = 18) +
theme(legend.position = "bottom")
```

2015 Trend of Births in the US

