

# edld652\_lab2

Michelle Cui

2025-01-15

```
# load packages
library(tidyverse)
library(here)
library(janitor)
```

## Load Data

```
# read in the data set
cdc <- read.csv(here("data", "US_births_1994-2003_CDC_NCHS.csv")) %>%
  filter(year < "2000")

ssa <- read.csv(here("data", "US_births_2000-2014_SSA.csv"))
```

## Wrangle Data

```
head(cdc)
```

```
##   year month date_of_month day_of_week births
## 1 1994     1             1           6   8096
## 2 1994     1             2           7   7772
## 3 1994     1             3           1  10142
## 4 1994     1             4           2  11248
## 5 1994     1             5           3  11053
## 6 1994     1             6           4  11406
```

```
head(ssa)
```

```
##   year month date_of_month day_of_week births
## 1 2000     1             1           6   9083
## 2 2000     1             2           7   8006
## 3 2000     1             3           1  11363
## 4 2000     1             4           2  13032
## 5 2000     1             5           3  12558
## 6 2000     1             6           4  12466
```

```
births_combine <- bind_rows(cdc, ssa)
```

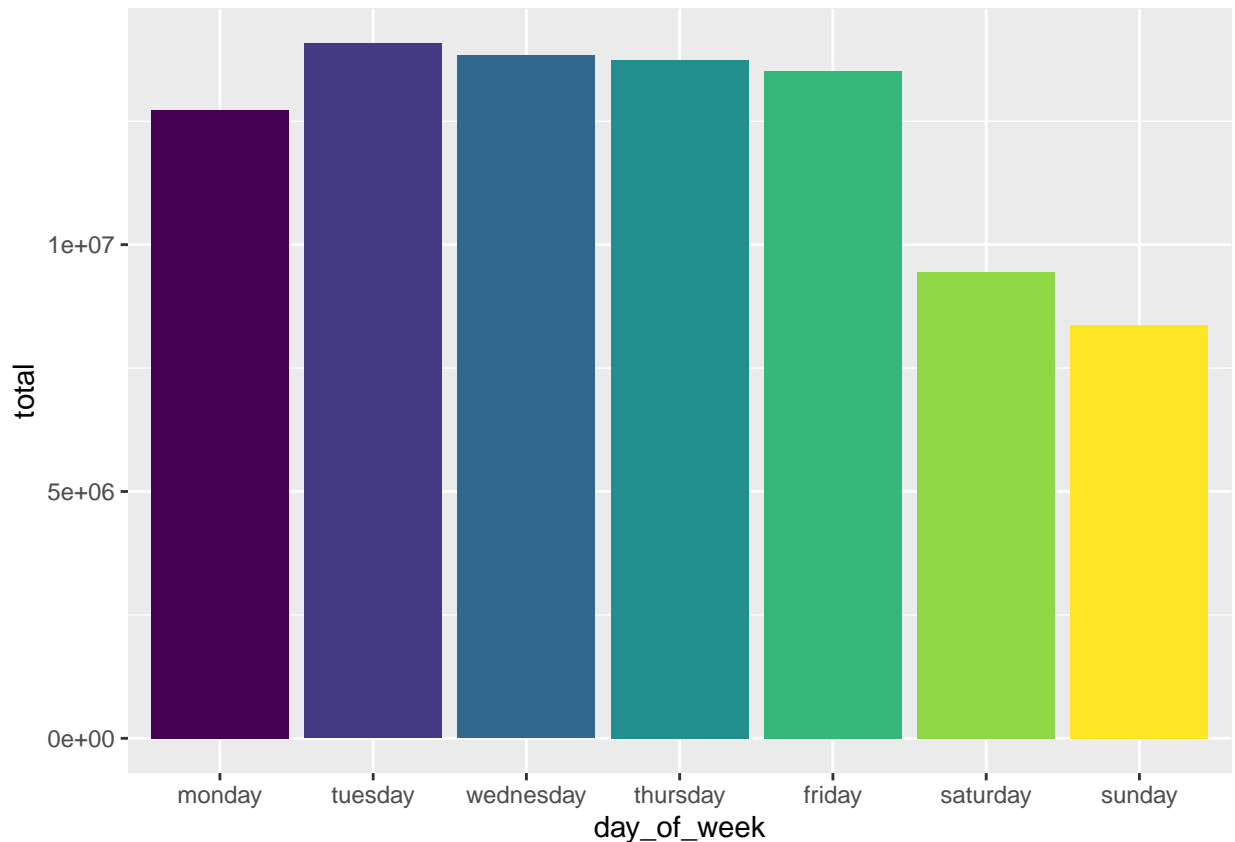
```
month_names <- c("january", "february", "march", "april", "may", "june", "july", "august", "september", "october", "november", "december")
```

```
day_names <- c("monday", "tuesday", "wednesday", "thursday", "friday", "saturday", "sunday")
```

```
birth <- births_combine %>%
  mutate(month = factor(month, labels = month_names, ordered = TRUE)) %>%
  mutate(day_of_week = factor(day_of_week, labels = day_names, ordered = TRUE)) %>%
  mutate(weekend = ifelse(day_of_week %in% c("saturday", "sunday"), TRUE, FALSE))
```

### Make a bar plot

```
total_births_weekdays <- birth %>%  
  group_by(day_of_week) %>%  
  summarize(total = sum(births))  
  
total_births_weekdays %>%  
  ggplot(aes(x = day_of_week, y = total, fill = day_of_week)) +  
  geom_col() +  
  guides(fill = "none")
```



### Make a point plot

```
total_births_weekdays %>%  
  mutate(total = total/1000000) %>%  
  ggplot(aes(day_of_week, total)) +  
  geom_point(color = "#1ed994",  
            size = 5) +  
  labs(x = "Day of Week",  
       y = "Total (Million)") +  
  expand_limits(y = c(0, 15)) +  
  theme_minimal() +  
  theme(axis.text.x = element_text(angle = 45))
```

