EDLD652 Lab 1

Hyeonjin Cha Rachael Latimer Tess Sameshima

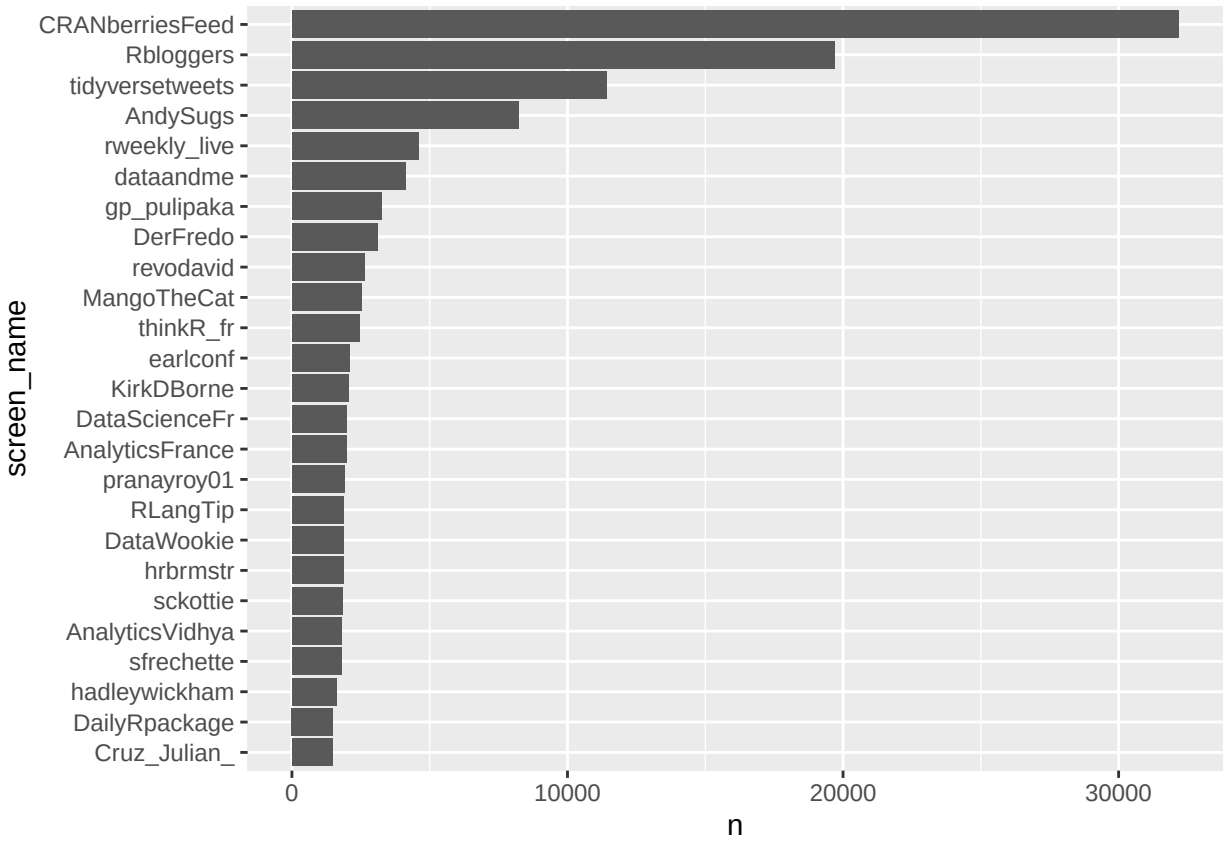
1/18/2021

Part 1. Load The Dataset

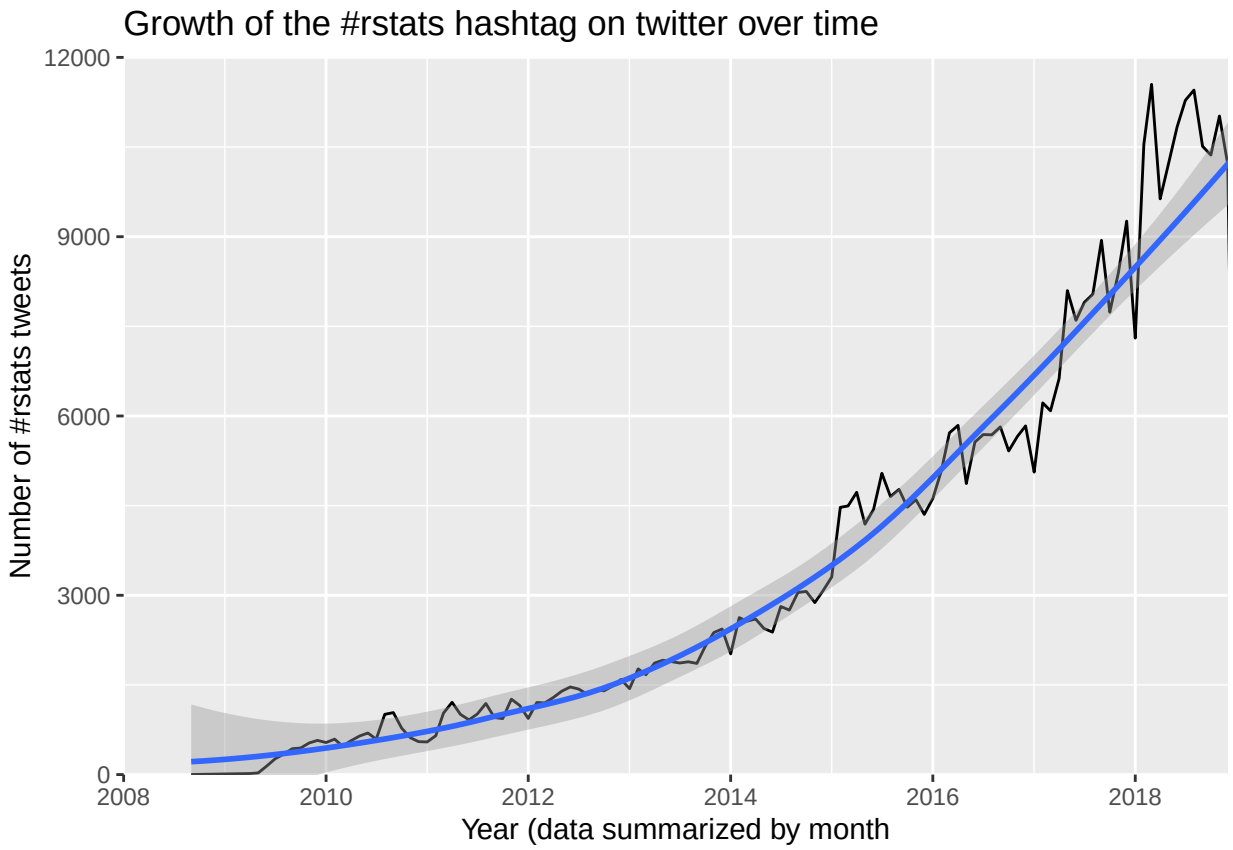
```
rstats <- import(here("data", "rstats_tweets.rds")) %>%
  clean_names() %>%
  as_tibble()
# glimpse(rstats)
# generally these sorts of things should not be printed in an RMarkdown doc
# unless you're trying to actually communicate the info with your audience
```

Part 2. Draft Plots

```
# In the below, I've collapsed your code into a single pipeline. I wouldn't
# recommend always doing this, but when the pipeline is small like this I think
# it helps make it more clear
data_plot1 <- rstats %>%
  count(screen_name, sort = TRUE) %>%
  mutate(screen_name = factor(screen_name),
         screen_name = fct_reorder(screen_name, n))  %>%
  slice_head(n = 25)

#how do we make sure this doesn't make up most of the knitted PDF? echo or eval won't do the job...

# In the above question, do you mean the figure is too large? You can adjust
# this with the chunk options `fig.width` and `fig.height`. The default is
# 7 inches for both. Also, you can make the size of the output vary depending on
# whether it's rendered to html and pdf. You would do this by having two chunks,
# one for each output, and include them conditionally on how it's rendered.
# see https://bookdown.org/yihui/rmarkdown-cookbook/latex-html.html for an
# example

# data_plot1 <- top25%>%
#   slice_head(n=25)

ggplot(data_plot1, aes(screen_name, n)) +
  geom_col() +
  coord_flip()
```

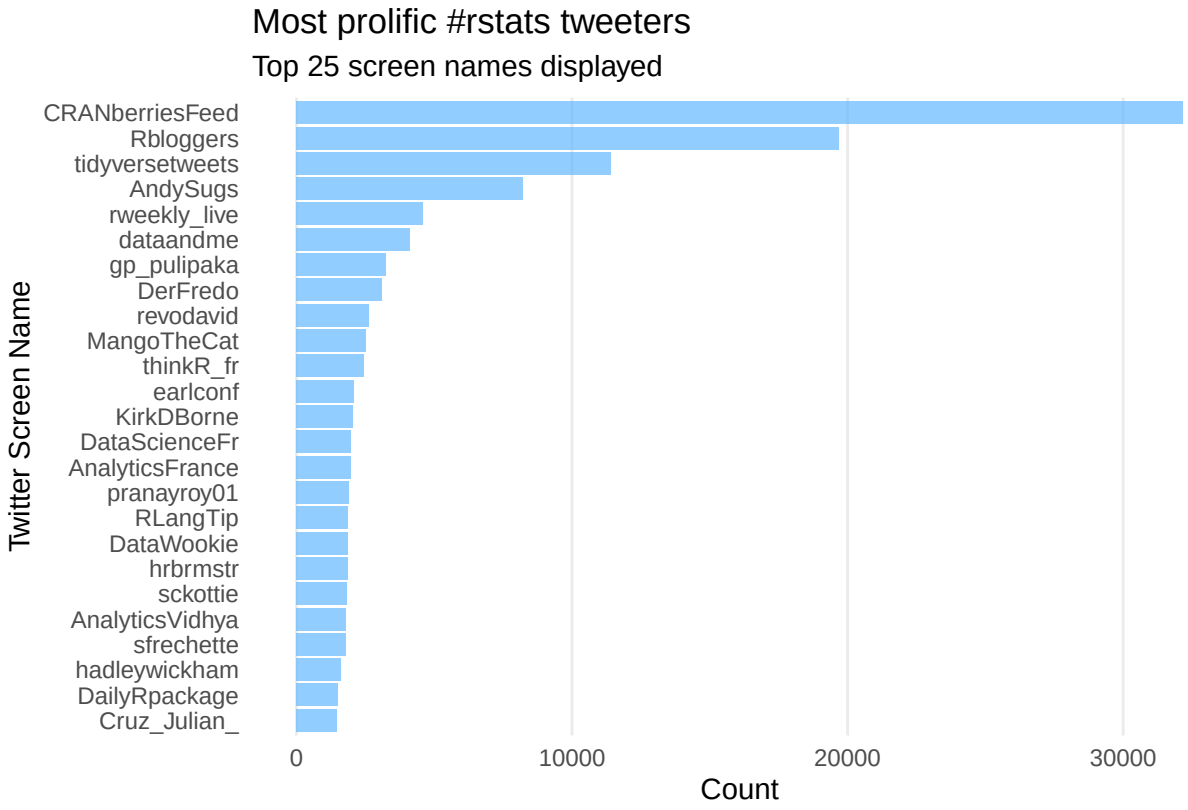



```
#create month variable from created_at column and cout the row numbers

data_plot2  <-
  rstats %>%
    count(month = round_date(created_at, "month"))

# You don't need to namespace lubridate below because you loaded it above
#plot draft of plot 2
ggplot(data_plot2, aes(month, n)) +
  geom_line() +
  geom_smooth() +
  xlab("Year (data summarized by month") +
  ylab("Number of #rstats tweets") +
  labs(title = "Growth of the #rstats hashtag on twitter over time") +
  coord_cartesian(xlim = as_datetime(c("2008-01-01","2018-12-01")),
                  ylim = c(0, 12000),
                  expand = FALSE)
```



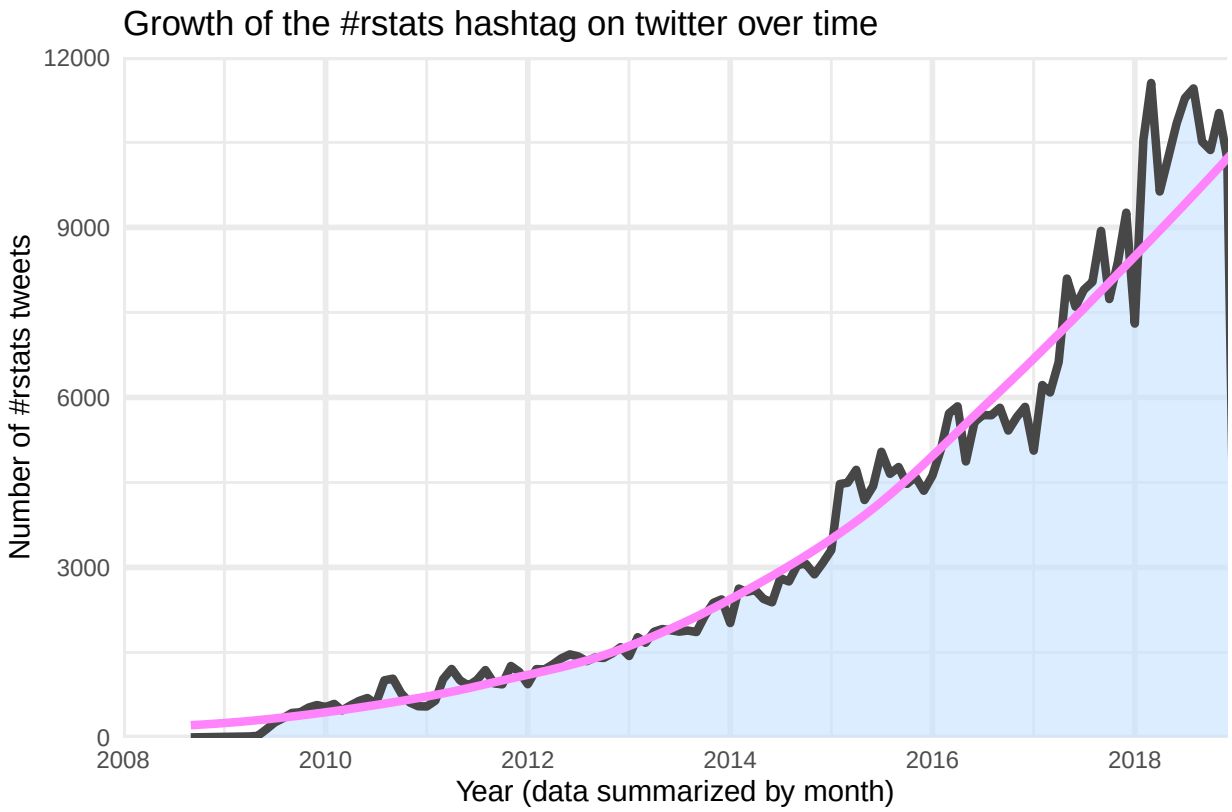
Part 3. Refinied Plots

```
#refined plot 1
plot1_refined <- data_plot1 %>%
  ggplot(aes(screen_name, n)) +
  geom_col(fill = "steelblue1", alpha = 0.7) +
  coord_flip() +
  labs(title = "Most prolific #rstats tweeters",
       subtitle = "Top 25 screen names displayed",
       x = "Twitter Screen Name",
       y = "Count",
       caption = "Data from Mike Kearny, distributed via #tidytuesday") +
  theme_minimal() +
  theme(panel.grid.minor = element_blank(),
        panel.grid.major.y = element_blank())
plot1_refined
```



plots-1.pdf

```
#refined plot 2
plot2_refined <- data_plot2 %>%
  ggplot(aes(x = month, y = n)) +
  geom_area(fill = "slategray1", alpha = 0.6) +
  geom_line(color = "gray28", size = 1.5) +
  geom_smooth(se = FALSE, color = "orchid1", size = 1.5) +
  coord_cartesian(
    xlim = as_datetime(c("2008-01-01", "2018-12-01")),
    ylim = c(0, 12000),
    expand = FALSE) +
  labs(title = "Growth of the #rstats hashtag on twitter over time",
       x = "Year (data summarized by month)",
       y = "Number of #rstats tweets",
       caption = "Data from Mike Kearny, distributed via #tidytuesday") +
  theme_minimal() +
  theme(panel.grid.major = element_line(size = 1),
        panel.grid.minor = element_line(size = 0.5))
plot2_refined
```



plots-2.pdf

```
# Notice I reordered the layers above
```