

# graphs.R

Glenn Tattersall

2019-03-27

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.2.1 --

## v ggplot2 3.1.0      v purrr  0.3.2
## v tibble  2.1.1      v dplyr  0.8.0.1
## v tidyr   0.8.3      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.4.0

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

factor_order = c('SPSS', 'R', 'SAS', 'STATA') # Needs to be common in both datasets
colors = c('red', 'darkgreen', 'blue', 'magenta', 'black', 'orange')

#####
# VISUALIZE SCHOLAR DATA #
#####

# Load the data. C for "citations"
C = read.csv('citations.csv') %>%
  filter(year >= 2010) %>%
  mutate(software = fct_relevel(software, factor_order))

# Plot
plot_citations = C %>%
  ggplot(aes(x=year, y=citations, color=software)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks=seq(1996, 2030, by=2)) +
  scale_y_continuous(breaks=seq(0, 4*10^5, 0.5*10^5), labels=scales::comma) +
  scale_colour_manual(values = colors) +
  labs(
    title = 'Scholar Citations',
    x = '',
    y = 'Citations'
  ) +
  theme_gray(13) +
  theme(
    axis.text.x = element_text(angle = 90, hjust = 1)
  )
#plot_citations

#####
# VISUALIZE GOOGLE TRENDS #
```

```
#####

# Load the data. P for "popularity"
P = read.csv('trends.csv') %>%
  # Yearly summary from 2010 in long format
  separate(Month, c('year', 'month'), '-') %>%
  gather('software', 'popularity', -year, -month) %>%
  filter(year >= 2010) %>%

  # Summarise it
  group_by(year, software) %>%
  summarise(
    popularity = mean(popularity),
  ) %>%

  # A bit of tidying
  ungroup() %>%
  mutate(
    year = as.numeric(as.character(year)),
    software = fct_relevel(software, factor_order),
    popularity = popularity / 100
  )

# Plot it
plot_trends = P %>%
  ggplot(aes(x=year, y=popularity, color=software)) +
  geom_line() +
  geom_point() +

  # Appearance stuff
  scale_x_continuous(breaks=seq(2010, 2030, by=2)) +
  scale_y_continuous(breaks=seq(0, 1, 0.2), labels = scales::percent_format(1)) +
  scale_colour_manual(values = colors) +
  labs(
    title = 'Google Trends',
    x = '',
    y = 'Relative search proportion'
  ) +
  theme_gray(13) +
  theme(
    legend.position = "none", # Remove legend
    axis.text.x = element_text(angle = 90, hjust = 1)
  )
#plot_trends

#####
# All together now #
#####
library(patchwork)
plot_trends +
  plot_spacer() +
```

```
plot_citations +
plot_layout(widths=c(0.48, 0.03, 0.51))
```

