

Analysis of life in the Good Place

Final project— PMAP 8551, Fall 2024

Eleanor Shellstrop

December 16, 2024

Table of contents

Executive summary	1
Data background	2
Data cleaning	2
Individual figures	3
Figure 1: Lollipop chart	3
Figure 2: Changes over time	4
Figure 3: Relationships	5
Final figure	7

Executive summary

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur

aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt. Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim ad minima veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur? Quis autem vel eum iure reprehenderit qui in ea voluptate velit esse quam nihil molestiae consequatur, vel illum qui dolorem eum fugiat quo voluptas nulla pariatur?

Data background

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Data cleaning

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

```
library(tidyverse)
library(scales) # For nicer scales

set.seed(1234) # Make all random draws the same
example_data <- tibble(x1 = rnorm(10000),
                      x2 = rnorm(10000),
                      y1 = sample(1:100, 10000, replace = TRUE),
                      y2 = sample(LETTERS[1:4], 10000, replace = TRUE),
                      y3 = sample(LETTERS[10:11], 10000, replace = TRUE),
                      year = sample(2010:2017, 10000, replace = TRUE)) |>
  arrange(y2, year)

# write_csv(example_data, "data/example_data.csv")
```

To make life easier, I created a custom ggplot theme that I can use in all my figures:

```
my_beautiful_fancy_theme <- theme_minimal(base_family = "Source Sans Pro") +
  theme(legend.position = "bottom",
        panel.background = element_rect(fill = "transparent", colour = NA),
        plot.background = element_rect(fill = "transparent", colour = NA),
        axis.title.x = element_text(margin = margin(t = 15)),
        axis.title.y = element_text(margin = margin(r = 15)),
        strip.text = element_text(family = "Source Sans Pro", face = "bold",
                                   size = rel(1.3)))
```

Individual figures

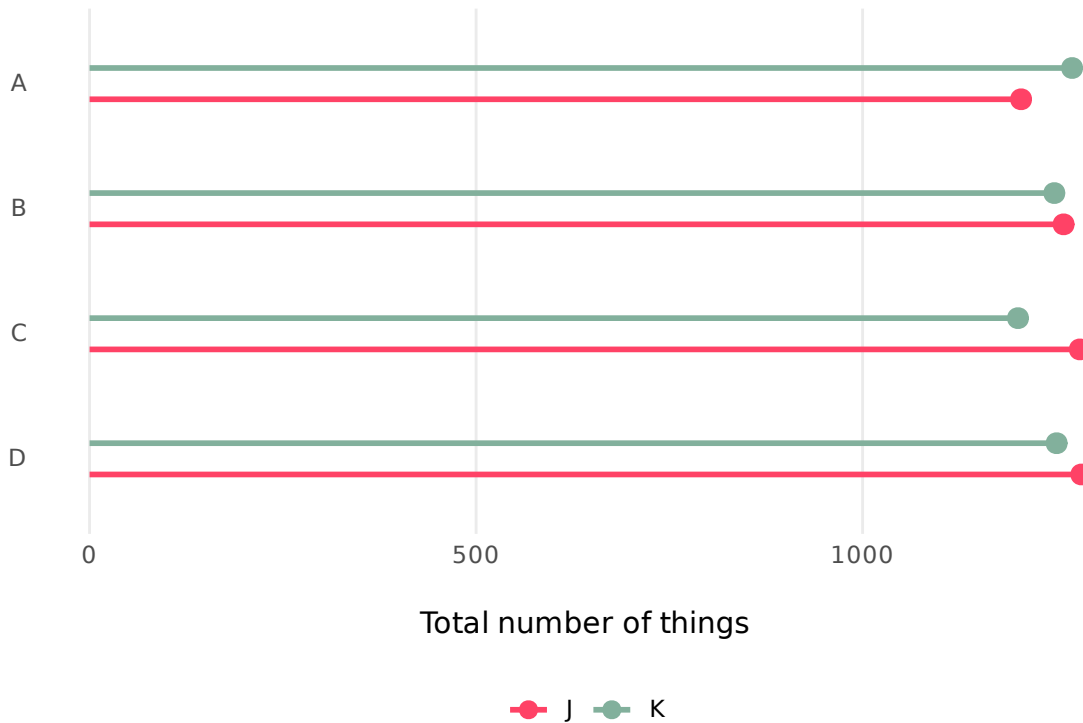
Figure 1: Lollipop chart

First, I was interested in blah because blah, so I created a lollipop chart to show blah. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

```
example_data_summarized <- example_data |>
  group_by(y2, y3) |>
  summarize(n = n())

figure1 <- ggplot(example_data_summarized, aes(x = n, y = fct_rev(y2), color = y3)) +
  geom_pointrange(aes(xmin = 0, xmax = n), position = position_dodge(width = 0.5),
                 linewidth = 1, fatten = 5) +
  labs(x = "Total number of things", y = NULL) +
  guides(color = guide_legend(title = NULL)) +
  scale_color_manual(values = c("#FF4266", "#82B09C")) +
  my_beautiful_fancy_theme +
  theme(panel.grid.minor = element_blank(),
        panel.grid.major.y = element_blank())

figure1
```



```
ggsave(figure1, filename = "output/figure1.pdf", device = cairo_pdf,
       width = 6, height = 4, units = "in", bg = "transparent")
```

Figure 2: Changes over time

Next, I wanted to see how things have changed over time, so I created a blah because blah. I found blah. Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

```
example_data_time <- example_data |>
  pivot_longer(cols = c(x1, x2), names_to = "x_names", values_to = "value") |>
  group_by(x_names, year, y2) |>
  summarize(x_avg = mean(value),
            error = sd(value) / sqrt(length(value))) |>
  ungroup() |>
  mutate(upper = x_avg + (1.96 * error),
         lower = x_avg - (1.96 * error)) |>
  mutate(x_names = recode(x_names,
```

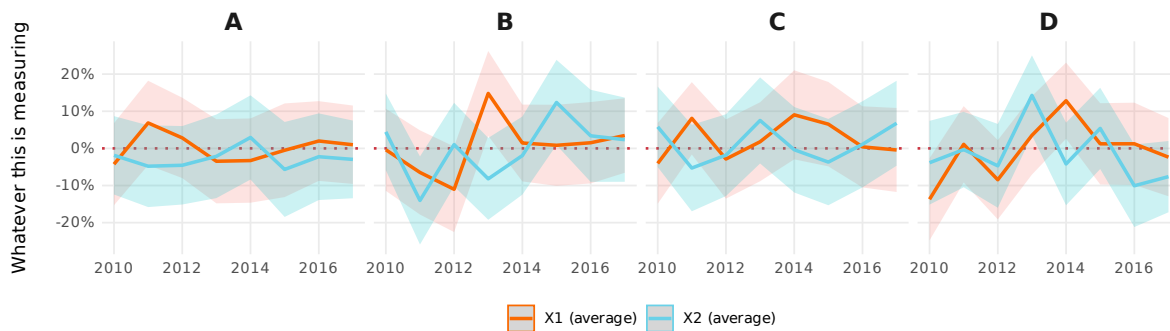
```

x1 = "X1 (average)",
x2 = "X2 (average)")

figure2 <- ggplot(example_data_time, aes(x = year, y = x_avg, color = x_names)) +
  geom_hline(yintercept = 0, linewidth = 0.75, color = "#CC3340", linetype = "dotted") +
  geom_ribbon(aes(ymin = lower, ymax = upper, fill = x_names, color = NULL), alpha = 0.2) +
  geom_line(linewidth = 1) +
  scale_color_manual(values = c("#FA6900", "#69D1E8")) +
  scale_y_continuous(labels = label_percent()) +
  guides(color = guide_legend(title = NULL), fill = "none") +
  labs(x = NULL, y = "Whatever this is measuring") +
  facet_wrap(~ y2, nrow = 1) +
  my_beautiful_fancy_theme +
  theme(panel.grid.minor = element_blank())

```

figure2



```

ggsave(figure2, filename = "output/figure2.pdf", device = cairo_pdf,
  width = 16, height = 3, units = "in", bg = "transparent")

```

Figure 3: Relationships

I was also interested in the relationship between blah and blah, so I blahed. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

```

# There are a lot of points here and they're all random and pointless, so I
# simplify this graphic by just taking a subset of them
example_data_subset <- example_data |>

```

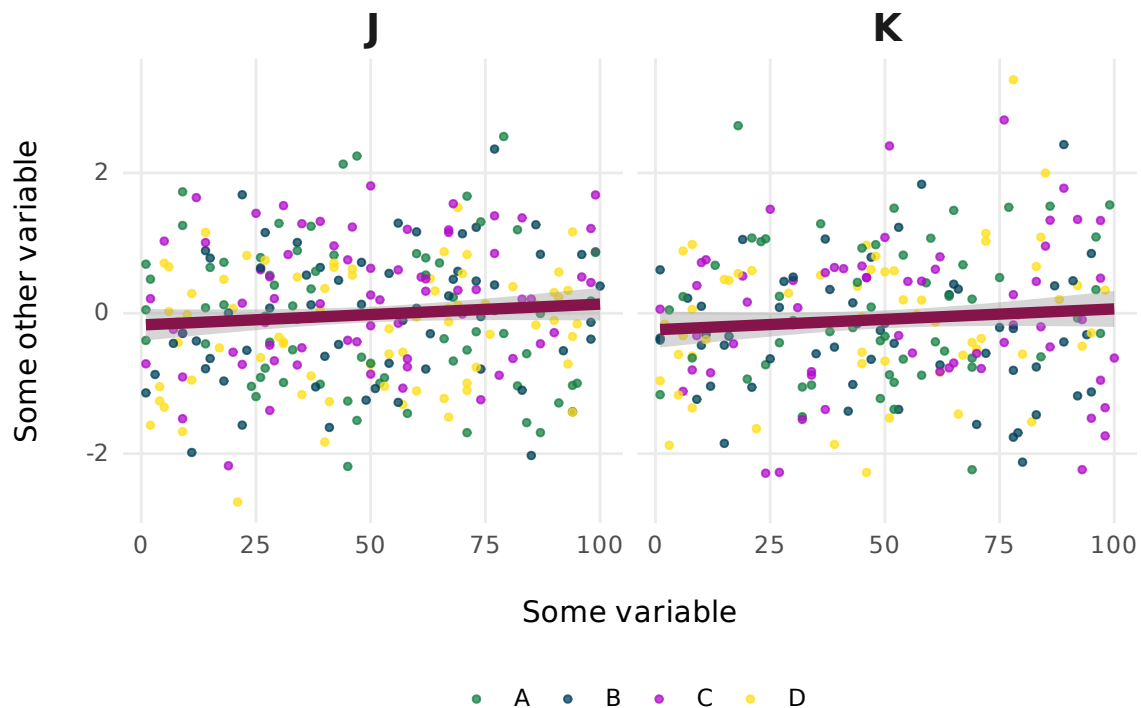
```

sample_n(500)

figure3 <- ggplot(example_data_subset, aes(x = y1, y = x2, color = y2)) +
  geom_point(size = 1, alpha = 0.75) +
  geom_smooth(method = "lm", color = "#85144A", linewidth = 2) +
  labs(x = "Some variable", y = "Some other variable") +
  guides(color = guide_legend(title = NULL)) +
  scale_color_manual(values = c("#188146", "#004259", "#B00DC9", "#FFE01C")) +
  facet_wrap(~ y3) +
  my_beautiful_fancy_theme +
  theme(panel.grid.minor = element_blank())

```

figure3



```

ggsave(figure3, filename = "output/figure3.pdf", device = cairo_pdf,
  width = 6, height = 4, units = "in", bg = "transparent")

```

Final figure

I took these three graphs and combined them and enhanced them in Illustrator. I chose the colors, fonts, alignment, etc. because blah and the final figure represents truth because of blah.

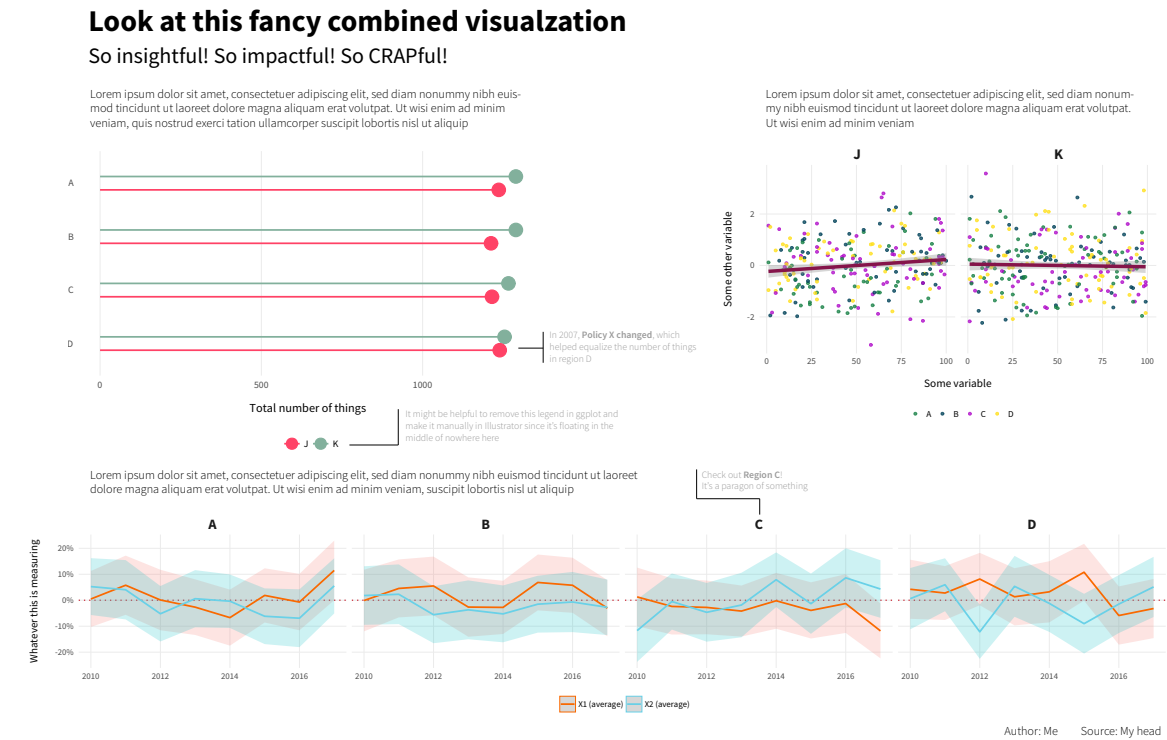


Figure 1: Final fancy visualization