Data Exploration: Contextual Influences

Your name here

November 11, 2021

In this Data Exploration assignment we will again be exploring the Nationscape dataset (Tausanovitch and Vavreck 2020), which was used in Reny and Newman's (2021) study of the effecs of the protests after George Floyd's killing.

Unlike previous assignments, however, you will be asked to take a bigger role in defining the research question and identifying the specific data that you would need to use. *This is practice for operationalizing questions of the type you will do for your research project.*

Throughout the assignment, we will provide a running example of how you might approach the tasks. For you own work, please do not use either this example or the Geroge Floyd protests.

Note: Because this assignment is a bit different, you are require to do all of the questions (although non-data science students can skip question 7). This is to ensure that you have enough material for you blog post.

If you have a question about any part of this assignment, please ask! Note that the actionable part of each question is **bolded**.

Developing a Research Question about Contextual Influences

Data Details:

- File Name: `vars_data.xlsx`

- Source: This file shows what variables are covered in each wave of the Nationscape Data Set (Tausanovitch and Vavreck 2020). We will be using data from the survey itself in other parts of the exercise, but which specific files and variables will be up to you! Therefore, we don't present them in depth here.

Variable Name	Variable Description
Date	The date of the wave of the Nationscape survey
response_id	This and all other variables are the names of variables in the Nationscape data; the cells are 1 if that variable was included in that week's survey and 0 otherwise

```
#Load the data summarizing variable availability
NationscapeVars_1 <- read_xlsx('vars_data.xlsx',sheet = 1) #we're using the read_xlsx function from the
NationscapeVars_2 <- read_xlsx('vars_data.xlsx',sheet = 2)
```

Now let's get the data from two sheets into one data set.

```
NationscapeVars <- full_join(NationscapeVars_1,NationscapeVars_2) %>% # the full_join function keeps al
  replace(is.na(.),0) # since we know that the NAs generated in the last command weren't asked in the w
```

Question 1

Contextual influences are all about the fleeting events that shift our attitudes and behavior. These can be something we personally experience, like encountering people on the street (Sands 2017) or voting at a school (Berger et al. 2008). But they can also be events we are exposed to by press coverage like Supreme Court decision (Tankard and Paluck 2017) or even emotions evoked by press coverage (as was experimentally modeled by Zeitzoff 2014). For this exercise we will think about events that people in a given state or across the country would plausibly have been exposed to via news coverage. **Think about events that happened between July 2019 and July 2020. Maybe this is something that made national news or maybe it was something that received a lot of coverage in your home state or region. Write down an example or two that you might be interested in considering. Use Google Trends to confirm that there was a spike in interest, as demonstrated by an increase in Google searches, in your event and include a screenshot or a hyperlink to your results.** Try entering a relevant search term and then using a "Custom time range" (one of the drop down options instead of the default "Past 12 months") to make your visualization.

Covid: https://trends.google.com/trends/explore?q=covid AAPI hate: https://trends.google.com/trends/explore?date=all&geo=US&q=aapi%20hate Amazon fires: https://trends.google.com/trends/explore?date=all&geo=US&q=amazon%20fires Trump impeachment: https://trends.google.com/trends/explore?date=all&geo=US&q=impeachment
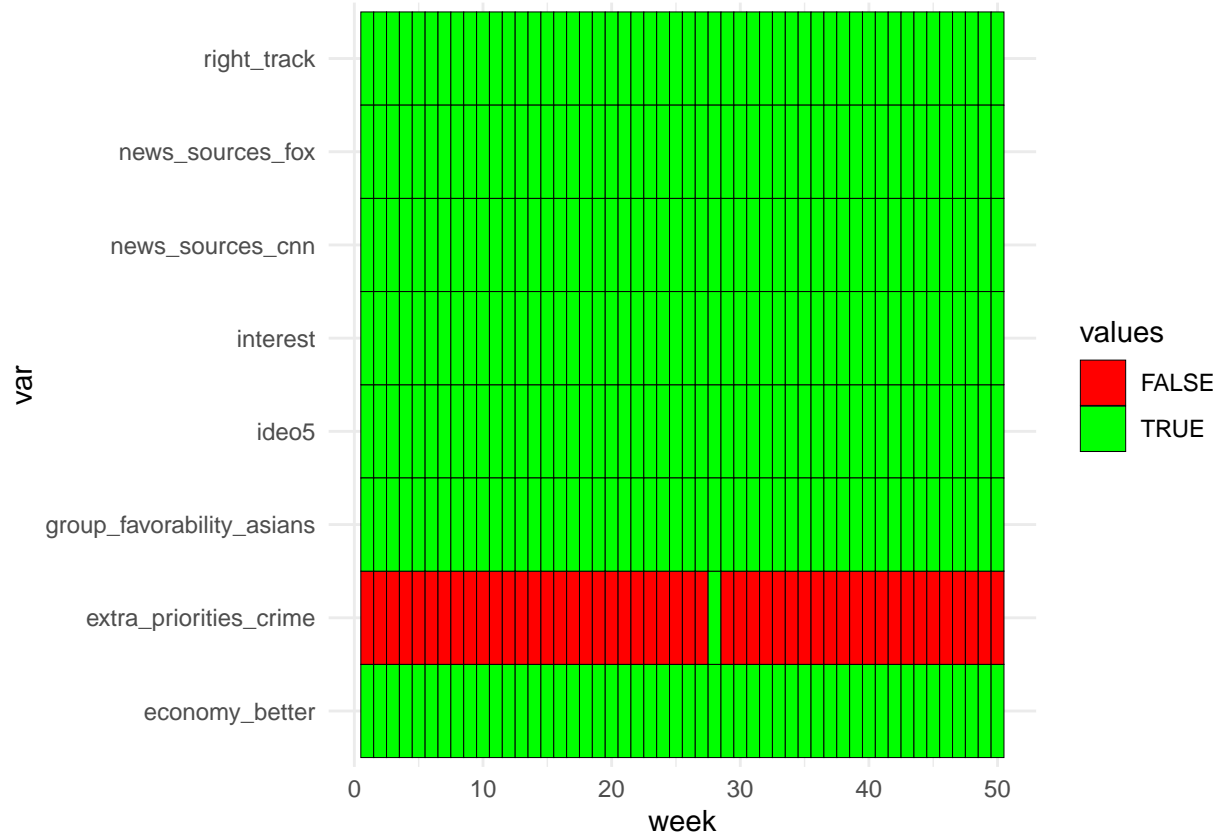
Question 2

Think about some outcomes of interest to you that might have been affected by the contextual influence of the event that you chose. Look in the Nationscape data for variables that fit your outcomes or are reasonable proxies for those outcomes. The variable names in the data you have loaded are pretty informative, but use the full data folder you downloaded earlier to look in the codebooks for more complete descriptions of the variables and how they are measured. There is a codebook in each week's folder; you can look at any week's codebook to get a sense of the variables that are common across the survey waves. **Make sure that your variables are present in the data for the time period in which you want to look for contextual effects. Present these results in a plot.**

Heatmap

```
heat_data_2 <- NationscapeVars %>%
  mutate(across(.cols = everything(), as.logical)) %>%
  select(ideo5, interest, right_track, economy_better, extra_priorities_crime, group_favorability_asian
         news_sources_fox)
```

```
heat_data_2 %>%
  mutate(week = row_number()) %>%
  pivot_longer(cols = ideo5:news_sources_fox, names_to = "var", values_to = "values") %>%
  ggplot(aes(x = week, y = var, fill = values)) +
    geom_tile(color = "black") +
    theme_minimal() +
    scale_fill_manual(values = c("red", "green"))
```

Question 3

Based on what you have thought about and the data you have found, clearly state a specific research question and a hypothesis. Which channel (or channels) through which situational factors can affect political behavior does your hypothesis implicate? (In class, we talked about rational choice, priming, and emotional channels.) The research question should not be obvious ahead of time (although you should have a theoretical expectation or competing hypotheses); it should be be descriptive, correlational, or causal in nature; and it should be answerable with the data you have available. Make sure your research question is specific; don't confuse the research question with a broader, motivating question that might be used to get people interested in your topic.

Research Question

Did the onset of the COVID pandemic lead to an increase in Favorability towards Asians? I'm also more broadly interested in how media consumption might have affected group favorability – using the date COVID was declared a pandemic, I'm going to use a regression discontinuity design to measure the contextual effects of COVID on group attitudes, and see how different sources of information might affect attitudes towards Asians.

If the progression of the pandemic is associated with rising Favorability towards Asians (as I'm predicting), this would be an example of how a situational factor affects political behavior through emotional channels. Media coverage of the pandemic might evoke negative feelings towards China, which could lead to more negative feelings towards Asians.

Example

Since we don't have enough data to consider crime as a national policy priority, we will focus on information search. Our research question is "Were the early August 2019 mass shootings associated with increased interest in and consumption of political news?" This might fit under a broader motivational question of "Does news coverage of violent events lead to information seeking by causing anxiety?" but we don't have the ability to answer such a broad, causal question using only the Nationscape data.

Our hypothesis is that the August 2019 mass shootings were associated with increased political interest and news consumption, especially in the states where those shootings took place. This is an example of how situational context could influence political behavior through the emotional channel, although we cannot directly test the role of anxiety.

Question 4

In academic and professional settings, peer feedback, especially early in a project, can force you to clarify your thinking and be an important source of ideas. It's also important to be able to give a quick 'elevator pitch' for your project (so named because it can be delivered in no more time than an elevator ride). We've randomly assigned you into groups to share your ideas so far and get your peers' input about sources you should read, different ways to approach your analysis, or questions about your hypotheses. **Get together in your groups, have everyone give their project's 'elevator pitch,' and gather feedback from your peers. Write at least one thing you took away from this session.** The next couple of questions will ask you to try to use the data to answer your research question and test your hypotheses, so be sure to brainstorm good ways to approach those tasks.

One problem with using the WHO declaration of the pandemic is that the pandemic is an ongoing event. Maybe instead of examining the contextual effects of a single event, it would be worth examining the severity of the pandemic over time (this would require data outside of the Nationscape survey). The data is also limited in its ability to capture the most recent/visible waves of Favorability towards Asians in the US, like the Atlanta spa shooting in March 2021 (searches for "AAPI hate" on Google spike in 2021).

Question 5

No research project exists in a vacuum. As you get ready for your final projects, we want you to practice finding, summarizing, and citing related literature. **Identify at least two academic articles that might provide some background for your research question. List the complete source citations and include links to the articles you found.** Google Scholar (https://scholar.google.com/) or Hollis (https://hollis.harvard.edu/) are good places to look for these.

References

Tessler, Hannah, Meera Choi, Grace Kao. "The Anxiety of Being Asian American: Hate Crimes and Negative Biases During the COVID-19 Pandemic." American Journal of Criminal Justice 45 (2020): 636–646. https://link.springer.com/article/10.1007/s12103-020-09541-5

Lantz, Brenan, Marin Wenger. "Are Asian Victims Less Likely to Report Hate Crime Victimization to the Police? Implications for Research and Policy in the Wake of the COVID-19 Pandemic." Crime and Delinquency (2021). https://journals.sagepub.com/doi/full/10.1177/00111287211041521

Ren, Jingnqiu, Joe Faegin. "Face mask symbolism in anti-Asian hate crimes." Ethnic and Racial Studies 44 (2020): 746-758. https://www.tandfonline.com/doi/full/10.1080/01419870.2020.1826553

Example

O'Brien and Taku (2022) find in an experiment on US undergraduates that reading news coverage about mass shootings increases anxiety. Joslyn and Haider-Markel (2018) show using survey evidence that people who experienced higher anxiety in the wake of the 2016 Orlando shooting changed their policy beliefs and perceptions of institutions. Our research examines the middle step in this causal chain: information search. Were the August 2019 mass shootings associated with increased interest in and consumption of political news?

(Your response can just be a list of articles, but feel free to expand on it as we did above if you so choose.)

References:

Joslyn, Mark R., and Donald P. Haider-Markel. "The Direct and Moderating Effects of Mass Shooting Anxiety on Political and Policy Attitudes." Research & Politics, (July 2018). https://doi.org/10.1177/2053168018794060.

O'Brien, Colin, and Taku, Kanako. "Alpha and beta changes in anxiety in response to mass shooting related information." Personality and Individual Differences, Volume 186, Part A, (2022). https://doi.org/10.1016/j.paid.2021.111326.

Question 6

Read in the data from the weeks surrounding your event of interest and test your hypothesis. This can be something straightforward like a difference-in-means or you can plot a visualization of the data. Just take one of the approaches we have used in class before to get an initial sense for if the data provide evidence of the contextual effects you theorized. Note that you might have to do a fair bit of data cleaning in order to do this. Pay particular attention to how missing data are coded.

```r
# Reading data with Yao's code in Slack

# Collect he file names and select which ones we want
file_names_1 <- list.files("Nationscape-DataRelease_WeeklyMaterials_DTA/phase_1_v20200814/") %>% .[1:24]
file_names_2 <- list.files("Nationscape-DataRelease_WeeklyMaterials_DTA/phase_2_v20200814/") %>% .[1:26]

# Reading in all phase 2 weeks
phase_1 <- map_dfr(.x = file_names_1,
                   ~read_dta(file = str_c("Nationscape-DataRelease_WeeklyMaterials_DTA/phase_1_v2020081
                     remove_all_labels()) %>%
  mutate(across(.cols = everything(), ~na_if(., 999))) %>%
  mutate(across(.cols = everything(), ~na_if(., 888)))

phase_2 <- map_dfr(.x = file_names_2,
                   ~read_dta(file = str_c("Nationscape-DataRelease_WeeklyMaterials_DTA/phase_2_v2020081
                     remove_all_labels()) %>%
  mutate(across(.cols = everything(), ~na_if(., 999))) %>%
  mutate(across(.cols = everything(), ~na_if(., 888)))

full_data <- full_join(phase_1, phase_2)

groups <- c("group_favorability_asians", "group_favorability_blacks", "group_favorability_whites", "dis
bio <- c("gender", "pid3", "race_ethnicity")

# Use WHO's decleration of COVID as global pandemic as the beginning of the
# treatment. COVID declared a pandemic on March 11, 2020.
```

```r
full_data_clean <- full_data %>%
  mutate(day_running = as_numeric(date(start_date) - date("2020-03-11")),
         news = case_when(news_sources_fox < news_sources_cnn ~ "Fox",
                          news_sources_fox > news_sources_cnn ~ "CNN",
                          news_sources_fox + news_sources_cnn == 2 ~ "Both",
                          T ~ NA_character_),
         pid3 = case_when(pid3 == 1 ~ "Democrat",
                          pid3 == 2 ~ "Republican",
                          pid3 == 3 ~ "Independent",
                          T ~ NA_character_)) %>%
  select(start_date, day_running, bio, groups, news) %>%
  drop_na()
```

```r
difference_in_means(formula = group_favorability_asians ~ treated,
       data = full_data_clean %>%
          mutate(treated = ifelse(day_running < 0, F, T)) %>%
          select(group_favorability_asians, treated))
```

```
## Design:  Standard
##          Estimate  Std. Error  t value     Pr(>|t|)   CI Lower   CI Upper
## treated 0.07930809 0.004596845 17.25272 1.347873e-66 0.07029833 0.08831786
##              DF
## treated 95418.32
```

```r
cohen.d(formula = group_favorability_asians ~ treated,
       data = full_data_clean %>%
          mutate(treated = ifelse(day_running < 0, F, T)) %>%
          select(group_favorability_asians, treated))
```

```
## Warning in cohen.d.formula(formula = group_favorability_asians ~ treated, :
## Cohercing rhs of formula to factor
```

```
##
## Cohen's d
##
## d estimate: -0.09738908 (negligible)
## 95 percent confidence interval:
##       lower       upper
## -0.10819939 -0.08657877
```

Difference in means shows there is a significant difference in favorability towards asians before and after the beginning of the pandemic, but the effect size is negligible.

```r
difference_in_means(formula = group_favorability_asians ~ treated,
       data = full_data_clean %>%
          mutate(treated = ifelse(day_running < 0, F, T)) %>%
          filter(news == "Fox") %>%
          select(group_favorability_asians, treated))
```

```
## Design:  Standard
```

```
##          Estimate Std. Error  t value     Pr(>|t|)  CI Lower  CI Upper
## treated 0.09985013 0.00916979 10.88903 1.522157e-27 0.08187671 0.1178235
##            DF
## treated 22591.22
```

```r
cohen.d(formula = group_favorability_asians ~ treated,
        data = full_data_clean %>%
          mutate(treated = ifelse(day_running < 0, F, T)) %>%
          filter(news == "Fox") %>%
          select(group_favorability_asians, treated))
```

```
## Warning in cohen.d.formula(formula = group_favorability_asians ~ treated, :
## Cohercing rhs of formula to factor
```

```
##
## Cohen's d
##
## d estimate: -0.1226388 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.1441606 -0.1011169
```

```r
difference_in_means(formula = group_favorability_asians ~ treated,
        data = full_data_clean %>%
          mutate(treated = ifelse(day_running < 0, F, T)) %>%
          filter(pid3 == "Republican") %>%
          select(group_favorability_asians, treated))
```

```
## Design:  Standard
##          Estimate Std. Error  t value     Pr(>|t|)  CI Lower  CI Upper
## treated 0.1044672  0.0080659 12.95171 2.812944e-38 0.08865775 0.1202766
##            DF
## treated 35086.77
```

```r
cohen.d(formula = group_favorability_asians ~ treated,
        data = full_data_clean %>%
          mutate(treated = ifelse(day_running < 0, F, T)) %>%
          filter(pid3 == "Republican") %>%
          select(group_favorability_asians, treated))
```

```
## Warning in cohen.d.formula(formula = group_favorability_asians ~ treated, :
## Cohercing rhs of formula to factor
```

```
##
## Cohen's d
##
## d estimate: -0.1230223 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.1412979 -0.1047466
```

Filtering for respondents whose only news source is Fox or respondents who indicate they're Republican
doesn't change the results.

Question 7: DATA SCIENCE QUESTION

Extend your work from the previous question to consider other factors, like the possibility of heterogen ous treatment effects, confounding variables, or use a more sophisticated approach to statistical inference, like regression discontinuity in time.

```r
# group_favorability_asians for all groups

full_data_clean %>%
  group_by(day_running) %>%
  summarize(group_favorability_asians = mean(group_favorability_asians),
            .groups = "drop") %>%
  mutate(treatment = fct_relevel(as.factor(ifelse(day_running < 0, "Before", "After")),
                                 "Before", "After")) %>%
  ggplot(aes(x = day_running, y = group_favorability_asians)) +
    geom_point(aes(color = treatment), alpha = 0.5, size = 0.5) +
    geom_smooth(aes(color = treatment), method = "lm", se = F) +
    theme_light() +
    scale_color_manual(name = "",
                       values = c("Before" = "#00bfc4", "After" = "#f8766d")) +
    labs(title = "Favorability towards Asians over time") +
    geom_vline(xintercept = 0, linetype = "dashed")
```

`geom_smooth()` using formula 'y ~ x'
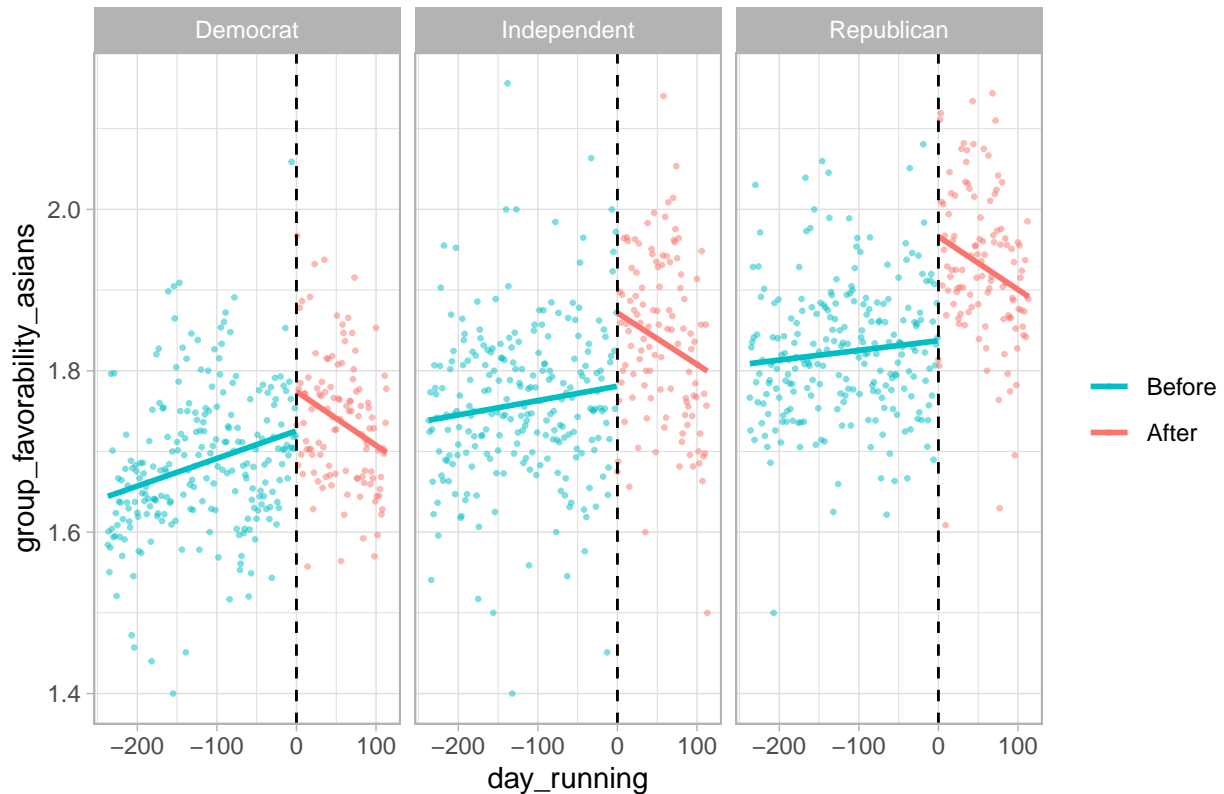

Favorability towards Asians over time
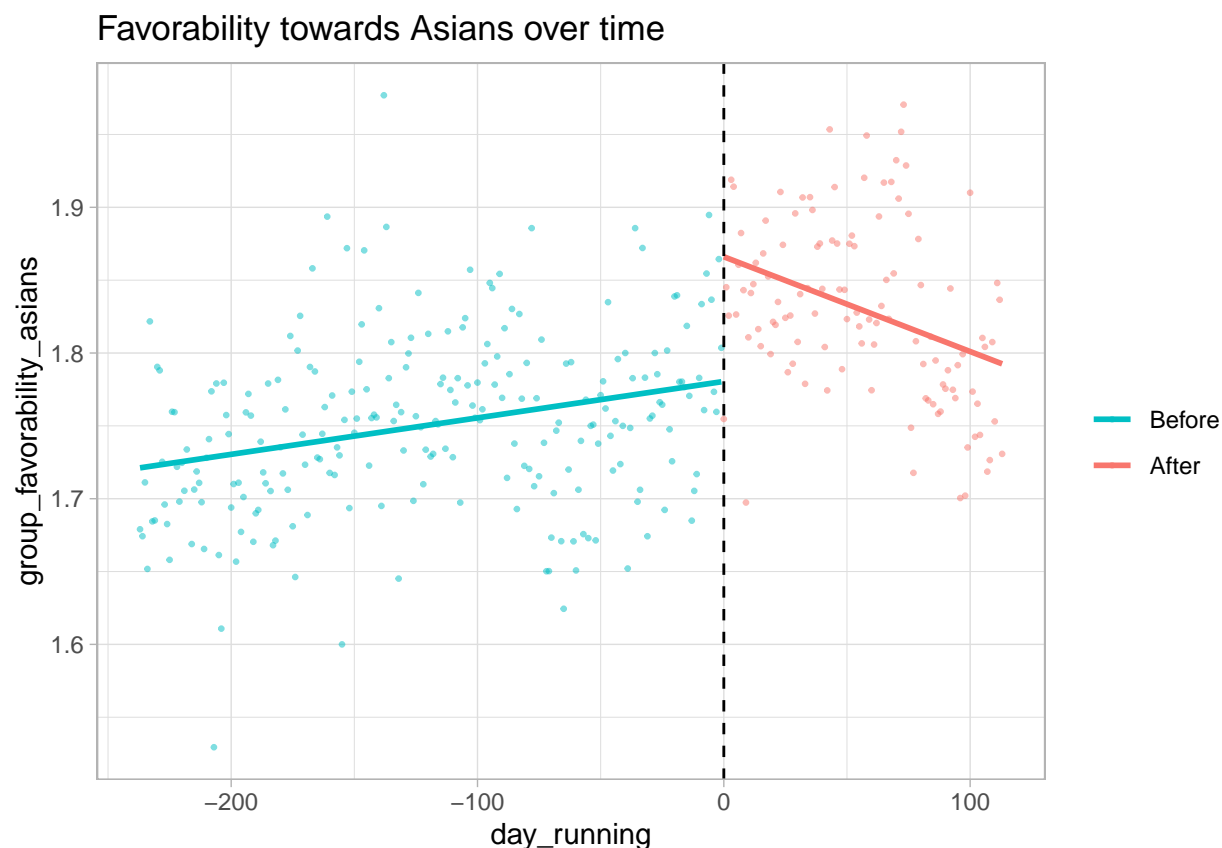
```r
#group_favorability_asians by party

full_data_clean %>%
  group_by(day_running, pid3) %>%
  summarize(group_favorability_asians = mean(group_favorability_asians),
            .groups = "drop") %>%
  mutate(treatment = fct_relevel(as.factor(ifelse(day_running < 0, "Before", "After")),
                                 "Before", "After")) %>%
  ggplot(aes(x = day_running, y = group_favorability_asians)) +
    geom_point(aes(color = treatment), alpha = 0.5, size = 0.5) +
    geom_smooth(aes(color = treatment), method = "lm", se = F) +
    facet_wrap("pid3") +
    theme_light() +
    scale_color_manual(name = "",
                       values = c("Before" = "#00bfc4", "After" = "#f8766d")) +
    labs(title = "Favorability towards Asians by party") +
    geom_vline(xintercept = 0, linetype = "dashed")
```

`geom_smooth()` using formula 'y ~ x'

Favorability towards Asians by party

```r
#group_favorability_asians by news source

full_data_clean %>%
  group_by(day_running, news) %>%
  summarize(group_favorability_asians = mean(group_favorability_asians),
```
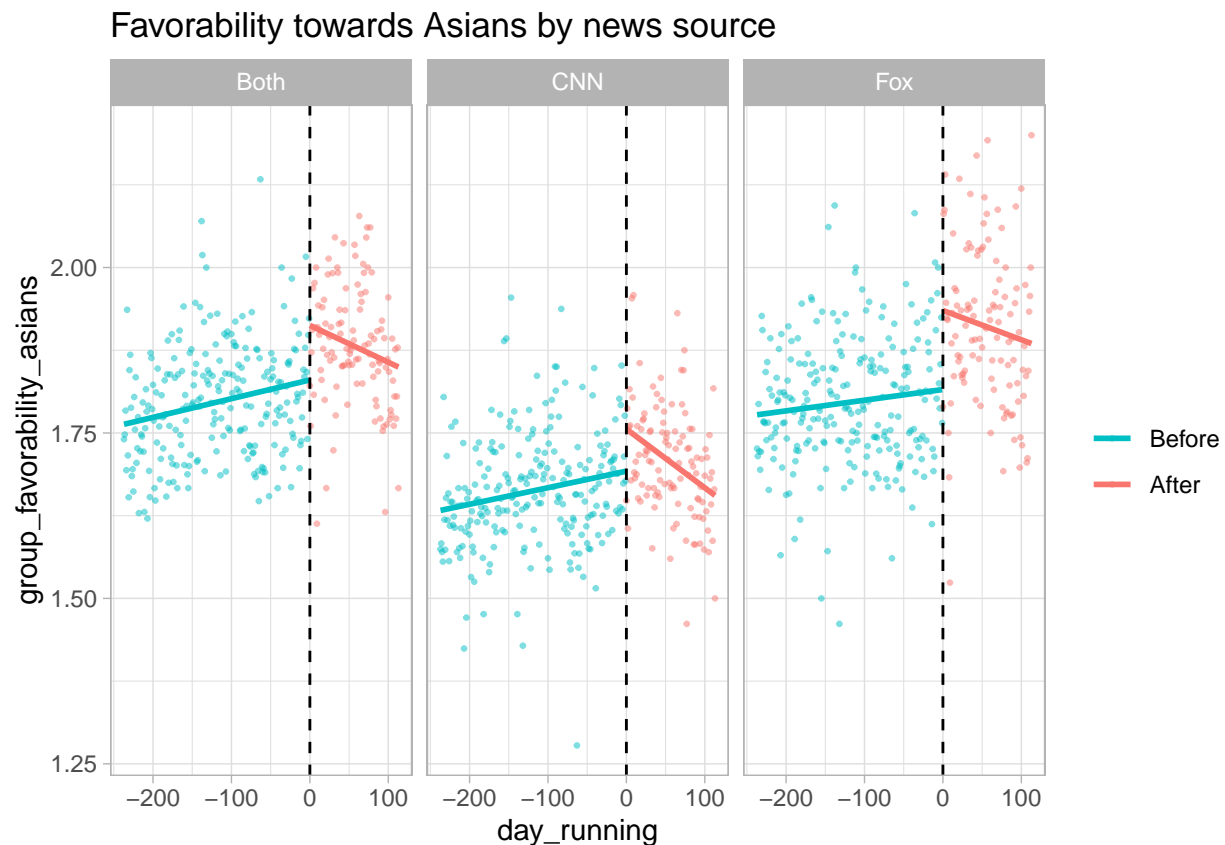
```
                      .groups = "drop") %>%
  mutate(treatment = fct_relevel(as.factor(ifelse(day_running < 0, "Before", "After")),
                                 "Before", "After")) %>%
  ggplot(aes(x = day_running, y = group_favorability_asians)) +
    geom_point(aes(color = treatment), alpha = 0.5, size = 0.5) +
    geom_smooth(aes(color = treatment), method = "lm", se = F) +
    facet_wrap("news") +
    theme_light() +
    scale_color_manual(name = "",
                       values = c("Before" = "#00bfc4", "After" = "#f8766d")) +
    labs(title = "Favorability towards Asians by news source") +
    geom_vline(xintercept = 0, linetype = "dashed")
```

`geom_smooth()` using formula 'y ~ x'

Favorability towards Asians by news source

```
# group_favorability_asians by racial group

full_data_clean %>%
  filter(race_ethnicity %in% c(1, 2, seq(4, 11, 1))) %>%
  mutate(race_ethnicity = case_when(race_ethnicity == 1 ~ "White",
                                    race_ethnicity == 2 ~ "Black",
                                    T ~ "Asian")) %>%
  group_by(day_running, race_ethnicity) %>%
  summarize(group_favorability_asians = mean(group_favorability_asians),
```
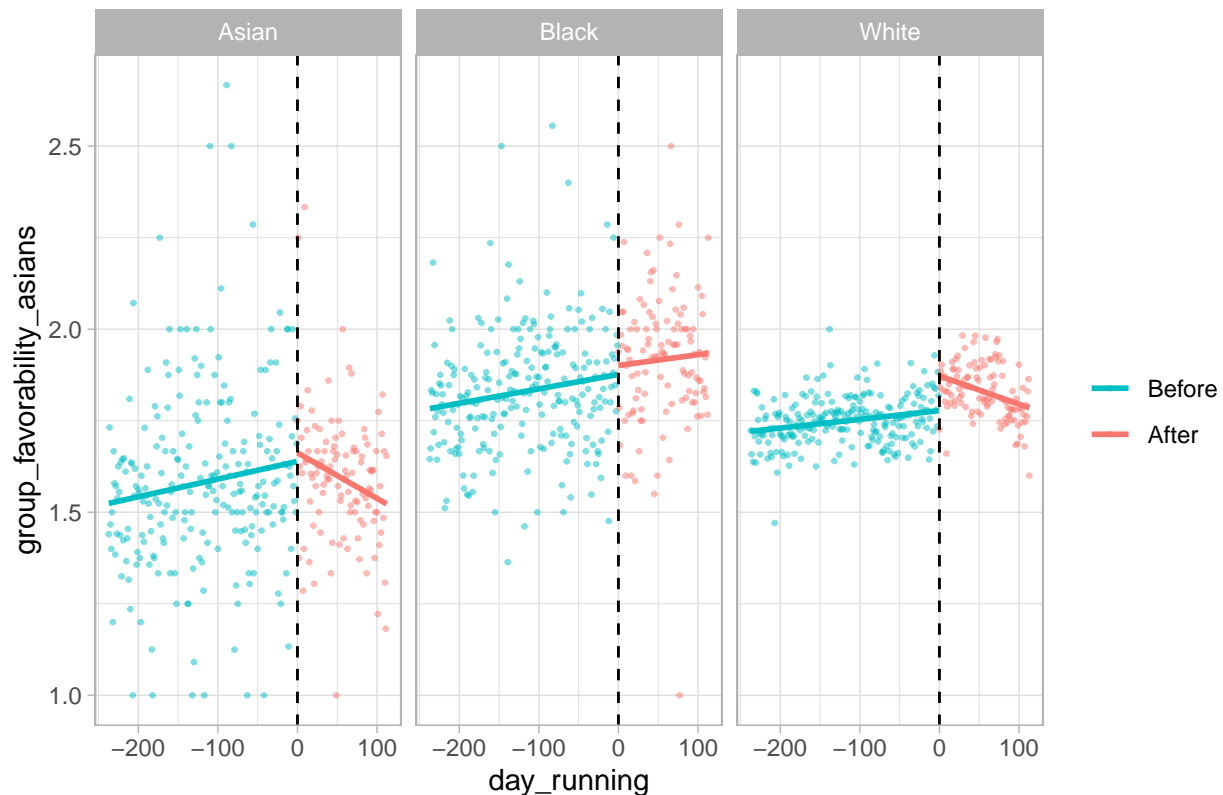
```
                group_favorability_blacks = mean(group_favorability_asians),
                .groups = "drop") %>%
  mutate(treatment = fct_relevel(as.factor(ifelse(day_running < 0, "Before", "After")),
                                 "Before", "After")) %>%
  ggplot(aes(x = day_running, y = group_favorability_asians)) +
    geom_point(aes(color = treatment), alpha = 0.5, size = 0.5) +
    geom_smooth(aes(color = treatment), method = "lm", se = F) +
    facet_wrap("race_ethnicity") +
    theme_light() +
    scale_color_manual(name = "",
                       values = c("Before" = "#00bfc4", "After" = "#f8766d")) +
    labs(title = "Favorability towards Asians by racial group") +
    geom_vline(xintercept = 0, linetype = "dashed")
```

`geom_smooth()` using formula 'y ~ x'

Favorability towards Asians by racial group

```
# Do Asian respondents report less favorability towards white or Black people
# after beginning of pandemic? (does subsequent rise in anti-Asian hate crimes
# evoke negative feelings towards white/Black people from Asians?)

full_data_clean %>%
  filter(race_ethnicity %in% seq(4, 11, 1)) %>%
  group_by(day_running) %>%
  summarize(group_favorability_blacks = mean(group_favorability_blacks),
```

```
              group_favorability_whites = mean(group_favorability_whites),
              .groups = "drop") %>%
  mutate(treatment = fct_relevel(as.factor(ifelse(day_running < 0, "Before", "After")),
                                 "Before", "After")) %>%
  pivot_longer(cols = c(group_favorability_blacks, group_favorability_whites),
               values_to = "values",
               names_to = "var") %>%
  ggplot(aes(x = day_running, y = values)) +
    geom_point(aes(color = treatment), alpha = 0.5, size = 0.5) +
    geom_smooth(aes(color = treatment), method = "lm", se = F) +
    facet_wrap("var") +
    theme_light() +
    scale_color_manual(name = "",
                       values = c("Before" = "#00bfc4", "After" = "#f8766d")) +
    labs(title = "Asian respondents' outgroup favorability") +
    geom_vline(xintercept = 0, linetype = "dashed")
```

```
## 'geom_smooth()' using formula 'y ~ x'
```


Asian respondents' outgroup favorability
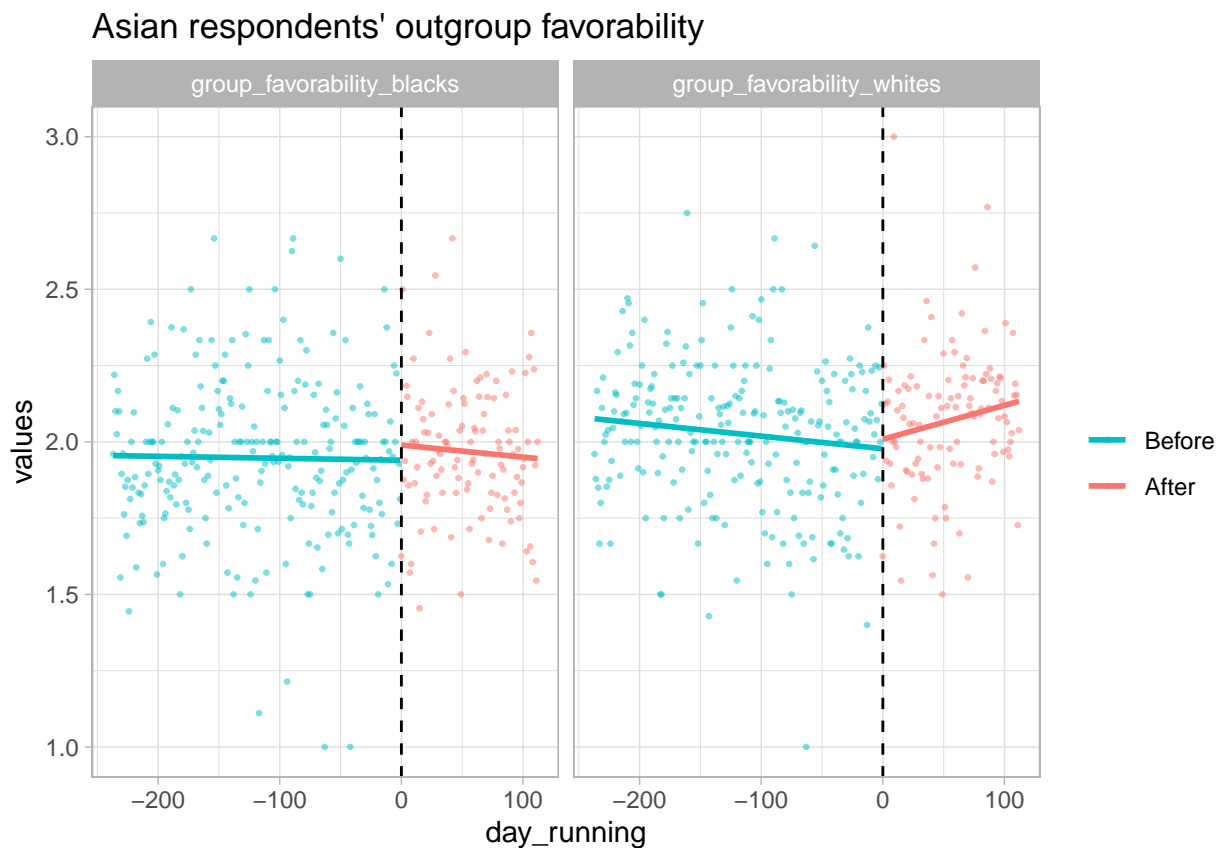
```
model_data <- full_data_clean %>%
  mutate(fox = ifelse(news == "Fox", T, F),
         treatment = ifelse(day_running < 0, F, T)) %>%
  select(day_running, pid3, race_ethnicity, fox, group_favorability_asians, treatment)
```

```r
m_total <- lm(group_favorability_asians ~ treatment + day_running + fox,
              data = model_data)

m_before <- lm(group_favorability_asians ~ day_running + fox + fox*day_running,
               data = model_data %>% filter(day_running < 0))

m_after <- lm(group_favorability_asians ~ day_running + fox + fox*day_running,
              data = model_data %>% filter(day_running >= 0))

stargazer(m_total, m_before, m_after, type = "text")
```

```
##
## ================================================================================================
##                                              Dependent variable:
##                     ----------------------------------------------------------------------------
##                                             group_favorability_asians
##                             (1)                       (2)                       (3)
## -----------------------------------------------------------------------------------------------
## treatment                0.061***
##                          (0.008)
##
## day_running             0.0001***                 0.0002***                  -0.001***
##                         (0.00004)                 (0.00004)                   (0.0001)
##
## fox                      0.077***                  0.059***                   0.099***
##                          (0.005)                   (0.011)                    (0.018)
##
## day_running:fox                                    -0.0001                    -0.0001
##                                                    (0.0001)                   (0.0003)
##
## Constant                 1.743***                  1.760***                   1.847***
##                          (0.005)                   (0.006)                    (0.009)
##
## -----------------------------------------------------------------------------------------------
## Observations             145,786                   95,608                     50,178
## R2                        0.004                     0.002                      0.003
## Adjusted R2               0.004                     0.002                      0.003
## Residual Std. Error  0.814 (df = 145782)       0.791 (df = 95604)        0.854 (df = 50174)
## F Statistic       191.291*** (df = 3; 145782) 58.981*** (df = 3; 95604) 50.562*** (df = 3; 50174)
## ================================================================================================
## Note:                                                            *p<0.1; **p<0.05; ***p<0.01
```

```r
# Define prediction range
pred_data <- data.frame(day_running = seq(-100, 300, 1))

# Define before-pandemic model for people who consume CNN or CNN and Fox.
m_before <- lm(group_favorability_asians ~ day_running, data = model_data %>% filter(day_running < 0, f

# Predict using before-pandemic model
preds_before <- predict(m_before, pred_data, se.fit = T)
preds_before <- tibble(group_favorability_asians = preds_before$fit,
                       day_running = seq(-100, 300, 1),
                       treatment = as_factor("Before"),
```

13

```
                           se = preds_before$se.fit)

# Define post-pandemic model for people who consume CNN or CNN and Fox.
m_after <- lm(group_favorability_asians ~ day_running, data = model_data %>% filter(day_running >= 0, f
preds_after <- predict(m_after, pred_data, se.fit = T)
preds_after <- tibble(group_favorability_asians = preds_after$fit,
                      day_running = seq(-100, 300, 1),
                      treatment = as_factor("After"),
                      se = preds_after$se.fit)

# Combine preds
preds_combined <- rbind(preds_before, preds_after) %>%
  mutate(treatment = fct_relevel(treatment, "Before", "After"))

# Visualize
preds_combined %>%
  ggplot(aes(x = day_running, y = group_favorability_asians)) +
    geom_ribbon(aes(ymin = group_favorability_asians - (1.96*se),
                    ymax = group_favorability_asians + (1.96*se),
                    fill = treatment),
              alpha = 0.5) +
    geom_path(aes(color = treatment)) +
    geom_vline(xintercept = 91, lty = "dashed") +
    scale_color_manual(name = "",
                       values = c("Before" = "#00bfc4", "After" = "#f8766d")) +
    scale_fill_manual(name = "",
                      values = c("Before" = "#00bfc4", "After" = "#f8766d")) +
    theme_light() +
    labs(title = "Treatment decay",
         subtitle = "Modeling decay for respondents' who consume CNN or CNN and Fox")
```
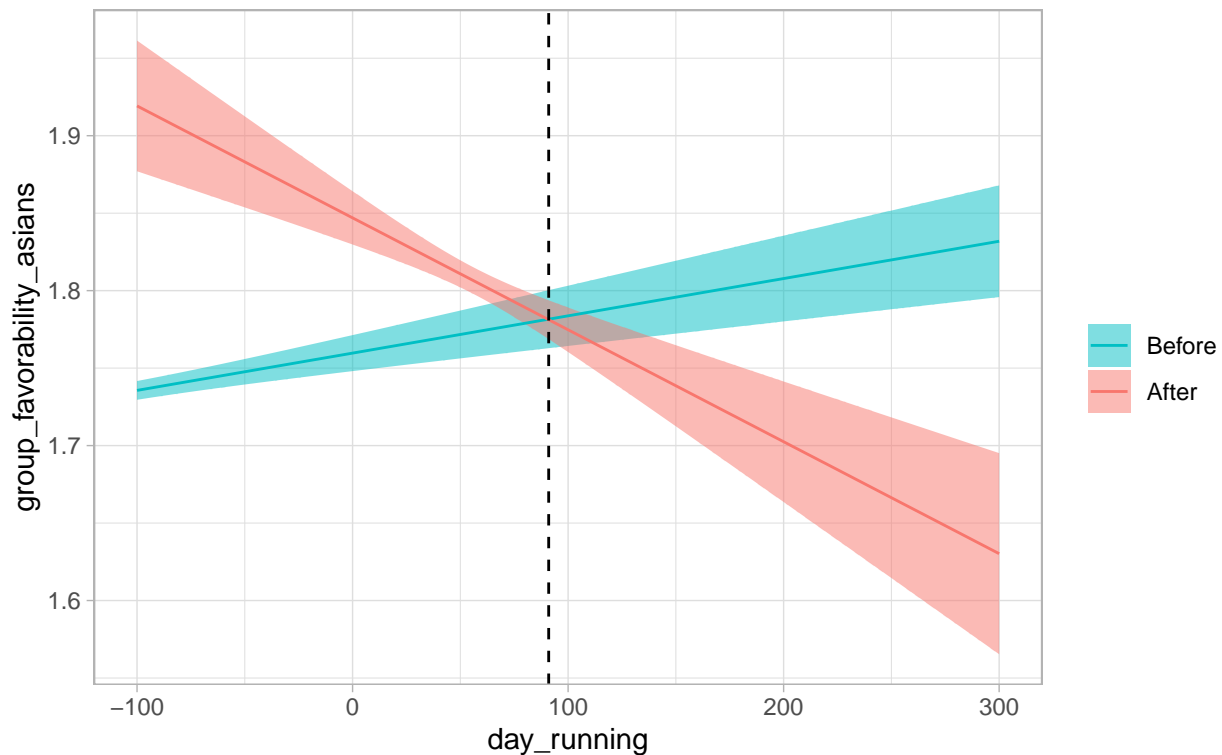
Treatment decay

Modeling decay for respondents' who consume CNN or CNN and Fox


```r
# Define prediction range
pred_data <- data.frame(day_running = seq(-100, 300, 1))

# Define before-pandemic model for people who consume only Fox.
m_before <- lm(group_favorability_asians ~ day_running, data = model_data %>% filter(day_running < 0, f

# Predict using before-pandemic model
preds_before <- predict(m_before, pred_data, se.fit = T)
preds_before <- tibble(group_favorability_asians = preds_before$fit,
                       day_running = seq(-100, 300, 1),
                       treatment = as_factor("Before"),
                       se = preds_before$se.fit)

# Define post-pandemic model for people who consume only Fox
m_after <- lm(group_favorability_asians ~ day_running, data = model_data %>% filter(day_running >= 0, f
preds_after <- predict(m_after, pred_data, se.fit = T)
preds_after <- tibble(group_favorability_asians = preds_after$fit,
                      day_running = seq(-100, 300, 1),
                      treatment = as_factor("After"),
                      se = preds_after$se.fit)

# Combine preds
preds_combined <- rbind(preds_before, preds_after) %>%
  mutate(treatment = fct_relevel(treatment, "Before", "After"))

# how long does it take for complete effect decay?
```

```
# preds_combined %>%
#   pivot_wider(names_from = treatment, values_from = group_favorability_asians) %>%
#   mutate(dif = 'FALSE'- 'TRUE') %>%
#   view()

# Day 133

# Visualize
preds_combined %>%
  ggplot(aes(x = day_running, y = group_favorability_asians)) +
    geom_ribbon(aes(ymin = group_favorability_asians - (1.96*se),
                    ymax = group_favorability_asians + (1.96*se),
                    fill = treatment),
                alpha = 0.5) +
    geom_path(aes(color = treatment)) +
    geom_vline(xintercept = 133, lty = "dashed") +
    scale_color_manual(name = "",
                       values = c("Before" = "#00bfc4", "After" = "#f8766d")) +
    scale_fill_manual(name = "",
                      values = c("Before" = "#00bfc4", "After" = "#f8766d")) +
    theme_light() +
    labs(title = "Treatment decay",
        subtitle = "Modeling decay for respondents' who consume only Fox")
```


16