AE 07: Federal judges and the Ivy League

Suggested answers

APPLICATION EXERCISE ANSWERS

MODIFIED
January 29, 2025

Important

These are suggested answers. This document should be used as reference only, it's not designed to be an exhaustive key.

```r
library(tidyverse)
library(scales)
library(skimr)
```

Article III judges

The Federal Judicial Center maintains detailed biographical data on all **Article III judges** in the United States.[1] This data is organized in a **relational structure**, with tables organized by:

- **Demographics**: Basic biographical information on each judge
- **Federal judicial service**: Information on each judge's service in the federal judiciary
- **Other federal service**: Information on each judge's service in other federal judicial offices
- **Education**: Information on each judge's education
- **Professional career**: Information on each judge's professional career
- **Other nominations/recess appointments**: Information on individual's other nominations[2] and recess appointments[3]

Data

The data is available in `fjc-judges.RData`.

`.Rdata`

`.RData` files are binary files that store one or more R objects. We can load the data using the `load()` function.

```r
load("data/fjc-judges.RData")
```

Relational structure

The data is organized in a **relational structure**. This means that the data is organized in multiple tables, and the tables are linked by common identifiers. For example, the `demographics` table contains basic biographical information on each judge, and the `education` table contains information on each judge's education. The two tables are linked by the `nid` variable, which is a unique identifier for each judge.

> **Identify the key columns**
>
> Every table contains a `nid` column, which is a unique identifier for each judge. For some tables such as `demographics`, there is one row for each judge. In other tables judges may have multiple rows, such as `education` (where each row represents a degree earned by the judge) or `service_fjs` (where each row represents a single appointment to the judiciary, and some judges have been appointed to multiple positions). For tables where judges appear in multiple rows, the `sequence` column uniquely identifies each row for a given judge. For example, the first appointment to the judiciary for a judge will have `sequence = 1`, the second appointment will have `sequence = 2`, and so on. While they have the same name, these columns do not necessary represent the same variable across tables.

Ivy League influence

In recent decades, scholars and policymakers have expressed concern about the lack of diversity in political institutions such as the federal judiciary. One aspect of this concern is the overrepresentation of graduates from a small number of elite colleges and universities.[4] [5] [6]

The **Ivy League** is a group of eight private colleges and universities in the northeastern United States. The Ivy League is known for its academic excellence and social prestige, and include:

- Brown University
- Columbia University
- Cornell University
- Dartmouth College
- Harvard University
- University of Pennsylvania
- Princeton University
- Yale University

We will use this data to answer the question: **What percentage of judges appointed to the federal judiciary attended an Ivy League school?**

To clarify the scope of our analysis, we will focus on judges appointed to the federal judiciary since 1945 (the post-World War II era). Furthermore we will limit our analysis to judges appointed to the **Courts of Appeals** and the **District Courts**.[7] And finally, we want to examine how this trend has evolved over time by examining the percentage of Ivy League graduates among judges appointed in each presidential administration. This will also allow us to consider the role of **presidential partisanship**. Are Democratic or Republican presidents more likely to make appointments from the Ivy League?

Joining the data

Your turn: First we need to identify the relevant tables for our analysis. Use the `skimr::skim()` function to examine the structure of the all the tables in the `fjc-judges.RData` file. Based on the output of the `skim()` function, identify the tables that contain the relevant information for our analysis, an appropriate `*_join()` function to combine them for our analysis, and the relevant key column(s).

```
skim(demographics)
skim(service_fjs)
skim(service_other)
skim(education)
skim(professional)
skim(other_appointments)
```

```
left_join(
  x = education,
  y = service_fjs,
  by = join_by(nid, judge_name),
  relationship = "many-to-many"
)
```

```
# A tibble: 9,369 × 35
       nid sequence.x judge_name school degree degree_year sequence.y court_type
     <dbl>      <dbl> <chr>      <chr>  <chr>        <dbl>      <dbl> <chr>
 1  1.38e7          1 Abelson, … Princ… B.A.          2005          1 U.S. Dist…
 2  1.38e7          2 Abelson, … New Y… J.D.          2010          1 U.S. Dist…
 3  1.39e6          1 Abrams, R… Corne… B.A.          1990          1 U.S. Dist…
 4  1.39e6          2 Abrams, R… Yale … J.D.          1993          1 U.S. Dist…
 5  1.38e6          1 Abruzzo, … Brook… LL.B.         1910          1 U.S. Dist…
 6  1.37e7          1 Abudu, Na… Colum… B.A.          1996          1 U.S. Cour…
 7  1.37e7          2 Abudu, Na… Tulan… J.D.          1999          1 U.S. Cour…
 8  1.38e6          1 Acheson, … Washi… B.A.          1845          1 U.S. Dist…
 9  1.38e6          1 Acheson, … Washi… B.A.          1845          2 U.S. Circ…
10  1.38e6          1 Acheson, … Washi… B.A.          1845          3 U.S. Cour…
# ℹ 9,359 more rows
# ℹ 27 more variables: court_name <chr>, appointment_title <chr>,
#   appointing_president <chr>, party_of_appointing_president <chr>,
#   reappointing_president <chr>, party_of_reappointing_president <chr>,
#   aba_rating <chr>, seat_id <chr>, statute_authorizing_new_seat <chr>,
#   recess_appointment_date <date>, nomination_date <date>,
#   committee_referral_date <date>, hearing_date <date>, …
```

We should use `education` to determine if a judge earned a degree from an Ivy League institution and `service_fjs` to determine the type of court to which the judge was appointed. We can use a `left_join()` or `inner_join()` to combine these tables. The appropriate key columns are `nid` and `judge_name`. `sequence` is not a key column since it represents different variables in each table.

Determine if a judge has an Ivy League degree

We will use the `education` table to determine if a judge earned a degree from an Ivy League institution. The `school` column contains the name of the school where the judge earned their degree.

Your turn: We need to know for every judge whether or not at least one of their degrees (e.g. undergraduate, law, etc.) was earned at an Ivy League institution. Create a data frame with one-row-per-judge and a variable `ivy_league` that is `TRUE` if the judge earned a degree from an Ivy League institution and `FALSE` otherwise.

Identifying Ivy League schools in the dataset

Unfortunately the names of schools in the database are not standardized, so we need to use a list of Ivy League schools to identify which judges attended an Ivy League institution. You can use this character vector to identify Ivy League schools:

```r
ivy_league <- c(
  # Brown
  "Brown University", "Rhode Island College (now Brown University)",
  # Columbia
  "Columbia University", "Columbia Law School", "King's College (now Columbia University)",
  # Cornell
  "Cornell University", "Cornell Law School", "Cornell University Department of Law",
  # Dartmouth
  "Dartmouth College",
  # Harvard
  "Harvard Law School", "Harvard University", "Harvard University, Kennedy School of Government",
  "Harvard College",
  # UPenn
  "University of Pennsylvania", "University of Pennsylvania, Wharton School",
  "College of Philadelphia (now University of Pennsylvania)",
  "University of Pennsylvania Law School (now Carey Law School)",
  # Princeton
  "Princeton University", "College of New Jersey (now Princeton University)",
  "Princeton University, Woodrow Wilson School of Public and International Affairs (now Princeton School of Public and International Affairs)",
  # Yale
  "Yale University", "Yale College", "Yale Law School", "Yale School of Architecture",
  "Yale School of Organization and Management"
)
```

```r
ivy_league <- c(
  # Brown
  "Brown University", "Rhode Island College (now Brown University)",
  # Columbia
  "Columbia University", "Columbia Law School", "King's College (now Columbia University)"
  # Cornell
  "Cornell University", "Cornell Law School", "Cornell University Department of Law",
  # Dartmouth
  "Dartmouth College",
  # Harvard
  "Harvard Law School", "Harvard University", "Harvard University, Kennedy School of Government",
  "Harvard College",
  # UPenn
  "University of Pennsylvania", "University of Pennsylvania, Wharton School",
  "College of Philadelphia (now University of Pennsylvania)",
  "University of Pennsylvania Law School (now Carey Law School)",
  # Princeton
  "Princeton University", "College of New Jersey (now Princeton University)",
  "Princeton University, Woodrow Wilson School of Public and International Affairs (now Princeton School of Public and International Affairs)",
  # Yale
  "Yale University", "Yale College", "Yale Law School", "Yale School of Architecture",
  "Yale School of Organization and Management"
)

education |>
  # determine if a judge earned a degree from an Ivy League institution
  mutate(ivy_league = school %in% ivy_league) |>
  # summarize the data frame so we have exactly one row per judge
  summarize(
    ivy_league = mean(ivy_league) > 0,
    .by = c(nid, judge_name)
  )
```

```
# A tibble: 3,997 × 3
        nid judge_name                ivy_league
      <dbl> <chr>                     <lgl>
 1 13761857 Abelson, Adam Ben         TRUE
 2  1393931 Abrams, Ronnie            TRUE
 3  1376976 Abruzzo, Matthew T.       FALSE
 4 13651551 Abudu, Nancy Gbana        TRUE
 5  1376981 Acheson, Marcus Wilson    FALSE
 6  1376986 Acker, William Marsh, Jr. TRUE
 7  1376991 Ackerman, Harold Arnold   FALSE
 8  1376996 Ackerman, James Waldo     FALSE
 9  1377001 Acosta, Raymond L.        FALSE
10  1377006 Adair, J[ackson] Leroy    FALSE
# ℹ 3,987 more rows
```

Combine with judicial appointments

Your turn: Combined the summarized education data with information on each appointment to the federal judiciary.

```
education |>
  mutate(ivy_league = school %in% ivy_league) |>
  summarize(
    ivy_league = mean(ivy_league) > 0,
    .by = c(nid, judge_name)
  ) |>
  # join with appointment table
  left_join(y = service_fjs)
```

```
# A tibble: 4,690 × 32
      nid judge_name ivy_league sequence court_type court_name appointment_title
    <dbl> <chr>      <lgl>         <dbl> <chr>      <chr>      <chr>
 1 1.38e7 Abelson, … TRUE              1 U.S. Dist… U.S. Dist… Judge
 2 1.39e6 Abrams, R… TRUE              1 U.S. Dist… U.S. Dist… Judge
 3 1.38e6 Abruzzo, … FALSE             1 U.S. Dist… U.S. Dist… Judge
 4 1.37e7 Abudu, Na… TRUE              1 U.S. Cour… U.S. Cour… Judge
 5 1.38e6 Acheson, … FALSE             1 U.S. Dist… U.S. Dist… Judge
 6 1.38e6 Acheson, … FALSE             2 U.S. Circ… U.S. Circ… Judge
 7 1.38e6 Acheson, … FALSE             3 U.S. Cour… U.S. Cour… Judge
 8 1.38e6 Acker, Wi… TRUE              1 U.S. Dist… U.S. Dist… Judge
 9 1.38e6 Ackerman,… FALSE             1 U.S. Dist… U.S. Dist… Judge
10 1.38e6 Ackerman,… FALSE             1 U.S. Dist… U.S. Dist… Judge
# ℹ 4,680 more rows
# ℹ 25 more variables: appointing_president <chr>,
#   party_of_appointing_president <chr>, reappointing_president <chr>,
#   party_of_reappointing_president <chr>, aba_rating <chr>, seat_id <chr>,
#   statute_authorizing_new_seat <chr>, recess_appointment_date <date>,
#   nomination_date <date>, committee_referral_date <date>,
#   hearing_date <date>, judiciary_committee_action <chr>, …
```

We can combine with `service_fjs` to determine the type of court to which the judge was appointed, using a `left_join()` (or `inner_join()`) to combine the two tables.

> **Note**
>
> Since we summarized the `education` data, the `sequence` column no longer exists so we can let `*left*_join()` use the default `by` argument.

Filter the data

Your turn: Filter the data to limit our analysis to judges appointed:

- To the **Courts of Appeals** and the **District Courts**,
- Since 1945, and
- By a Democratic or Republican president[8]

```
education |>
  mutate(ivy_league = school %in% ivy_league) |>
  summarize(
    ivy_league = mean(ivy_league) > 0,
    .by = c(nid, judge_name)
  ) |>
  left_join(y = service_fjs) |>
  mutate(
    year = year(commission_date)
  ) |>
  filter(
    # look at appointments between 1945 and 2025
    between(year, 1945, 2025),
    # look at appointments to Courts of Appeals and District Courts
    court_type %in% c(
      "U.S. Court of Appeals",
      "U.S. District Court"
    ),
    # only appointments by Democratic or Republican presidents
    party_of_appointing_president %in% c(
      "Democratic",
      "Republican"
    )
  )
```

```
# A tibble: 3,224 × 33
      nid judge_name ivy_league sequence court_type court_name appointment_title
    <dbl> <chr>      <lgl>         <dbl> <chr>      <chr>      <chr>
 1 1.38e7 Abelson, … TRUE              1 U.S. Dist… U.S. Dist… Judge
 2 1.39e6 Abrams, R… TRUE              1 U.S. Dist… U.S. Dist… Judge
 3 1.37e7 Abudu, Na… TRUE              1 U.S. Cour… U.S. Cour… Judge
 4 1.38e6 Acker, Wi… TRUE              1 U.S. Dist… U.S. Dist… Judge
 5 1.38e6 Ackerman,… FALSE             1 U.S. Dist… U.S. Dist… Judge
 6 1.38e6 Ackerman,… FALSE             1 U.S. Dist… U.S. Dist… Judge
 7 1.38e6 Acosta, R… FALSE             1 U.S. Dist… U.S. Dist… Judge
 8 1.38e6 Adams, Ar… TRUE              1 U.S. Cour… U.S. Cour… Judge
 9 1.38e6 Adams, He… FALSE             1 U.S. Dist… U.S. Dist… Judge
10 1.39e6 Adams, Jo… FALSE             1 U.S. Dist… U.S. Dist… Judge
# ℹ 3,214 more rows
# ℹ 26 more variables: appointing_president <chr>,
#   party_of_appointing_president <chr>, reappointing_president <chr>,
#   party_of_reappointing_president <chr>, aba_rating <chr>, seat_id <chr>,
#   statute_authorizing_new_seat <chr>, recess_appointment_date <date>,
#   nomination_date <date>, committee_referral_date <date>,
#   hearing_date <date>, judiciary_committee_action <chr>, …
```

Calculate proportion of Ivy League appointments

Your turn: Calculate the proportion of judicial appointments that were Ivy League graduates, for each Courts of Appeals and District Courts, president, and presidential party.

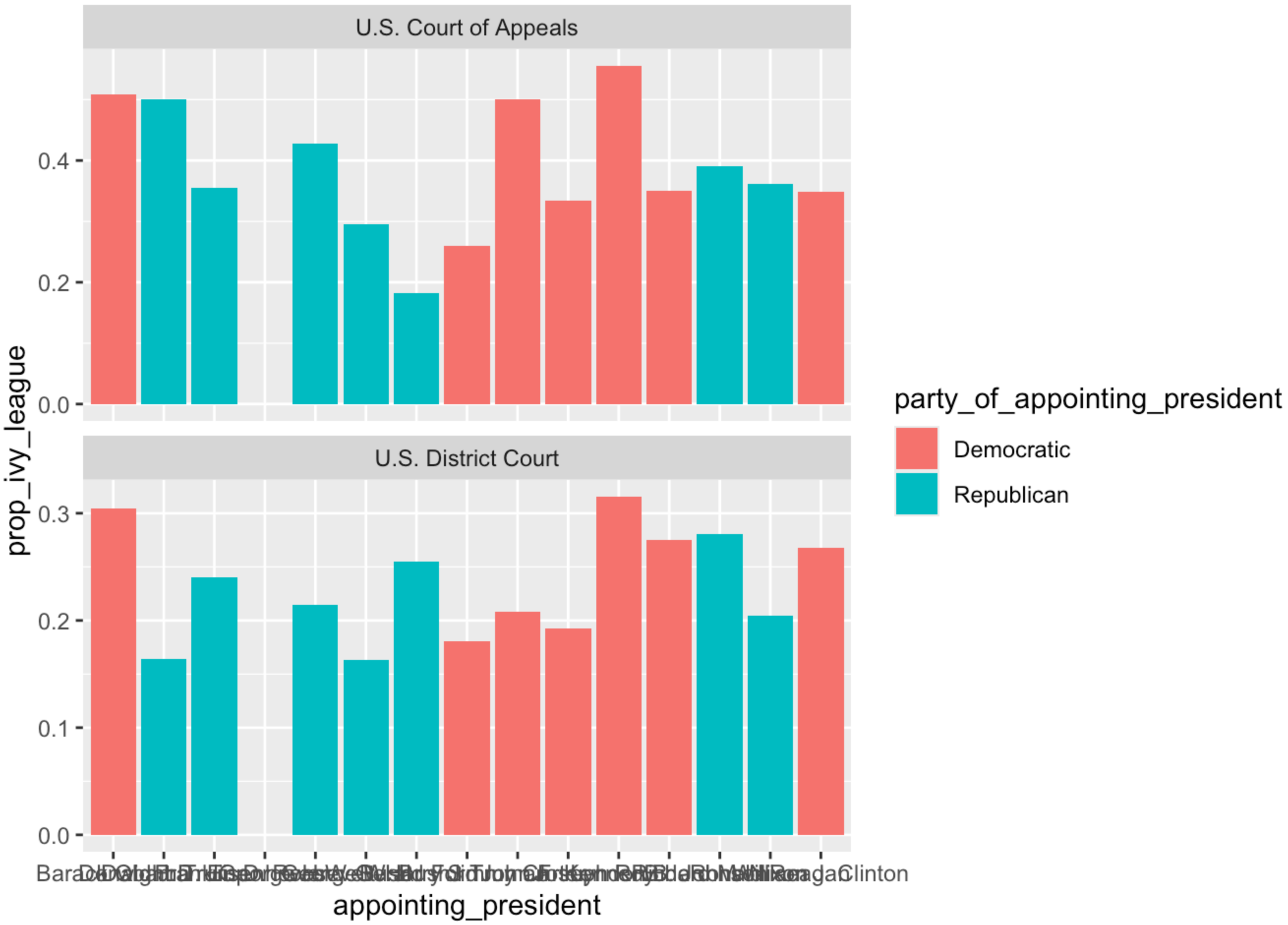
```
education |>
  mutate(ivy_league = school %in% ivy_league) |>
  summarize(
    ivy_league = mean(ivy_league) > 0,
    .by = c(nid, judge_name)
  ) |>
  left_join(y = service_fjs) |>
  mutate(
    year = year(commission_date)
  ) |>
  filter(
    between(year, 1945, 2025),
    court_type %in% c(
      "U.S. Court of Appeals",
      "U.S. District Court"
    ),
    party_of_appointing_president %in% c(
      "Democratic",
      "Republican"
    )
  ) |>
  summarize(
    # calculate the proportion of Ivy League appointments
    prop_ivy_league = mean(ivy_league),
    # grouped by appointing president, court type, and party of appointing president
    .by = c(appointing_president, court_type, party_of_appointing_president)
  )
```

```
# A tibble: 29 × 4
   appointing_president court_type        party_of_appointing_…¹ prop_ivy_league
   <chr>                <chr>             <chr>                            <dbl>
 1 Joseph R. Biden      U.S. District Co… Democratic                       0.316
 2 Barack Obama         U.S. District Co… Democratic                       0.305
 3 Joseph R. Biden      U.S. Court of Ap… Democratic                       0.556
 4 Ronald Reagan        U.S. District Co… Republican                       0.204
 5 Jimmy Carter         U.S. District Co… Democratic                       0.208
 6 Gerald Ford          U.S. District Co… Republican                       0.255
 7 Richard M. Nixon     U.S. Court of Ap… Republican                       0.391
 8 William J. Clinton   U.S. District Co… Democratic                       0.268
 9 George W. Bush       U.S. District Co… Republican                       0.163
10 Donald J. Trump      U.S. District Co… Republican                       0.164
# ℹ 19 more rows
# ℹ abbreviated name: ¹party_of_appointing_president
```

Visualize the data

Demo: Let’s generate an initial draft of our plot. Since we want to visualize proportions, a **bar chart** is a reasonable starting point.

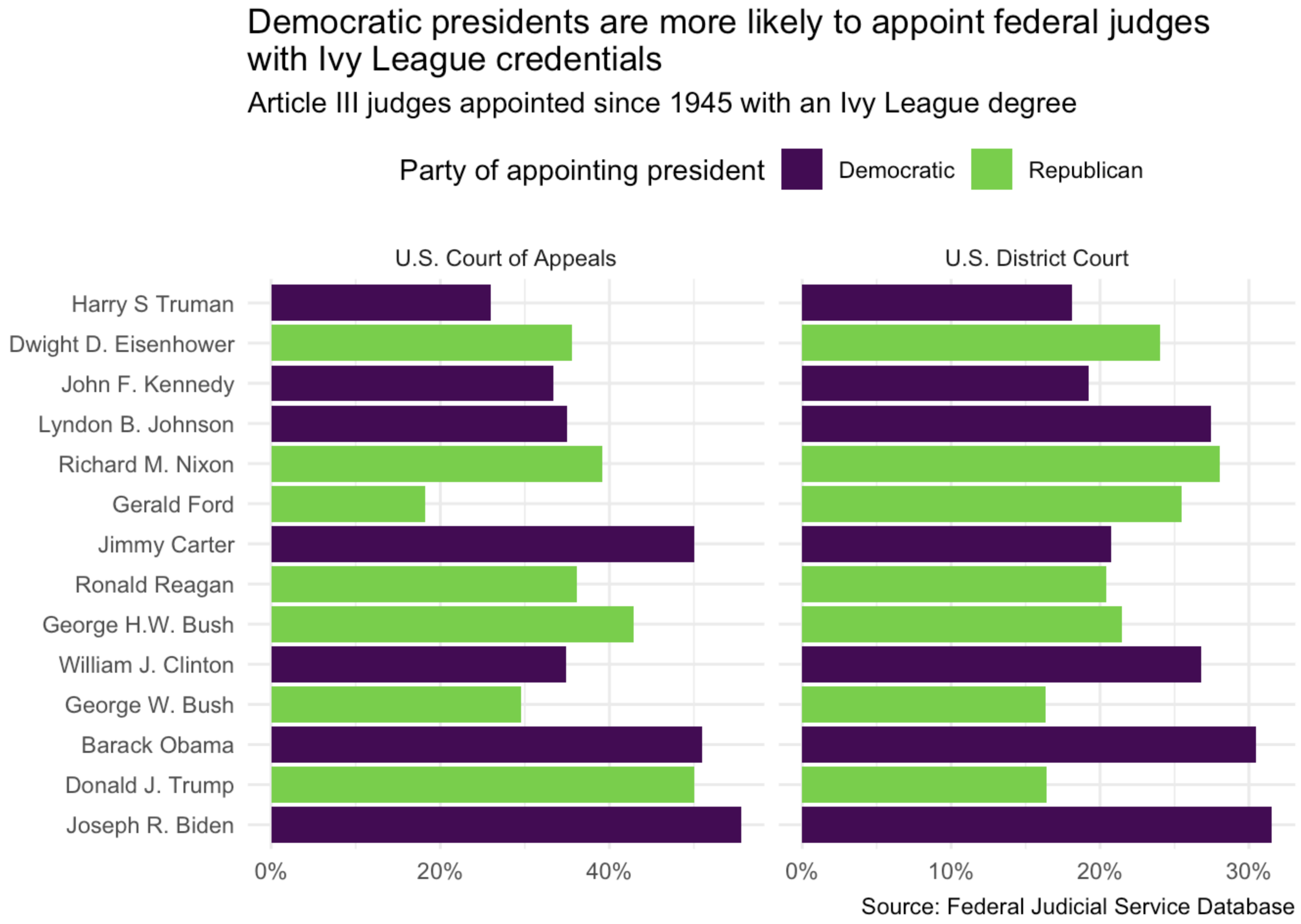
```
education |>
  mutate(ivy_league = school %in% ivy_league) |>
  summarize(
    ivy_league = mean(ivy_league) > 0,
    .by = c(nid, judge_name)
  ) |>
  left_join(y = service_fjs) |>
  mutate(
    year = year(commission_date)
  ) |>
  filter(
    between(year, 1945, 2025),
    court_type %in% c(
      "U.S. Court of Appeals",
      "U.S. District Court"
    ),
    party_of_appointing_president %in% c(
      "Democratic",
      "Republican"
    )
  ) |>
  summarize(
    prop_ivy_league = mean(ivy_league),
    .by = c(appointing_president, court_type, party_of_appointing_president)
  ) |>
  # standard bar plot
  ggplot(mapping = aes(
    x = appointing_president, y = prop_ivy_league,
    fill = party_of_appointing_president
  )) +
  geom_col() +
  # facet by court type, have the panels top and bottom
  facet_wrap(facets = vars(court_type), scales = "free_y", ncol = 1)
```



It’s a decent first step. Some things we need to improve still:

- Reverse the axes to make it a horizontal bar chart. Ensures all presidential names are visible.
- Logical ordering of the presidents. We should order them chronologically based on their terms in office.
- Remove Franklin D. Roosevelt. He died in office in early 1945 and only made a couple of appointments that year, so we don’t really have enough data to justify including him on the chart.
- Optimal color palette
- Better labels
- Move the legend to the top to provide more horizontal space for the plot

```
education |>
  mutate(ivy_league = school %in% ivy_league) |>
  summarize(
    ivy_league = mean(ivy_league) > 0,
    .by = c(nid, judge_name)
  ) |>
  left_join(y = service_fjs) |>
  arrange(commission_date) |>
  mutate(
    year = year(commission_date),
    appointing_president = fct_inorder(f = appointing_president) |>      2
      fct_rev()
  ) |>
  filter(
    between(year, 1945, 2025),
    court_type %in% c(
      "U.S. Court of Appeals",
      "U.S. District Court"
    ),
    party_of_appointing_president %in% c(
      "Democratic",
      "Republican"
    )
  ) |>
  summarize(
    prop_ivy_league = mean(ivy_league),
    .by = c(appointing_president, court_type, party_of_appointing_president)
  ) |>
  filter(appointing_president != "Franklin D. Roosevelt") |>      3
  ggplot(mapping = aes(
    x = prop_ivy_league, y = appointing_president,      1
    fill = party_of_appointing_president
  )) +
  geom_col() +
  scale_x_continuous(labels = label_percent()) +      4
  scale_fill_viridis_d(end = 0.8) +      5
  facet_wrap(facets = vars(court_type), scales = "free_x", ncol = 2) +
  labs(      6
    x = NULL,
    y = NULL,
    title = "Democratic presidents are more likely to appoint federal judges\nwith Ivy Lea
          credentials",
    subtitle = "Article III judges appointed since 1945 with an Ivy League degree",
    fill = "Party of appointing president",
    caption = "Source: Federal Judicial Service Database"
  ) +
  theme_minimal() +      7
  theme(legend.position = "top")
```



Democratic presidents are more likely to appoint federal judges with Ivy League credentials

Article III judges appointed since 1945 with an Ivy League degree

Source: Federal Judicial Service Database

How does this answer our original question? We have a few key pieces of information.

1. Judges appointed to the Courts of Appeal are more likely to have a Ivy League degree than judges appointed to the District Courts.
2. The proportion of appointees with Ivy League degrees has (somewhat) risen over time.
3. Since Jimmy Carter, Democratic presidents have been more likely to appoint judges with Ivy League degrees than Republican presidents.

Session information

Footnotes

1. Article III judges are federal judges appointed under the authority of Article III of the United States Constitution. They are appointed for life and possess significant independence from the other branches of government. They include justices of the Supreme Court, as well as judges on the Courts of Appeals, the District Courts, and several other specialized courts. ↩
2. Including failed nominations which were never confirmed by the U.S. Senate. ↩
3. Individuals appointed to the judiciary while the Senate was in recess - these appointments are temporary and expire at the end of the Senate's next session. ↩
4. Harvard & Yale Still Dominate as Biden Focuses on Diversifying the Judiciary ↩
5. Supreme Court shouldn't be covered in Ivy, 2 lawmakers say ↩

[6] The New Diversity Crisis in the Federal Judiciary ↩

[7] While the Supreme Court is the most influential court in the United States, it is also the smallest, with only nine justices at any given time. The Courts of Appeals and the District Courts are the workhorses of the federal judiciary, and are the primary focus of our analysis. ↩

[8] Some appointments in the dataset were **reassignments** from one judicial seat to another, and are classified under the `party_of_appointing_president` as “None (reassignment)”. ↩

Made with R and Quarto.