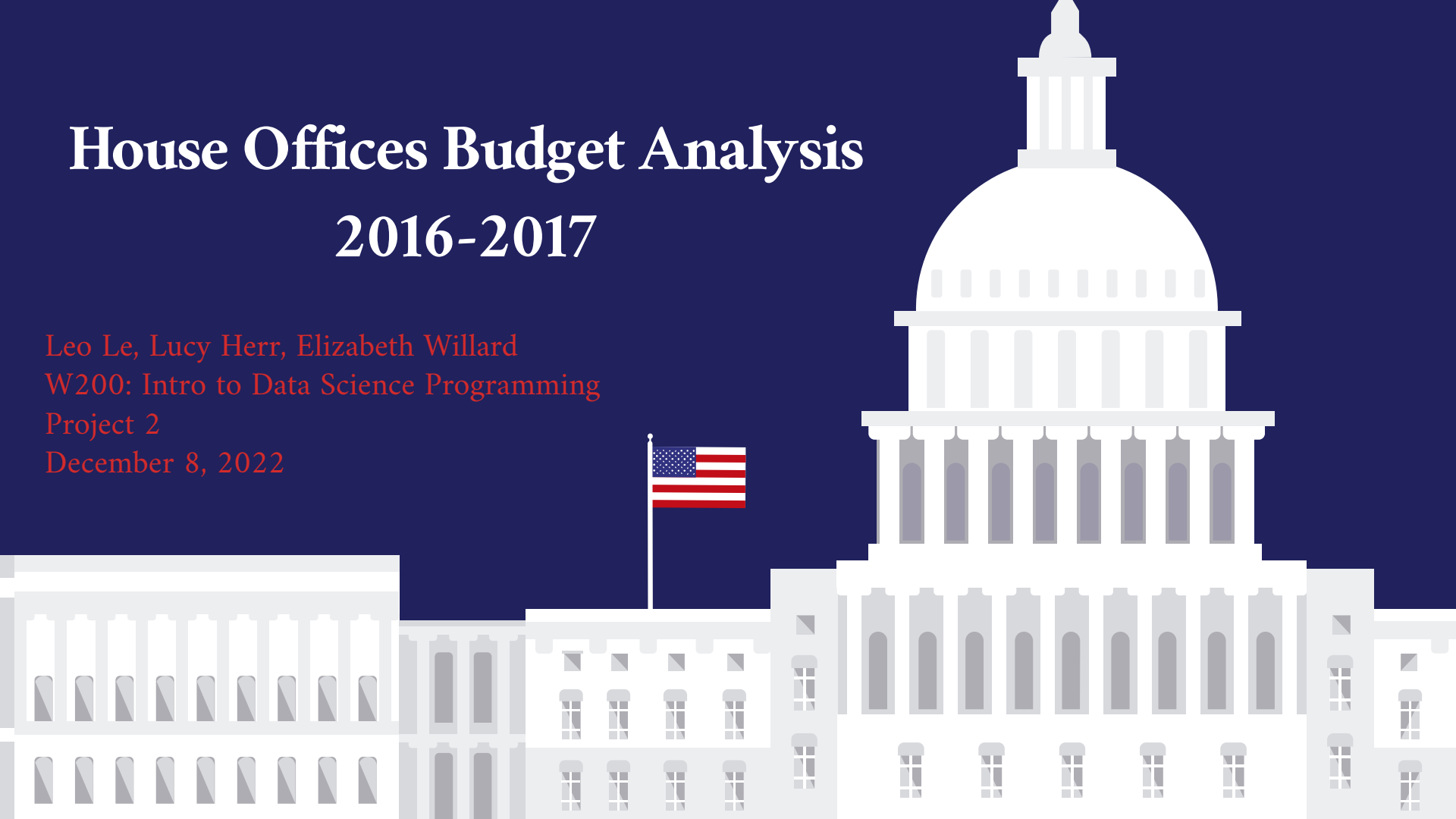
House Offices Budget Analysis 2016-2017

Leo Le, Lucy Herr, Elizabeth Willard
W200: Intro to Data Science Programming
Project 2
December 8, 2022



Background

- The House of Representatives is one of two legislative bodies in the federal U.S. government
- Members of the House of Representatives receive an annual budget to spend at their discretion (with some restrictions)
- Focusing on the 2016-2017 time period when former president Donald J. Trump was elected to office

Goal: Explore the spending habits of House of Representatives members

Data

Main dataset, version 1: Summary Expenditures (Propublica)

- 'Bioguide_ID' – the official ID of members of the House (unique/primary key to be joined with supplementary dataset for member party details)
- 'Description' - Expenditure categories
- 'Purpose' - Details of the categories expenditure
- 'Amount' - amount spent by office in that category and quarter

Main dataset, version 2: Detailed Expenditures (Propublica)

- 'Office' - House of Representatives offices
- 'Category' - Expenditure categories
- 'Purpose' - Details of the categories expenditure
- 'Payee' - Recipient of the payment
- 'Amount' - amount spent by office in that category

Supplementary dataset: ID data for 2015-2019 U.S. House Representatives (https://bioguide.congress.gov/)

Research Questions

- **Are there any identifiable patterns in amounts spent on different expenditure categories and periods (quarter and/or year)?**
  - Which categories yield the most interesting trends?
  - In particular, how does personnel spending vary between offices over time? Is there a relationship between spending on personnel and office characteristics (such as turnover)?
- **What, if any, differences are there in the spending breakdowns (by category and overall amounts) for Republican and Democratic house members between 2016 and 2017?**
  - How do these trends vary on a quarterly and/or annual basis?
  - Depending on the answers to the above questions, which categories yield the most interesting trends or patterns in terms of partisan spending differences?

Analytical Process & Assumptions

Cleaning Overview:

- Understand and standardize data context, categories, and approaches
- Combine quarterly summary expenditures datasets for years 2016-2017 (8 in total)
- Clean combined summary dataset: check column types, address null values, identify subsets and/or columns to drop, convert between data types
- Wherever possible, use outside research to interpret decisions about the data

Assumptions

- Excluded negative expenses (amounts)
- Excluded offices with not enough quarterly financial reports
- Use non-total amounts from summary
- We are confident that our data captures the larger patterns, but the underlying raw data are messy.

Spending categories analysis

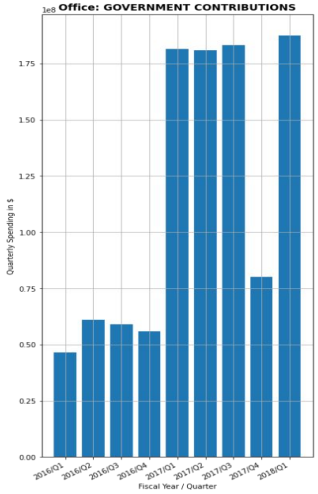
Goal: Understand how House Office's spending habits throughout different categories and offices

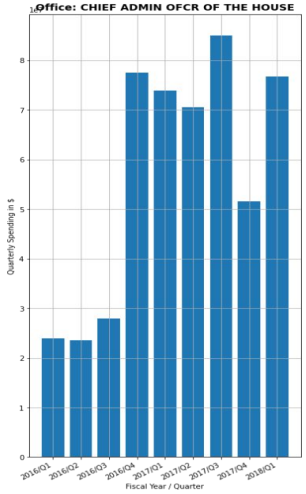
Assumptions/Process:

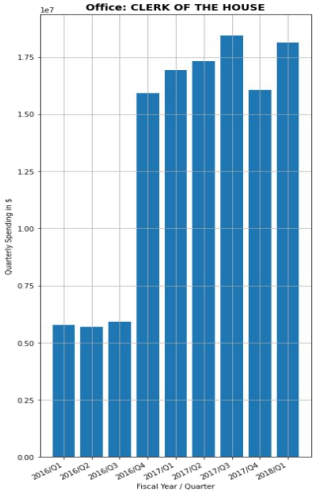
- Removed Offices of Singular People (i.e. Honorable So-and-So)
- Removed Offices that did not have enough quarterly data (less than 5 quarters worth of data)
- Converted dollar amounts into numerical values since they were originally strings
- Removed "expense total:" row from each office as that was the cumulative spending for the quarter, and would skew the aggregation

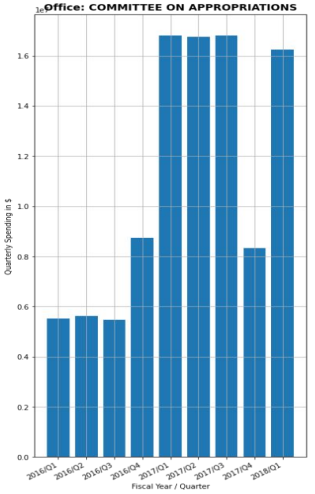
10 Highest Spending Offices / Largest Spending Category

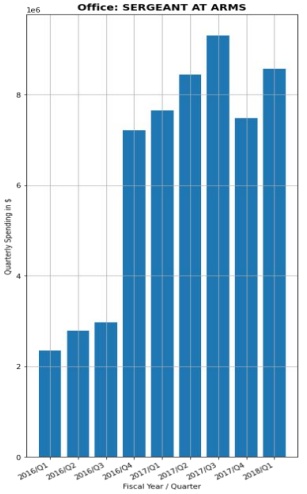
GOVERNMENT CONTRIBUTIONS - Personnel Benefits
CHIEF ADMIN OFFICER OF THE HOUSE - Personnel Compensation
CLERK OF THE HOUSE - Personnel Compensation
COMMITTEE ON APPROPRIATIONS - Personnel Compensation
SERGEANT AT ARMS - Personnel Compensation
COMMITTEE ON ENERGY & COMMERCE - Personnel Compensation
JOINT COMMITTEE ON TAXATION - Personnel Compensation
NET EXPENSES TELECOMMUNICATION - Office Expenditures (Supplies, Materials, Rent, Communication Utilities)
LEGISLATIVE COUNSEL - Personnel Compensation
COMMITTEE ON WAYS AND MEANS - Personnel Compensation

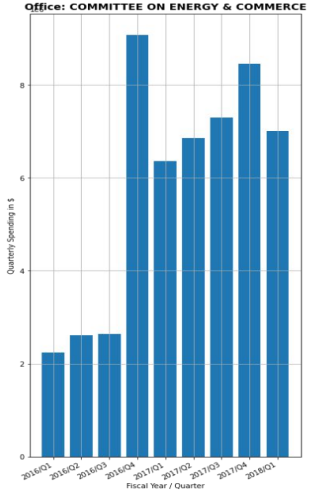

Office: GOVERNMENT CONTRIBUTIONS
1e8
Quarterly Spending in $
Fiscal Year / Quarter
2016/Q1
2016/Q2
2016/Q3
2016/Q4
2017/Q1
2017/Q2
2017/Q3
2017/Q4
2018/Q1

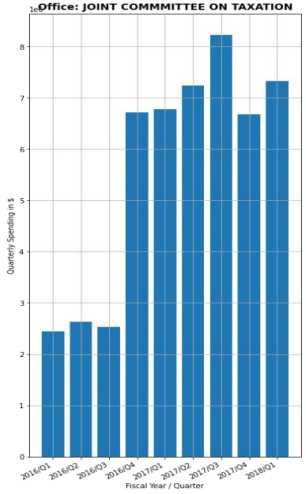

Office: CHIEF ADMIN OFCR OF THE HOUSE
1e7
Quarterly Spending in $
Fiscal Year / Quarter
2016/Q1
2016/Q2
2016/Q3
2016/Q4
2017/Q1
2017/Q2
2017/Q3
2017/Q4
2018/Q1

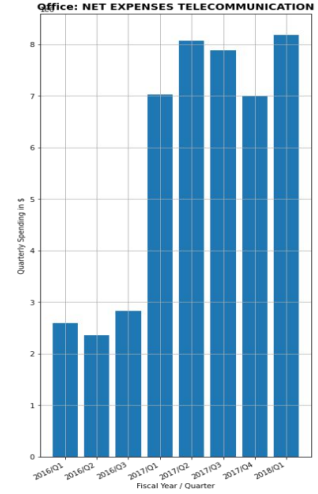

Office: CLERK OF THE HOUSE
1e7
Quarterly Spending in $
Fiscal Year / Quarter
2016/Q1
2016/Q2
2016/Q3
2016/Q4
2017/Q1
2017/Q2
2017/Q3
2017/Q4
2018/Q1

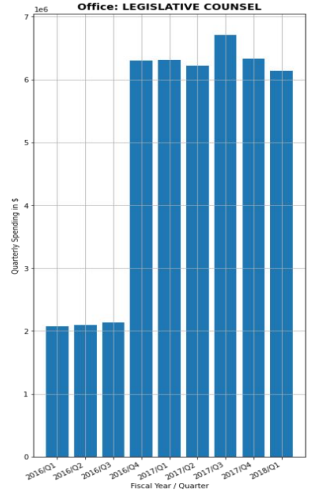

Office: COMMITTEE ON APPROPRIATIONS
1e7
Quarterly Spending in $
Fiscal Year / Quarter
2016/Q1
2016/Q2
2016/Q3
2016/Q4
2017/Q1
2017/Q2
2017/Q3
2017/Q4
2018/Q1

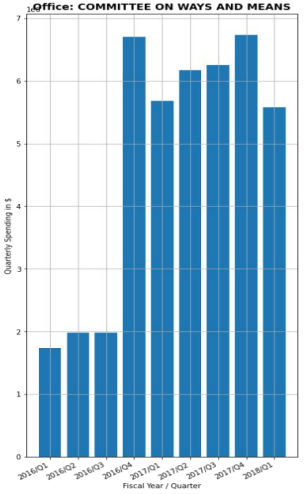

Office: SERGEANT AT ARMS
1e6
Quarterly Spending in $
Fiscal Year / Quarter
2016/Q1
2016/Q2
2016/Q3
2016/Q4
2017/Q1
2017/Q2
2017/Q3
2017/Q4
2018/Q1


Office: COMMITTEE ON ENERGY & COMMERCE
1e7
Quarterly Spending in $
Fiscal Year / Quarter
2016/Q1
2016/Q2
2016/Q3
2016/Q4
2017/Q1
2017/Q2
2017/Q3
2017/Q4
2018/Q1


Office: JOINT COMMMITTEE ON TAXATION
1e7
Quarterly Spending in $
Fiscal Year / Quarter
2016/Q1
2016/Q2
2016/Q3
2016/Q4
2017/Q1
2017/Q2
2017/Q3
2017/Q4
2018/Q1

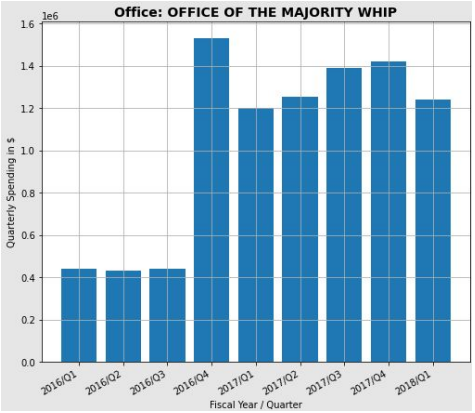

Office: NET EXPENSES TELECOMMUNICATION
1e7
Quarterly Spending in $
Fiscal Year / Quarter
2016/Q1
2016/Q2
2016/Q3
2016/Q4
2017/Q1
2017/Q2
2017/Q3
2017/Q4
2018/Q1

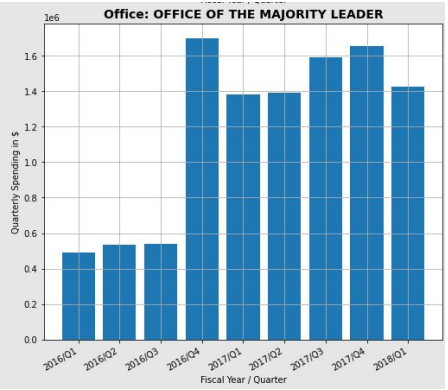

Office: LEGISLATIVE COUNSEL
1e6
Quarterly Spending in $
Fiscal Year / Quarter
2016/Q1
2016/Q2
2016/Q3
2016/Q4
2017/Q1
2017/Q2
2017/Q3
2017/Q4
2018/Q1


Office: COMMITTEE ON WAYS AND MEANS
1e7
Quarterly Spending in $
Fiscal Year / Quarter
2016/Q1
2016/Q2
2016/Q3
2016/Q4
2017/Q1
2017/Q2
2017/Q3
2017/Q4
2018/Q1

Points of Interest with Office Data

Both components of the Minority Office spent more than the Majority Office.





Comparison of Expenditures by Party

Goal: Compare spending trends over time and category for Republican versus Democrat House Representatives. Are there any interesting patterns or differences to evaluate?

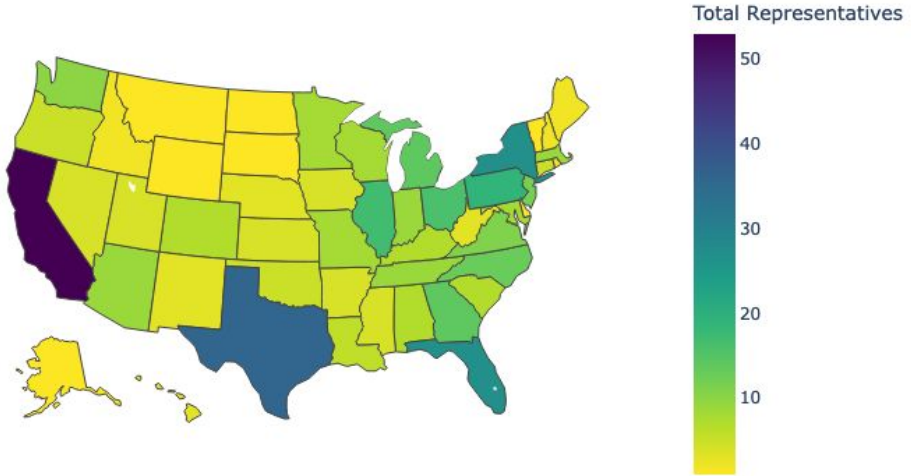
Process & Assumptions:

- Supplementary dataset: 2015-2019 U.S. House Representatives Bioguide (https://bioguide.congress.gov/)
  - Parse "congresses" column (json/dictionary type) to obtain party data

```
{'position': 'Representative',
 'congressNumber': 114,
 'stateName': 'MI',
 'parties': ['Democrat']},
```

  - Join supplementary data for each Rep. with 'bioguide id' (primary key/unique identifier) in Expenditures dataset
- Subset singular Representative or leadership offices (exclude groups/committees that can't be linked to specific party)
- Impute party information for atypical cases (e.g., District or Territory members) where applicable

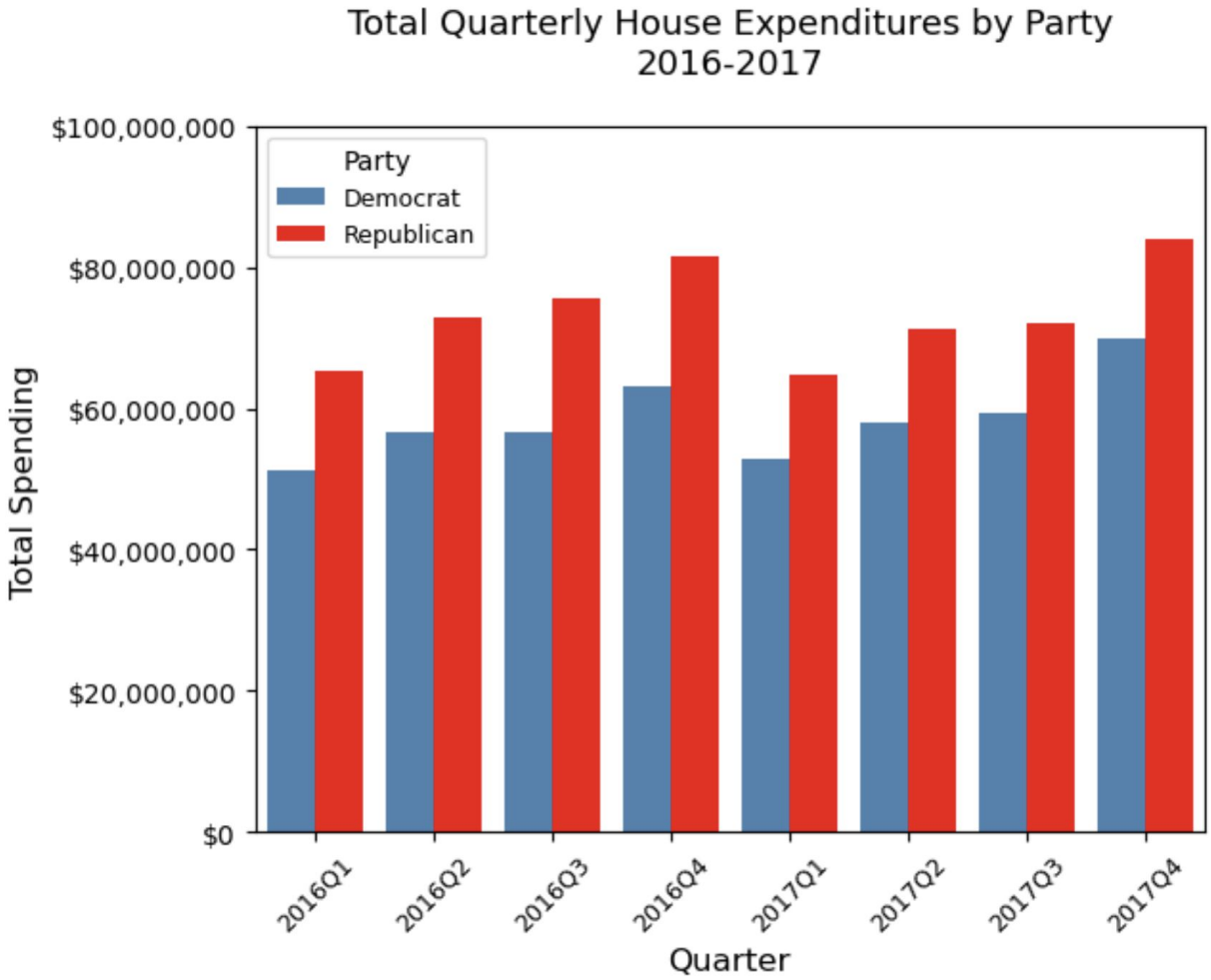
Number of U.S. House Representatives by State/District/Territory*



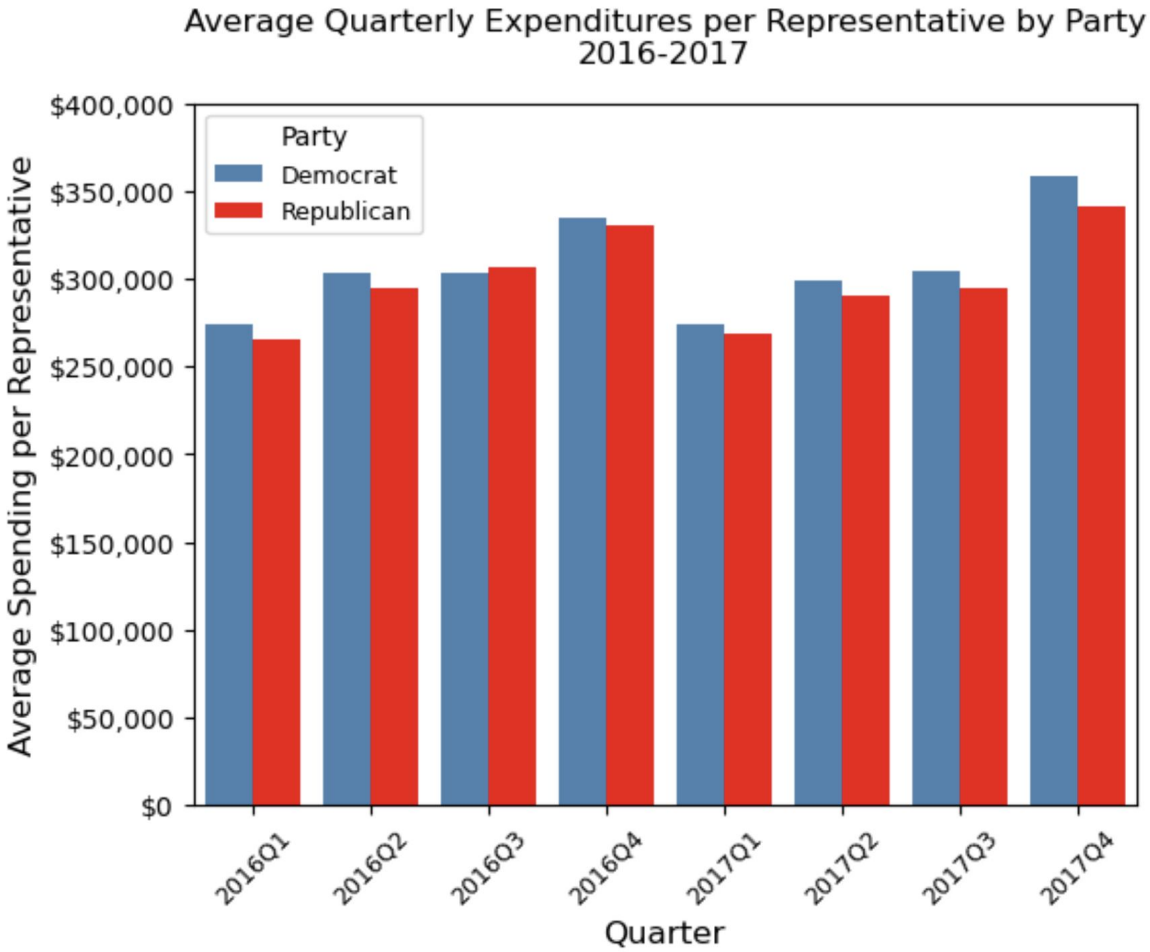
District/Territory	2016 Members	2017 Members
DAL American Samoa	1	1
DAL DC	1	1
DAL Guam	1	1
DAL Northern Mariana Islands	1	1
DAL Puerto Rico	1	1
DAL Virgin Islands	1	1

*It's more complicated than you'd think

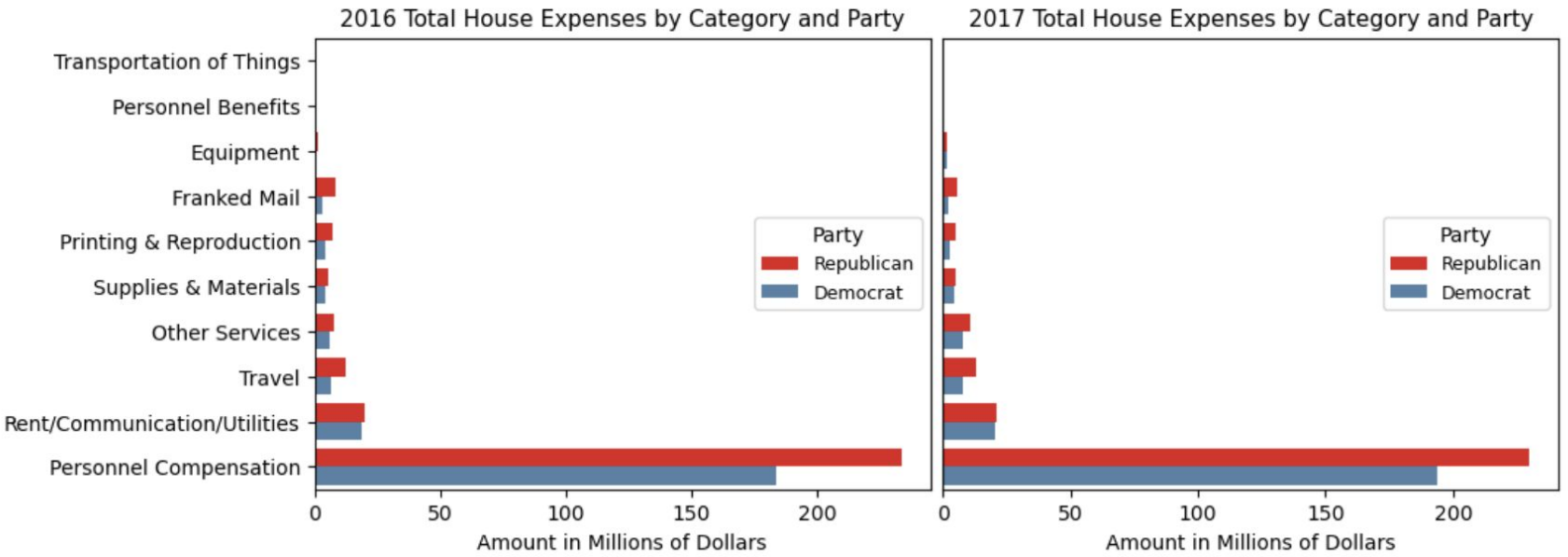
Quarterly TOTAL Spending: Republicans vs. Democrats



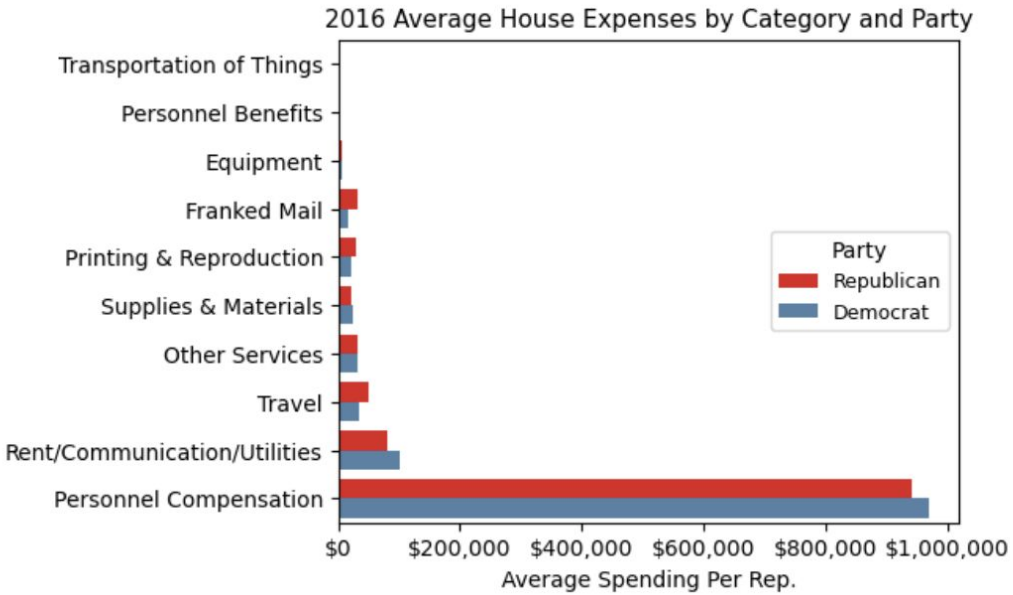
Quarterly AVERAGE Spending: Republicans vs. Democrats

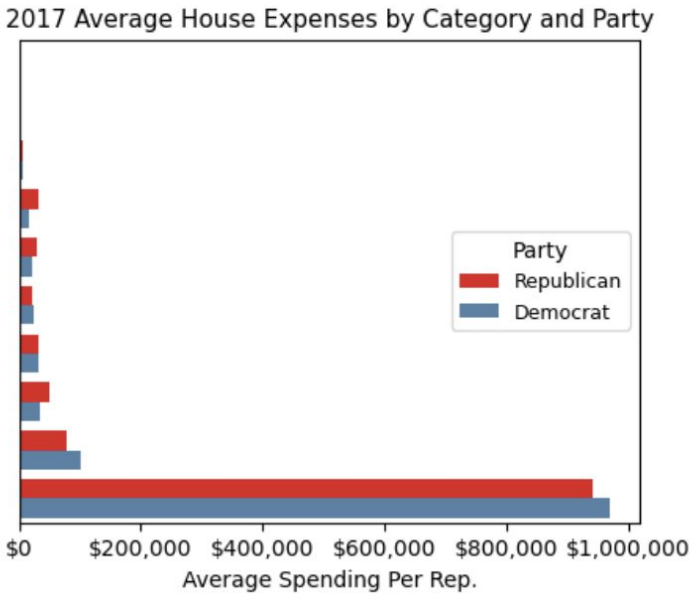


Total Yearly Spending by Category: Republicans vs. Democrats



Average Yearly Spending by Category: Republicans vs. Democrats





Head count and retention analysis

Goal: Understand how House Office manages labor distribution.

Assumptions/Process:

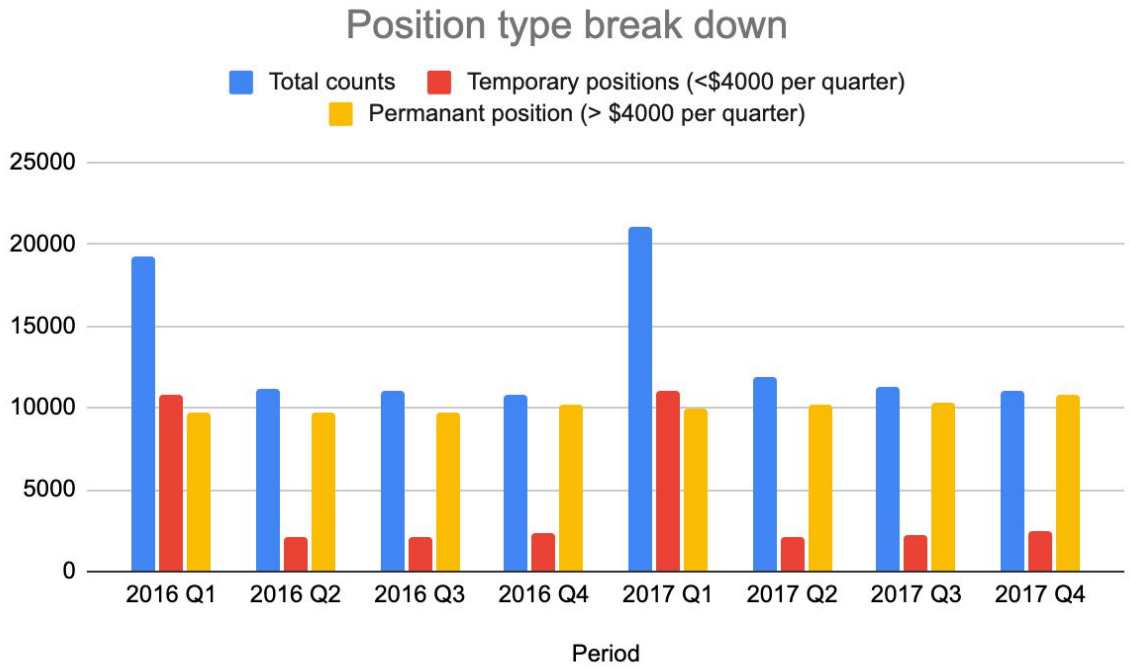
-The biggest challenge in aggregating the data is that different House offices classify expenses in different ways, so an extensive and rigorous analysis was done to identify unique positions:

Identifying overlapping wordings

Cleaning up grammar

Converting values to appropriate data type

Headcount analysis



Retention Analysis

	Mean retention Rate
2016	47.38%
2017	80.69%
2016-2017	41.19%
2016 Q3 - 2016 Q4	50.00%

Average retention rate during the 2008 recession was 64.2%!

Conclusion

- Most of our time and effort was spent on understanding, cleaning and standardizing the data.
- Personnel Compensation accounted for the majority of the spending
- Retention rate 2016-2017 averaged at 41.19% compared to 64.2% during the 2008 recession
- Comparing party spending: at the summary level, the limited patterns that appear depend a lot on the spending metric (total vs. average)

THANKS!