Congressional Stock Trading

Neill Killgore, Rajiv Kamal, Eissa Qader,
Bryce Meyering, Kha Tran

Team 062

Georgia Tech

Background

- **Congressional Stock Trading:**
  - Congressional members has made many headlines in the US recently due to insider trading and violation of the STOCK act. In 2021, Congress made an estimated total of $355 million worth of stock trades.

- **STOCK Act:**
  - **Stop Trading on Congressional Knowledge of 2012,** is an Act of Congress designed to combat insider trading.
  - The Act requires Congressmen and Congresswomen to file disclosure within 45 days of transactions

Main Objectives

- **Main objectives:**
  - **Leverage publicly available daily stock price data** and Congressional stock transaction disclosures to **identify ‘suspicious’ trades within Congress**.
  - **Identify the best and worst traders** within each respective house of congress and/or political party by computing their individual returns compared to broader market indices
  - **Identify any distinct clusters of congressional members** based on similar trading **patterns, position, held, cumulative returns.**

Data Source

- Live database containing all the disclosed stock purchases/sales of all U.S. Congress members from Quiver Quant.

- Daily stock data for all the unique ticker symbols found in our congressional trading data.

- Congressional member metadata using ProPublica API

Georgia Tech

Dataset Description

TransactionDate	Ticker	Representative	Transaction	Range
2021-05-13 00:00:00	RVLV	Josh Gottheimer	Purchase	$1,001-$15,000
2019-01-23 00:00:00	QGEN	Donna Shalala	Sale	$1,001-$15,000
2021-03-15 00:00:00	LMND	Marie Newman	Purchase	$1,001-$15,000
2018-12-03 00:00:00	BUD	Sheldon Whitehouse	Sale	$15,001 - $50,000

- **Transaction Date:** Date of the transaction
- **Ticker:** Ticker symbol of the transaction
- **Transaction:** type of transaction, “Sale” or “Purchase”
- **Amount:** Lower bound of the transaction range
- **Representative:** The name of the congressional member
- **Range:** The range of the transaction value

Sample of market data

symbol	date	open	high	low	close	volume	adjusted
AAPL	2021-12-27	177.090	180.420	177.0700	180.3300	74919600	180.10054
AAPL	2018-08-10	51.840	52.275	51.6675	51.8825	98444800	50.12737
AAPL	2020-05-26	80.875	81.060	79.1250	79.1825	125522000	78.32063
AAPL	2021-02-23	123.760	126.710	118.3900	125.8600	158273000	125.11638

- **Symbol:** Stock ticker symbol
- **Date:** Date of the observation
- **Volume:** The number of shares traded
- **Adjusted:** Closing price after accounting for any corporate actions

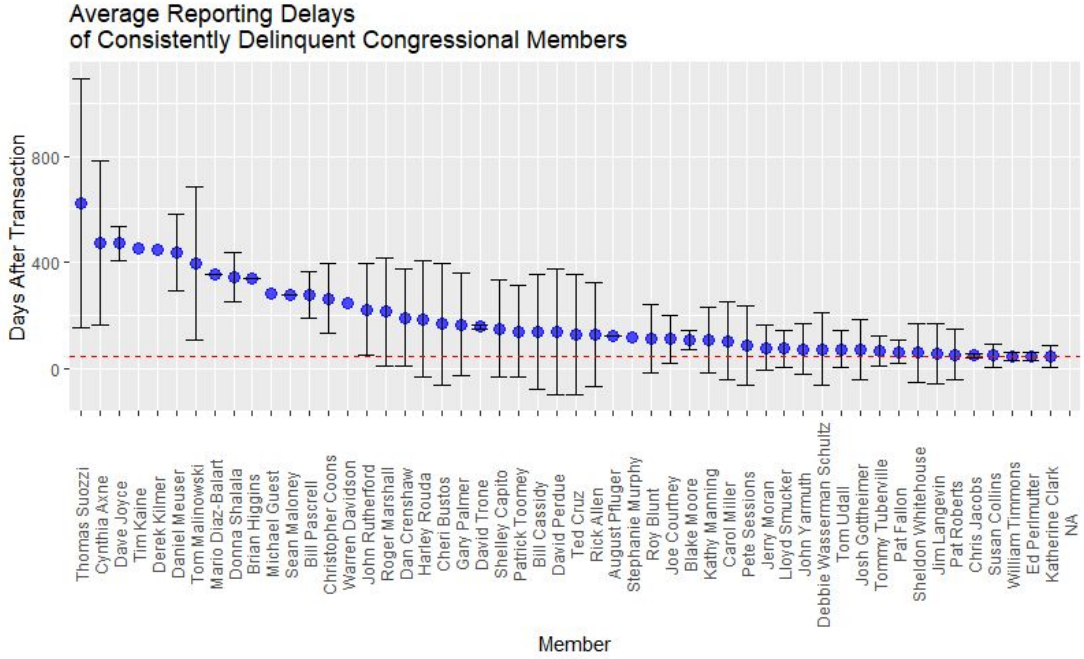
Methodology

Our test steps are as followed:

1. Data cleaning
2. Exploratory Data Analysis (EDA)
3. Clustering
4. Calculating Member Returns
5. Identify Suspicious Trades

Exploratory Data Analysis

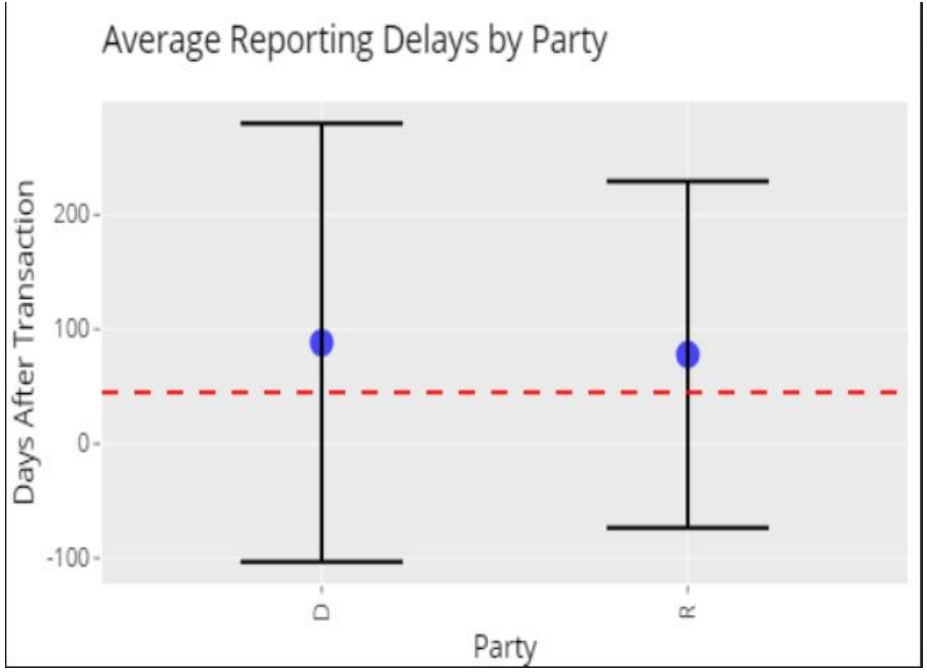
- Worst violator of the STOCK act is Thomas Suozzi
- This could be caused by weak fines/punishments for violating the act ($200 fine)



Georgia Tech

Exploratory Data Analysis

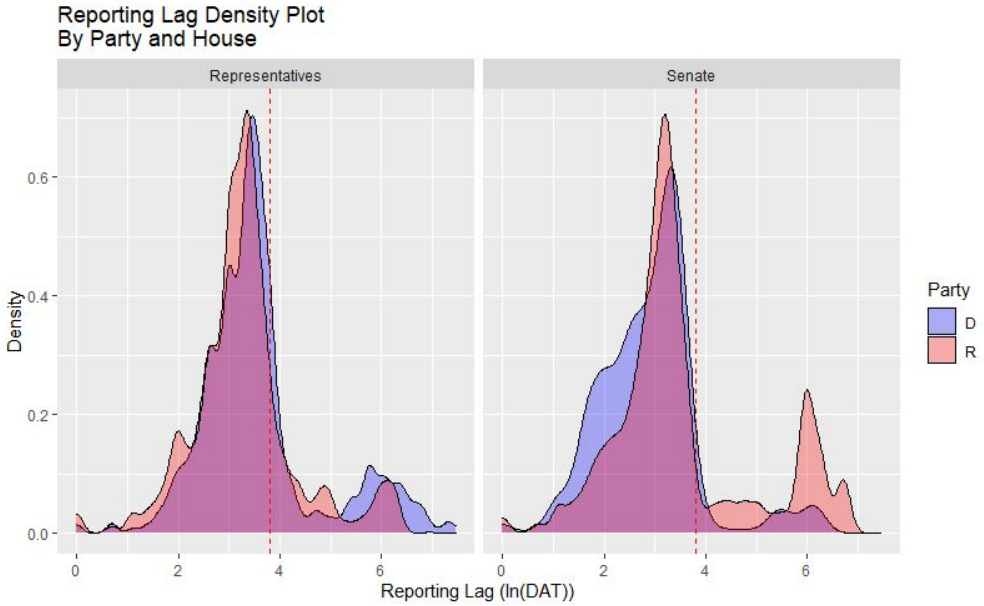
- Both parties average around double the allotted time to report in the STOCK act.
- Democrats average 88 days of lag
- Republicans average 78 days of lag



Georgia Tech

Exploratory Data Analysis

Reporting lag according to STOCK Act, based on Party and House



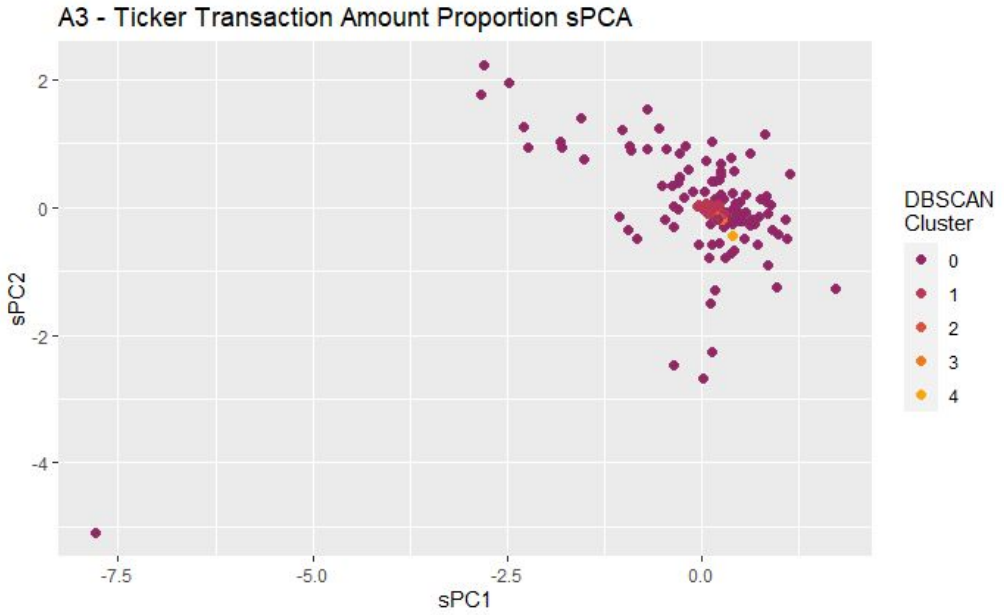
Georgia Tech

Clustering

- Several set of sparse matrices to use in cluster analysis.
- Two basic categories: tickers and transaction dates
- Tickers category:
  - Basic matrix with transactional history for each stock
  - Proportional amount of transactions for each stock
  - Proportional transaction amount for each stock
- Transaction category
  - Transaction date history for each member and calculated matrices for the
    - total transaction amount per day
    - transaction count proportion per day
    - transaction amount proportion per day for each member

Clustering

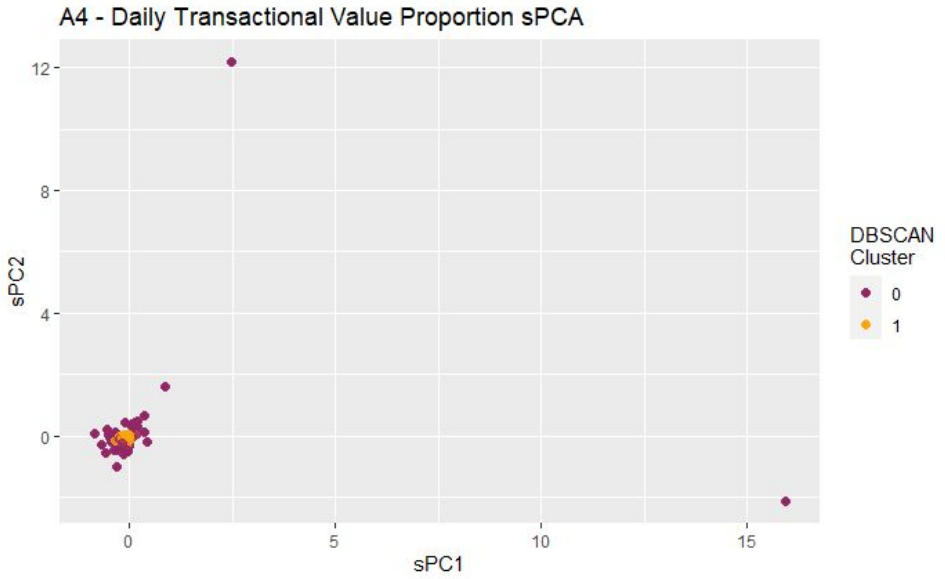
A majority of the congress trade the same stocks with some outliers.



Georgia Tech

Clustering

A majority of congress trades within the same ballpark amounts at similar times, but some whaling officials exist



Georgia Tech

Calculating Member Returns

Assumption:

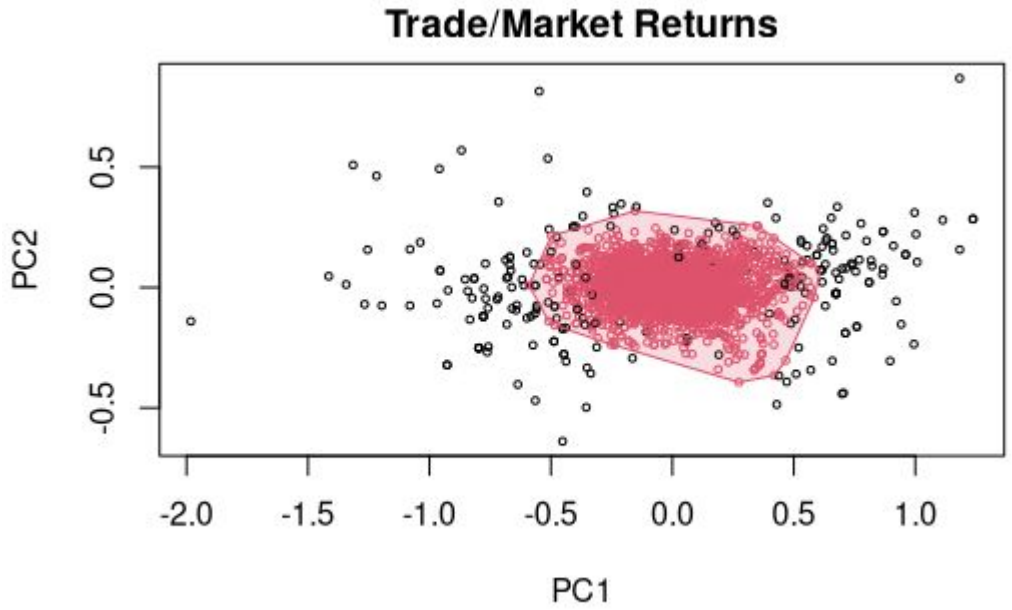
- No member has ever has a negative balance of either stocks or cash
- All executed trades occur at the lower bound of the specified range

Calculating Method:

- Transformed the prices table to a data structure with each column representing a stock ticker and each row represent a date.
- The cells contained the value from the adjusted_close column of the original prices table.
- Added a signed_amount column to the congressional trades dataframe to represent sell orders with a negative number and buy orders with a positive numbers
- Use *tidyquant* and *PerformanceAnalytics* libraries were used to calculate daily returns and alphas

Suspicious Trades

- DBSCAN clustering for individual trades



Georgia Tech

What’s next?

Optimizing Clustering Algorithms

Identify the Best Traders

Suspicious Trading Detection

Georgia Tech

References

- https://www.barrons.com/articles/microsoft-stock-suan-del benes-husband-buys-up-shares-51553191241
- https://www.marketwatch.com/story/u-s-lawmakers-traded-an-estimated-355-million-of-stock-last-year-these-were-the-biggest-buyers-and-sellers-11643639354