Twitter and the Moscow City Duma Election: Collecting Data

EPPS 6302: Methods of Data Collection and Production

November 21, 2019

Introduction

2019 Moscow City Duma Election

- Duma – legislative body of Moscow
- Multiparty system
- 45 constituencies
- Moscow City Election Commission (MCEC)
- Election day – 9/8
- Protests during registration procedures for candidates

Russian Twitter

- Social media platform
- Usage differs based on society and culture
- Interaction between political figures and general public
- Not all political figures or political parties have Twitter accounts

Research Statement

- Original focus: political polarization and sentiment analysis
- Explore whether twitter interactions are impacted by discussions surrounding specific political parties

Research Statement:

Determine whether amount of tweets or discussions concerning different political parties impacts social media platform connections (such as favorites, replies, retweets).

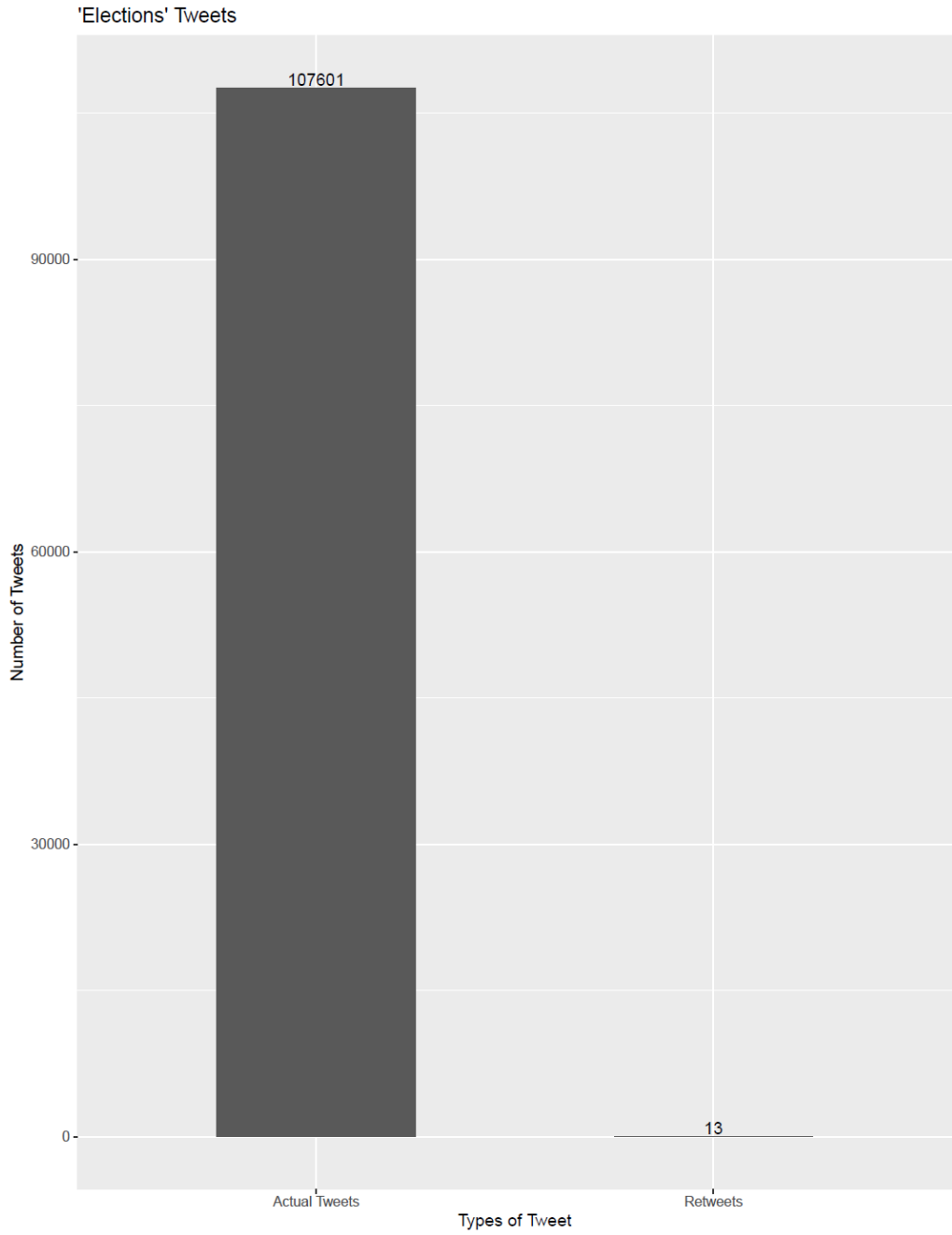
Data Collection Methods

- GetOldTweets3 query search
  - Find Tweets using the word: elections
  - Quantitative and qualitative data
- Time range – 6/5/19 to 9/9/19
  - Twitter API not an option due to limit on time range
- Limitations:
  - Query search of specific election related terms
  - Geographic range

Next step: understand data

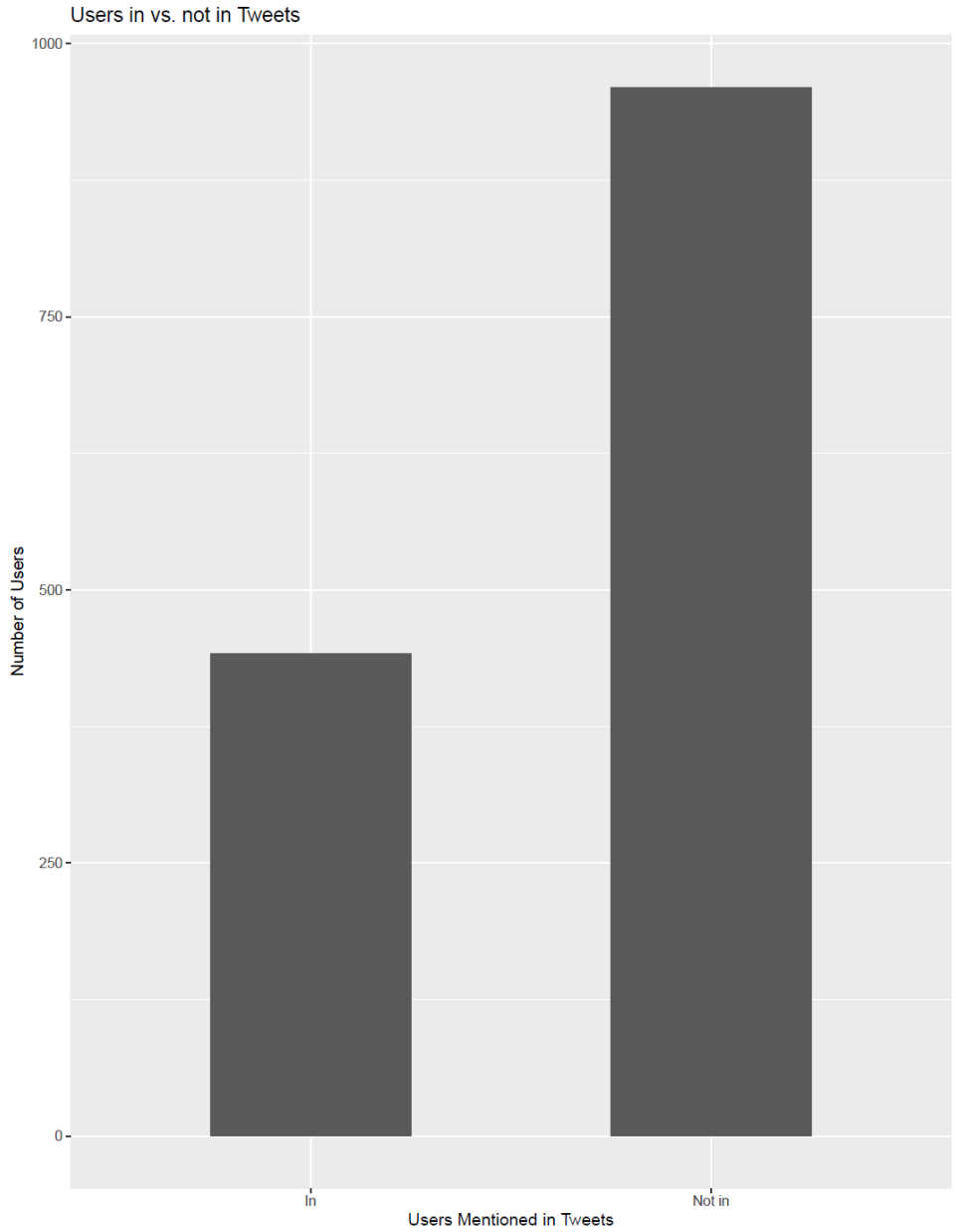
Tweets

Unique Tweets	Total Tweets
100,725.00	107,614.00



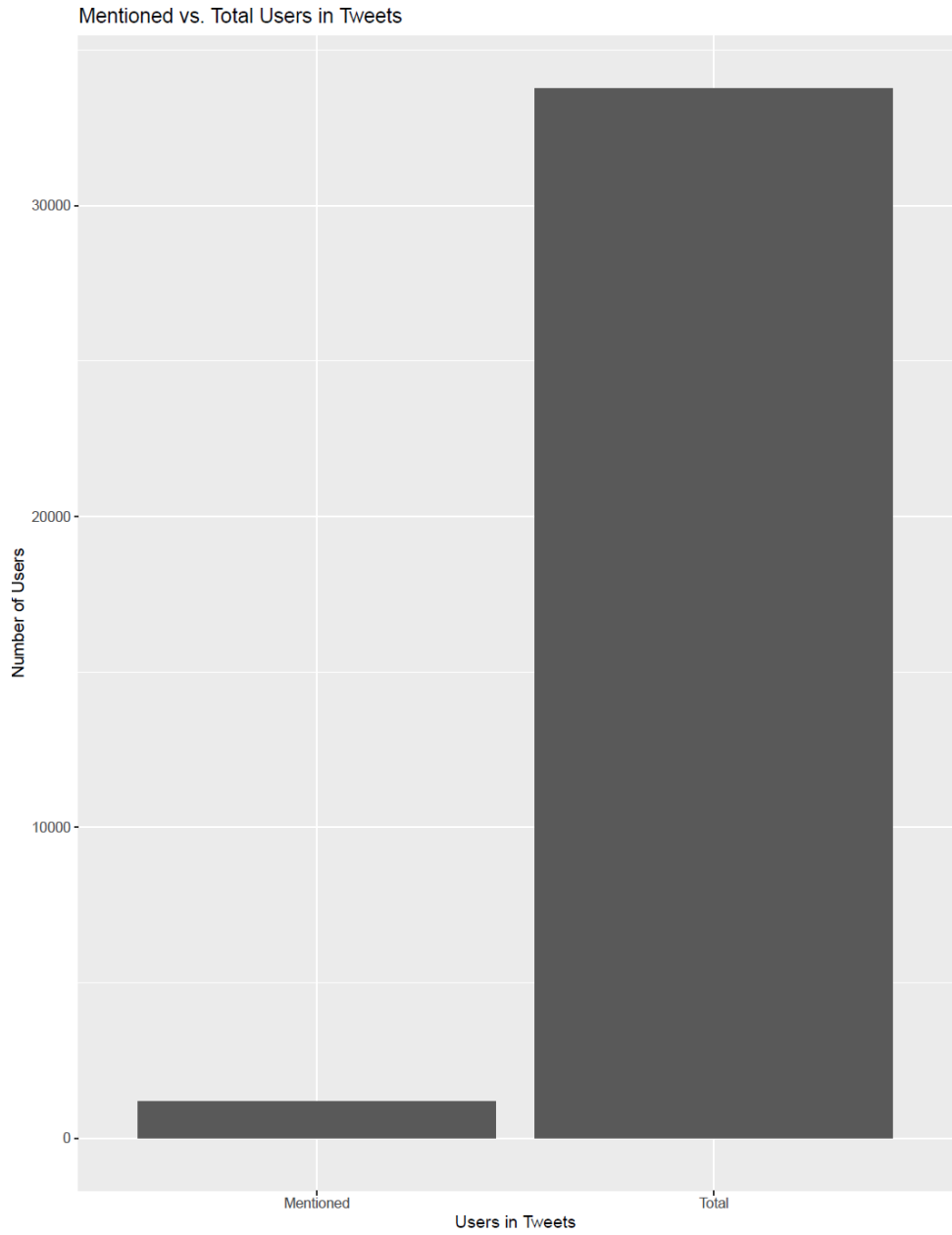
Users

Unique Username	Total Usernames
34,147.00	107,614.00



Mentioned Users

mentioned_users	volume
meduzaproject	304
navalny	181
SobolLubov	167
MosSobyanin	116
CIKRussia	84



Data Manipulation and Production

- Import CSV file into R
  - Cyrillic script
    - Text mining and sentiment analysis unachievable
- Identify users discussing political parties
- Create rows and columns for party affiliation
- Pivot table to process data by reorganizing it
- 10 parties
  - United Russia, Communist Party (CPRF), Yabloko, A Just Russia, LDPR, Rodina, Communists of Russia, The Greens, Civilian power, Party of Growth
- 3 forms of Twitter activity (replies, retweets, favorites)

Data Overview

username	Count.of.text	Sum.of.Twit	Sum.of.party
105	1	0	0
115446	3	2	0
185790	1	6	0
430014	4	17	0
580947	1	0	0
600427	6	3	0

Analysis: Sum.of.Twit ~ Sum.of.party + Count.of.text

Poisson Regression

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.6893	0.0008	4428.76	0.0000
Sum.of.party	0.7908	0.0005	1601.83	0.0000
Count.of.text	0.0108	0.0000	1435.42	0.0000

Quasi-Poisson Regression

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.6893	0.0682	54.06	0.0000
Sum.of.party	0.7908	0.0404	19.55	0.0000
Count.of.text	0.0108	0.0006	17.52	0.0000

Model Dispersion

- Residual Deviance is greater than the degrees of freedom in the Poisson regression, which means that over-dispersion exists
- Residual deviance: 13035397 on 34143 degrees of freedom
- Therefore, the estimates made by the model are correct, but the standard errors/standard deviation are wrong and unaccounted for by the model

Conclusion

- Predictor variables are significant…
  - Number of unique Tweets
  - Number of different political parties discussed
- Response variable requires further study
  - Number of Twitter interactions (replies, retweets, favorites)

- Explore impact of different political parties onto different Twitter interactions

Bibliography

Bhadane, Chetashri, Hardi Dalal, and Heenal Doshi. 2015. “Sentiment Analysis: Measuring opinions.” *Procedia Computer Science* 45: 808-814.

Bodrunova, Svetlana S. 2013. “Fragmentation And Polarization Of The Public Sphere In The 2000s: Evidence from Italy and Russia.” *Global Media Journal-German Edition* 3(1).

Cha, Meeyoung, Fabrício Benevenuto, Hamed Haddadi, and Krishna Gummadi. 2012. “The world of connections and information flow in twitter.” *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 42(4): 991-998.

Colton, Timothy J., and Michael McFaul. 2000. “Reinventing Russia's Party of Power: “Unity” and the 1999 Duma Election.” *Post-Soviet Affairs* 16(3): 201-224.

Conover, Michael D., Jacob Ratkiewicz, Matthew Francisco, Bruno Gonçalves, Filippo Menczer, and Alessandro Flammini. 2011. “Political polarization on twitter.” In *Fifth international AAAI conference on weblogs and social media*: 89-96.

Fiorina, Morris P., and Samuel J. Abrams. 2008. “Political polarization in the American public.” *Annu. Rev. Polit. Sci.* 11(June): 563-588. doi: 10.1146/annurev.polisci.11.053106.153836

Frye, Timothy. 2002. “The Perils Of Polarization: Economic performance in the postcommunist world." *World politics* 54(3): 308-337.

Jackson, Linda A., and Jin-Liang Wang. 2013. “Cultural Differences In Social Networking Site Use: A comparative study of China and the United States.” *Computers in Human Behavior* 29(3): 910-921. doi: 10.1016/j.chb.2012.11.024

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning with Applications in R*. New York City, NY: Springer.

Kolossov, V. A. 1997. “Political Polarization At The National And The Intra-Urban Levels: The role of Moscow in Russian politics and socio-political cleavages within the city.” *GeoJournal* 42(4): 385-401.

Kunicova, Jana, and Thomas Frederick Remington. 2008. “Mandates, Parties and Dissent: Effect of Electoral Rules on Parliamentary Party Cohesion in the Russian State Duma, 1994—2003.” *Party Politics* 14(5): 555-574.

McAllister, Ian, and Stephen White. 2008. “‘It's the Economy, Comrade!’ Parties and Voters in the 2007 Russian Duma Election.” *Europe-Asia Studies* 60(6): 931-957.

Tan, Chenhao, Lillian Lee, Jie Tang, Long Jiang, Ming Zhou, and Ping Li. 2011. “User-level sentiment analysis incorporating social networks.” In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining* (August): 1397-1405. doi: 10.1145/2020408.2020614

Thames Jr, Frank C. 2001. “Legislative Voting Behaviour In The Russian Duma: Understanding the effect of mandate.” *Europe-Asia Studies* 53(6): 869-884.

Yardi, Sarita, and Danah Boyd. 2010. “Dynamic Debates: An analysis of group polarization over time on Twitter.” *Bulletin of science, technology & society* 30(5): 316-327. doi: 10.1177/0270467610380011