

Different Types of Application for Polls and Participation in Toronto

Qingya Li

6/02/2022

Abstract

The dataset reports the public polls conducted by the City between 2015-2021. Studying the dataset can help researchers know how the residents/tenants/owners feel and react to the issues. In this paper, we summarize the number of potential voters and the ballots returned in the polls. We also create plots to find the relationship between application types and the response rate met/final polls results. We conclude that the number of ballots returned is less than the number of potential voters, meaning fewer people cast a vote than expected. We should consider encouraging their participation and conducting the polls more significantly and efficiently. Meanwhile, more polls met the response rate than not meet and more favorable results than opposed among the applications. Besides, the plots show the number of applications for front yard parking is higher than others, the government may need to consider how to better the situations.

The main paragraphs are shown below¹

1 Introduction

Voting is necessary for daily life. The voting results can significantly reflect what people like and dislike, both personal and social (André and Jean-François 2020). As citizens, we have the political right and responsibility to vote (André and Jean-François 2020). The City Clerk's Office responded for the voting in Toronto, including the polls for appeal-front yard parking, front yard parking, boulevard cafe, and boulevard parking ("City Clerk's Office" 2021). Therefore, if people need to require the front yard parking, boulevard parking, and any other type of permissions, they shall submit their applications. The City Clerk's Office would mail the ballots to the residents, tenants, and owners in the polling area ("Information Regarding Cafe, Parking & Traffic Polls" 2019). The results will depend on the number of ballots returned and the number of submitted ballots marked as in favour ("Information Regarding Cafe, Parking & Traffic Polls" 2019).

We need to know the voting results in the city. Because the results can help people know what kind of issue may exist in their community and also help the government know what type of questions people most care about and how people react to these issues. More importantly, to help the government consider implementing some relevant policies to reduce the issues that happen in the future. Besides, by observing the number of ballots and other variables. Government shall consider how to improve the participation of the poll and make the vote more efficient.

In the report, we investigate the polls in the city of Toronto. The data contains information from 2015 to 2021. First, we conduct a table of the potential voters, calculating the sum, maximum, minimum, mean, median, and standard deviation. Second, we create another table that summarizes the number of ballots returned, and we can compare the results to the first table. We find the number of potential voters less than ballots returned. Third, we draw a bar plot to find how the application types relate to the response rate.

¹The files are available at:<https://github.com/QingyaLi/sta304-paper1>

The outcome reflects that participants most care about the front yard parking. Fourth, we draw another bar plot to tell the relationship between application types and the final poll results. The bar plot supports most of the polls passed. Lastly, we conduct a scatter plot to demonstrate the year and the number of ballots returned, grouped by poll results. The plot shows, on average, the number of ballots is below 100 among the years, also most of the results are favorable. The data analyzing process for this report is in **R** (R Core Team 2020). We primarily use the **tidyverse** (Wickham et al. 2019) and **dplyr** (Wickham et al. 2021). The graphs in the report are created with the packages of **ggplot2** (Wickham 2016), the tables are made with the **knitr** (Xie 2021).

2 Data

2.1 Data Source

The data set is published by The City Clerk’s Office, which aims to build public trust in the government (“City Clerk’s Office” 2021). The Office provides various types of support and services. Including managing the municipal election, making the information to the public with privacy protection, organizing the decision-making process for the Council, and so forth (“City Clerk’s Office” 2021). The Office mail the ballots to the residents and owners and provide information about the poll (“Information Regarding Cafe, Parking & Traffic Polls” 2019). People can cast to vote by submitting their ballots (“Information Regarding Cafe, Parking & Traffic Polls” 2019). The data set contains information about the poll relevant to numerous topics regarding property owners, residents, businesses in the neighborhoods. It is collected from April 1, 2015, updated after the closed date and certification of the polls. Recently, the data source is available on the City of Toronto Open Portal (“Polls Conducted by the City” 2022) with the **opendatatoronto** package (Gelfand 2020).

2.2 Data Collection

The Polls Conducted by the City data set includes polls data from 2015 to 2021 (“Information Regarding Cafe, Parking & Traffic Polls” 2019). The population is all residents, tenants, owners in Toronto (“Information Regarding Cafe, Parking & Traffic Polls” 2019). The participants in the sample are required to be over 18 years old (“Information Regarding Cafe, Parking & Traffic Polls” 2019). The applications involved appeal-front yard parking, commercial boulevard parking, boulevard cafe, and front yard parking. The data set gives various details to describe the polls, such as the location, the number of ballots returned, the number of ballots received as “In favour” or “opposed”, etc. The data is updated daily. However, the data set does not explain the difference between the appeal - front yard parking and front yard parking. Readers may be confused about these two types of application polls. In addition, non-response bias may exist in the designed polls. Because some respondents are neither in favour nor against, they decided not to submit the ballots. In this case, the response rate may fail, leading to insignificant and unreliable results. The results cannot represent the opinions of all the people nor reflect the situations. Thus, we should consider the bias in future studies.

2.3 Data Characteristic and Exploratory Data Analysis

Overall, the data set has 25 variables and 1054 observations. There are 15 categorical and 10 numerical variables. This report mainly focuses on analyzing several variables including **APPLICATION_FOR** (application type), **OPEN_DATE** (the open date of the polls), **BALLOTS_CAST** (the number of ballots returned), **RESPONSE_RATE_MET** (the number of ballots returned that has met the response rate), **POLL_RESULT** (final polls results), and **POTENTIAL_VOTERS** (the number of voters living in the poll boundary range). Specifically, 25% of the ballots need to be returned in order to meet the response rate, and 50% plus 1 of the ballots have to be marked as in favor to pass the poll (“Information Regarding Cafe, Parking & Traffic Polls” 2019). A cleaning data set process is maintained. All the missing observations

are removed by `na.omit()` function and 710 observations are remaining. We convert the **OPEN_DATE** variable from categorical to date with `as.Date()` function since it is easier to plot the variable in the following. We create a new variable named **OPEN_YEAR** from **OPEN_DATE** with `as.numeric()` function. **OPEN_YEAR** indicates the year the poll opened to the public. The exploratory Data Analysis (EDA) process is shown below.

2.3.1 Tables

Table 1 summarizes the number of potential voters who live in the poll boundary range. The table is grouped by application types of the polls and year. We are interested in analyzing the potential voters since we intend to find out how many people can be influenced directly by the results of the polls. We can compare the outcome to the number of people who voted.

Table 1: Summary Table for The Number of Potential Voters

APPLICATION_FOR	YEAR	sum	min	max	median	mean	sd
Appeal - Front Yard Parking	2015	3947	34	216	96.0	101.20513	48.011610
Appeal - Front Yard Parking	2016	9280	21	331	103.0	111.80723	51.640988
Appeal - Front Yard Parking	2017	5436	24	228	131.0	123.54545	50.252597
Appeal - Front Yard Parking	2018	4155	16	902	94.0	129.84375	152.511472
Appeal - Front Yard Parking	2020	80	80	80	80.0	80.00000	NA
Boulevard Cafe	2015	1931	23	341	105.5	107.27778	76.050916
Boulevard Cafe	2016	2872	12	201	114.5	102.57143	51.787621
Boulevard Cafe	2017	2807	24	239	89.0	103.96296	59.973380
Boulevard Cafe	2018	2773	56	345	147.0	154.05556	82.722408
Boulevard Cafe	2019	9413	29	7009	112.0	427.86364	1470.911357
Commercial Boulevard Parking	2015	604	54	333	108.5	151.00000	124.137021
Commercial Boulevard Parking	2016	299	40	96	81.5	74.75000	24.157470
Commercial Boulevard Parking	2018	110	19	91	55.0	55.00000	50.911688
Commercial Boulevard Parking	2019	128	16	63	49.0	42.66667	24.131584
Commercial Boulevard Parking	2021	93	40	53	46.5	46.50000	9.192388
Front Yard Parking	2015	3150	29	194	90.0	90.00000	44.661472
Front Yard Parking	2016	7851	21	483	97.0	109.04167	67.989526
Front Yard Parking	2017	7820	12	1069	93.0	124.12698	164.246421
Front Yard Parking	2018	7592	17	292	96.0	106.92958	53.964889
Front Yard Parking	2019	4642	21	248	91.5	96.70833	49.864480
Front Yard Parking	2020	2708	28	165	109.0	100.29630	34.845610
Front Yard Parking	2021	7076	11	734	97.0	105.61194	89.839749

We notice that the sum of the potential participants for the appeal-front yard parking, boulevard cafe, and front yard parking are much greater than commercial boulevard parking. Moreover, the minimum and maximum number of potential respondents among these applications are not too different, except the maximum number of potential voters for boulevard cafe in 2019 is extremely high, meaning the many voters can be affected by the result of this poll. And the mean and standard deviation are also high for this row. Lastly, one missing value appears in the standard deviation for the appeal - front yard parking in 2020, because there is only one poll for this application in the year.

Table 2 summarizes the number of ballots returned, including the sum, minimum, maximum, median, mean, and standard deviation, by the application types and occurred year.

Table 2: Summary Table for The Number of Ballots Returned

APPLICATION_FOR	YEAR	sum	min	max	median	mean	sd
Appeal - Front Yard Parking	2015	1490	12	76	37.0	38.20513	16.557999
Appeal - Front Yard Parking	2016	3338	7	99	39.0	40.21687	17.712506
Appeal - Front Yard Parking	2017	1775	6	94	35.0	40.34091	19.535231
Appeal - Front Yard Parking	2018	1095	5	57	36.5	34.21875	14.998353
Appeal - Front Yard Parking	2020	25	25	25	25.0	25.00000	NA
Boulevard Cafe	2015	494	0	104	29.0	27.44444	24.149588
Boulevard Cafe	2016	640	0	61	21.5	22.85714	15.927415
Boulevard Cafe	2017	581	2	48	20.0	21.51852	14.409853
Boulevard Cafe	2018	504	13	48	27.5	28.00000	9.305280
Boulevard Cafe	2019	930	14	275	28.0	42.27273	53.196117
Commercial Boulevard Parking	2015	58	10	26	11.0	14.50000	7.681146
Commercial Boulevard Parking	2016	33	1	23	4.5	8.25000	10.045729
Commercial Boulevard Parking	2018	27	8	19	13.5	13.50000	7.778175
Commercial Boulevard Parking	2019	38	4	18	16.0	12.66667	7.571878
Commercial Boulevard Parking	2021	37	17	20	18.5	18.50000	2.121320
Front Yard Parking	2015	1137	13	63	31.0	32.48571	12.980464
Front Yard Parking	2016	3086	7	96	41.0	42.86111	18.401091
Front Yard Parking	2017	2177	5	74	36.0	34.55556	14.982547
Front Yard Parking	2018	2641	8	91	36.0	37.19718	14.858975
Front Yard Parking	2019	1835	11	75	36.0	38.22917	16.406070
Front Yard Parking	2020	1189	11	70	42.0	44.03704	15.710069
Front Yard Parking	2021	2555	7	77	38.0	38.13433	17.714825

By observing the sum columns, we notice that generally, the total number of ballots returned for appeal-front yard parking and front yard parking are greater than boulevard cafe and commercial boulevard parking. The polls of (appeal -) front yard parking have more than 1,000 ballots returned, except appeal - front yard parking in 2020 only have 25 returned. The results illustrate that the front yard parking issue exists, and participants care more about this problem than other types of applications. We can also conclude that the actual participants are less than the potential voters. The minimum and maximum columns demonstrate the boulevard cafe poll received the least ballots returned in both 2015 and 2016 (0 ballots returned), the most ballots returned in 2019. On average, the number of ballots returned for each application is between 20-40, except for the commercial boulevard parking. Its average is between 10-20. Besides, the standard deviation of appeal - front yard parking in 2020 is missing, because there is only one observation. Meaning in 2020 had one poll for the front yard parking. We cannot calculate the standard deviation when the mean is the one observation and the variance is zero.

2.3.2 Figures

Figure 1 draws the relationships between application types and the response rate met (yes/no).

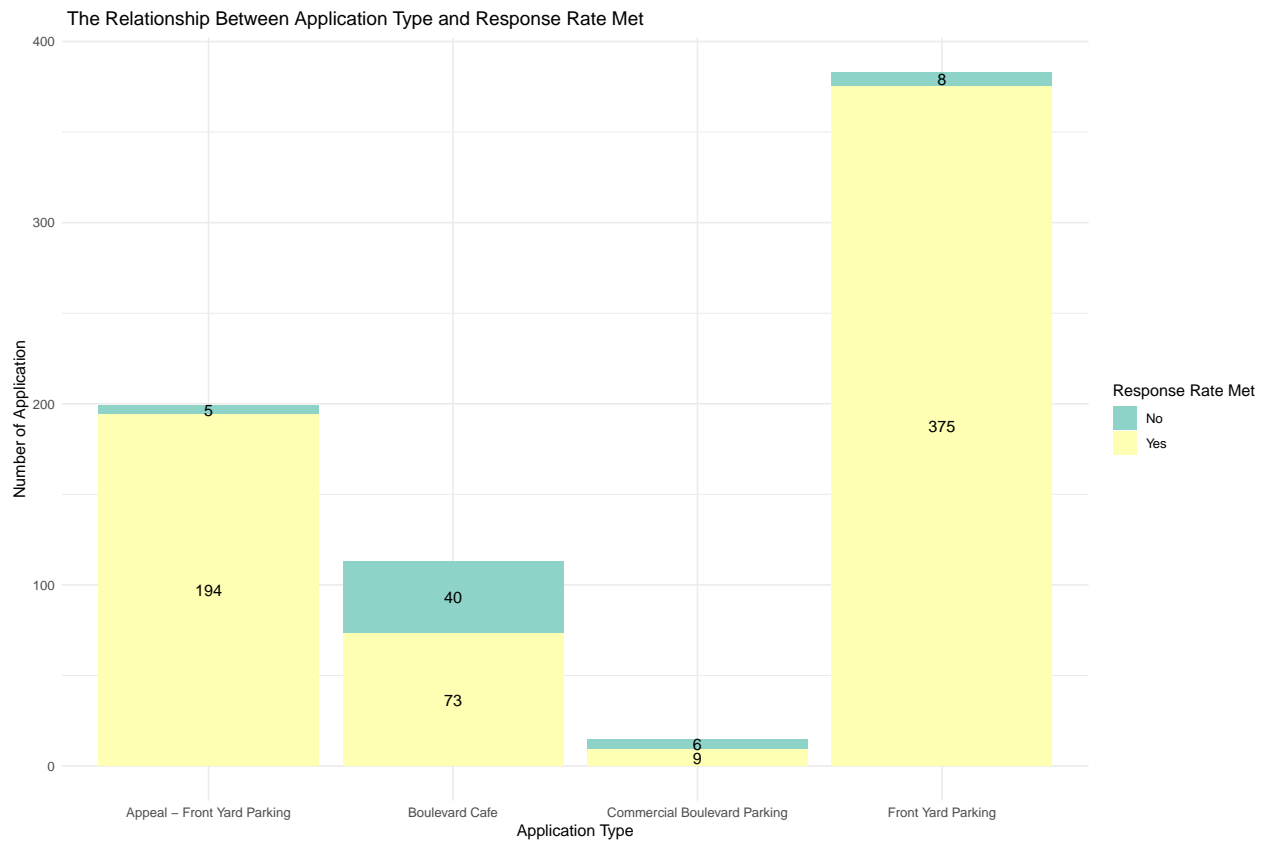


Figure 1: Comparing the application types by response rate met

Overall, we can see response rate met in all the applications is higher than the rate did not meet. Specifically, the response rate met for the front yard parking application is the highest (375 polls met and only 8 polls did not meet the response rate), the response rate met for the commercial boulevard parking is the lowest (9 met and 6 did not meet). Once meet the response rate, the poll is considered valid.

Figure 2 shows the relationship between the final results of the polls (in favor, opposed, response rate not met) and application types.

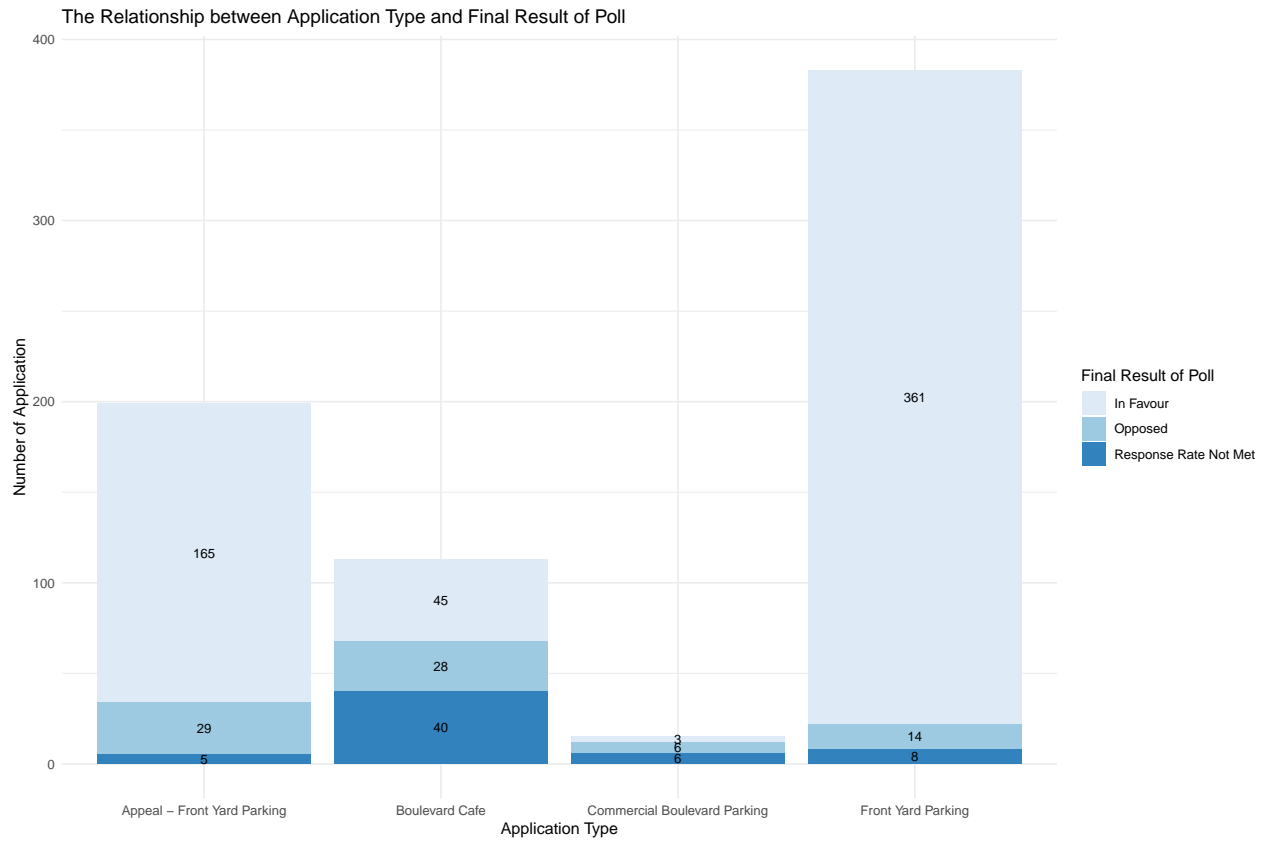


Figure 2: Comparing application types by the final results

Note that generally, we have more in favour results than opposed results among the applications, except the commercial boulevard parking (3 in favour, and 6 against). Front yard parking polls have the most in favour results. And appeal - front yard parking pools have the most opposed results.

Figure 3 We conduct a scatter plot to indicate the relationship between the year and number of ballots returned, by poll results.

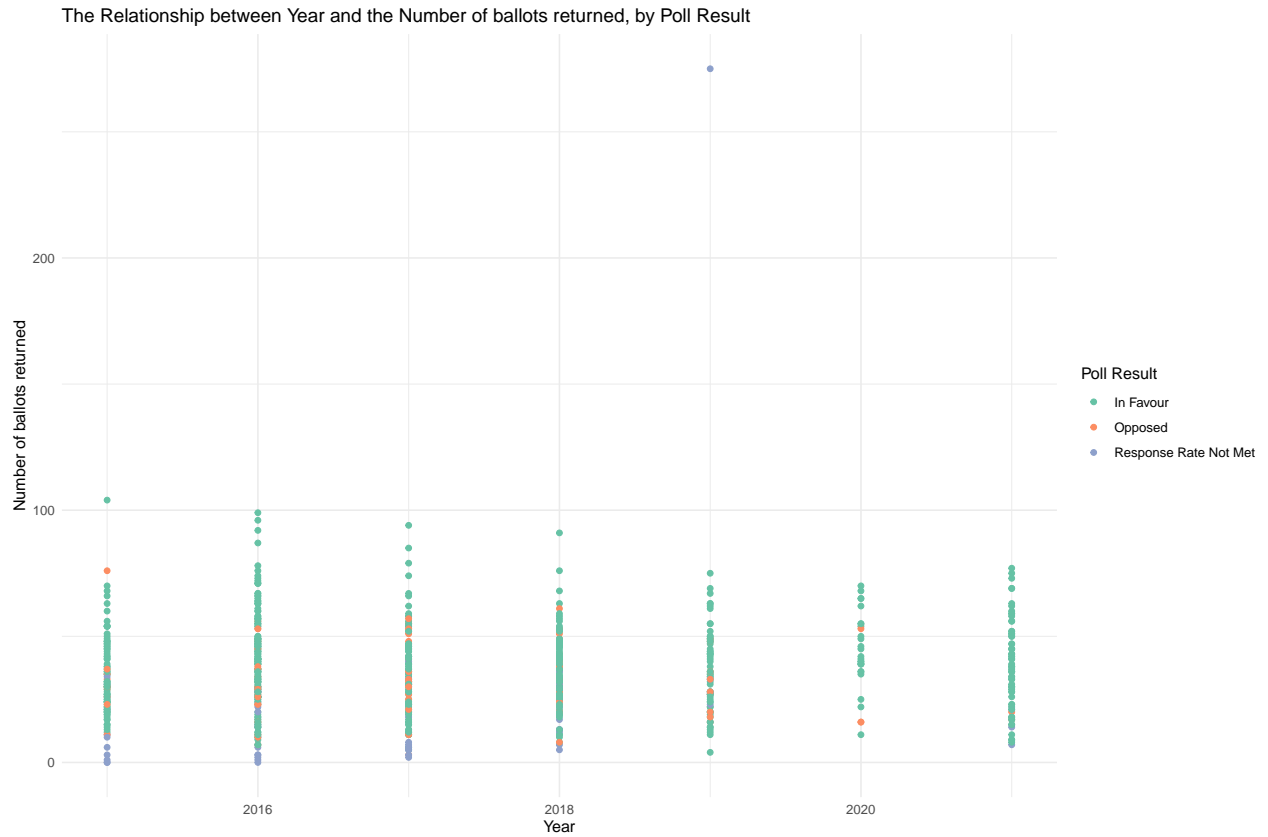


Figure 3: The relationship between the year and number of ballots returned, by poll results

We can see an outlier in 2019, which the number of ballots returned is high but it seems did not meet the response rate. This case may be discussed in further study. Most of the polls are approved instead of against. We can also see that on average, the number of ballots returned is below 100 among these years.

3 Reference

- André, Blais, and Daoust Jean-François. 2020. *The Motivation to Vote: Explaining Electoral Participation*. UBC Press. <https://books-scholarsportal-info.myaccess.library.utoronto.ca/uri/ebooks/ebooks5/upress5/2020-04-09/1/9780774862691>.
- “City Clerk’s Office.” 2021. City of Toronto. <https://www.toronto.ca/city-government/accountability-operations-customer-service/city-administration/staff-directory-divisions-and-customer-service/city-clerks-office/>.
- Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://sharlagelfand.github.io/opendatatoronto/,%20https://github.com/sharlagelfand/opendatatoronto/>.
- “Information Regarding Cafe, Parking & Traffic Polls.” 2019. City of Toronto. <https://www.toronto.ca/city-government/planning-development/polls-regarding-changes-in-a-neighbourhood/information-regarding-cafe-parking-traffic-polls/>.
- “Polls Conducted by the City.” 2022. City Clerk’s Office. <https://open.toronto.ca/dataset/polls-conducted-by-the-city/>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Xie, Yihui. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in R*. <https://yihui.org/knitr/>.