Procurement Dynamics: Analyzing Federal Contract Awards to Microsoft*

Trends and Insights of Microsoft's Future in Canadian Government from 2020

Deyi Kong

November 26, 2024

This paper analyzes the federal procurement contracts awarded to Microsoft, focusing on trends in award values, contract categories, and the relationships between buyers and Microsoft. The analysis reveals that Microsoft dominates high-value contracts and demonstrates consistent partnerships with key government departments. This paper utilizes a linear model to predict the future contract amount trend for Microsoft. By examining these procurement dynamics, this paper providing insights into the broader implications of tech-industry participation in government procurement.

Table of contents

---

*Code and data are available at: https://github.com/eeeee-cmd/Procurement/.

1 Introduction

Government procurement serves as a critical avenue for public sector innovation, economic stimulus, and the fulfillment of strategic objectives. Within this landscape, the relationship between major technology providers and governmental bodies has become increasingly prominent. Among these, Microsoft has emerged as a pivotal partner in driving IT modernization and infrastructure development. Despite the growing reliance on tech giants, limited research has comprehensively examined the dynamics of such partnerships, particularly in the context of federal procurement contracts.

This paper addresses this gap by conducting a detailed analysis of federal procurement contracts awarded to Microsoft between 2020 and September 2024. Using publicly available data from the Investigative Journalism Foundation, the study explores trends in contract values, contract types, and Microsoft's relationships with key government buyers. Emphasis is placed on identifying patterns in contract allocation and the implications of Microsoft's dominance in high-value contracts.

The main estimand in this study is the amount of federal procurement contracts awarded to Microsoft. This metric is analyzed across multiple dimensions, including contract categories, durations, and partnerships, to quantify Microsoft's role in federal procurement comprehensively. The findings reveal that Microsoft plays a dominant role in IT infrastructure and modernization, securing the majority of high-value contracts and establishing consistent partnerships with departments such as Employment and Social Development Canada and National Defence. These insights highlight the broader implications of the private tech sector's influence on public sector procurement, raising questions about competitiveness, dependency, and strategic alignment.

This study contributes to understand the procurement landscape, providing valuable information for policymakers, industry stakeholders, and academics. It underscores the importance of transparency, strategic planning, and balancing innovation with equity in government contracts.

The structure of the paper is as follows: Section 2 outlines the data sources and variables considered, followed by the model setup in Section 3.1 and justification. The results in Section 4 presents the key findings of the analysis, with a discussion on the implications. Section 5 then discusses potential limitations and suggestions for future research. Section 6 provides additional detailed information about the data, model and methodology.

2 Data

2.1 Overview

The data used in this analysis comes from publicly available procurement data (Investigative Journalism Foundation 2024a). The analysis uses the statistical programming language R (R Core Team 2023) and several libraries, including `tidyverse` (Wickham et al. 2019), `janitor` (Firke 2023), `knitr` (Xie 2024), `dplyr` (Wickham et al. 2023), `arrow` (Richardson et al. 2024), `purrr` (Wickham and Henry 2023), and `here` (Müller 2020) for data manipulation. `ggplot2` (Wickham 2016) and `kableExtra` (Zhu 2024) for visualization. The dataset covers the information of chosen Federal Procurement Supplier - Microsoft - capturing the contracts from 2020 to 2024 (up to September).

2.2 Measurement

The dataset used in this analysis originates from Canadian governments of all levels post solicitations for companies to search for business opportunities, as documented in Investigative Journalism Foundation (2024b), then collects data from each of these platforms and unifies into standard format.

The measurement process delegates contracts routed through Public Service and Procurement Canada (PSPC), and blanket disclosures organized by Treasury Board of Canada Secretariat (TBS) to understand the federal procurement. The methodologies of federal solicitations are either competitive or non-competitive. Competitive contracts accept bids from all or some subset of suppliers. Non-competitive contracts do not, citing certain exceptions, including a "pressing emergency in which delay would be injurious to the public interest," or when the estimated cost to the buyer is below certain thresholds defined for some goods and services.

More information about solicitation method can be found in Section 6.

2.3 Data Cleaning

After imported the raw dataset using `read_csv` function, I first select specific columns to focus on relevant information towards this study by omitting redundant information, such as the contract link and same value of region. Then I rename key variables from raw data for clarity. The key variables of interest in this analysis include Contract, StartDate, AwardDate, EndDate, Buyer, and Amount.

For any missing data (NA) of StartDate, I infer the StartDate from the AwardDate to replace the missing value as the value of AwardDate. Then I standardize the date format of all date related datas for proper parsing by using `mdy()` function and change the dollar format of amount into numeric numbers for future calculation purposes. Due to the error with unknown reason that EndDate is before StartDate or AwardDate, those entries were excluded from this paper.

PreparatoryPhase, PhaseDays, and ContractDays variables were created for the purpose of analyzing preparatory phase and performance phase corresponding to each contract. PreparatoryPhase calculates the days from AwardDate to StartDate. ContractDays calculates the days from StartDate to EndDate. PhaseDays take the absolute value of PreparatoryPhase because of the minor effect of either prepare before contract being awarded or prepare after contract being awarded.

This dataset contains only two entries for 2019 federal procurement contracts awarded to Microsoft. So, I filter out the contracts before 2020 to focus on the records between 2020 and September 2024 as desired. Then I check the character type variables, Supplier and Contract, using `table()` function for the purpose of manipulating input error. I standardize the supplier's name which combine the similar name due to capitalization or some notation differences, and format them with first letter of each word capitalized.

After completing the cleaning, we saved the final dataset in both Parquet and CSV formats for later analysis.

2.4 Outcome Variables

The main outcome variable of interest is the amount of federal procurement contracts awarded to Microsoft. This variable is need to represent the financial magnitude of Microsoft with the federal government and measurement of evaluating procurement dynamics. Figure 1 illustrates the annual total amounts of federal procurement contracts awarded to Microsoft from 2020 to 2024 (up to September). The histogram reveals a generally increasing trend of total value of each year's contract. This may reflect the growing of technology in government operations, including software, cloud solutions, and IT infrastructure. The amount in 2023 doubles the amount in 2021, which suggests a heightened investment in federal procurement funding and large-scale projects.
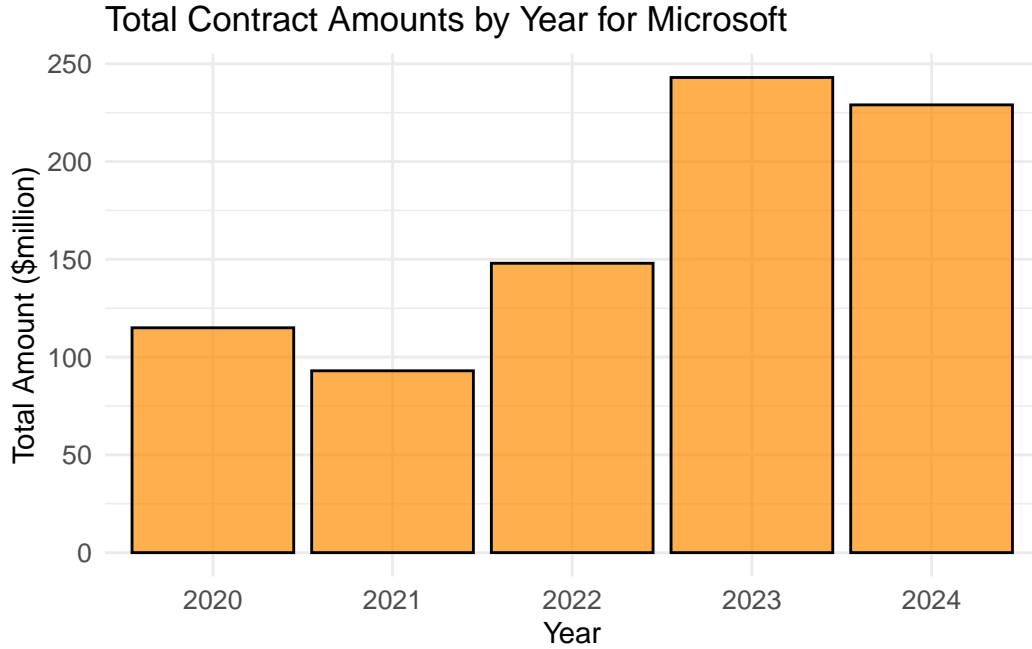
Figure 1: The histogram shows the distribution of the total amount($million) of each year's federal contract (since 2020).

2.5 Predictor Variables

In this analysis, several key predictors were identified to evaluate procurement data effectively. The first variable, **Contract**, is a character-type variable that identifies each procurement contract awarded to Microsoft. Table 1 [SOME MODIFICATION OF CONTRACT CATEGORIES CLEANING PROCESS MAY APPLY]

Table 1: The table summarizes the contract type which has over 25 contracts with the total dollar amount, and average dollar amount per contract.

Contract Type	Counts	Total Amount	Average
License/Maintenance fees for client software related to Distributed Computing Environment (DCE)	188	$77,057,325	$409,879
License/Maintenance fees for application software (including COTS) and application development and delivery software	112	$77,896,809	$695,507
Client software related to Distributed Computing Environment (DCE)	85	$18,065,532	$212,536

5

License/Maintenance fees for operating system and utility software related to servers, storage, peripherals and components	77	$145,171,773	$1,885,348
Computer services (includes IT solutions/deliverables as well as IT managed services)	70	$71,226,748	$1,017,525
Information technology and telecommunications consultants	65	$64,241,246	$988,327
Application software (including COTS) and application development and delivery software	44	$81,845,321	$1,860,121
Computer services	31	$31,537,735	$1,017,346

Buyer is a character-type variable representing the federal government department or agency who initiate and award the contract. Table 2 shows the top 10 buyer's name who has over 25 contracts with Microsoft, which identifies key partners and understand the distribution of contracts across sectors. The table highlights that Employment and Social Development Canada (ESDC) and National Defence account over 50% of Microsoft's total awards. [CALCULATED 52.02%, BUT THINKING IF THIS SHOULD BE ADDED TO THE TABLE]

Table 2: The table summarizes top 10 buyer's name who has over 25 contracts with Microsoft in total with contract count, total dollar amount, and average dollar amount per contract.

Buyer	Contract	Total Amount	Average
Employment and Social Development Canada	117	$238,875,858	$2,041,674
Global Affairs Canada	71	$30,011,093	$422,691
National Defence	64	$81,913,883	$1,279,904
Natural Resources Canada	35	$12,856,204	$367,320
National Research Council Canada	32	$4,436,990	$138,656
Transport Canada	31	$16,297,432	$525,724
Indigenous Services Canada	28	$5,000,374	$178,585
Natural Sciences and Engineering Research Council of Canada	26	$4,626,898	$177,958
Shared Services Canada	26	$17,588,415	$676,478
Social Sciences and Humanities Research Council of Canada	26	$4,626,898	$177,958

Supplier is a character-type variable includes all the affiliated subsidiary companies of Microsoft. Table 3 shows the top 5 buyer's name who has awarded over 25 contracts. [DELETE THIS VARIABLE? OR MAYBE MOVE IT TO APPDX.]

Table 3: The table summarizes the main Microsoft supplier's name who has over 25 contracts with buyers in total with contract count, total dollar amount, and average dollar amount per contract.

Suppllier	Contract	Total Amount	Average
Microsoft Canada Inc.	693	$597,013,345	$861,491
Microsoft Corp.	117	$41,508,411	$354,773
Microsoft Canada	67	$44,881,593	$669,875
Microsoft	35	$47,462,149	$1,356,061
Microsoft Canada Co.	29	$87,516,151	$3,017,798

This dataset contains a series of date-type variables capturing key milestones in the procurement process. **Award Date** is the date when the contract was officially granted to Microsoft. **Start Date** is the date when the contracted services or goods provision begins. **End Date** is the date when the contract is completed or terminated. These variables are essential for analyzing procurement timelines and trends over the study period.

PhaseDays is a numeric variable demonstrates the days between AwardDate and StartDate. In Figure 2, there is a high concentration around 0 indicates rapid procurement processes or emergency software services. Figure 2 also shows a few outlying contracts with a mobilization period of over one year. These long mobilization periods reflects issues such as delays in the procurement process, a long negotiation period, or administrative approvals.
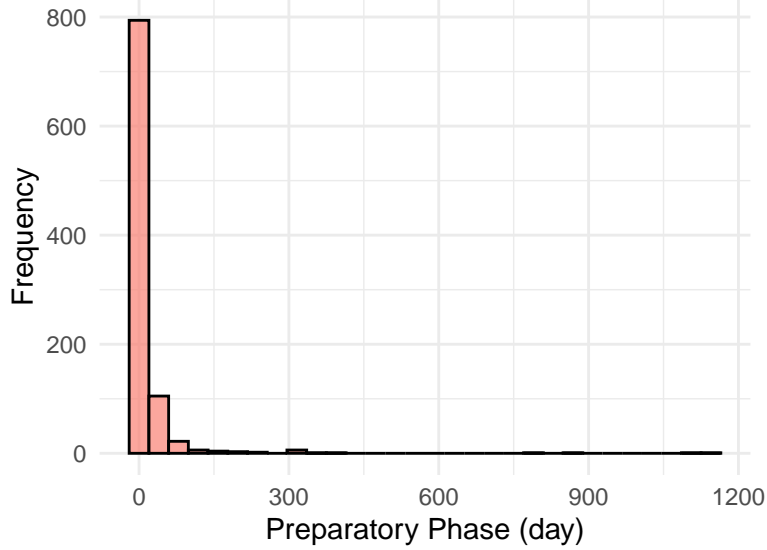

Figure 2: The histogram shows the distribution of preparatory period.

ContractDays is a numeric variable which is the days between StartDate and EndDate. Figure 3 shows the distribution of performance period. The right skewed distribution in Figure 3 indicates that the majority of contract duration is within a year. There is an over 5 years performance period in Figure 3 implies a massive project that requires extensive preparation.
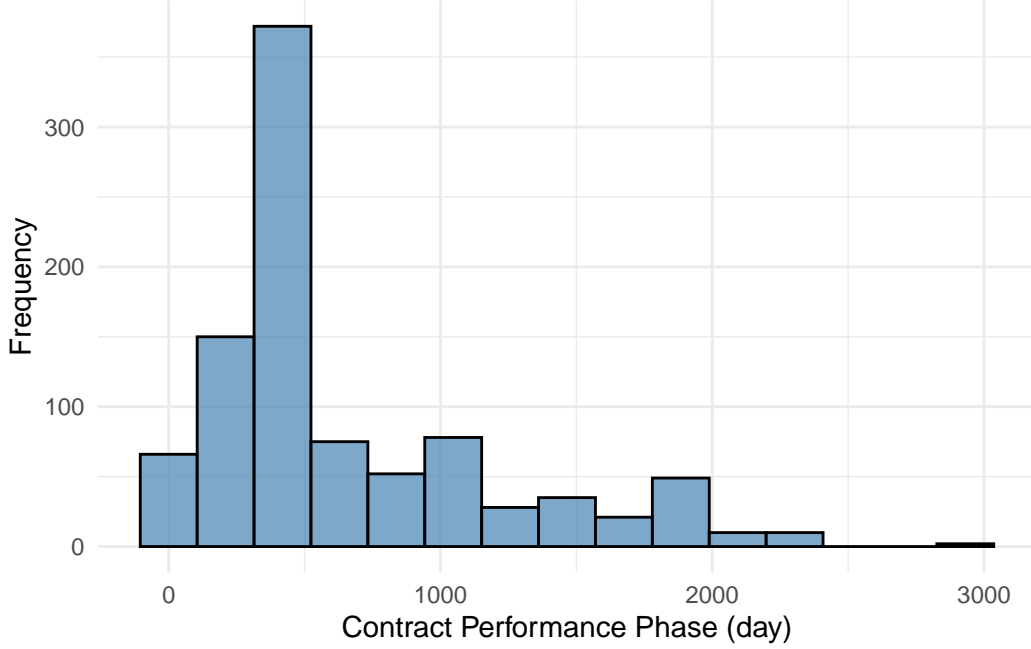

Figure 3: The histogram shows the distribution of performance period, which demonstrates the days between start and end date.

3 Model

To understand the factors influencing the amount of federal procurement contracts awarded to Microsoft. I construct a linear regression model to explains variations in contract values.

3.1 Model Set-Up

This paper utilizes a linear regression model to examine the relationships between the dependent variable.

$$Amount = \beta_0 + \beta_1 Contract + \beta_2 Buyer + \beta_3 ContractDays + \beta_4 PhaseDays + \epsilon$$

As defined in Section 2, each predictor variable has been chosen to reflect characteristics that influence the contract amount.

3.2 Assumptions and Limitations

The model takes several assumptions, including the relationship between the predictors and the outcome is linear, each contract is independent, the variance of the error term is constant, and predictor variables are not highly correlated with each other.

While the model presented offers valuable insights into the dynamics of Microsoft's federal procurement contracts, there are some limitations including categorization of contract types, generalizability, data quality, potential for endogeneity, and unobserved variables. These limitations will be discussed in Section 5.

3.3 Model Validation?

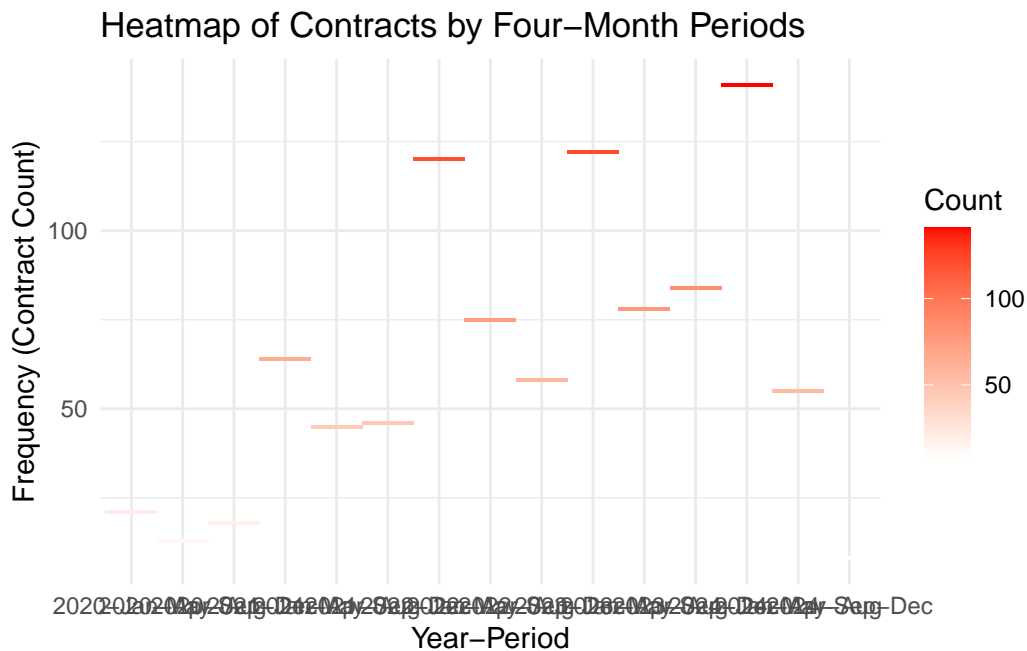R-squared, adjusted R-squared

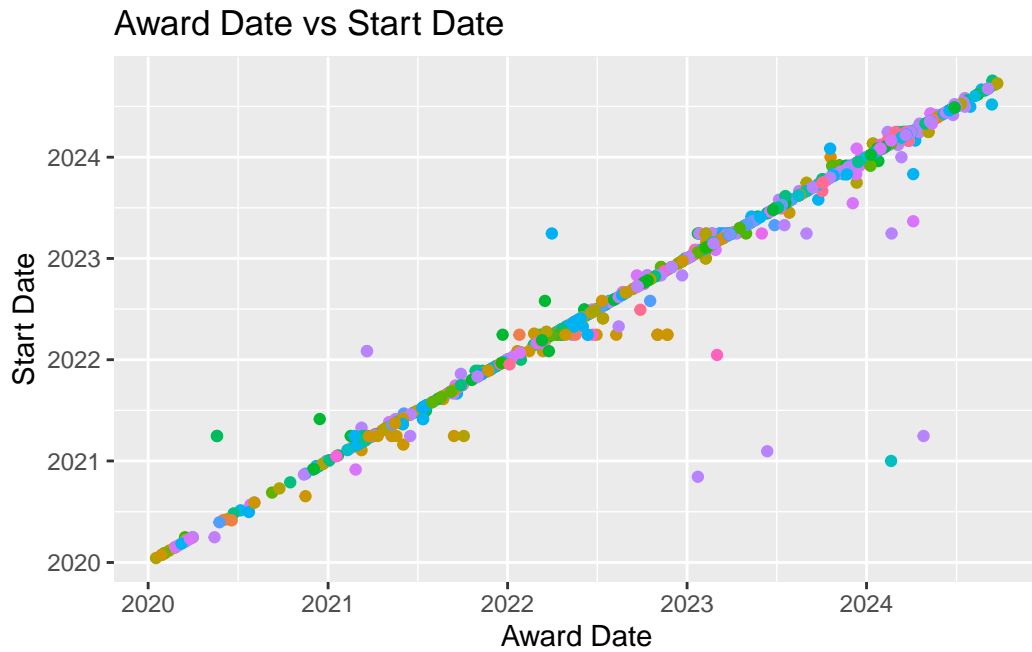tests (e.g. VIF, AIC, BIC).

4 Results

Figure 4: The heatmap ….

Award Date vs Start Date

Figure 5: The scatter plot shows ...
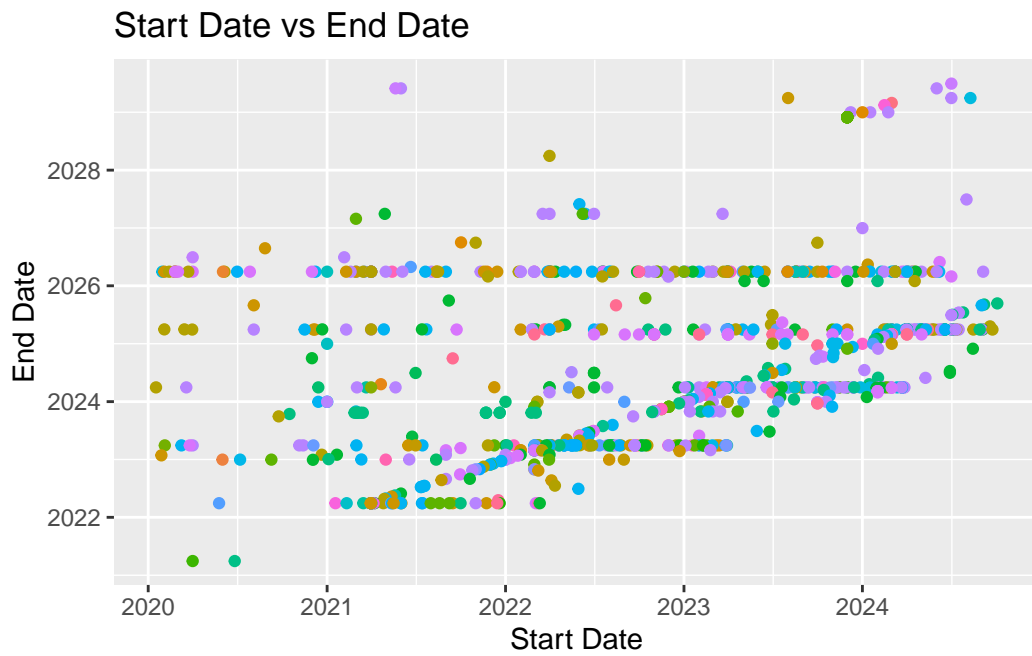
Start Date vs End Date

Figure 6: The scatter plot shows ....

5 Discussion

6 Appendix

6.1 Data Cleaning Notes

6.2 Methodology

Federal solicitation methods are either competitive or non-competitive. Competitive contracts accept bids from all or some subset of suppliers. Non-competitive contracts do not, citing certain exceptions, including a "pressing emergency in which delay would be injurious to the public interest," or when the estimated cost to the buyer is below certain thresholds defined for some goods and services. (Investigative Journalism Foundation 2024c)

6.3 Idealized Methodology

References

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data.* https://CRAN.R-project.org/package=janitor.

Investigative Journalism Foundation. 2024a. "Procurement – Microsoft." Toronto, Canada. https://theijf.org/procurement/supplier/microsoft?region=Federal.

———. 2024b. "Procurement Handbook Introduction." Toronto, Canada. https://theijf.org/procurement-introduction.

———. 2024c. "Procurement Sources." Toronto, Canada. https://theijf.org/procurement-sources#solicitation-method.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files.* https://CRAN.R-project.org/package=here.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'.* https://CRAN.R-project.org/package=arrow.

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 Grolemund, 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, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation.* https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, and Lionel Henry. 2023. *Purrr: Functional Programming Tools.* https://CRAN.R-project.org/package=purrr.

Xie, Yihui. 2024. *Knitr: A General-Purpose Package for Dynamic Report Generation in r.* https://yihui.org/knitr/.

Zhu, Hao. 2024. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax.* https://CRAN.R-project.org/package=kableExtra.