# My title*

**My subtitle if needed**

Mandy He

November 27, 2024

This study investigates supplier equity within Toronto's public procurement system, focusing on the representation of small businesses and historically underrepresented sectors. Leveraging data from the Toronto Bids Awarded Contracts dataset, the analysis examines award distributions, contract values, and supplier diversity. The study aims to identify potential disparities in supplier representation and evaluate the extent to which public contracts promote equitable opportunities. Recommendations are provided to enhance supplier inclusion and foster a fairer procurement environment, emphasizing transparency, outreach, and support for underrepresented businesses.

## Table of contents



*Code and data are available at: https://github.com/MandyHe7/Toronto-Bids-Awarded-Contracts

# 1 Introduction

Equitable public procurement is a cornerstone of inclusive economic development. The City of Toronto, through its competitive contracting process, allocates significant resources to suppliers for goods and services. These contracts serve as vital economic lifelines, particularly for small businesses and sectors that have traditionally been underrepresented in public procurement. However, ensuring equitable access to these opportunities remains a challenge, as systemic barriers may limit the participation of diverse suppliers.

This study explores the equity of supplier representation in Toronto's public contracts, focusing on small businesses and underrepresented sectors. Using the Toronto Bids Awarded Contracts dataset, the research seeks to answer critical questions: Are small businesses and diverse suppliers equitably represented in contract awards? What trends emerge in terms of award amounts and supplier participation? By addressing these questions, the study aims to provide actionable insights to promote fair and inclusive procurement practices.

The research contributes to a growing discourse on supplier diversity and public accountability, underscoring the importance of fostering inclusive economic opportunities in municipal contracting. This is particularly relevant in a diverse city like Toronto, where equitable procurement practices can significantly impact local communities and drive sustainable development. Through this investigation, the study highlights current gaps and offers recommendations to strengthen supplier equity in Toronto's public procurement system.

The remainder of this paper is structured as follows. Section 2 will discuss the data used in the study, highlighting key aspects and limitations. **?@sec-Results** will focus on generating graphs using the data presented in Section 2. **?@sec-discussion** will elaborate on the graphs from **?@sec-Results**, providing interpretations and possible explanations for the findings.

# 2 Data

## 2.1 Overview

The data used in this paper is access from Open Data Toronto and the particular data set used was the Toronto Bids Awarded Contracts (Toronto 2024). To analysis the data and creating graphs using the data, following package that was build in the statistical programming language R (R Core Team 2023) was used: tidyverse ((**tidyverse?**)), dplyr ((**dplyr?**)), lubridate (Wickham and Grolemund (2023)), scales (Wickham and Chang (2023)), knitr (Xie (2023)), and ggplot2 (Wickham (2023)). LLM (OpenAI (2023)) was also use in term of helping with coding in R to graph data, simulating data, cleaning data, and revising writing.

Table 1 is a preview of the clean data that will be used through out the analysis of this paper.

Table 1: Data Preview

Table 1: Data Preview

| Unique ID | Document Number | RFx Type | High Level Category | Successful Supplier | Awarded Amount ($) | Award Date | Division | Small Business |
|---|---|---|---|---|---|---|---|---|
| b347eaf5-9744-44b3-a5d2-a2f9a8a09395_1 | Doc3718847.1 | RFQ24 | Goods and Services | Asplundh Canada Ulc | 5952260 | 2023-07-11 | Parks, Forestry & Recreation | FALSE |
| b347eaf5-9744-44b3-a5d2-a2f9a8a09395_2 | Doc3718847.1 | RFQ24 | Goods and Services | Diamond Tree Care & Consulting Inc. | 4941090 | 2023-07-11 | Parks, Forestry & Recreation | FALSE |
| b347eaf5-9744-44b3-a5d2-a2f9a8a09395_3 | Doc3718847.1 | RFQ24 | Goods and Services | Weller Tree Service Ltd | 31395898 | 2023-07-11 | Parks, Forestry & Recreation | FALSE |
| b347eaf5-9744-44b3-a5d2-a2f9a8a09395_4 | Doc3718847.1 | RFQ24 | Goods and Services | Davey Tree Expert Co | 12054355 | 2023-07-11 | Parks, Forestry & Recreation | FALSE |
| 59b756d8-1a3d-4d46-a361-8d7f41fcab63_1 | Doc3220683.1 | RFP24 | Professional Services | Eplansoft | 1580950 | 2023-05-04 | City Manager's Office | FALSE |

| Unique ID | Document Number | RFx Type | High Level Category | Successful Supplier | Awarded Amount ($) | Award Date | Division | Small Business |
|---|---|---|---|---|---|---|---|---|
| e49e2059-4f41-44f9-a24b-83c49b4df1d7_1 | Doc3872270 | RFQ | Professional Services | Enterprise Rent A Car Canada | 7396946 | 2023-10-01 | Fleet Services | FALSE |

## 2.2 Measurement

To understand how a real-world phenomenon transitions into a structured dataset, it is essential to examine the lifecycle of data collection, organization, and analysis. In the case of the Toronto Bids Awarded Contracts dataset, the phenomenon begins with the City's procurement process—a critical operation where goods, services, and construction are solicited and acquired to meet public needs. This process involves multiple stages, from identifying the need for procurement, issuing requests for proposals or tenders, evaluating bids, and awarding contracts to successful suppliers. Each of these steps reflects a real-world activity that, when documented, forms the foundation for the dataset.

The transition starts with the City identifying a need, such as infrastructure development, service improvement, or administrative support. This leads to the creation of a solicitation document, outlining the scope of work, evaluation criteria, and requirements. Suppliers submit their bids, and the City evaluates them based on predefined criteria, such as cost, quality, and adherence to specifications. Once a contract is awarded, key details—including the awarded amount, supplier name, and the division requesting the procurement—are recorded.

This information is entered into the City's procurement management system, capturing the structured attributes necessary for documentation and analysis. For instance, details like the solicitation type (e.g., RFP, RFQ), high-level category (e.g., Construction Services, Professional Services), and award date are meticulously logged. To enhance the dataset's utility, additional attributes such as the division involved, buyer contact information, and a summary of the solicitation are included. The creation of a dummy variable, such as the classification of small businesses, involves further data processing to derive new insights from raw attributes like the awarded amount.

Ultimately, the dataset represents a structured snapshot of a complex real-world process. Each entry reflects decisions, negotiations, and activities that unfolded in the real world, carefully abstracted into numerical and categorical variables. This process of translation allows researchers, policymakers, and the public to analyze the data to uncover patterns, monitor equity, and assess the efficiency and fairness of the City's procurement system. The dataset, while detailed, is always an abstraction—capturing key facets of the phenomenon while omitting others deemed less relevant to the analysis. ## Outcome variables

### 2.3 Predictor variables

The dataset includes information on competitive procurement processes and awarded contracts by the City of Toronto. Key variables are:

- Unique Row Identifier (_id): A unique identifier for each record in the Open Data database.
- Composite Key (unique_id): A unique identifier combining related procurement details.
- Document Number: A reference number used by the City to track the solicitation during the procurement process.
- RFx Type (Solicitation Type): The procurement method used for the solicitation, such as Request for Proposal (RFP), Request for Quotation (RFQ), Tender (RFT), and others. This indicates the process used to seek bids or proposals from suppliers.
- High-Level Category: Groups the goods or services being procured, such as Goods and Services, Professional Services, or Construction Services.
- Successful Supplier: The name of the supplier awarded the contract.
- Awarded Amount: The total monetary value awarded to the supplier for the contract.
- Award Date: The date the contract between the City and the supplier was finalized.
- Division: The City Division requesting the procurement. The awarded supplier will have a contract with this division.
- Buyer Details: Includes the buyer's name, email, and phone number as the primary purchasing contact for the solicitation or contract.
- Solicitation Document Description: A summary of the solicitation, providing additional context about the procurement request.

In addition, there was a small business classification was created for its role in the analysis.

- Small Business (Dummy Variable): A binary variable indicating whether a supplier qualifies as a small business. Businesses with awarded contract amounts under $500,000 are classified as small (TRUE), while others are classified as large (FALSE).

## 3 Model

The goal of our modelling strategy is twofold. Firstly,...

Here we briefly describe the Bayesian analysis model used to investigate... Background details and diagnostics are included in Appendix B.

### 3.1 Model set-up

Define $y_i$ as the number of seconds that the plane remained aloft. Then $\beta_i$ is the wing width and $\gamma_i$ is the wing length, both measured in millimeters.

$$y_i|\mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$
$$\mu_i = \alpha + \beta_i + \gamma_i \tag{2}$$
$$\alpha \sim \text{Normal}(0, 2.5) \tag{3}$$
$$\beta \sim \text{Normal}(0, 2.5) \tag{4}$$
$$\gamma \sim \text{Normal}(0, 2.5) \tag{5}$$
$$\sigma \sim \text{Exponential}(1) \tag{6}$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of (**rstanarm?**). We use the default priors from `rstanarm`.

#### 3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular...

We can use maths by including latex between dollar signs, for instance $\theta$.

## 4 Results

The "RFT" type accounts for the highest awarded amount Figure 1, exceeding $1.2 billion. This indicates that tenders dominate procurement expenditure in Toronto's public contracts. This could reflect a preference for competitive, price-driven bidding processes for large-scale projects like construction and infrastructure.

The pie chart Figure 2 demonstrates a significant disparity, with the majority of contracts awarded to large businesses (marked as FALSE for small business). This indicates that small businesses have minimal representation in the awarding process, suggesting a lack of equity.

Looking at Figure 3, categories like "Construction Services" and "Professional Services" show a wide range of awarded amounts, with some outliers receiving significantly higher awards. Smaller categories (e.g., "Offer to Purchase") have fewer instances and seem to award comparatively smaller amounts. This indicates potential concentration of funding in certain categories, which may not benefit small businesses.

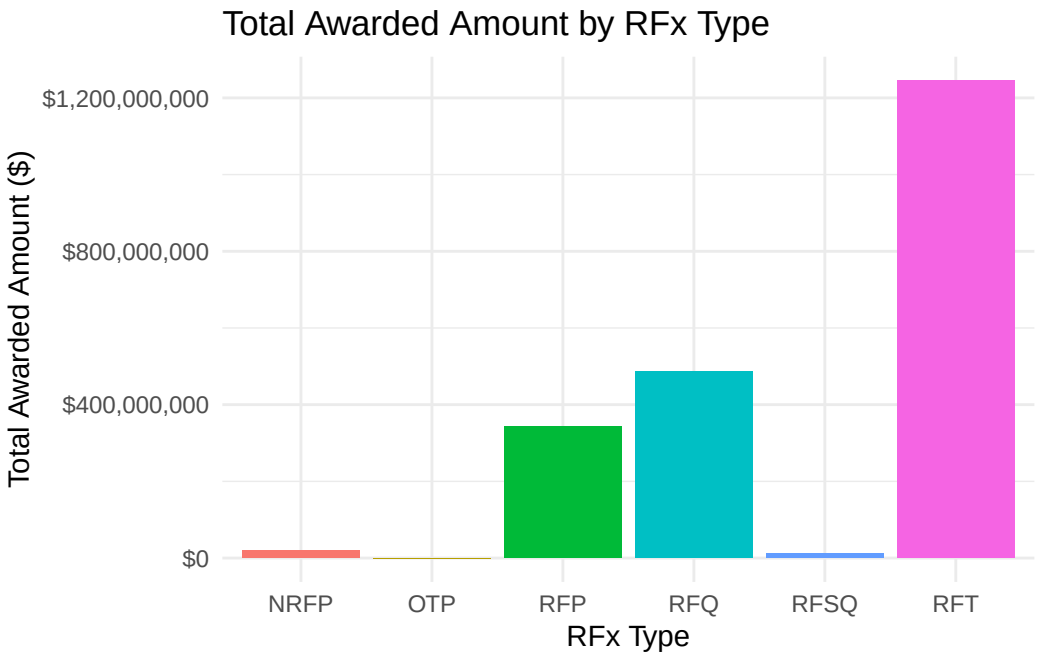


Figure 1: Bar graph

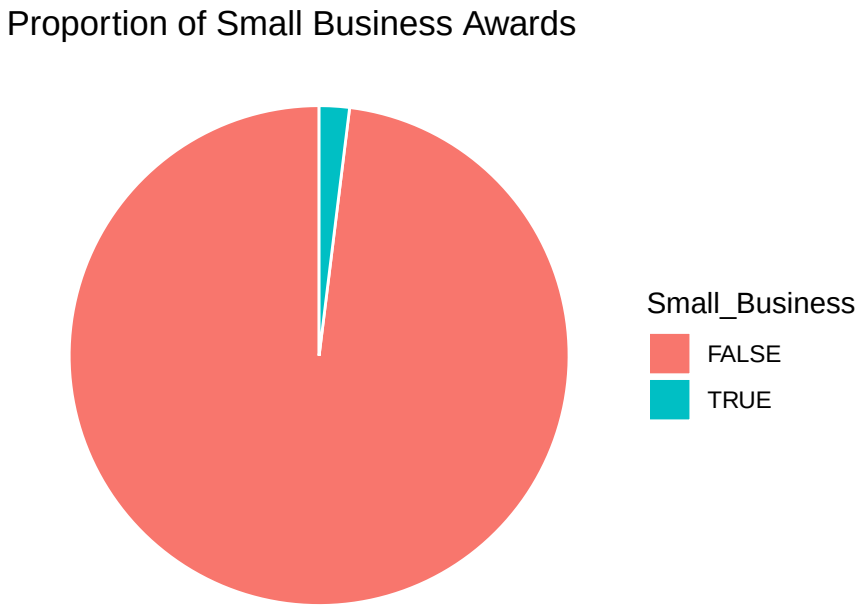


Figure 2: Proportion of Small Business Awards

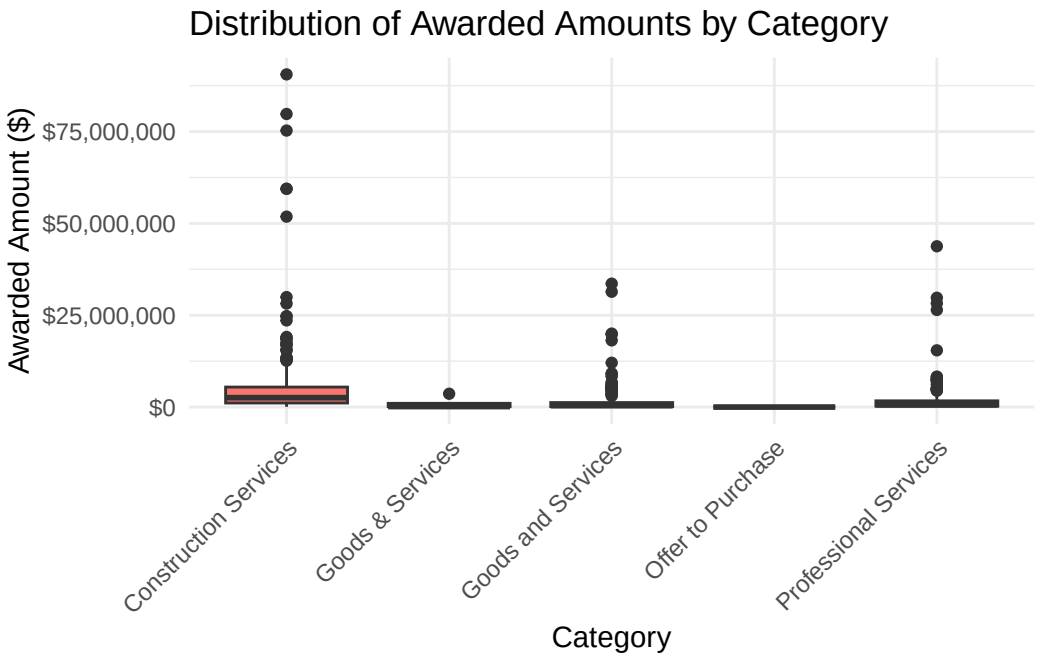


Figure 3: Awarded Amounts by High-Level Category

Looking at Figure 4, large businesses consistently receive significantly higher awarded amounts compared to small businesses. The awarded amounts to small businesses remain relatively stable over time, showing minimal fluctuation. This highlights the dominance of large businesses in the procurement process.

**?@fig-TimeseriesSmall**, the scatterplot shows awarded amounts to small businesses distributed across time, with no clear upward or downward trend. The shaded confidence interval around the regression line suggests high uncertainty in predicting changes over time. Overall, there is no strong evidence of growth or decline in the awarded amounts to small businesses.

```
#| label: fig-TimeseriesSmall
#| fig-cap: Awarded Amounts Over Time
#| echo: false
#| warning: false
#| message: false

small_business_data <- Contract_data %>%
  filter(Small_Business == TRUE) %>%
  mutate(Month = floor_date(Award_Date, "month"))

# Aggregate awarded amounts by month for small businesses
```

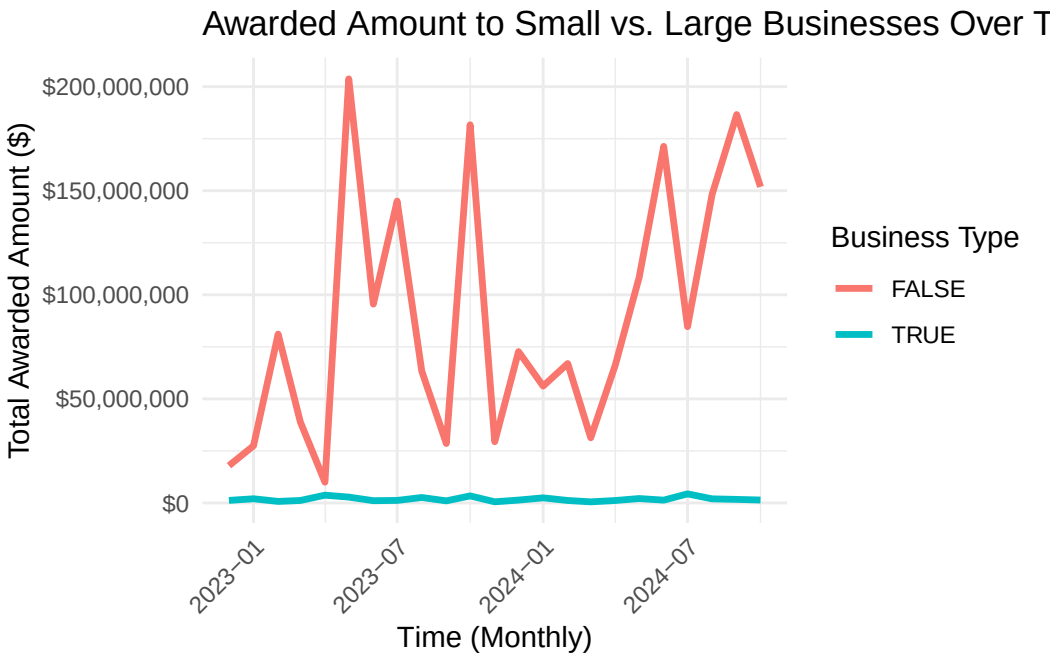


Figure 4: Awarded Amounts Over Time

```r
regression_data <- small_business_data %>%
  group_by(Month) %>%
  summarise(Total_Awarded = sum(Awarded_Amount, na.rm = TRUE)) %>%
  ungroup()

# Add a numeric time variable for regression
regression_data$Time <- as.numeric(regression_data$Month)

# Fit a linear regression model
model <- lm(Total_Awarded ~ Time, data = regression_data)
summary(model)
```

```
Call:
lm(formula = Total_Awarded ~ Time, data = regression_data)

Residuals:
     Min       1Q   Median       3Q      Max
-1288443  -645831  -472147   469260  2497513
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -3901349.0 21142635.2  -0.185    0.855
Time              289.3     1075.3   0.269    0.790

Residual standard error: 1042000 on 21 degrees of freedom
Multiple R-squared:  0.003436,  Adjusted R-squared:  -0.04402
F-statistic: 0.07241 on 1 and 21 DF,  p-value: 0.7905
```

```r
# Plot the results
ggplot(regression_data, aes(x = Month, y = Total_Awarded)) +
  geom_point(color = "blue", size = 2) +
  geom_smooth(method = "lm", color = "red", se = TRUE) +
  labs(title = "Regression Analysis: Awarded Amount to Small Businesses Over Time",
       x = "Time (Monthly)",
       y = "Total Awarded Amount ($)") +
  theme_minimal() +
  scale_y_continuous(labels = scales::dollar_format()) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

```
`geom_smooth()` using formula = 'y ~ x'
```

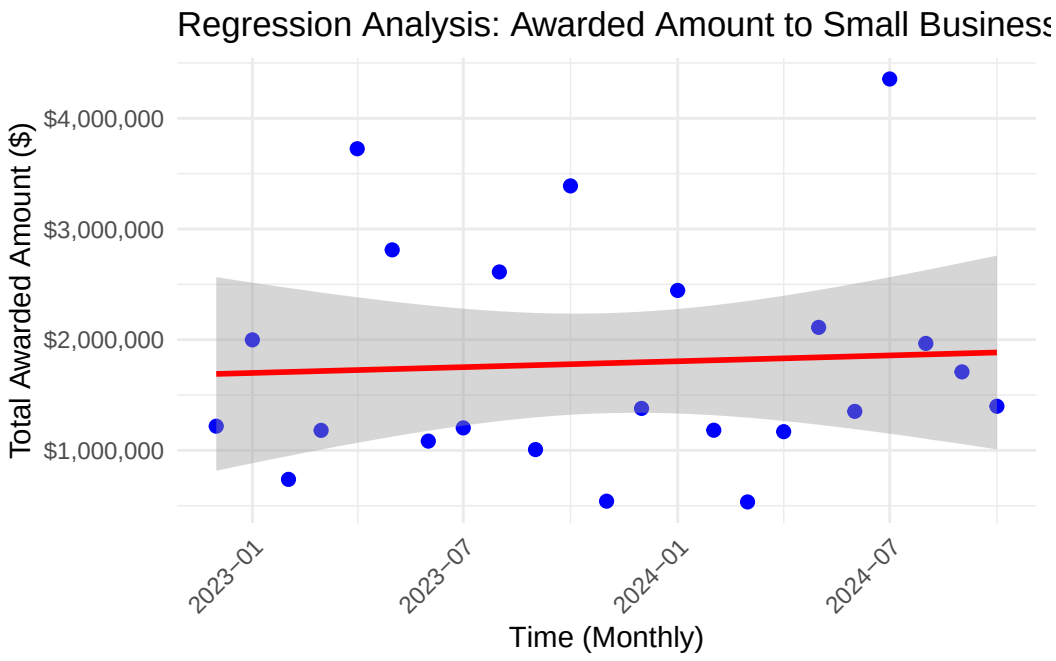

Figure 5, Across most categories, the representation of small businesses is minimal compared to large businesses.Notably, in the "Offer to Purchase" category, there seems to be equal representation of small and large businesses. This could imply specific policies or practices in that category that favor small businesses.

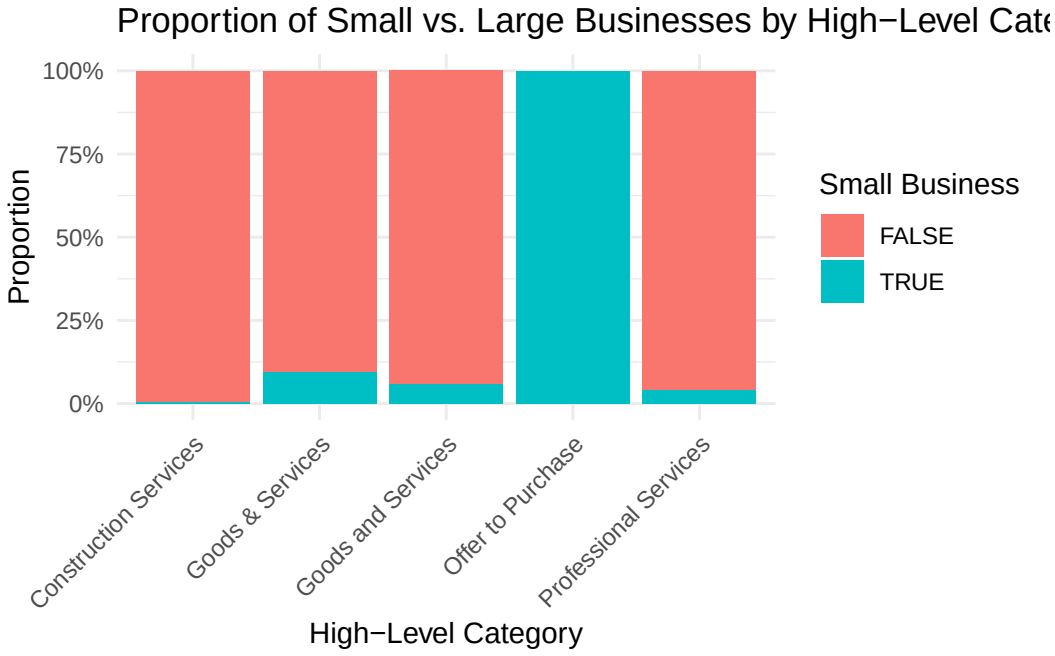


Figure 5: Awarded Amounts Over Time

Figure 6, we can observe that for "RFT," where the largest contracts are awarded, small businesses are virtually excluded. Similarly, "RFQ" and "RFP" see small businesses struggling to compete against larger counterparts.

A logistic regression model can help assess the factors influencing whether a contract is awarded to a small business.

Coefficients indicate how the likelihood of a contract being awarded to a small business changes with variables like Awarded_Amount, RFx_Type, and High_Level_Category. Significant p-values for predictors suggest they have an impact on the outcome.

```
Call:
glm(formula = Small_Business ~ RFx_Type + High_Level_Category +
    Awarded_Amount, family = binomial, data = Contract_data)

Coefficients: (1 not defined because of singularities)
```

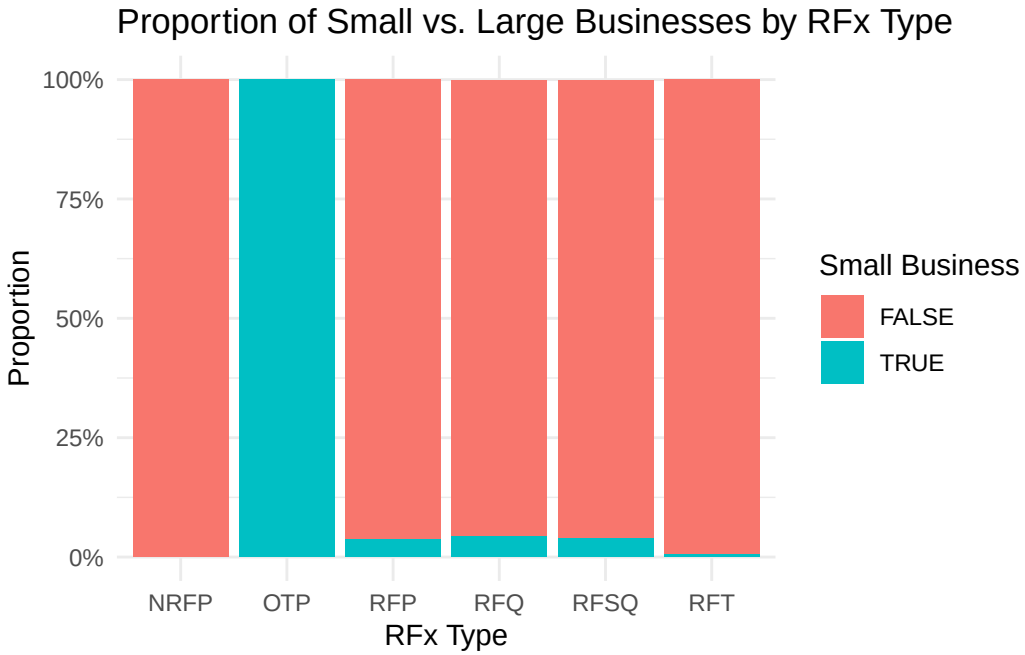


Figure 6: Awarded Amounts Over Time

| | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| (Intercept) | 1.635e+04 | 4.143e+05 | 0.039 | 0.969 |
| RFx_TypeOTP | -1.632e+04 | 5.463e+05 | -0.030 | 0.976 |
| RFx_TypeRFP | -1.573e+04 | 4.094e+05 | -0.038 | 0.969 |
| RFx_TypeRFQ | -1.577e+04 | 4.091e+05 | -0.039 | 0.969 |
| RFx_TypeRFSQ | -1.400e+04 | 2.537e+07 | -0.001 | 1.000 |
| RFx_TypeRFT | -1.588e+04 | 4.113e+05 | -0.039 | 0.969 |
| High_Level_CategoryGoods & Services | -4.337e+02 | 1.907e+05 | -0.002 | 0.998 |
| High_Level_CategoryGoods and Services | -1.683e+02 | 3.543e+04 | -0.005 | 0.996 |
| High_Level_CategoryOffer to Purchase | NA | NA | NA | NA |
| High_Level_CategoryProfessional Services | -2.171e+02 | 2.589e+04 | -0.008 | 0.993 |
| Awarded_Amount | -8.192e-04 | 1.043e-02 | -0.079 | 0.937 |

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 7.3219e+02  on 572  degrees of freedom
Residual deviance: 3.3456e-05  on 563  degrees of freedom
AIC: 20

Number of Fisher Scoring iterations: 25

# 5 Discussion

General Inferences: Equity Gaps: There is a clear lack of equity in the distribution of contracts between small and large businesses. Most of the contracts and larger awards seem to favor larger corporations. Category-Specific Opportunities: Some categories like "Offer to Purchase" might be worth further investigation to understand what policies support small business equity and whether they can be applied to other categories. Implications for Policy: These findings could motivate policy adjustments to enhance opportunities for small businesses across all categories, potentially leading to more diverse supplier representation.

## 5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

## 5.2 Second discussion point

Please don't use these as sub-heading labels - change them to be what your point actually is.

## 5.3 Third discussion point

## 5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

# Appendix

# A Additional data details

# B Model details

## B.1 Posterior predictive check

In **?@fig-ppcheckandposteriorvsprior-1** we implement a posterior predictive check. This shows…

In **?@fig-ppcheckandposteriorvsprior-2** we compare the posterior with the prior. This shows…

## B.2 Diagnostics

**?@fig-stanareyouokay-1** is a trace plot. It shows… This suggests…

**?@fig-stanareyouokay-2** is a Rhat plot. It shows… This suggests…

## References


OpenAI. 2023. "ChatGPT: A Language Model." https://www.openai.com/chatgpt.

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

Toronto, City of. 2024. "TOBids Awarded Contracts." https://open.toronto.ca/dataset/tobids-awarded-contracts/.

Wickham, Hadley. 2023. *Ggplot2: Elegant Graphics for Data Analysis*. https://ggplot2.tidyverse.org/.

Wickham, Hadley, and Winston Chang. 2023. *Scales: Scale Functions for Visualization*. https://CRAN.R-project.org/package=scales.

Wickham, Hadley, and Garrett Grolemund. 2023. *Lubridate: Make Dealing with Dates a Little Easier*. https://CRAN.R-project.org/package=lubridate.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.