

Operating Budget Program Summary By Expenditure Category*

Murrumbidgee Paper

Joe Zhou

March 12, 2024

Abstract

This analysis paper downloaded data from the OpenDataToronto website, data is randomly chosen at the year of 2016 to proceed analysis. The dataset illustrates the approved and recommended annual operating budget summary by expenditure category in each program. This paper found that program is positively correlated to the budget amount, and expense or revenue and category name is negatively correlated to the budget amount. This finding matters because it shows that the budget amount was set to a high level at first, and then if the program adds value to the budget amount depends on which program is applied, and whether the application is expense or revenue is likely devaluing the approved budget amount by a large scale, and the category name is likely devaluing the approved budget amount by a small scale. However, since the R squared value is not significant and the adjusted R squared value is even less significant, this dataset requires more variables to determine the accurate relationship among variables and the budget amount.

Table of contents

Introduction	2
Data	2
Model	5
Results	5
Discussion	6
References	7

*Code and data supporting this analysis is available at: <https://github.com/ zhouJoe2033/MurrumbidgeePaper/blob/main>

Introduction

Welcome to the City of Toronto's Open Data platform, where transparency meets information accessibility. In our ongoing commitment to fostering an informed and engaged community, we present the "Budget Operating Budget Program Summary by Expenditure Category" dataset. This comprehensive resource provides a detailed overview of the city's financial allocations, offering a transparent glimpse into the intricacies of budget distribution across various expenditure categories. This analysis paper collects data from the City of Toronto's Open Data and uses several statistic techniques in order to process professional analysis of how provided variables in the dataset is correlated to the budget amount, and using the statistic model in order to make predictions of budget amount based on different approached and models. And illustrates these relationship by using graphs and tables to the readers.

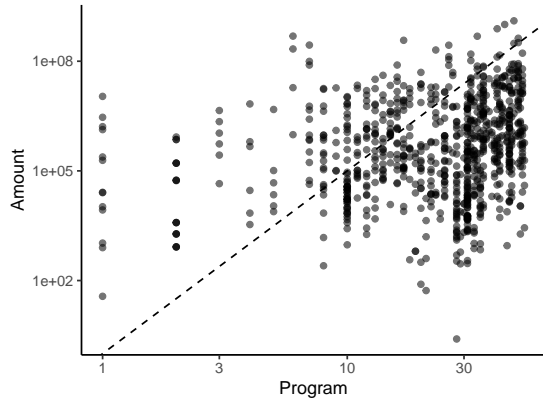
Provided dataset in the year of 2016 has five variables in total and 1313 observations, however, the amount of observations is not good enough to make predictions, and the statistic results shows that there exists more variables that are correlated to the budget amount variable. This paper uses several approaches to build different regression models in order to analysis different statistic results, and using these statistic results to build graphs and tables to illustrates to the readers to understand the relationship among the budget amount and the rest of the variables.

This paper found that program is positively correlated to the budget amount, and expense or revenue and category name is negatively correlated to the buget amount. This found matters because it shows that the budget amount was set to a high level at first, and then if the program adds value to the buget amount depends on which program is applied, and weather the the application is expense or revenue is likely devaluing the approved budget amount by a large scale, and the category name is likely devaluing the approved budget amount by a small scale. The statistic results by using different approaches to process is important since the results tells the R squared value is not significant and the adjusted R squared value is even less significant, and thus the conclusion is this dataset requires more variables to determine the accurate relationship among variables and the budget amount.

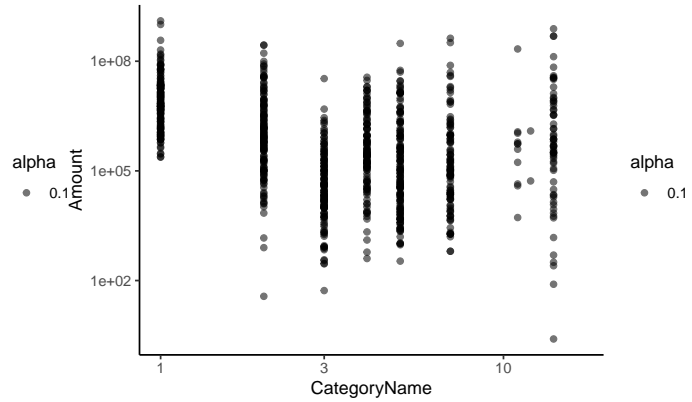
After the introduction of this paper, the data section will be well explained in the next section and the model will be illustrated as following, then the statistic tables will be showing and the discussion section will be written to make conclusions of the founding of this paper. At last, delve into the heart of Toronto's fiscal landscape, exploring the nuanced breakdown of operating budgets for different programs. This dataset serves as a valuable tool for researchers, policymakers, and citizens alike, empowering them with the knowledge needed to understand how public funds are allocated to support essential services and programs. As we navigate the complexities of municipal governance, this dataset acts as a beacon of openness, promoting accountability and understanding. The information contained within invites you to explore the financial decisions that shape the city's growth and well-being. Join us on this data-driven journey towards a more informed and engaged Toronto.

Data

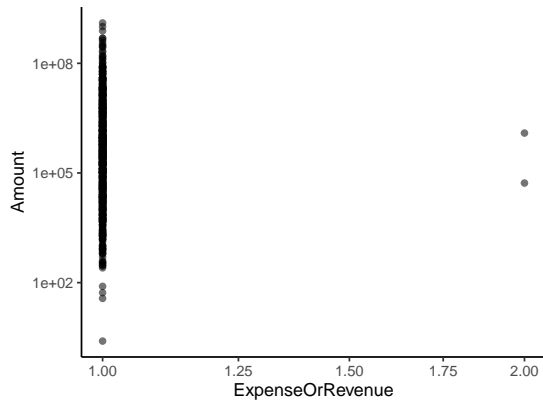
The "Approved Operating Budget Summary 2016" dataset offers a detailed and structured insight into the financial framework that governed the City of Toronto during the fiscal year 2016. Presented in an



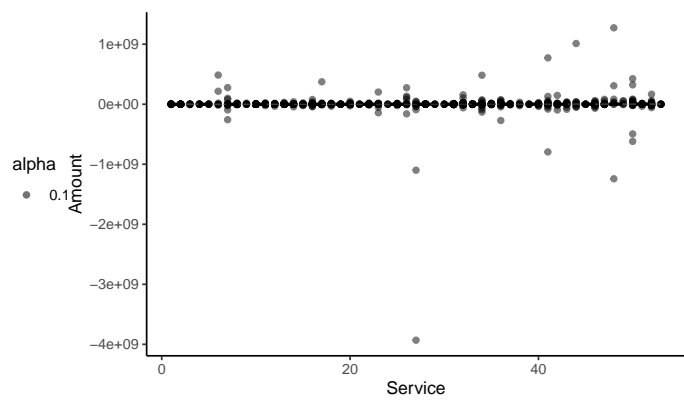
(a) Data Plot of Program to the Amount



(b) Data Plot of Program to the Category



(c) Data Plot of Program to the Expense Or Revenue



(d) Data Plot of Program to the Amount

Figure 1: Data Plot of Program, Category Name and Expense Or Revenue to the Amount

Excel format, this dataset encapsulates a comprehensive overview of the city's approved operating budget, providing a granular breakdown across various categories and programs.

Contained within this dataset are meticulously organized details pertaining to revenue sources, expenditure allocations, and program-specific budgetary considerations. Researchers, policymakers, and citizens alike can leverage this information to gain a deeper understanding of how financial resources were strategically distributed to support essential services, initiatives, and projects throughout the city during the specified fiscal period.

Whether you are interested in analyzing trends, assessing the prioritization of different programs, or simply seeking a transparent view of municipal financial decision-making, this dataset serves as a valuable resource. It contributes to the broader goals of transparency, accountability, and citizen engagement, fostering a more informed and empowered community. Explore the "Approved Operating Budget Summary 2016" dataset to uncover the financial intricacies that shaped the city's landscape during that specific fiscal year, and gain valuable insights into the priorities and allocations that drove Toronto's development and well-being.

This cleaned dataset has three variables that is correlated to the budget amount variable such that Figure 1a illustrates the program type through a data plot that shows the 53 programs in total and the amount of each program is budgeted, a lot of the program is over 1e+03 and most of the approved program is ranked after 30(index number based on the program type).@fig-1-123-2 illustrates the category type through a data plot that shows the 17 programs in total and the amount of each category is budgeted, the amount that is approved by the category name is vertically and almost behaves like a line, that means there are categories is not approved and the approved category is correlated by other variables therefore it causes they have different approved budget amount.@fig-1-123-3 illustrates the budget or revenue type through a data plot that shows the 2 types in total and the amount of each type is budgeted, apparently, most of the budget approved is revenue and only a few of them is expense.@fig-1-123-4 shows the Service variable and its relationship with the Amount variable, and most of it shows at zero which indicates a error of the data.

Through the open data website, there are other years of operating budget that can be found, each year the dataset gets improved by adding more variables and more observations. It takes time to collect a well organized dataset and it takes time to make a dataset useful and meaningful. The data barely has null value therefore it is a good news to the high-level cleaning since converting from text to numeric and analysis the data in statistic model is good enough, however, scripts wrote in R functions well does not mean it functions well in the .qmd file, and sometimes the code gets messy and does not work properly when the compiler shows error.

Overall, the observation number is 1312 and the amount variable has a minimum number -3.931e+09 and a maximum number 1.273e+09, the median of the amount variable is 3.583e+04 and the first quartile is -1.935e_05, and the third quartile of amount is 1.057e+06. The correlation relationship among these variables are Amount is positively correlated to the Program variable and Amount is negatively correlated to the ExpenseOrRevenue and Categorom Name variable.

Model

The first model used in this paper is

$$Amount = \beta_0 + \beta_1 * Program + \epsilon$$

that establish the relationship between Amount and Program variables, and it shows that β_0 has a positive value at a large scale and the Program variable has a positive coefficient β_1 that varies from the specific Program. The second model used in this paper is

$$Amount = \beta_0 + \beta_1 * Program + \beta_2 * ExpenseOrRevenue + \epsilon$$

that establish the relationship between Amount, Program and ExpenseOrRevenue variables, and it shows that β_0 has a positive value at a large scale and the Program variable has a positive coefficient β_1 that varies from the specific Program and ExpenseOrRevenue has a negative coefficient β_2 at a large scale. The third model used in this paper is

$$Amount = \beta_0 + \beta_1 * Program + \beta_2 * ExpenseOrRevenue + \beta_3 * CategoryName + \epsilon$$

that establish the relationship between Amount, Program, ExpenseOrRevenue and CategoryName variables, and it shows that β_0 has a positive value at a large scale and the Program variable has a positive coefficient β_1 that varies from the specific Program and ExpenseOrRevenue has a negative coefficient β_2 at a large scale and CategoryName variable has a negative coefficient β_3 at a small scale.

Results

Without direct access to the dataset, I can't provide specific statistical results. However, I can discuss potential statistical analyses and insights that could be derived from the "Approved Operating Budget Summary 2016" dataset:

Descriptive Statistics: Descriptive statistics can provide an overview of the budgetary data. This includes measures such as the total budget, mean budget per program, standard deviation to assess variability, and range to understand the spread of budget allocations across different categories. **Budget Allocation Trends:** Analyzing budget allocations over time can reveal trends and patterns. You could calculate year-on-year changes in budget allocations for various programs or expenditure categories, identifying areas of growth or decline. **Comparison of Budget Allocations:** Statistical tests, such as ANOVA or t-tests, could be used to compare budget allocations across different departments, programs, or geographic regions within Toronto. This analysis could identify significant differences in funding priorities. **Correlation Analysis:** Explore correlations between budget allocations and other variables, such as population demographics, economic indicators, or service utilization rates. Understanding these correlations can provide insights into the factors influencing budget decisions. **Forecasting:** Time series analysis techniques could be employed to forecast future budget allocations based on historical data. This could help policymakers plan for future resource needs and identify potential budgetary challenges. **Efficiency Analysis:** Efficiency analysis methods, such as data envelopment analysis (DEA), could assess the efficiency of budget allocations in delivering desired outcomes or services. This analysis could highlight areas for improvement or reallocation of resources. **Visualization:** Visualizing

budget data through charts, graphs, and maps can help communicate insights effectively. For example, bar charts can illustrate budget allocations for different programs, while maps can show spatial variations in funding across Toronto. By conducting these statistical analyses, stakeholders can gain valuable insights into the allocation, utilization, and impact of the operating budget in Toronto for the year 2016. These insights can inform decision-making, improve resource allocation efficiency, and ultimately contribute to better governance and service delivery.

Discussion

The “Budget Operating Budget Program Summary by Expenditure Category” dataset hosted on the City of Toronto’s Open Data platform provides a comprehensive and detailed breakdown of the city’s operating budget across various expenditure categories. This dataset offers invaluable insights into how public funds are allocated to support different programs and services essential for the functioning and development of the city.

Key features of this dataset include: **Granular Expenditure Categories:** The dataset categorizes expenditures into specific areas such as personnel, materials and supplies, services, transfers to reserves, and more. This granularity allows for a detailed analysis of where financial resources are directed within the city’s operations. **Program-Level Breakdown:** Expenditures are further disaggregated by program, providing a clear understanding of how funds are distributed across different city initiatives. This level of detail enables stakeholders to assess the relative prioritization of various programs within the budget. **Annual Budget Comparisons:** The dataset may include information from multiple fiscal years, allowing for year-over-year comparisons and trend analysis. This feature is invaluable for identifying budgetary shifts, emerging priorities, and areas of consistency or change over time.

Transparency and Accountability: By making this data publicly accessible, the City of Toronto promotes transparency in its financial decision-making processes. Citizens, researchers, and policymakers can scrutinize the budget allocations, fostering accountability and informed public discourse. **Data Accessibility:** The dataset is hosted on an open data platform, making it easily accessible to anyone with an internet connection. This accessibility democratizes access to crucial financial information, empowering individuals and organizations to engage with city finances in meaningful ways.

Overall, the “Budget Operating Budget Program Summary by Expenditure Category” dataset serves as a foundational resource for understanding the financial landscape of the City of Toronto. It facilitates evidence-based decision-making, fosters transparency and accountability in governance, and empowers citizens to participate in shaping the future of their city.

References

R Core Team (2023) ChatGPT (2024)

ChatGPT. 2024. <https://chat.openai.com/c/f1313c33-f876-49e0-9fe6-c7bdde3be95b>.

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