

Salaries and Rising Police Force Budgets*

An Exploratory Analysis on Expenditures of Toronto Police Force

Meha Grewal

27 April 2022

Abstract

In this paper we will explore the budget of the Toronto Police Service and examine the implications of their fiscal responsibilities. Through the understanding of the central influences of the police force, we can determine how to effectively distribute resources and funding. These findings provide rationale for policy decisions.

1 Introduction

The Toronto Police Service (TPS) are supposed to be the upholders of law and order. They provide a sense of safety to the community and work to prevent and solve criminal cases. The budget for the the TPS is highly significant as it allows them the ability to safely govern the city. To identify the most influential aspects of the TPS budget and thus the parts of policing that they deem as priority, I will analyze the relevant variables and model them. I explore their expenditures and fiscal responsibilities for 2021 and its implications in providing safety for residents of Toronto.

Conducting an exploratory analysis on the budget for TPS would indicate areas that the government prioritizes. Utilizing the dollar amount as the response variable, we can conclude the largest predictors for police force expenditures and its implications. This information can be applied by government officials and policy makers to minimize government spending and allocate finances to areas that most necessitate it.

Beginning with a general overview of the variables provided in the data, I visually identified where majority of the budget is spent on. Through a multiple linear regression model, I was able to determine the most statistically significant variables. This indicates what portions of the police force have the most influence over the budget and the priorities of the force. By analyzing each predictor to understand its importance, I then considered the broader societal implications. The linear regression indicates how finances are allocated and which districts and units relieve the largest portions of the budget. By determining what districts and categories within the police force require the most funding, we can evaluate the future of the force. This analysis provides the determining factors of the fiscal responsibilities for the police and showcases what areas require more resources. Policy decisions influence the acceptable expenditures of the TPS and this paper aims to understand the ramifications of the budget distribution.

The data set provided by the city of Toronto showcases information regarding actual expenditures from the police force in 2021. By examining the relevant summary statistics and plots, I am able to identify key predictors for our model. I then model the data using the amount for the budget as the response variable and conduct a multiple linear regression. This was done to predict the budget and the crucial components. My discussion includes the understanding of our model and how it was assembled. Lastly, I consider the social implications of the model and how the budget has an impact on safety in Toronto.

*Code and data are available at: <https://github.com/meha-g/FinalPaper>.

Table 1: Top ten budget expenditures for Toronto Police Budget

Pillar Name	Unit Name	Feature Category	Amount
Detective Operations	Other	Salaries	55734700
East Field Command	Court Services	Salaries	49314100
West Field Command	Traffic Services	Salaries	33167600
East Field Command	Division 55	Salaries	31265300
Centralized Service Charges	CCC - Employee Benefits	Benefits	27717200
East Field Command	Communication Services	Salaries	27563100
West Field Command	Division 51	Salaries	26530300
West Field Command	Division 14	Salaries	26264800
East Field Command	Division 43	Salaries	25093100
Centralized Service Charges	CCC - Central Paid Duty	Salaries	24667000

2 Data

2.1 Data Source

This data was obtained from the Open Data Toronto portal (Gelfand 2020), which is accessible due to the City of Toronto. It is a breakdown of the approved budget for TPF and actual expenditures. The data was collected in the fiscal year of 2021 and was published by Toronto Police Services. The budget indicates expenses for the operation of TPF approved by City Council and the Toronto Police Services Board. The raw data includes 2,755 observations of 12 variables, and the data was last refreshed on October 20, 2021. An exploratory data analysis was done on the data using **R statistical programming** (R Core Team 2020). The data was cleaned and manipulated for modelling using **tidyverse** (Wickham et al. 2019), **knitr** (Xie 2022), **dplyr** (Wickham et al. 2022), and **modelsummary** (Arel-Bundock 2022).

2.2 Variables

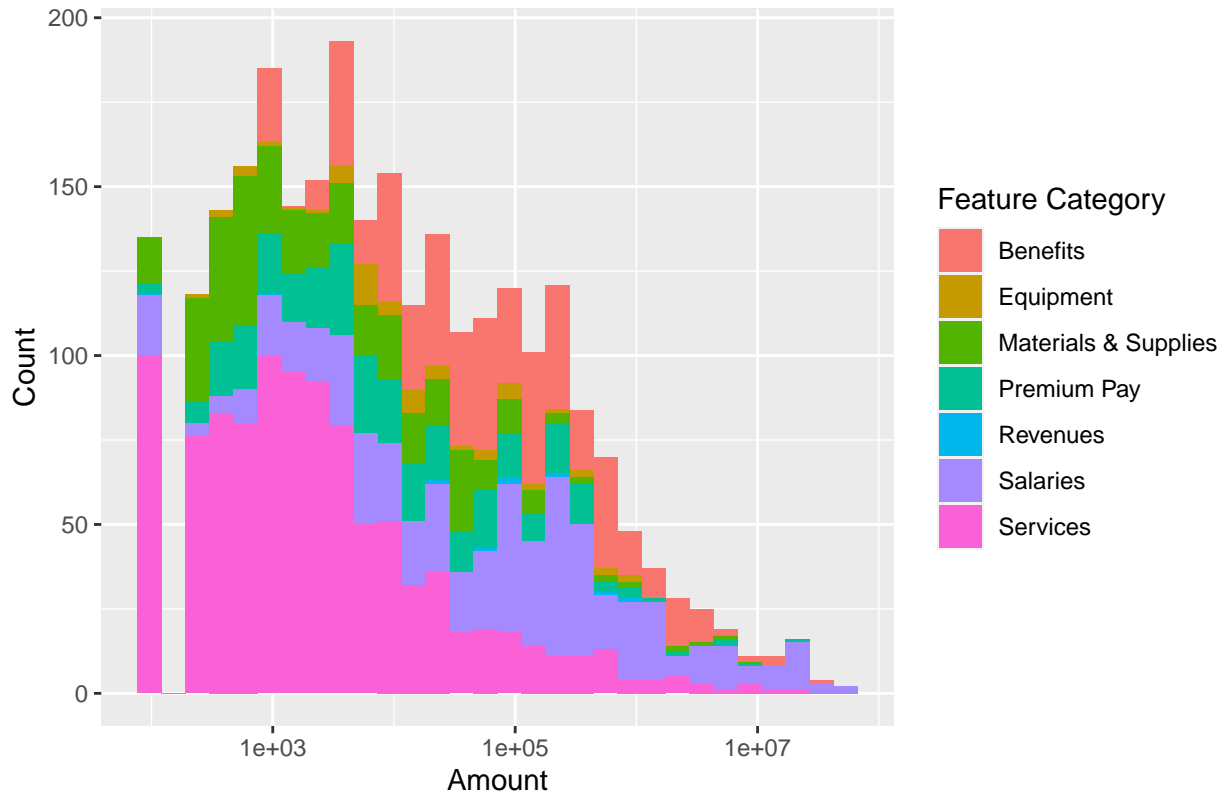
The data collected by the Toronto Police Service includes many variables that explain the budget amount, the organization within TPS, name of the command, pillar, district, and unit. It also includes the feature category which classifies the expenditure, and the cost elements. The dollar amount variable is the most significant sub-section of the data I will be analyzing. The other notable variables will be used as predictors for the budget.

To begin with, I removed the variables for budget type and fiscal year as each data point was the same, all the points in this data set were from the fiscal year 2021 and were approved for the budget. Furthermore, the column for command name and district type were removed as they were similar to pillar name. (Table 1) showcases a preview of the data but only including the pillar name, unit name, feature category, and budget amount. The table showcases the top ten budget expenditures for the TPS in the descending order. Since it is just a brief preview of the data, only the most essential columns are included.

Based off (Table 1), a large portion of budget expenses can be attributed to the salaries of those employed by the police force. Only one data point in the top ten budget expenditures for feature category is associated with benefits while every other expense is due to salaries. Additionally, pillar name includes many points in this table from the East Field Command or the West Field Command, indicating that commandments and particularly those two, require higher budgets.

2.3 Features

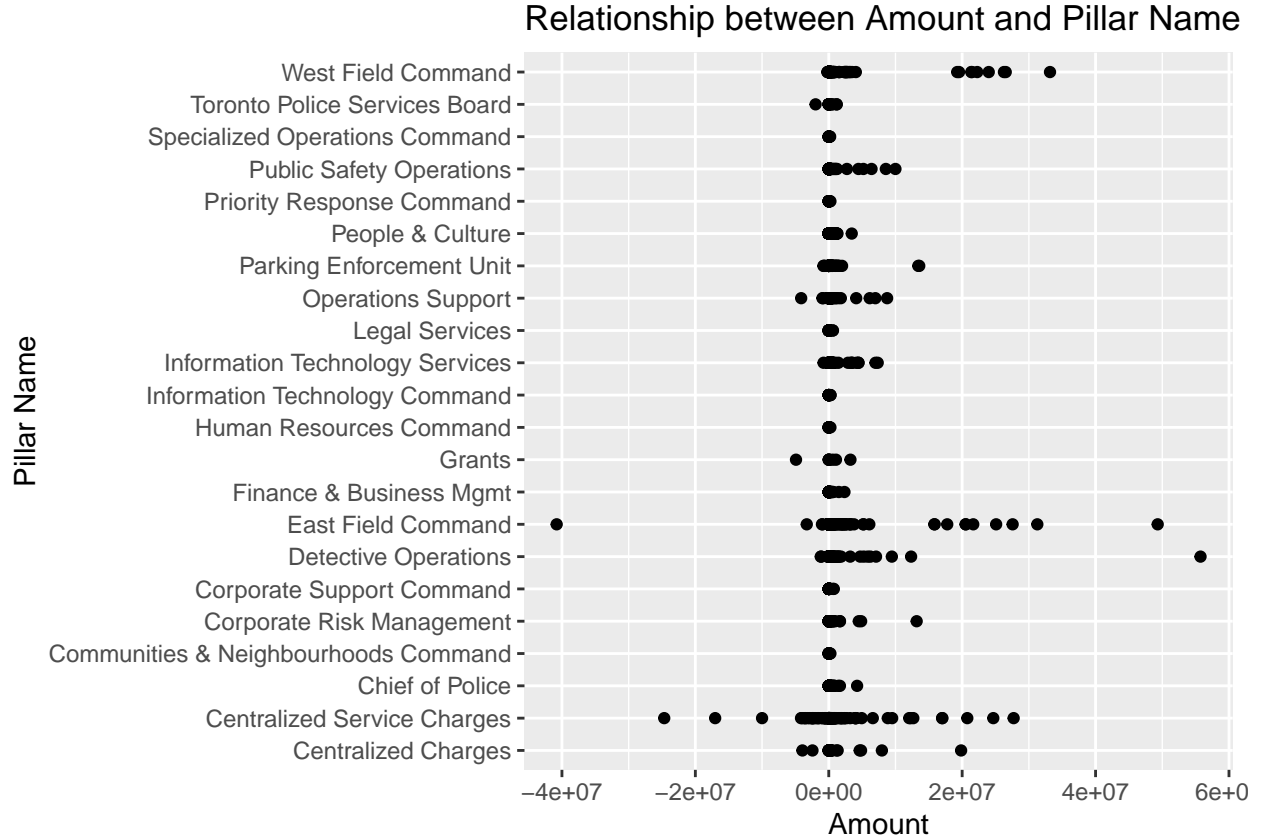
Distribution of Amount and Feature Category



(Figure ??) displays the distribution of budget expenditures for the TPS, grouped by feature category. The amount of expenses is right skewed and follows a unimodal distribution. Expenses within the salaries category has the largest distribution as it reaches the highest amount, this corresponds with (Table 1) where the highest budgets were for salaries.

Table 2: Summary of discrete variables for TPS budget

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max
Amount	1043	0	411 215.0	2 923 700.0	-40 790 800.0	6100.0	55 734 700.0
is_command	2	0	0.4	0.5	0.0	0.0	1.0
organization_num	3	0	1.1	0.4	1.0	1.0	3.0
rev_or_sal	2	0	0.1	0.3	0.0	0.0	1.0



(Figure ??) is a scatter plot of the relationship between budget amount and pillar name which determines where the expense is for. The East Field Command and Detective Operations appear to have outliers, with East Field Command also having a large negative outlier. The largest expenditures derive from West Field Command, East Field Command, Detective Operations, and Centralized Service Charges. This information is useful in determining which sections of the police force require the most or least funding, it can be used for future budget determinations.

2.4 Methodology

To model this data for an exploratory data analysis, the data set was manipulated. I created numerical variables that would allow for a multiple linear regression to be done and statistical significance be tested. I constructed a dummy variable that indicates whether the budget expense was for a command, or other service. Similarly, I built a dummy variable that determined whether the budget was spent on salaries or another category. These variables were created for the model to see whether salaries or commands have a strong impact on the budget. Another variable was made to just assign a numerical value to the organization entity, ranging from 1 to 3.

Table 3: Explaining TPF budget based on whether it is a command, organization entity, and feature category

	Model 1
(Intercept)	461 880.660 (198 199.561)
is_command	118 590.829 (112 987.133)
organization_num	−4663.404 (123 446.891)
rev_or_sal	2 882 793.173 (235 822.807)
Feature_CategoryEquipment	−401 584.043 (385 700.313)
Feature_CategoryMaterials & Supplies	−437 204.835 (198 495.342)
Feature_CategoryPremium Pay	−309 342.804 (206 948.526)
Feature_CategoryRevenues	−3 163 292.946 (389 221.838)
Feature_CategorySalaries	−357 846.198 (209 512.692)
Feature_CategoryServices	−388 355.574 (154 753.092)
Num.Obs.	2751
R2	0.108
R2 Adj.	0.105
AIC	89 428.8
BIC	89 494.0
Log.Lik.	−44 703.417
F	36.939
RMSE	2 765 580.51

(Table 2) showcases the summary statistics for the discrete variables in the data set. This includes the amount from the budget, alongside the three new numerical variables constructed for the model.

3 Model

3.1 Multiple Linear Regression

To analyze the budget of the TPF and how different variables impact the expenditures of the police force, I created a model using a multiple linear regression.

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon_i \quad (1)$$

We can use maths by including latex between dollar signs, for instance θ .

Predictor variable: Amount Explanatory variables:

Validity:

4 Results

5 Discussion

5.1 Findings

If my paper were 10 pages, then should 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 Implications

5.3 Limitations and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional details

References

- Arel-Bundock, Vincent. 2022. *Modelsummary: Summary Tables and Plots for Statistical Models and Data: Beautiful, Customizable, and Publication-Ready*. <https://CRAN.R-project.org/package=modelsummary>.
- Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://CRAN.R-project.org/package=opendatatoronto>.
- 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, 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. 2022. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Xie, Yihui. 2022. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.