

Analysis of Salaries for Health Care Charities in Canada

A report prepared by:
Yann Guerin

Date: October 29th 2022

Summary

Data from the Canada Revenue Agency's website was scraped, cleaned, and analyzed in order to examine the relationship of various metrics to salaries. Clear relationships between higher total revenue and higher salaries are present. Analyzing other charities with similar metrics of total revenue, total expenses, total compensation, and average salary, shows two opposing perspectives on the data. With regards to total revenue and expenses, Lung Cancer Canada's top salary in the range of \$120k to \$160k is higher than other charities. But, with regards to the average compensation of employees at Lung Cancer Canada and the percentage of their expenses which go towards compensation, their top salary is significantly lower than other comparable charities.

The Data

All data used in this analysis comes from the Canada Revenue Agency's List of Charities website and each Charity's Quick View page(s). I extracted the data in two steps; first extracting a list of all charities (86000+) and then extracting the Quick View page(s) for each Health Care Charity. Through their search page I extracted a detailed list of all Registered Charities in Canada along with information about them such as: name, registration number, address, city, province, charity type, and category. I categorized all charities with the types related to health care as Health Care Charities and then extracted all relevant data from their Quick View Pages using Web Scraping. The Quick View pages contained lots of data: dates of reporting period, description of ongoing programs, breakdown of Revenue, breakdown of Expenses, Total Compensation, number of employees (part-time and full-time), professional and consulting fees, and the salary bins for up to the top 10 highest paid full-time positions at the charity. A more detailed list of the features of the data extracted can be found in the Dataset section along with a link to the complete dataset.

The Analysis

Preparing the Data

Before starting the analysis I first had to clean the data I extracted. I transformed the raw JSON data I extracted from the Quick View pages into clean columns in a CSV file. Each Charity comprises a row in the dataset, and missing values were encoded to be NaN (not a number) values. A few clearly incorrect values (using the value of Total Compensation as the number of employees for example) were manually corrected in order to remove them as outliers. A couple of other outliers were removed such as Alberta Health Services since the values for the number

of employees and revenue among others were orders of magnitude greater than the rest of the Charities. Then, both datasets (the detailed list and the quick view data) were merged into a single CSV file.

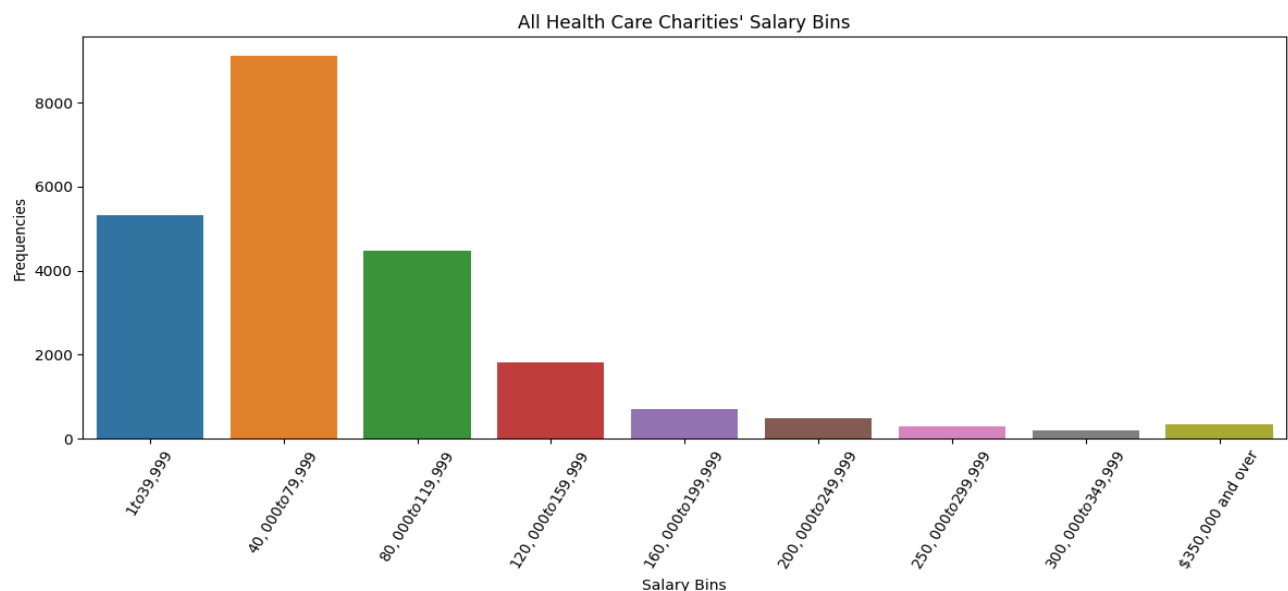
Exploring the Data

To begin my analysis I first explored the distributions (histograms) of relevant columns/data points to see what values were most common and what other outliers there were in the dataset. From this initial exploration I decided to restrict my analysis to charities with values that weren't very large: less than 10000 employees, and less than 10 million dollars in revenue for example. This removed the majority of outlier charities that would not be similar at all to Lung Cancer Canada. I then created a few extra columns by combining columns into more useful data points such as average compensation per employee, total number of all employees, ratio of Total Compensation to Total Expenses, and the ratio of Government Funding to Total Revenue. The analysis below focuses on the most recent reporting period.

Focusing on Compensation and Salaries

Next, I started to analyze the relationships of Total Compensation, Salary Bins, and Top Salary Bin to other data points to examine what trends showed up. This helped me better understand what data points act as stronger predictors of Salary data.

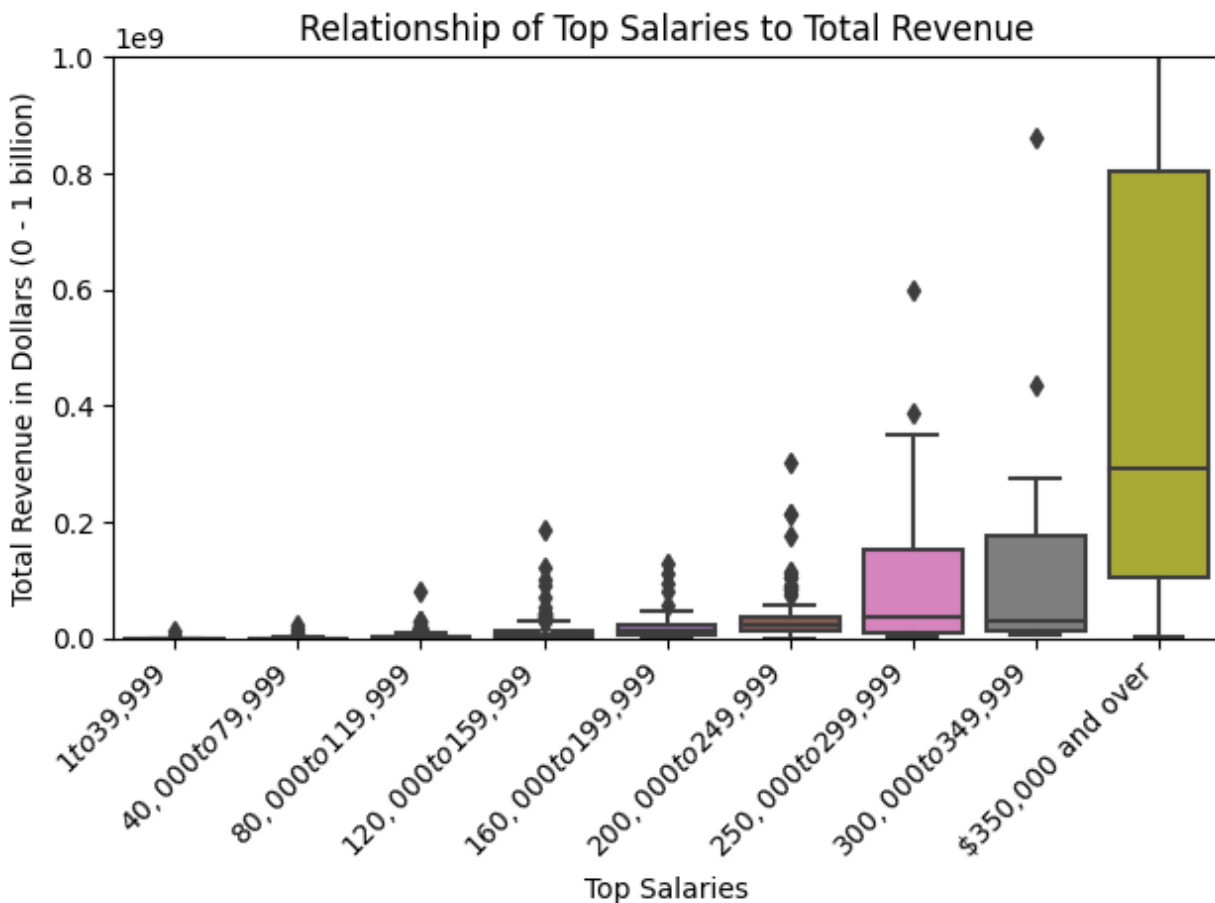
This graph shows the total number of salaries in each bin for all health care charities in Canada (that reported the salary bins).



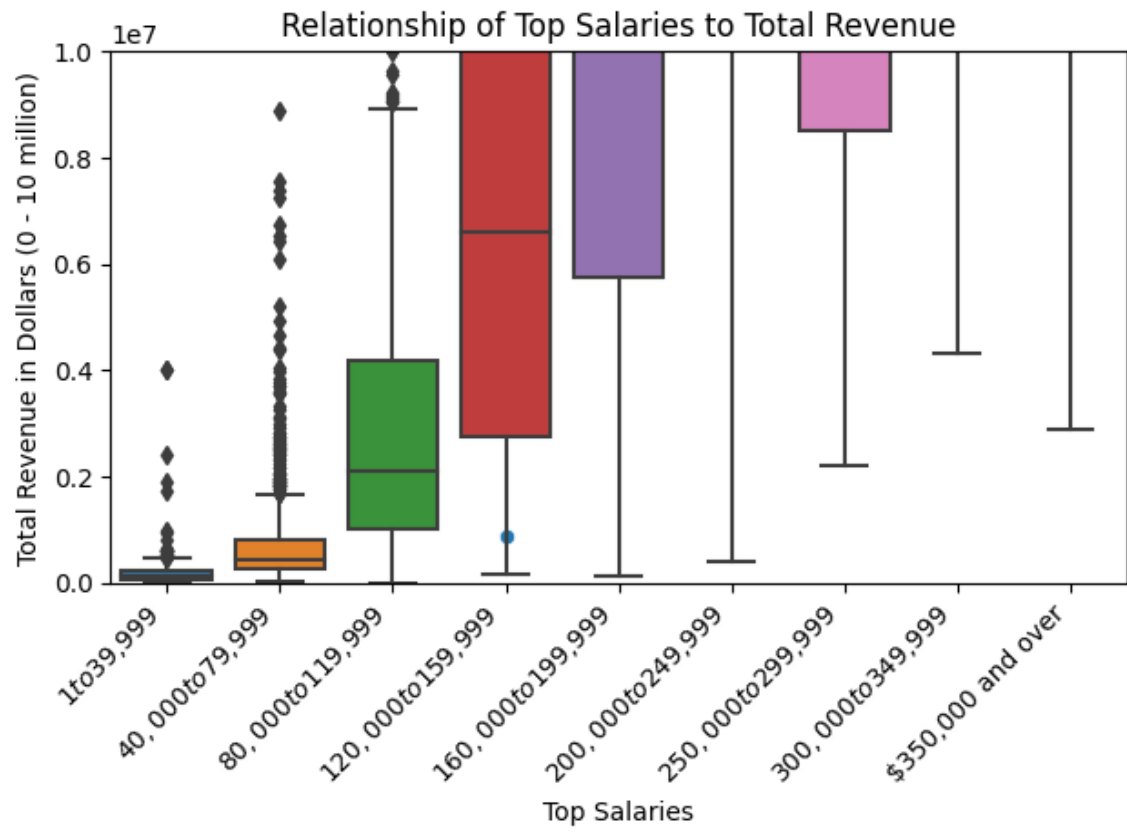
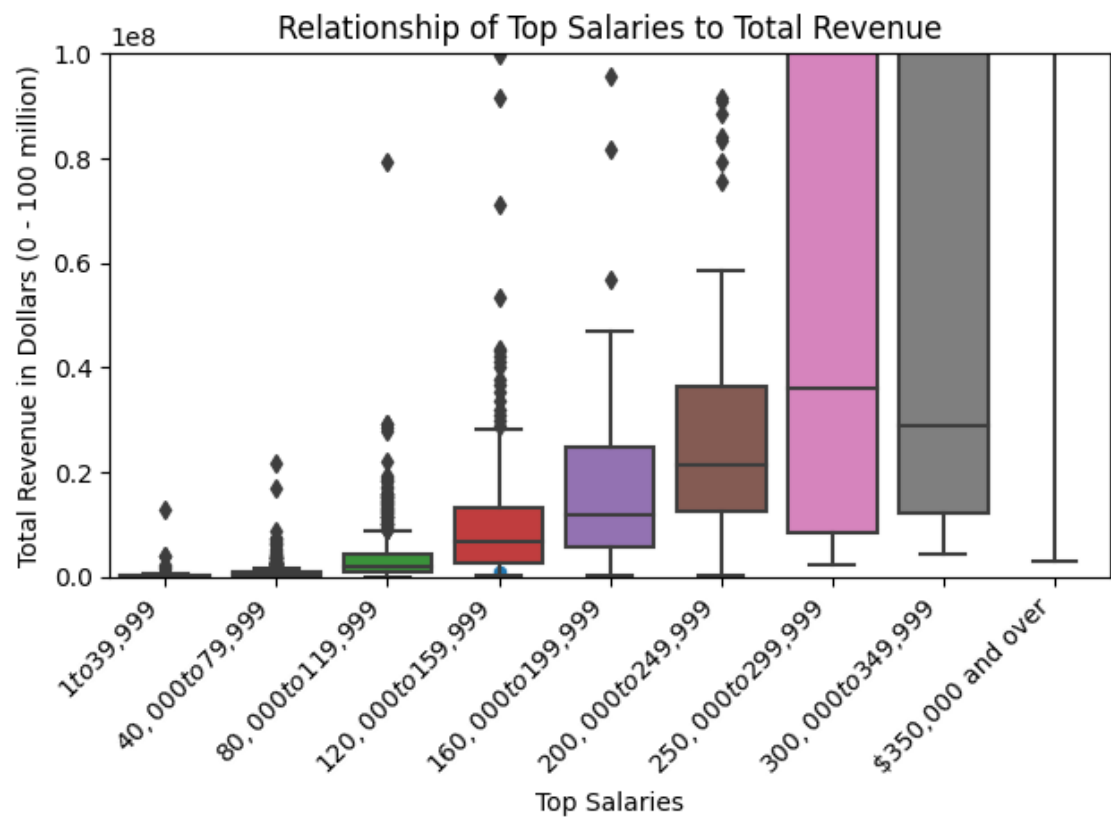
As you would expect, the vast majority of salaries are in the lower bins, with the peak being the \$40k to \$80k range, beyond which the numbers quickly decrease. At the bin for \$350k and over there is a small uptick in people with those salaries, this is due to the range extending, in principle, to infinity.

One of the strongest predictors of top salary is total revenue. The ranges of total revenue for charities with top salaries falling in each bin grows consistently as the salary gets larger.

Side Note: Box Plots are graphs which represent several key data points about a distribution in a simple plot. The black horizontal lines in the coloured boxes represent the median of the data, the boxes represent the interquartile range (IQR) which is 50% of the data points of the data (25% above/below the black line). The "T"s represent the range of the data. The small black diamonds represent outliers in the data.

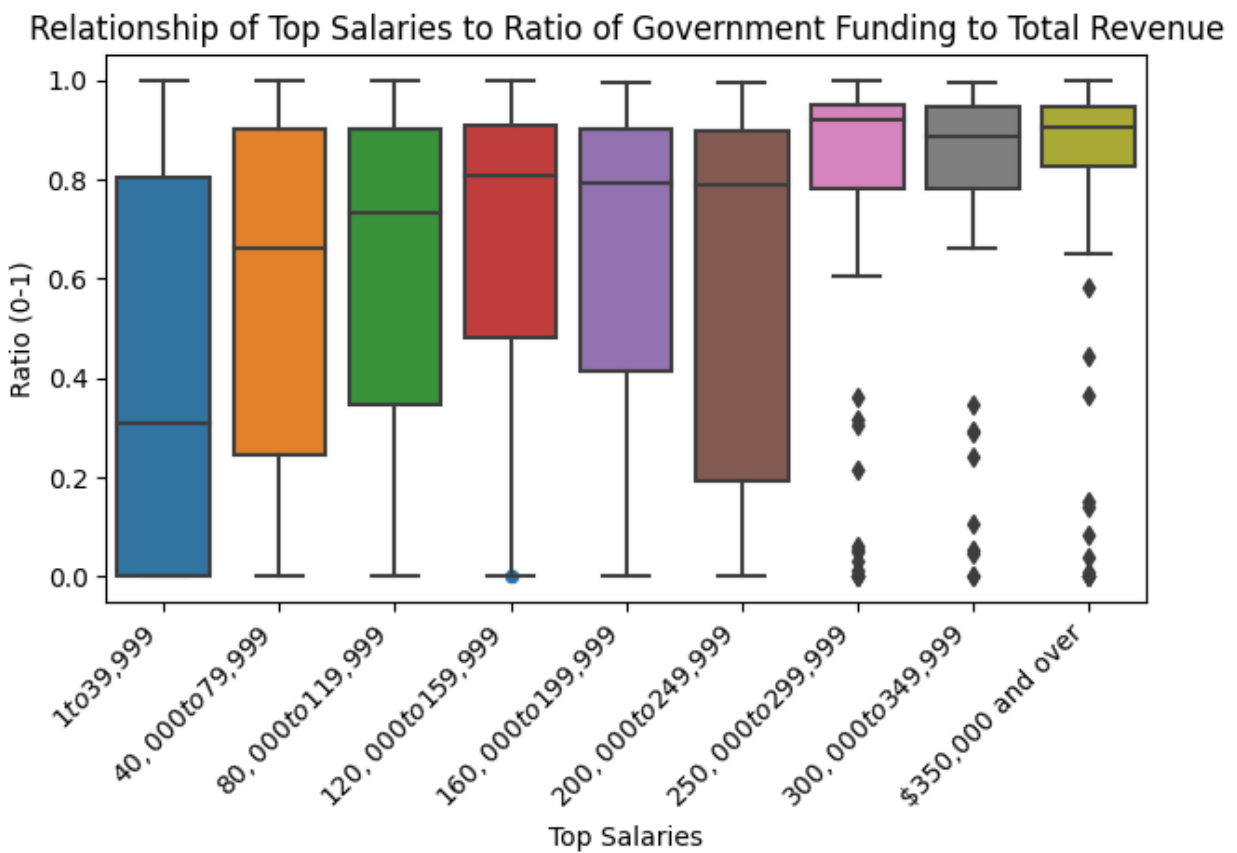


The two graphs below are of the same data and relationship as the one above, but zoomed in to more clearly see the distribution for lower salary bins. A blue dot was added to the \$120k to \$160k range to represent Lung Cancer Canada.



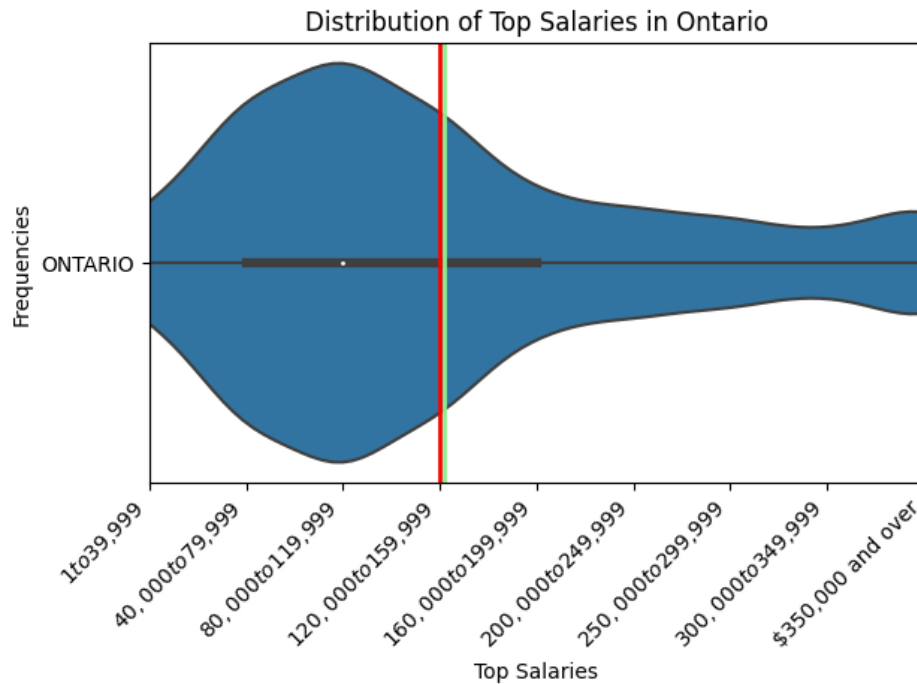
This relationship, of top salaries being more common for charities with higher total revenue, is present when looking at all monetary columns, and also the number of employees. The more money a charity makes or spends, the more people they can hire, the higher the top salary.

A more interesting relationship, in my opinion, is the relationship between the ratio of government spending to total revenue versus the top salary bins. The higher the percentage of total revenue that comes from the government, the more likely the charity is paying a larger top salary. This is shown through the medians in the box plot below.

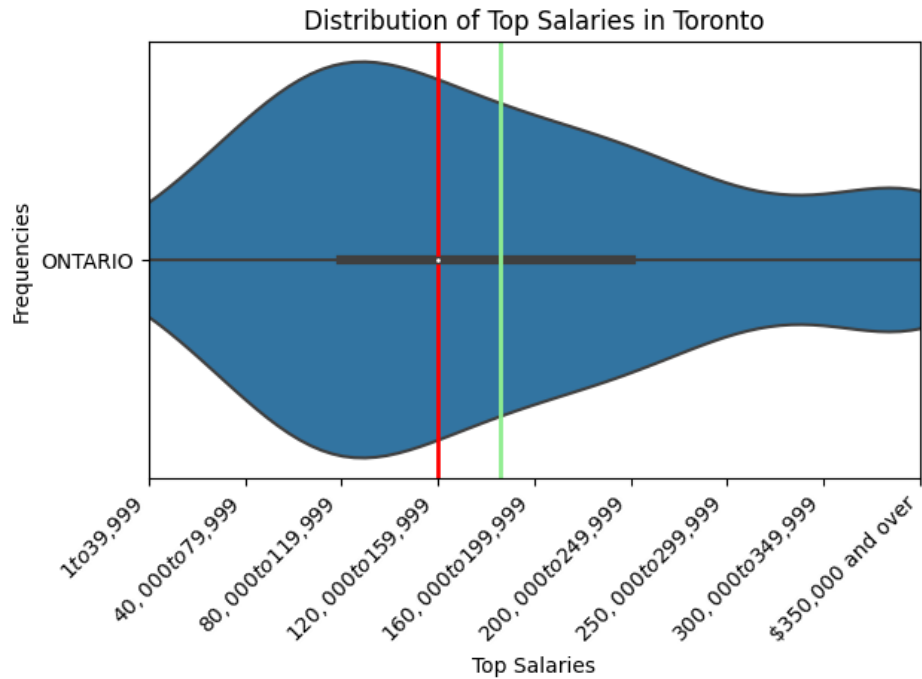


Focusing on Charities Similar to Lung Cancer Canada

From my initial exploration of the distributions of data in the dataset, along with noting the values for Lung Cancer Canada, I began to restrict my analysis to a subset of charities that have data similar to those of Lung Cancer Canada. I first looked at charities within Ontario and within Toronto.



Within Ontario, the average top salary is slightly above the \$120k to \$160k range.
Top Salaries of Charities in Toronto



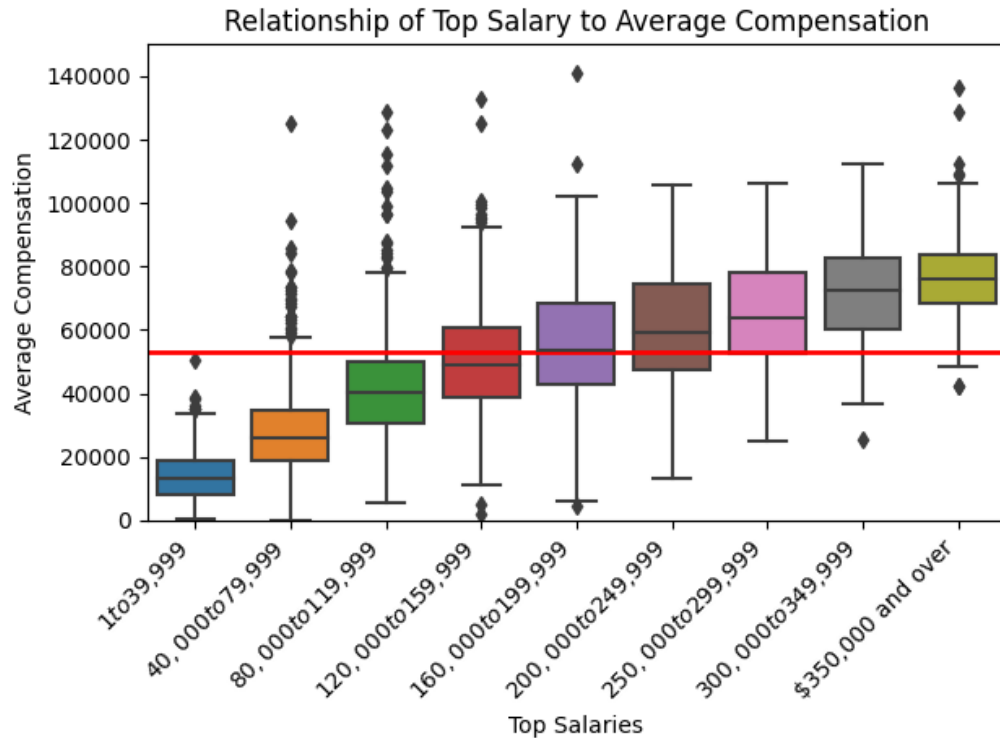
Within Toronto, however, the average top salary rises to close to the \$160k to \$200k range.



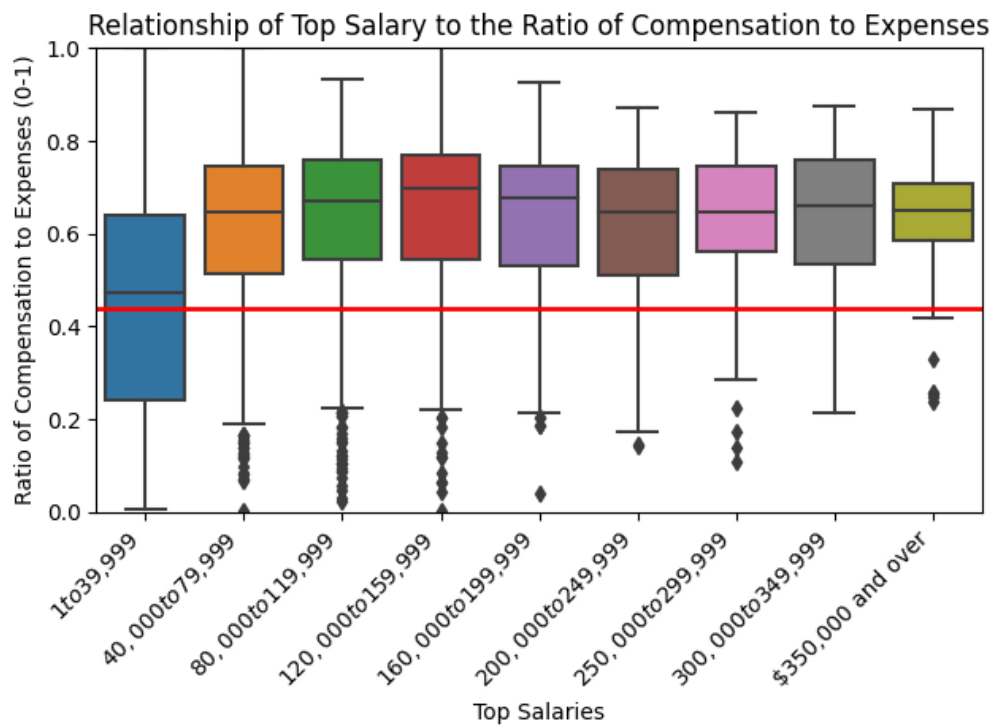
This graph shows the relationship of top salary to total revenue for charities in Ontario, zoomed in to only show total revenues up to \$10 million. Top salaries in Ontario are slightly higher as a function of total revenue compared to Canada as a whole.



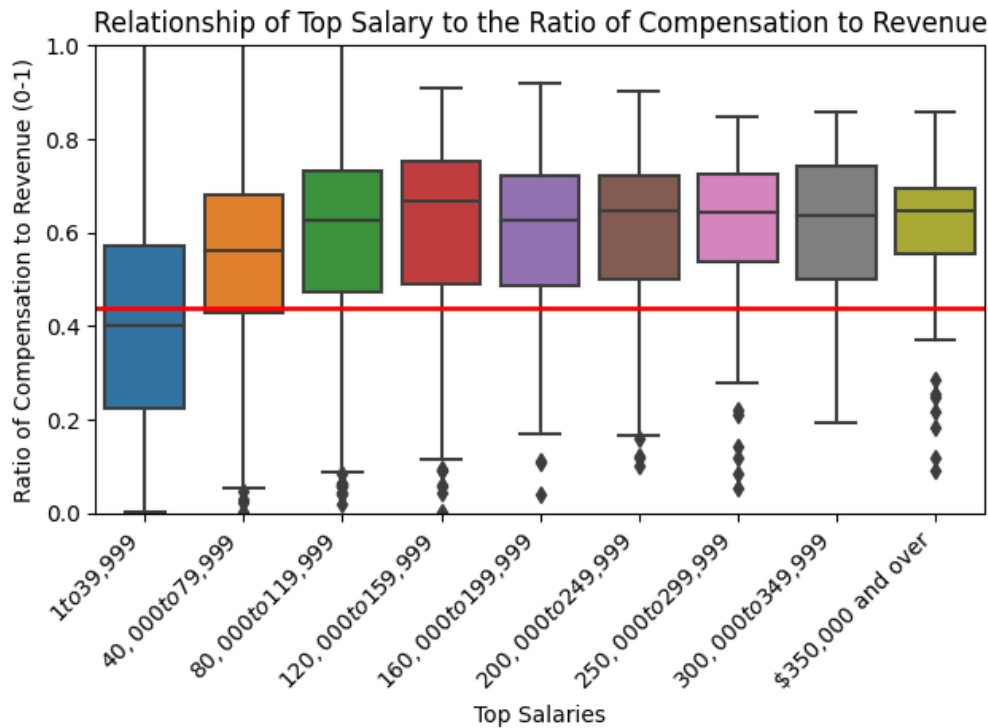
This graph shows the relationship of top salary to total revenue for charities in Toronto, zoomed in to only show total revenues up to \$10 million. Top salaries in Toronto are higher as a function of total revenue compared to Canada as a whole.



As top salaries grow higher, as expected, so does the average compensation per employee. However, with an average compensation for the employees at Lung Cancer Canada of over \$53k, this average compensation is nearly identical to the median average compensation for charities with top salaries in the range of \$160 to \$200k.



This graph shows the relationship between top salary and the ratio of total compensation to total expenses. With a ratio of total compensation to total expenses of 0.44, or 44% of total expenses go to compensation, the only comparable top salary distribution is for the range of \$0 to \$40k. As more of a charity's expenses go to compensating their employees, the more likely charities are to offer a higher top salary. This relationship is also present when looking at the relationship of top salaries to the ratio of total compensation to total revenue.



Conclusion

When compared to other health care charities in Canada, the top salary at Lung Cancer Canada at times appears to be too high, while at other times too low. If looking strictly at total revenue and expenses, with a total revenue of \$755k and total expenses of \$698k, Lung Cancer Canada's top salary would be expected to be between \$40k and \$120k. But, when taking a closer look at other metrics, the conclusion changes. Due to Lung Cancer Canada's high average compensation, and the low ratio of total compensation to both expenses and revenue, one would expect their top salary to be in the range of \$160k to \$250k dollars.