In recent years, Toronto has witnessed a transformative shift in its childcare landscape. The municipal and regional childcare systems have been stretched to support the rising demand for accessible and affordable childcare spaces, compounded by the need to ensure quality care that supports the developmental needs of children. The urgency of this situation has called for efforts to investigate and rectify the underlying issues of capacity and resource allocation within Toronto's childcare infrastructure.

This report explores the impact of operational type, subsidy provision, and geographical location on the availability of childcare space in Toronto, with an emphasis on policy implications and management of childcare center operations. The dataset used for this analysis is obtained from Open Data Toronto, and contains information and operational capacity of 1063 licensed childcare centers in Toronto, including center name, auspice type (Commercial, Non-Profit, and Public), ward, total childcare space available, and subsidy status.

This exploratory analysis is guided by three main research questions:

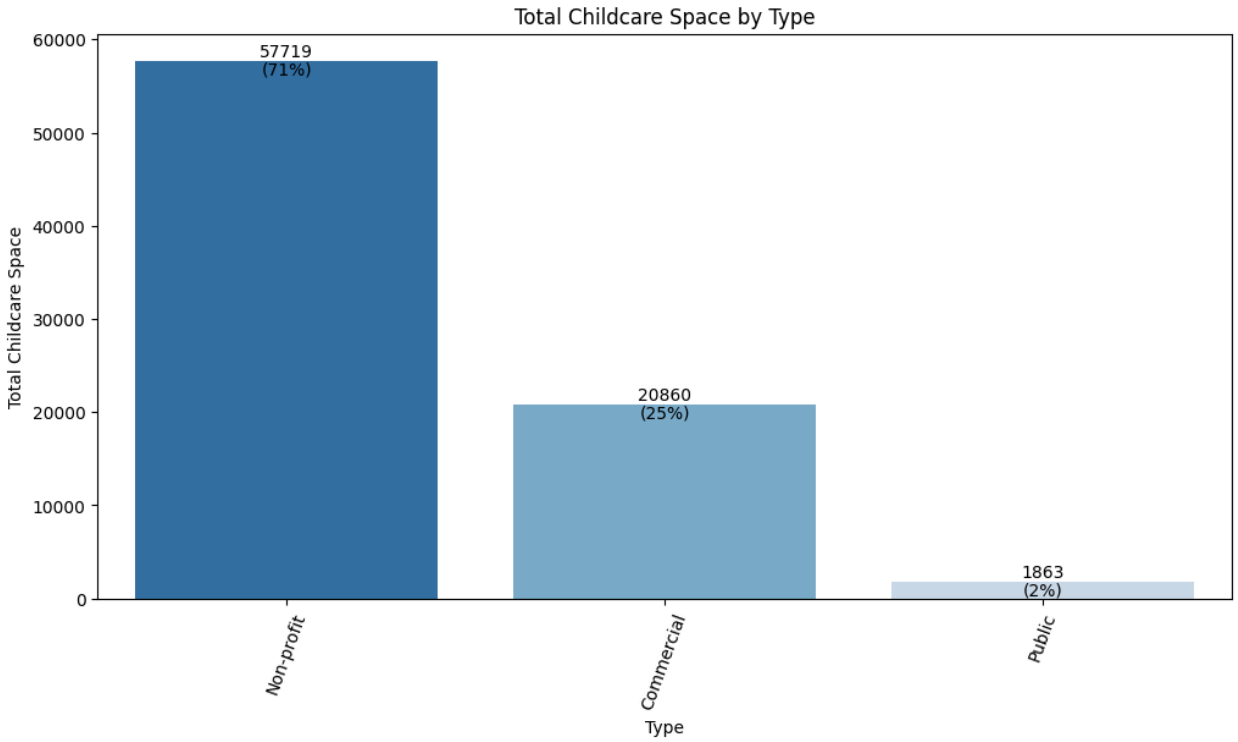
1. Is there a significant difference in total childcare space between different types of childcare centers (Commercial, Non-profit, Public)?
2. Does the provision of subsidies correlate with the amount of total childcare space in childcare centers? If so, how?
    a. Is there an interaction effect between the type of childcare center and subsidy status on the total childcare space?

I. Comparative Analysis of Childcare Space (total_space) by Auspice Type

Among all 80442 reported childcare spaces across 1063 childcare centers in 25 wards, the majority (71%) of these spaces are provided by non-profit childcare centers. This indicates that non-profit organizations play a crucial role in the childcare sector in Toronto, possibly due to a combination of government support, community trust, and potentially lower costs compared to commercial providers.

Commercial childcare centers account for a quarter (25%) of the total childcare spaces. This substantial portion underscores the role of the private sector in meeting the demand for childcare, likely catering to a variety of parental needs and preferences, possibly offering more diverse services or locations not covered by non-profit centers.Only a small fraction (2%) of childcare spaces are provided by public centers. This suggests that direct municipal or governmental operation of childcare centers is limited, possibly reflecting policy choices, funding priorities, or historical development of the childcare system in Toronto.

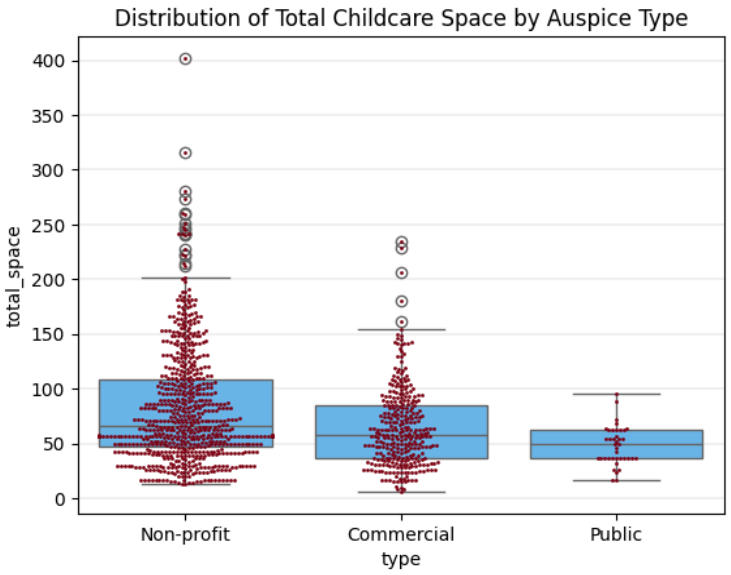
After seeing significant disparity in childcare provision models, it is important to investigate if there are truly more childcare spaces at non-profit childcare centers, or simply a large share of non-profit organizations.To statistically investigate whether the disparity in childcare provision models reflects genuinely more childcare spaces at non-profit centers or simply a larger share of non-profit organizations, you can conduct a hypothesis test. The key question is whether the average number of spaces per center differs by type of provider (non-profit, commercial, public).



Hypotheses:

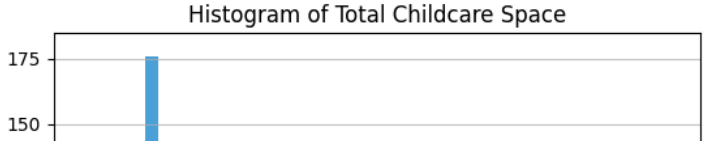
- **Null Hypothesis (H0)**: There is no difference in the average number of childcare spaces provided by non-profit, commercial, and public centers.
- **Alternative Hypothesis (H1)**: There is a significant difference in the average number of childcare spaces provided among non-profit, commercial, and public centers.

This boxplot visualization of distribution of total childcare space among three types of childcare centers: Non-profit, with a median of 82 spaces, Commercial, with a median of 65 spaces, and Public, with a median of 48 spaces, suggests that: while non-profit centers dominate the provision of childcare spaces, there is a high variability in the capacity of these centers.



Some non-profit centers offer many more spaces than the median, while others offer significantly fewer. Commercial centers also show variability, with a number of outliers indicating that certain commercial centers have a large capacity, though generally, they have less space than non-profits. Public centers tend to have the least variability and typically fewer spaces than both nonprofit and commercial centers. To statistically verify if the observed variability in childcare spaces across different types of centers, a one-way ANOVA is needed for further analysis.

To check if data meets the assumptions for an ANOVA test, the frequency distribution of total childcare spaces by center was visualized using a histogram.



The distribution is right-skewed, with the majority of childcare centers offering a smaller number of spaces (in the range of 1 to 50) and fewer centers offering a larger number of spaces. The right-skewness of the histogram could imply that while the capacity to accommodate a large number of children is available in the system, it is not the norm, signaling a need to address the potential shortage of spaces in larger centers or the need to support the sustainability of smaller centers.In addition, childcare spaces significantly decrease beyond a total space of 100, with very few centers offering more than 250 spaces. Upon further examination, all seven (7) childcare centers with over 250 spaces are non-profit centers.

To test the normality of the data, a Shapiro-Wilk test was run. The test statistic w is 0.901775598526001, which is less than 1. In a perfectly normal distribution, w would be approximately equal to 1. The p-value is very small (1.4964898448030214e-25), which is practically zero. This indicates the null hypothesis of the Shapiro-Wilk test, which states that the data is normally distributed, should be rejected. With the p-value being extremely small, it is concluded that the residuals of your ANOVA model do not follow a normal distribution. This suggests that the normality assumption of ANOVA has been violated, which can potentially invalidate the results of the ANOVA.

In addition, homogeneity of variances was tested using Levene's test. The p-value from Levene's test is extremely small (2.2036889276494363e-08), which means there is enough evidence to reject the null hypothesis of equal variances across the three types of childcare centers (Commercial, Non-profit, and Public). This suggests that the variances of 'total_space' are significantly different between at least two of the groups of childcare types. This inequality of variance could affect the validity of the ANOVA results.

However, ANOVA is considered robust to violations of normality with large sample sizes and could provide important insights. The sample size of the data (1063) is large enough to apply the Central Limit Theorem, which states that the distribution of the sample means will approximate a normal distribution, even if the underlying data do not.

As a result, a one-way ANOVA was run with caution and consideration for the limitations and produced the following results.

Table: One-Way ANOVA Results for Total Childcare Space by Type

Source	df	Sum of Squares	Mean Square	F-Value	p-value
Type	2	96,112.11	48,056.06	21.84305	<0.00001
Residual	1060	2,332,065	2,200.06		

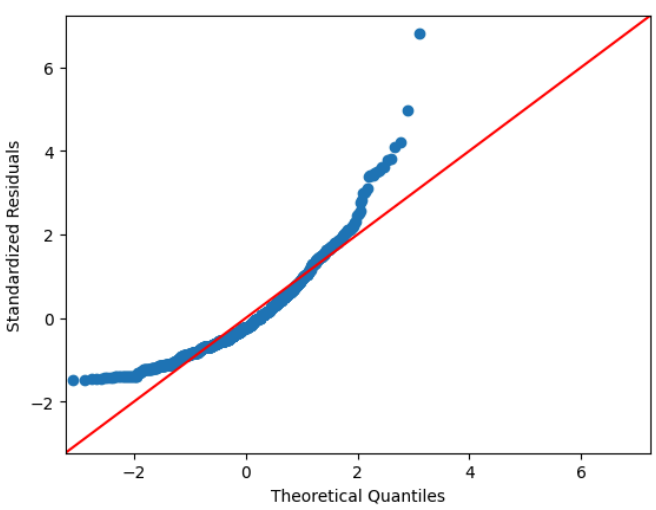
The p-value is approximately 5.06e-10, which is much lower than 0.05, indicating that there is a statistically significant difference in the total childcare space among the different auspice types of agencies.

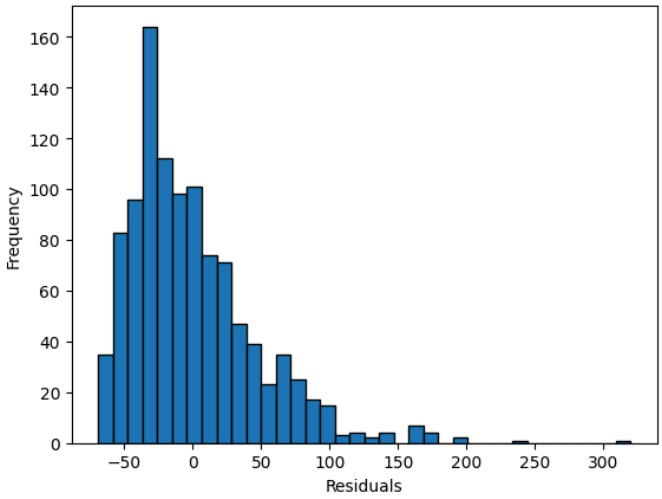
To further examine differences between groups, a Post-hoc Tukey test was conducted and produced the following results:

- Commercial vs. Non-profit: The mean difference is 17.1194, with the p-value adjusted for multiple comparisons being less than 0.001. This result is statistically significant, meaning that non-profit childcare centers tend to have, on average, 17 more spaces than commercial centers
- Commercial vs. Public: The mean difference is -7.2152, but the adjusted p-value is 0.0779, which means no statistically significant difference detected

- Non-profit vs. Public: The mean difference is -34.3346, with an adjusted p-value of less than 0.001. This result is statistically significant, indicating that non-profit childcare centers have on average 34 more spaces than public ones

The Q-Q plot and histogram of residuals of this one-way ANOVA showed non-standardized residuals obtained from ANOVA and confirmed the violation of assumptions.





Considering the violation of ANOVA assumptions, in addition to a traditional one-way ANOVA, a non-parametric ANOVA (Kruskal-Wallace) was used to further study the differences in means across the various types of childcare centers in the dataset, and confirmed that there is a statistically significant difference in the total childcare spaces between at least some of the groups defined by auspice type (Commercial, Non-profit, and Public). A Games Howell test was run as a post-hoc test for non-parametric ANOVA and showed the following results.

	A	B	mean(A)	mean(B)	diff	se	T	df	pval	hedges
0	Commercial	Non-profit	64.984424	82.103841	-17.119417	2.829010	-6.051380	850.881429	6.449011e-09	-0.359051
1	Commercial	Public	64.984424	47.769231	17.215193	3.514136	4.898840	84.038233	1.371187e-05	0.489742
2	Non-profit	Public	82.103841	47.769231	34.334610	3.465785	9.906734	80.999508	3.108624e-15	0.676255

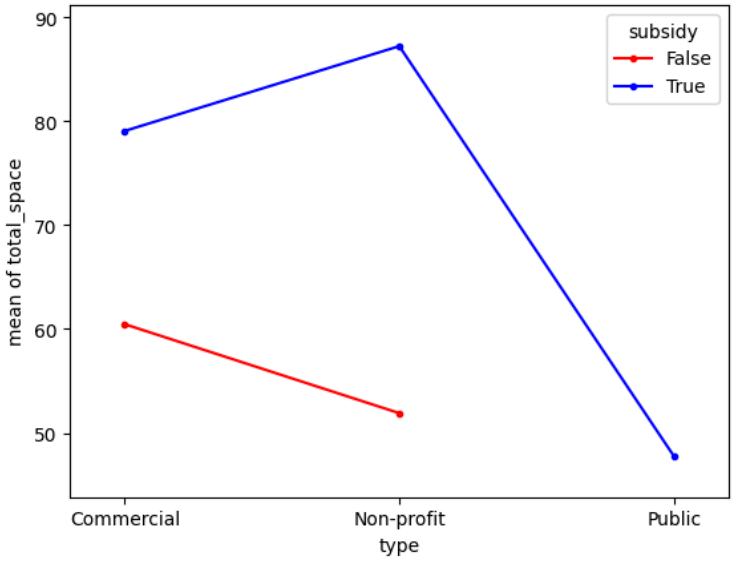
In conclusion, there is a significant difference in the number of total childcare spaces between non-profit childcare centers, and commercial and public centers respectively, with non-profit agencies having more spaces. Non-profit organizations seem to be the leading providers of childcare spaces, which might relate to their potential for government support. This calls for further investigation to examine factors such as funding models and subsidy status for non-profit as well as commercial and public childcare providers.

II. Subsidy

The averages of childcare space by subsidy status indicates a significant difference in the average total childcare space between centers that have a fee subsidy contract (84.2) and those that do not (57.9) (Appendix 1). This suggests that subsidy status might be an important factor influencing the availability of childcare space.

A Kruskal-Wallace Test was run and indicated there is a statistically significant difference in the mean total childcare space between centers that have subsidy contracts and those that don't, with subsidized centers having more space on average. The effect size (Hedges' g) suggests a moderate to large difference between the groups, indicating that the difference is not only statistically significant but also practically significant.

To explore other variables that might influence total space or interact with subsidy status, a two-way ANOVA would be helpful to understand if there is an interaction effect between auspice type and subsidy on total childcare space.



The interaction plot illustrates a clear interaction effect between the type of childcare center and subsidy status on the mean total childcare space. Specifically, non-profit centers exhibit a substantial increase in total space when subsidized compared to commercial ones, while public centers show a significant decrease in total space when subsidized, suggesting that subsidies have a differential impact on childcare space depending on the type of center.

A two-way ANOVA was conducted and produced the following results.

Table: Two-Way ANOVA Results for Childcare Space by Type and Subsidy

Source	Sum of Squares	df	F-Statistic	p-value
Type	8,567.996	2	2.057586	0.128273
Subsidy	83,527.440	1	40.117876	<0.0001
Type:Subsidy (Interaction)	56,034.450	2	13.456555	0.00000169
Residual	2,202,809	1058		

The significant main effect of subsidy status confirms that it is an important factor in determining total space, independent of the type of center. In addition the significant interaction effect suggests that the impact of subsidies on total space is not uniform across all types of childcare centers. Some types may benefit more from subsidies in terms of total space than others. Since there are no public, non-subsidized centers, the analysis of how subsidy status affects total space focused on the groups that do exist, by comparing non-subsidized public centers with subsidized and non-subsidized centers of other types. A Kruskal-Wallis Test was run for non-parametric data, followed by a Dunn's Test for post-hoc pairwise analysis (Appendix 2).

Table : Kruskal-Wallis Test for Differences in Total Childcare Space by Type

Test Statistic	Degrees of Freedom	p-value
110.74	2	<0.0000000000000001

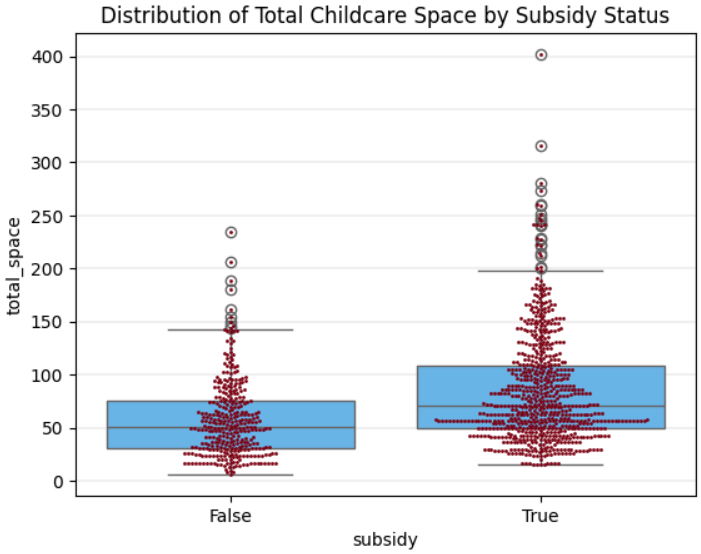
Subsidized Commercial vs. Subsidized Non-profit and Non-subsidized Commercial vs. Non-subsidized Non-profit both indicated no significant difference in the total space between subsidized or non-subsidized commercial and non-profit centers. However, subsidized public childcare centers vs subsidized commercial and subsidized non-profit centered both extremely small p-values, indicating statistically significant differences in the total space between subsidized public centers and other types of subsidized centers.

The effect of subsidy on commercial and non-profit childcare centers is not very significant. However, subsidized public centers differ significantly from both subsidized commercial and non-profit centers, which shows the need for further investigation into the factors driving these differences, potentially informing policy decisions regarding childcare subsidies and center operations.

This report highlights a significant variance in childcare spaces among non-profit, commercial, and public centers in Toronto, with non-profits providing the majority of spaces. Subsidy status significantly affects childcare space, with subsidized centers offering more spaces, especially among non-profits, underscoring the impact of funding policies on the city's childcare landscape.

Appendix

Appendix 1:



Appendix 2:

Table : Kruskal-Wallis and Post-Hoc Test Results for Childcare Space by Type and Subsidy Status

	Subsidized Public	**Subsidized Commercial**	**Non-subsidized Commercial**	**Subsidized Non-profit**	**Non-subsidized Non-profit**
Subsidized Public	-	p < 0.001	p = 0.746	p < 0.001	p = 1.000
Subsidized Commercial	p < 0.001	-	p < 0.001	p = 1.000	p < 0.001
Non-subsidized Commercial	p = 0.746	p < 0.001	-	p < 0.001	p = 0.360
Subsidized Non-profit	p < 0.001	p = 1.000	p < 0.001	-	p < 0.001
Non-subsidized Non-profit	p = 1.000	p < 0.001	p = 0.360	p < 0.001	-