

Exploring Toronto's Child Care Space Distribution

Introduction

The struggle for adequate child care in Ontario, particularly in its largest city, Toronto, is a multifaceted issue that combines economic, social, and governmental factors. High fees and low availability of spaces for children make it challenging for families to find both licensed and unlicensed child care options. According to Toronto Children’s Services, a staggering 75% of families cannot afford child care, highlighting a significant socioeconomic barrier. In response, the provincial government's commitment to creating 100,000 new child care spaces from 2016 to 2026 represents a strategic effort to address these challenges.

This analysis aims to delve into the distribution of child care spaces in Toronto, examining how factors such as geographic location and governmental policies play a role in shaping access to child care for families across the city. By understanding these dynamics, we can better assess the effectiveness of the government's initiatives and identify areas where further improvements are needed to make child care more accessible and affordable for all.

My analysis focus on two primary research questions:

Research Question 1: Determine if the total space available in child care facilities significantly varies among different auspices (e.g., public, commercial, non-profit).

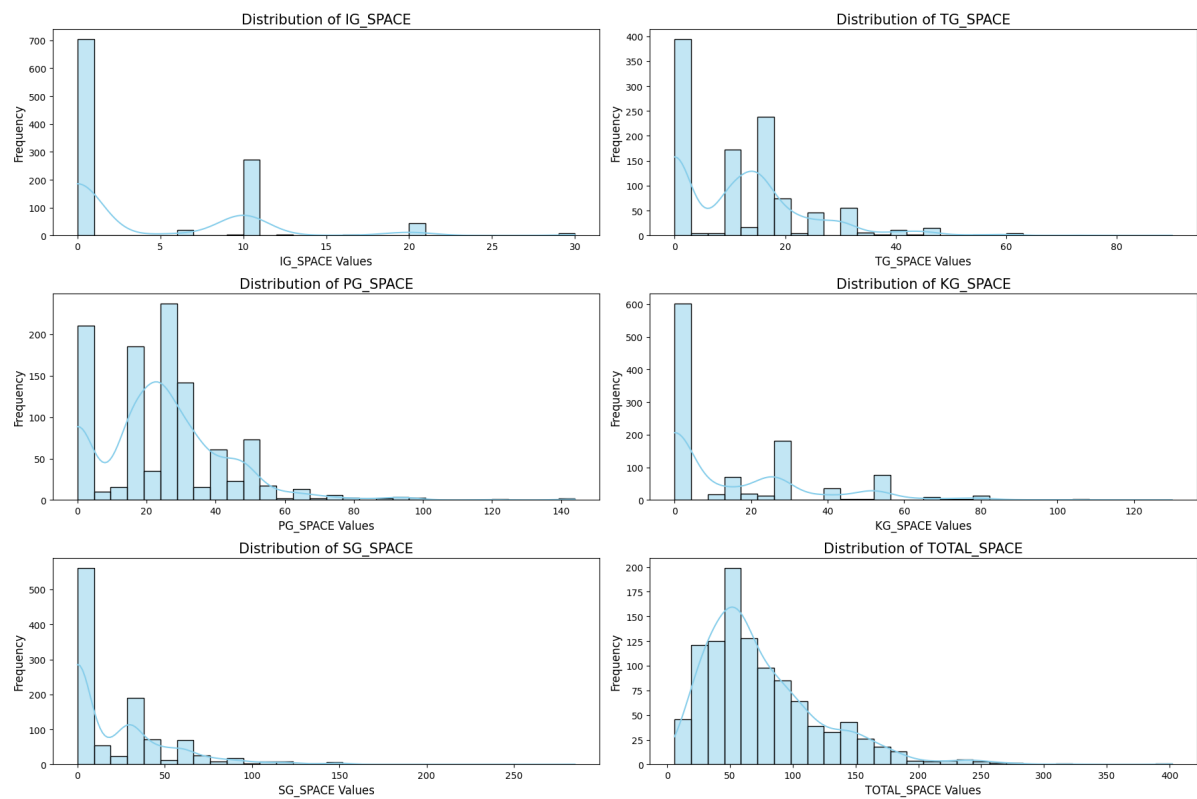
Research Question 2: Explore how the auspice of the child care facility and the presence of a subsidy contract individually and jointly affect the total space.

By understanding these dynamics, we can better assess the effectiveness of the government's initiatives and identify areas where further improvements are needed to make child care more accessible and affordable for all.

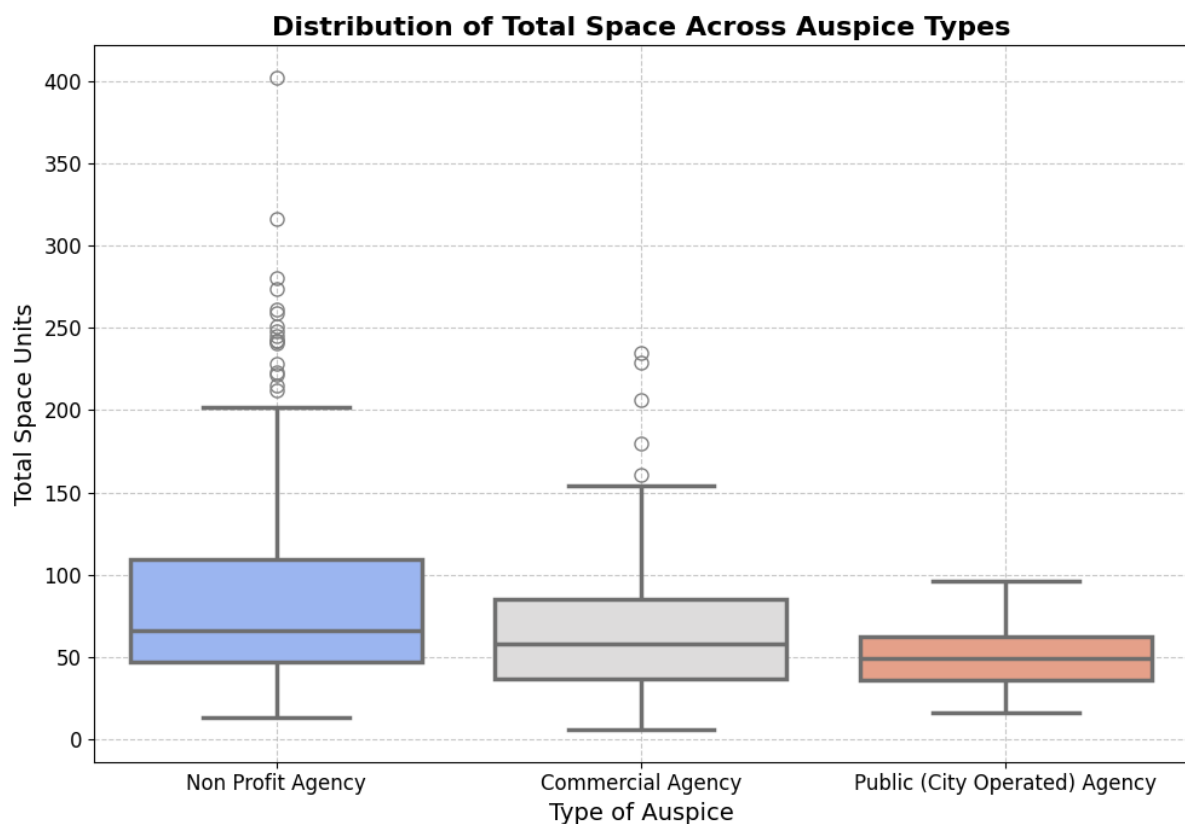
Descriptive Analysis

The dataset has 1063 rows and 17 columns. The dataset provided offers a comprehensive overview of child care facilities in Toronto, with complete data for essential variables such as `AUSPICE`, `TOTAL_SPACE`, and `SUBSIDY`, which are crucial for analyzing the distribution of child care spaces across different management types and the impact of subsidy contracts. The absence of missing values in these key areas ensures a robust foundation for statistical analysis. Although there are missing values in the `BUILDING_NAME` column, this does not impact the core objectives of our research, allowing us to proceed without concerns over data completeness.

index	count	mean	std	min	25%	50%	75%	max
IG_SPACE	1063.0	3.896519285042333	6.091782084537662	0.0	0.0	0.0	10.0	30.0
TG_SPACE	1063.0	11.60018814675447	12.088152359225928	0.0	0.0	10.0	15.0	90.0
PG_SPACE	1063.0	24.258701787394166	18.577416199873962	0.0	16.0	24.0	32.0	144.0
KG_SPACE	1063.0	14.257761053621826	20.49328982713816	0.0	0.0	0.0	26.0	130.0
SG_SPACE	1063.0	21.661335841956728	30.423502921928435	0.0	0.0	0.0	30.0	285.0
TOTAL_SPACE	1063.0	75.67450611476951	47.81651783325935	6.0	43.0	62.0	97.0	402.0



The descriptive analysis of child care spaces across different groups in Toronto reveals varying allocation patterns. Infant spaces average at 3.9 with a modest spread, indicating a limited provision. Toddler spaces are more available, averaging at 11.6, with facilities capable of accommodating up to 90 toddlers. Preschool spaces show the highest average allocation at 24.3, reflecting the broader need and focus on early childhood education, with some facilities offering as many as 144 spaces. Kindergarten and school-age spaces exhibit more significant variability, with averages of 14.3 and 21.7, respectively, and maximum capacities stretching up to 130 for kindergarten and 285 for school-age children. Overall, the total space across facilities averages at 75.7, with a substantial range from 6 to 402, highlighting a wide disparity in facility sizes and the varying focus on different age groups within Toronto's child care infrastructure.



This box plot illustrates the apparent differences between distributions of total space in child care facilities among auspice type: Non-Profit Agency, Commercial Agency, and Public (City Operated) Agency. It compares the variance, median, and outliers in space allocations across these categories. Non-Profit Agencies show a broad range with significant outliers, indicating a diversity in size. Commercial Agencies have a narrower spread but also feature outliers, while Public Agencies display the most consistent space distribution.

One-Way ANOVA

Null Hypothesis (H0): There is no significant difference in the mean total space available in child care facilities among the different auspice types (Non-Profit Agency, Commercial Agency, Public (City Operated) Agency).

Alternative Hypothesis (HA): There is a significant difference in the mean total space available in child care facilities among at least one pair of the different auspice types (Non-Profit Agency, Commercial Agency, Public (City Operated) Agency).

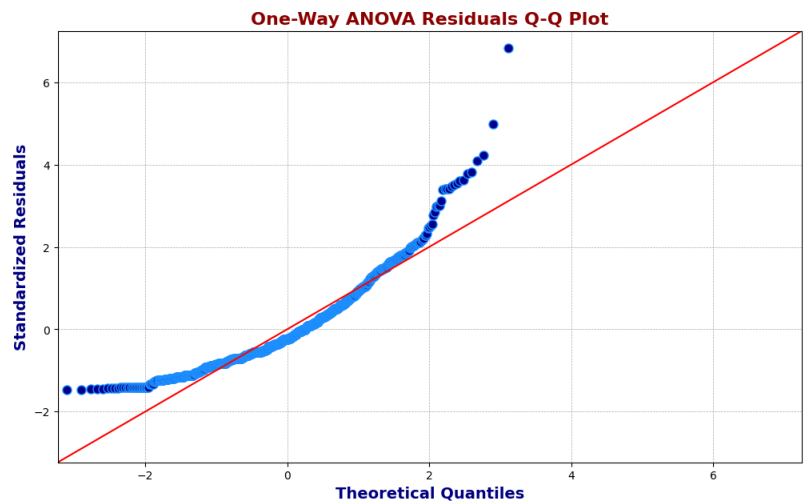
index	df	sum_sq	mean_sq	F	PR(>F)
C(AUSPICE)	2.0	96112.11429024815	48056.057145124076	21.843051025265375	5.057716322707518e-10
Residual	1060.0	2332065.2648254596	2200.0615705900564	NaN	NaN

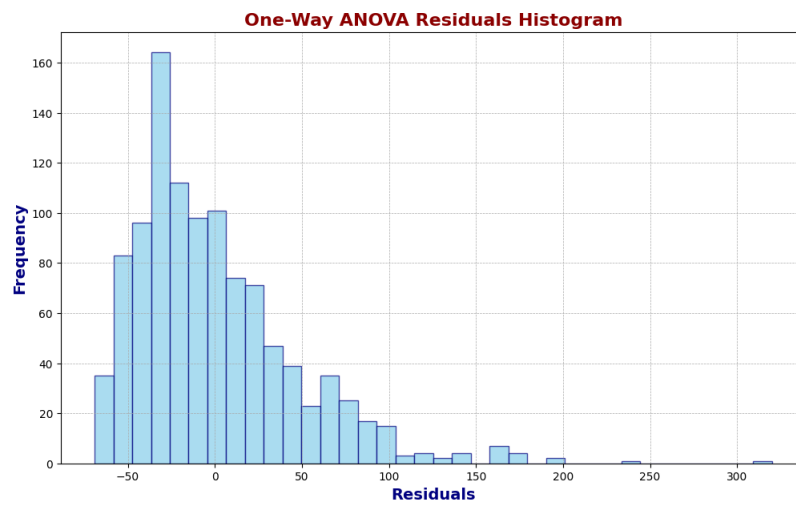
This ANOVA table assesses the effect of different child care facility auspice types (C(AUSPICE)) on total space. The table reveals:

F-value: 21.843, indicates the ratio of variance between the auspice groups to the variance within the groups. The high value suggests the groups' means significantly differ.

P-value (PR(>F)): 5.057e-10, significantly less than the typical alpha level of 0.05, provides strong evidence to reject the null hypothesis, indicating a significant effect of auspice type on total space.

The one-way ANOVA results strongly suggest that the total space available in child care facilities significantly differs across different auspice types, supporting the alternative hypothesis that not all auspice types offer the same amount of total space.





The first image is a One-Way ANOVA Residuals Q-Q Plot. The plot reveals that while many residuals closely follow the expected normal distribution (indicated by the red line), there's a noticeable deviation, especially in the tails. This suggests that the data might not perfectly meet the assumption of normality, with a slight right skew as residuals veer off the line in the positive direction.

The second image is a One-Way ANOVA Residuals Histogram. The histogram illustrates that the residuals are mostly centered around zero, indicating that the model is unbiased. However, the distribution is not perfectly symmetrical, with a tail extending to the right, which supports the observation from the Q-Q plot about the slight deviation from normality. This right skewness indicates that there are more large positive residuals than would be expected in a perfectly normal distribution.

Together, these plots suggest that while the ANOVA model is generally effective, there are some deviations from the assumptions of normality and homoscedasticity. This could affect the interpretation of the ANOVA results, particularly for data points that lie far from the mean.

Through the Shapiro-Wilk Test, with the Statistic: 0.9018, P-value: 1.496e-25. The test statistic of 0.9018 suggests a moderate fit to the normal distribution, but it's not very close to 1, indicating some deviation from normality. The extremely small p-value leading us to reject the null hypothesis of normality.

The Levene's test results with a statistic of 17.9271 and an extremely small p value is significant, which strongly suggest that there are statistically significant differences in variances across the three groups being compared. This indicates that the assumption of homogeneity of variances is violated, and alternative adjustments may be needed for further analysis.

index	group1	group2	Diff	Lower	Upper	q-value	p-value
0	Non Profit Agency	Commercial Agency	17.119417006775592	9.70359948372429	24.535234529826894	7.662434085053721	0.001
1	Non Profit Agency	Public (City Operated) Agency	34.33460991355728	16.224077335822503	52.44514249129206	6.292709906603918	0.001
2	Commercial Agency	Public (City Operated) Agency	17.215192906781688	-1.453145543277408	35.883531356840784	3.0608570206437835	0.07796644644932438

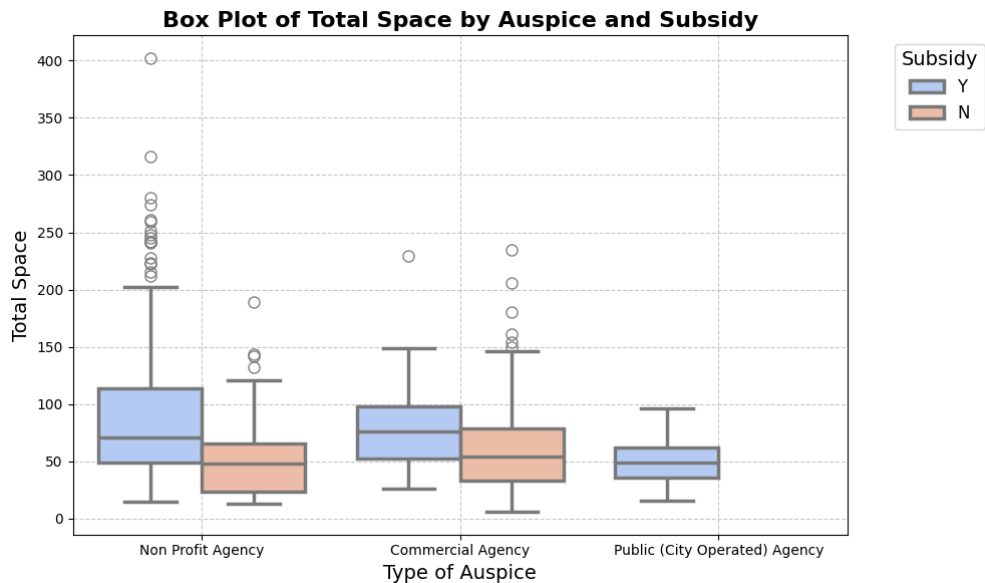
Comparison between Non Profit Agency and Commercial Agency: Shows a significant mean difference of 17.12 with a 95% confidence interval from 9.70 to 24.54 and a very small p-value (0.001), indicating a statistically significant difference between these two groups.

Comparison between Non Profit Agency and Public (City Operated) Agency: Reveals a larger significant mean difference of 34.33 with a 95% confidence interval from 16.22 to 52.45 and a p-value of 0.001, suggesting a very significant difference in means.

Comparison between Commercial Agency and Public (City Operated) Agency: Indicates a mean difference of 17.22, but the confidence interval includes negative values (-1.45 to 35.88), and the p-value of 0.07796644644932438 slightly exceeds the typical alpha level of 0.05. This suggests that while there is a mean difference, it is not statistically significant at the 0.05 level.

The Tukey HSD results provide detailed comparisons between each pair of groups, confirming significant differences between Non Profit Agencies and the other two types, and indicating that the difference between Commercial and Public Agencies might not be statistically significant at the conventional alpha level.

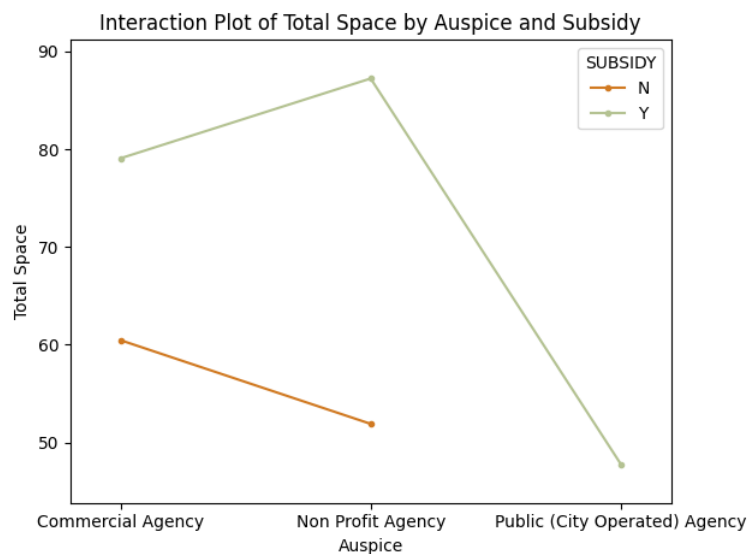
Two-Way ANOVA



The boxplot analysis reveals that subsidies correlate with larger median spaces and greater variability in both Non-Profit and Commercial Agencies, indicating that financial support likely influences facility expansion and diversity in space offerings. Notably, only subsidized Public (City Operated) Agencies are represented, lacking a comparison for non-subsidized counterparts, which could skew the analysis. This pattern suggests that subsidies play a crucial role in enhancing child care space.

index	sum_sq	df	F	PR(>F)
C(AUSPICE)	8567.996108745856	2.0	2.0575860838979505	0.1282729503898439
C(SUBSIDY)	83527.4417869589	1.0	40.117875793386546	3.529094165783121e-10
C(AUSPICE):C(SUBSIDY)	56034.454073261026	2.0	13.456555238426358	1.6942822007381886e-06
Residual	2202809.3876588224	1058	NaN	NaN

The significant interaction term indicates that the relationship between subsidy status and total space is dependent on the type of auspice, underscoring the importance of considering both factors in analyses of child care facility space. While subsidies generally increase space, the extent of this effect differs by auspice type, highlighting the complexity of factors influencing child care facility capacities.



Subsidy Impact Across Auspices: Facilities with subsidies (Y) generally have more total space across all types of auspices compared to those without subsidies (N). This suggests that subsidies are effective in increasing the total space available in child care facilities.

Differential Impact of Subsidies: The impact of subsidies on total space varies significantly by auspice type. Non Profit Agencies show a substantial increase in total space with subsidies, indicating that subsidies may be particularly effective for these agencies. Conversely, Commercial Agencies show a modest decrease in total space with subsidies, suggesting a nuanced impact of subsidies in different operational contexts.

Significant Drop for Public Agencies: There is a dramatic decrease in total space for Public (City Operated) Agencies receiving subsidies compared to those not receiving them. However, this could be indicative of data representation or the unique nature of public agency operations and funding mechanisms, rather than an actual decrease in space due to subsidies.

Multiple Comparison of Means - Tukey HSD, FWER=0.05						
group1	group2	meandiff	p-adj	lower	upper	reject
Commercial Agency,N	Commercial Agency,Y	18.5991	0.0153	2.3733	34.8249	True
Commercial Agency,N	Non Profit Agency,N	-8.5631	0.5037	-23.273	6.1469	False
Commercial Agency,N	Non Profit Agency,Y	26.7646	0.0	17.2862	36.243	True
Commercial Agency,N	Public (City Operated) Agency,Y	-12.6958	0.4893	-34.2034	8.8119	False
Commercial Agency,Y	Non Profit Agency,N	-27.1621	0.0008	-45.9161	-8.4082	True
Commercial Agency,Y	Non Profit Agency,Y	8.1655	0.5712	-6.8401	23.1711	False
Commercial Agency,Y	Public (City Operated) Agency,Y	-31.2949	0.0045	-55.747	-6.8427	True
Non Profit Agency,N	Non Profit Agency,Y	35.3277	0.0	21.9757	48.6796	True
Non Profit Agency,N	Public (City Operated) Agency,Y	-4.1327	0.9891	-27.6064	19.3409	False
Non Profit Agency,Y	Public (City Operated) Agency,Y	-39.4604	0.0	-60.0631	-18.8577	True
Multiple Comparison of Means - Tukey HSD, FWER=0.05						
group1	group2	meandiff	p-adj	lower	upper	reject
N	Y	26.2658	0.0	20.3236	32.208	True

The Tukey HSD results show significant differences in total space between groups, especially when comparing agencies with and without subsidies. Notably, Non-Profit Agencies with subsidies have significantly more space than those without, and similar trends are observed when comparing Commercial to Public Agencies with subsidies. The overall analysis confirms that subsidies play a crucial role in increasing total space across different types of agencies, with significant disparities between subsidized and non-subsidized facilities.