INF 2178-A2
Mingrui Fu 1010506551

Exploring Child Care Spaces in Toronto

1. Introduction

Child care accessibility and affordability are crucial factors impacting families, particularly in regions with high living costs like Toronto, Ontario. The challenge of finding licensed or unlicensed child care services is exacerbated by the combination of expensive fees and limited availability of spaces. Toronto Children's Services highlighted this issue, revealing that a significant portion of families, approximately 75%, struggle to afford child care services. In response to this pressing need, the provincial government pledged to create 100,000 new child care spaces between 2016 and 2026.

In this report, we delve into the landscape of licensed child care centers in Toronto, aiming to gain insights into their operation, capacity, and demographic characteristics. To achieve this, we utilize a comprehensive dataset titled "INF2178-A2-data.xlsx", which contains detailed information about these child care centers, including their locations, operating auspices, and daily capacity for various age groups. With data updated as of February 2024, this dataset offers a timely snapshot of the child care landscape in Toronto.

Through quantitative analysis using one-way and two-way ANOVAs (Analysis of Variance), we aim to uncover patterns and trends within the dataset. ANOVA is a powerful statistical technique that allows us to compare means across multiple groups, making it well-suited for exploring differences in child care center capacity based on categorical variables such as auspice type and opt in CWELCC (Child Welfare Early Learning and Child Care) or not.

In this assignment, we formulate two research questions to guide our analysis:
1. **Research Question #1:** What are the differences in child care center capacity among different auspice types?
2. **Research Question #2:** How does the auspice type of child care centers interact with the opt in CWELCC or not in influencing child care space provision?

By addressing these research questions and conducting rigorous quantitative analysis, we aim to shed light on the dynamics of child care provision in Toronto. Through this exploration, we hope to contribute valuable insights that can inform policymakers, child care providers, and families in their efforts to improve child care accessibility and affordability in the region.

2. Data Wrangling

Upon initial exploration of the dataset, we gained valuable insights into its structure, content, and quality. The dataset comprises 1063 entries with 17 columns, providing

comprehensive information about licensed child care centers in Toronto.

The dataset includes essential information such as the location ID, name, auspice type (i.e., operating auspice), address, postal code, city ward number, building type, and daily capacity for various age groups. Notably, some columns contain missing values, with 'BLDGNAME' having the highest number of missing entries (348).

To better understand the distribution and characteristics of child care center capacity, we calculated the proportions of capacity for each age group relative to the total capacity. This step involved dividing the capacities for infants (IGSPACE), toddlers (TGSPACE), preschoolers (PGSPACE), kindergarten children (KGSPACE), and grade one and up children (SGSPACE) by the total capacity (TOTSPACE) of each center. In this assignment, we only consider the child care spaces for grade one and up children (SGSPACE), its capacity proportion should be :

$$\text{SGSPACE\_prop} = \frac{Number\ of\ child\ care\ spaces\ for\ children\ grade\ one\ and\ up}{Total\ space\ capacity} \times 100\%$$

By computing these proportions, we obtained normalized measures that allow for meaningful comparisons across centers, facilitating our subsequent quantitative analysis using one-way and two-way ANOVAs. These proportions enable us to examine the distribution of child care spaces among different age groups and explore potential variations based on factors such as auspice type and CWELCC subsidy availability.

In the next steps of our analysis, we will delve into the quantitative assessment of these proportions using ANOVAs, aiming to uncover insights into the factors influencing child care space provision in Toronto. Through this rigorous analysis, we seek to contribute valuable insights that can inform policy decisions and improve child care accessibility and affordability for families in the region.

3. Quantitative Analysis Using One-way and Two-way ANOVAs

3.1 Research Question #1: What are the differences in child care center capacity among different auspice types?

To investigate the differences in child care center capacity among different auspice types, we conducted a one-way ANOVA analysis. This analysis allows us to determine whether there are statistically significant differences in mean proportions of child care spaces allocated among auspice types.

The ANOVA results revealed a significant effect of auspice type on child care center capacity ($F(2, 1060) = 107.68$, $p < 0.001$). The mean square between groups (14.43) was substantially higher than the mean square within groups (0.067), indicating a

strong influence of auspice type on capacity variation.

indext	df	Sum_square	Mean_square	F	Pr(>F)
auspice	2	14.43	7.21	107.68	2.68e-43
residual	1060	71.02	0.07	NaN	NaN

Table1: One-way anova for auspice type

Further examination through post-hoc tests using Tukey's Honestly Significant Difference (HSD) method confirmed significant differences between certain pairs of auspice types. Specifically, the mean proportion of child care spaces provided by Non-Profit Agencies differed significantly from both Commercial Agencies (mean difference = 0.25, $p < 0.001$) and Public (City Operated) Agencies (mean difference = 0.34, $p < 0.001$). However, there was no significant difference was observed between Commercial Agencies and Public (City Operated) Agencies ($p = > 0.05$).

Group 1	Group 2	Diff	Lower	Upper	Q-value	P-value
Non Profit Agency	Commercial Agency	0.25	0.18	0.32	12.14	0.001
Non Profit Agency	Public (City Operated) Agency	0.34	0.19	0.48	7.62	0.001
Commercial Agency	Public (City Operated) Agency	0.087	-0.071	0.24	1.83	0.40

Table2: Post-hoc test for auspice type

To ensure the validity of our one-way ANOVA results, we tested two key assumptions: normality of residuals and homogeneity of variances. Normality of Residuals: The Shapiro-Wilk test for normality of residuals yielded a significant p-value ($p < 0.001$), indicating that the residuals are not normally distributed. Violation of this assumption may affect the reliability of our ANOVA results.

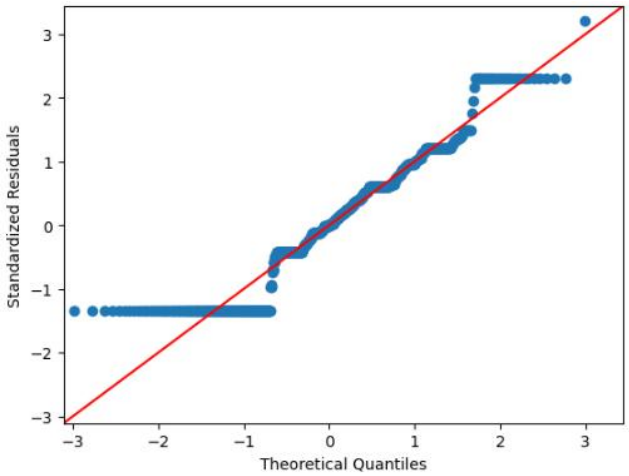

Figure1: QQ plot for residules

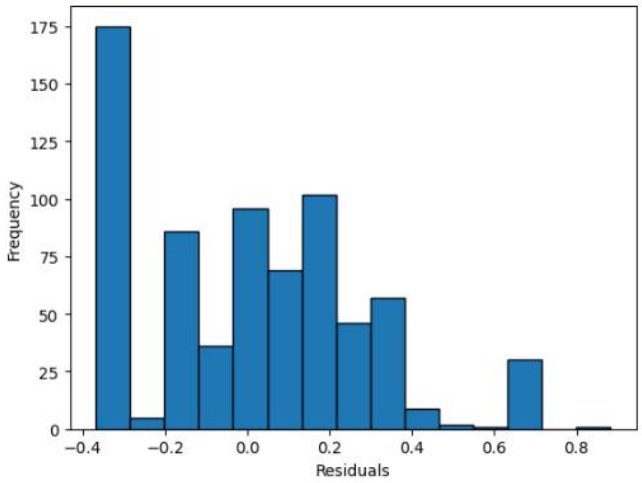

Figure2: Histogram for residuals

Homogeneity of Variances: We conducted Levene's test for homogeneity of variances, resulting in a significant p-value ($p < 0.001$). This indicates that the variances across groups are not equal, violating the assumption of homogeneity of variances.

The results of our analysis provide valuable insights into the distribution of child care center capacity across different auspice types. While significant differences in capacity were observed between certain auspice types, the validity of our findings may be impacted by violations of key ANOVA assumptions. Alternative approaches, such as non-parametric tests or transformations of the data, may be considered to address these violations and ensure the validity of the analysis. Further exploration and refinement of the statistical approach may be necessary to obtain reliable insights into the differences in child care center capacity among auspice types.

3.2 Research Question #2: How does the auspice type of child care centers interact with the opt in CWELCC or not in influencing child care space provision?

To further explore the differences in child care center capacity, we conducted a two-way ANOVA analysis considering auspice type and subsidy availability (represented by the cwelcc_flag variable). This analysis allows us to assess the main effects of auspice type and subsidy availability, as well as their interaction effect, on child care space provision.

The two-way ANOVA results revealed significant main effects of both auspice type ($F(2, 1058) = 183.31$, $p < 0.001$) and subsidy availability ($F(1, 1058) = 5.53$, $p = 0.019$). Additionally, the interaction effect between auspice type and subsidy availability approached significance ($F(2, 1058) = 2.91$, $p = 0.055$).

indext	df	Sum_square	Mean_square	F	Pr(>F)
AUSPICE	2	24.40	12.20	183.31	1.24e-38

cwelcc_flag	1	0.37	0.37	5.53	0.0189
AUSPICE:cwelcc_flag	2	0.39	0.19	2.91	0.0549
Residual	1058	70.42	0.067	NaN	NaN

Table3: Two-way anova for auspice type and cwelcc_flag

Further examination through post-hoc tests using Tukey's HSD method indicated significant differences in mean proportions of child care spaces between certain groups. Specifically, when comparing the presence (Y) and absence (N) of cwelcc_flag within each auspice type, there is a statistically significant difference ($p = 0.001$), with a mean difference of 0.12312, indicating that the presence of cwelcc_flag is associated with higher mean proportions of child care spaces. Additionally, within the Non Profit Agency auspice type, significant differences in mean proportions of child care spaces are observed between various combinations of cwelcc_flag availability (Y and N) and Commercial Agency or Public (City Operated) Agency auspice types, suggesting that the presence of cwelcc_flag influences the capacity of child care centers differently across auspice types.

Group 1	Group 2	Diff	Lower	Upper	Q-value	P-value
Y	N	0.12	0.053	0.19	4.87	0.001

Group 1	Group 2	Diff	Lower	Upper	Q-value	P-value
(Non Profit Agency, Y)	(Non Profit Agency, N)	0.04	-0.11	0.18	1.11	0.9
(Non Profit Agency, Y)	(Commercial Agency, Y)	0.24	0.14	0.34	9.93	0.001
(Non Profit Agency, Y)	(Commercial Agency, N)	0.27	0.13	0.41	7.85	0.001
(Non Profit Agency, Y)	(Public (City Operated) Agency, Y)	0.34	0.16	0.52	7.66	0.001
(Non Profit Agency, Y)	(Public (City Operated) Agency, N)	0	-inf	inf	0	0.9

Table 4：Two-way anova Post-hoc test for auspice type and cwelcc_flag

Interaction plots were generated to visualize the interaction effect between auspice type and subsidy availability on child care space provision. From figure below we can see that when CWELCC_flag is no, the line from non-profit agency type to public agency type is missed, this is because the number of data in this case is very less. Although the interaction effect approached significance, further investigation may be warranted to fully understand the nature of this interaction.

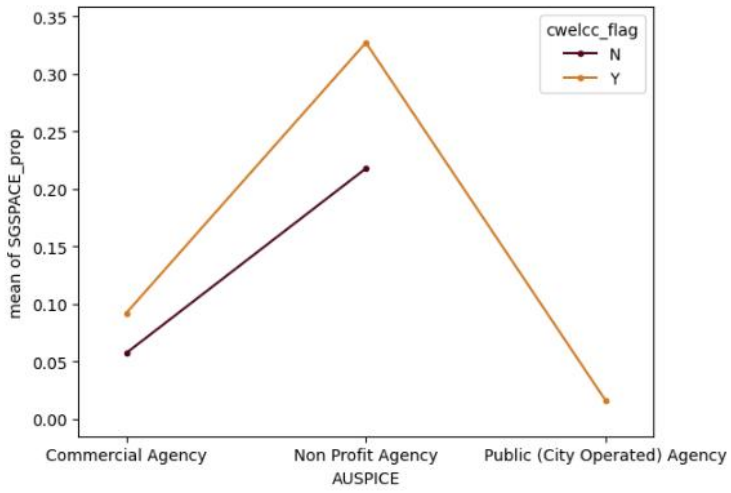

Figure3: Two-way anova interaction plot

Assumption testing for the two-way ANOVA revealed violations of the normality of residuals, indicated by a Shapiro-Wilk test statistic of 0.9285 and a p-value < 0.001. Additionally, Levene's test for homogeneity of variances could not be computed due to invalid calculations, suggesting potential issues with the assumption of homogeneity of variances.

Given these violations of ANOVA assumptions, caution should be exercised in the interpretation of results. Alternative statistical approaches, such as non-parametric tests or transformations of the data, may be considered to address these violations and ensure the validity of the analysis. Further exploration and refinement of the statistical approach may be necessary to obtain reliable insights into the differences in child care center capacity among auspice types and subsidy availability.

4. Conclusion

In conclusion, our analysis of licensed child care centers in Toronto reveals significant differences in capacity among auspice types, with Non-Profit Agencies demonstrating higher capacity. Additionally, the presence of CWELCC subsidy is associated with higher mean proportions of child care spaces, particularly within Non-Profit Agencies. However, caution is warranted due to violations of ANOVA assumptions, suggesting the need for alternative statistical approaches for more reliable insights. These findings provide valuable information for policymakers, child care providers, and families to address challenges in child care accessibility and affordability, promoting the well-being of children in Toronto. Further research and collaborative efforts are essential to explore solutions tailored to the diverse needs of families in the region.