Name: Qi Li

Student ID: 1005299302

Course: INF2178-Experimental Design for Data Science

Instructor: Shion Guha

Content: Technical Assignment II

Introduction

I aim to tackle three core research questions, helping to understand the concerning patterns exist in childcare establishments that cause inefficient allocation:

1. Are there significant differences in the number of childcare spaces offered to different age groups in Toronto licensed childcare centers, if any, between which pairs of age groups?
2. Are there statistically significant main effects of age group and subsidy in determining the number of childcare spaces available in Toronto licensed childcare centers, as well as the potential interaction effect?
3. How does the effect of subsidy interact with the effect of age group that influences the childcare space availability in Toronto licensed childcare centers?

By answering these questions, I wish to gain valuable insights into the effectiveness of current childcare establishments and provide implications on policy makings.

One-way ANOVA

I am interested in the variation of childcare availability among different age groups. The variables involved here are IGSPACE (Child care spaces for infants 0-18 months), TGSPACE (Child care spaces for toddlers 18-30 months), PGSPACE (Child care spaces for preschoolers 30 months up until they enter grade one), KGSPACE (Child care spaces for children in full-day kindergarten), SGSPACE (Child care spaces for children grade one and up). In other words, treatments are the type of capacity that differed by age group and the response variable is the number of spaces.

The first step is to generate a boxplot to see the distribution of the data by children in different age groups.
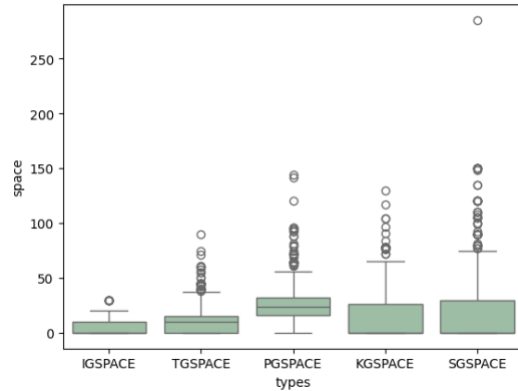
Fig 1. Box plot of childcare spaces across different age groups

As can be seen in Fig 1, there are differences between the average availabilities of spaces among different ages. Hence, it's worthwhile to test if such differences are statistically significant or not using one-way ANOVA. The corresponding null hypotheses would be:

Null hypothesis: There is no significant difference in childcare space between different age groups.
Alternative hypothesis: There is a significant difference in childcare space for at least two age groups.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F-value	P-value
Between Groups	2.821233e+05	4.0	70530.816839	188.190768	4.517383e-151
Error	1.990101e+06	5310.0	374.783617		

Table 1. One-way ANOVA Table

From Table 1, we see that the p-value is much smaller than the conventional significant level, of 0.05. Therefore, the null hypothesis is rejected and concludes that we have enough evidence to say there is a significant difference in childcare space for at least two age groups. Then, Tukey's Honestly Significant Difference test provides pairwise comparisons to determine which pairs of specific age groups exhibit statistically significant differences.

group1	group2	Diff	Lower	Upper	q-value	p-value
IGSPACE	TGSPACE	7.703669	5.412308	9.995029	12.974001	0.001000
IGSPACE	PGSPACE	20.362183	18.070822	22.653543	34.292619	0.001000
IGSPACE	KGSPACE	10.361242	8.069881	12.652602	17.449707	0.001000
IGSPACE	SGSPACE	17.764817	15.473456	20.056177	29.918310	0.001000
TGSPACE	PGSPACE	12.658514	10.367153	14.949874	21.318618	0.001000
TGSPACE	KGSPACE	2.657573	0.366213	4.948933	4.475706	0.013527
TGSPACE	SGSPACE	10.061148	7.769787	12.352508	16.944309	0.001000
PGSPACE	KGSPACE	10.000941	7.709580	12.292301	16.842912	0.001000
PGSPACE	SGSPACE	2.597366	0.306006	4.888726	4.374309	0.017028
KGSPACE	SGSPACE	7.403575	5.112214	9.694935	12.468603	0.001000

Table 2. Pairwise comparison of childcare spaces in different age groups (one categorical variable model)

Table 2 presents the outcomes of pairwise comparisons between every two age groups. The interpretation is that, for example, there is a significant difference between childcare spaces for

infants of 0-18 months and for toddlers of 18-30 months with an average difference of 7.704 as the p-value is approximately 0.01 which is less than the significance level of 0.05. In general, significant differences exist in childcare spaces across all age group pairings, as evidenced by the p-values below the 0.05 significance level for all.

Implications from One-way ANOVA and Future Research

The observed significant variations suggest that childcare centers may prioritize or allocate resources differently based on the age groups. To answer the current inefficiency of childcare policies in Toronto, further analysis of data on actual occupancy during the same period can be performed to answer whether such differences are reasonably aligned with the variations in demand. It's important to consider the specific needs of different age cohorts when making funding allocations. It highlights the necessity of regularly assessing and adjusting childcare policies to ensure that they effectively meet the evolving needs of families.
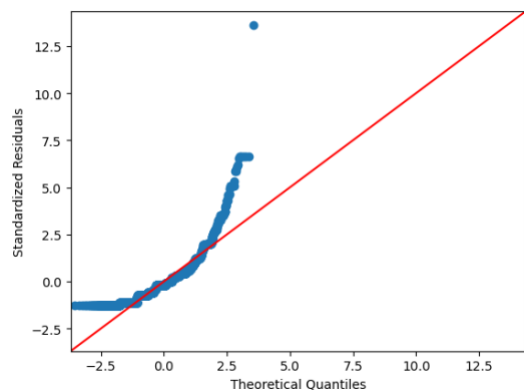
Testing One-Way ANOVA Assumptions
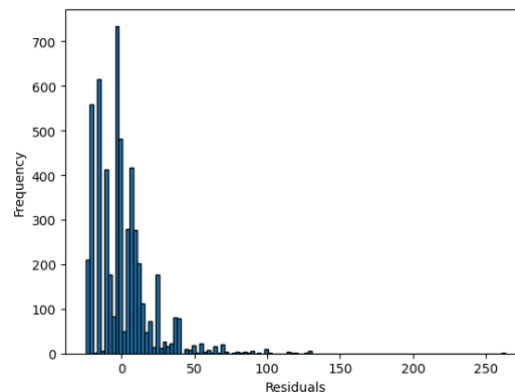

Fig 2.Q-Q plot for One-way ANOVA   Fig 3. Residual histogram for One-way ANOVA

To assess the normality assumption, the Shapiro-Wilk test, along with the visual test such as the Q-Q plot and histogram of the residual, was employed. The Shapiro-Wilk test resulted in a p-value < 0.001, providing strong evidence against the null hypothesis that states the sample (i.e. residual) came from a normally distributed population. It's worth noting that the Shapiro-Wilk test may be less sensitive for very large sample sizes (i.e. N >5000). When combined with visual tests in Fig 2 and Fig 3, all indicate the normality assumption is violated. Despite this violation, we can still proceed with ANOVA as it's known to be robust to deviations from normality (particularly in the context of large sample sizes) and test the consistency of variance assumption through Levene's test.

The p-value obtained from Levene's test is less than 0.001, indicating a rejection of the null hypothesis, thereby suggesting the assumption of consistent variances across groups is not met. It can affect the Type I error and reduce the power of analysis. Violations of both assumptions can lead to biased estimates and inaccurate conclusions. To enhance the validity of the study, future analyses should carefully consider non-parametric approaches or transformations of the data.

Two-way ANOVA

While the one-way ANOVA answers my first research question, I'm going to use two-way ANOVA to answer the second and third research questions. By integrating the additional 'subsidy' variable into the model, I aim to study how the presence or absence of government subsidy influences childcare capacity relative to age distinctions.
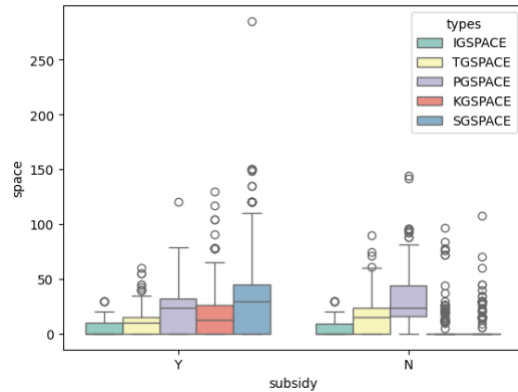

Fig 4. Box plot of childcare spaces across age groups and subsidy status

Before proceeding with ANOVA, a boxplot can provide an initial glimpse into the distribution. In Fig 4, distinct variations emerge across various age groups with and without subsidy. Notably, childcare spaces for children in full-time kindergarten (KGSPACE) and those in grade one and up (SGSPACE) appear to be influenced by subsidies. To explore this further, a two-way ANOVA, along with its post-hoc test will be employed to explore the main effects of age group and subsidy, as well as potential interaction effects.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F-value	P-value
Subsidy	3.215308e+04	1.0	32153.078590	98.168364	6.073586e-23
Types (i.e. Age)	2.821233e+05	4.0	70530.816839	215.341584	1.758733e-171
Subsidy*Types	2.204016e+05	4.0	55100.395296	168.230101	8.040792e-136
Error	1.737546e+06	5305.0	327.529943		

Table 3. Two-way ANOVA Table

Based on the findings from Table 3, one can conclude that all main effects of subsidy and age group, as well as the interaction effect between them, are statistically significant as the p-values are less than the significance level of 0.05.

group1	group2	Diff	Lower	Upper	q-value	p-value
IGSPACE	TGSPACE	7.703669	5.561623	9.845715	13.878379	0.001000
IGSPACE	PGSPACE	20.362183	18.220137	22.504228	36.683051	0.001000
IGSPACE	KGSPACE	10.361242	8.219196	12.503288	18.666072	0.001000
IGSPACE	SGSPACE	17.764817	15.622771	19.906862	32.003823	0.001000
TGSPACE	PGSPACE	12.658514	10.516468	14.800559	22.804673	0.001000
TGSPACE	KGSPACE	2.657573	0.515527	4.799619	4.787693	0.006429
TGSPACE	SGSPACE	10.061148	7.919102	12.203193	18.125444	0.001000
PGSPACE	KGSPACE	10.000941	7.858895	12.142987	18.016979	0.001000
PGSPACE	SGSPACE	2.597366	0.455320	4.739412	4.679229	0.008380
KGSPACE	SGSPACE	7.403575	5.261529	9.545621	13.337751	0.001000

Table 4. Pairwise comparison of childcare spaces in different age groups (two categorical variables model)

Table 4 presents the outcomes of pairwise comparisons between every pairing of age groups. It highlights the significant variations in childcare spaces among different age groups because all the p-values are less than the significance level of 0.05 (similar interpretation as the one-way ANOVA's post hoc test).

group1	group2	Diff	Lower	Upper	q-value	p-value
(Y, IGSPACE)	(Y, TGSPACE)	5.754875	2.731709	8.778040	8.520647	0.001
(Y, IGSPACE)	(Y, PGSPACE)	16.777159	13.753993	19.800324	24.840202	0.001
(Y, IGSPACE)	(Y, KGSPACE)	15.137883	12.114717	18.161049	22.413097	0.001
(Y, IGSPACE)	(Y, SGSPACE)	26.257660	23.234495	29.280826	38.877000	0.001
(Y, IGSPACE)	(N, IGSPACE)	0.486202	-3.266154	4.238558	0.579977	0.900

group1	group2	Diff	Lower	Upper	q-value	p-value
Y	N	5.253166	4.213767	6.292566	14.012021	0.001

Table 5. Pairwise comparison of childcare spaces with or without subsidy   Table 6. Pairwise comparisons of interaction terms

Table 5 demonstrates that the disparity in the childcare spaces between subsidized and non-subsidized facilities is statistically significant as the p-value equals 0.001 which is less than 0.05.
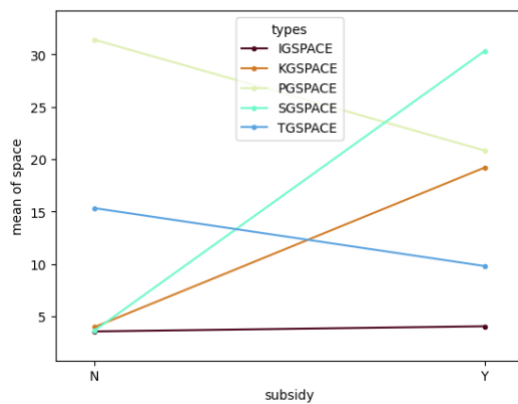

Fig 5. Interaction plot

Table 6 displays a portion of the pairwise comparison results for different combinations of subsidy status and age groups, contributing to the variations in childcare space availability. For instance, there is no significant difference in the mean of space for the IGSPACE (infants 0-18 months) either with or without subsidy as the p-value equals 0.9 which is greater than the confidence level of 0.05. We failed to reject the null hypothesis that states the means of space are equal. The rest of the pairwise comparisons show statistically significant differences as the p-values are less than 0.05.

Analysis of Interaction Effect Through Plot

Fig 5 depicts the interaction plot, a visual representation of the interaction effect of subsidy and types (age) with the x-axis representing subsidy status, the y-axis representing the means of space and different type/age groups represented by different colour lines. We observe that as variable subsidy changes from 'No' (N) to 'Yes' (Y), the response variable represented by the mean of space is increasing for the SGSPACE type. However, for the TGSPACE/PGSPACE types, we see a reversal in this trend where the mean of space decreases as the subsidy variable goes from N to Y. This crisscrossing pattern suggests a significant interaction effect between the subsidy and types variables, specifically for the types of SGSPACE and TGSPACE/PGSPACE. A similar crisscrossing pattern is evident between types of KGSPACE and TGSPACE that can be explained as a significant interaction effect.

In contrast, as variable subsidy changes from N to Y, the response variable represented by the mean of space shows an identical trend between the types of PGSPACE and TGSPACE. The relationships between the two types remain unchanged, with mean space decreasing in a parallel manner as the variable of subsidy goes from N to Y. This suggests no significant interaction effect between subsidy and types, particularly for the types of PGSPACE and TGSPACE.

Implication From Two-way ANOVA

By understanding how subsidy status influences childcare space availability across different age groups, policymakers can subsidize childcare centers more effectively. For example, when there is a greater need for subsidies, such as SGSPACE (child care spaces for children grade one and up), to increase space availability, policymakers can prioritize the allocation of subsidies accordingly to meet the heightened demands of those groups. Such a strategy ensures the subsidy is directed where it is most needed, maximizing the return out of the dollars invested.

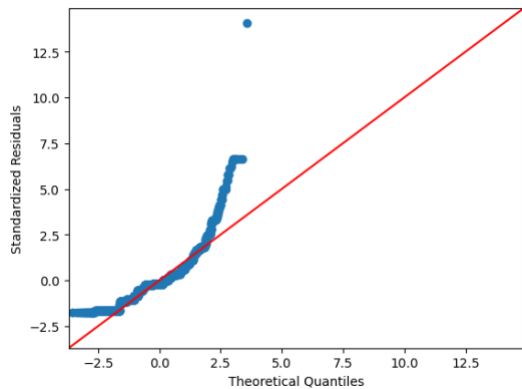## Testing One-Way ANOVA Assumptions

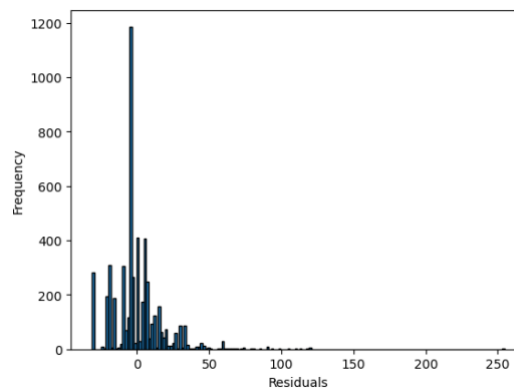Fig 6. Q-Q plot for Two-Way ANOVA   Fig 7. Residual histogram for Two-way ANOVA

The Shapiro-Wilk test and Levene's test revealed violations of both normality and homogeneity of variance assumptions, with p-values less than 0.001. Additionally, Fig 6 and Fig 7 visually demonstrate deviations from normal distribution. These findings are akin to the conclusion reached in the One-way ANOVA.