

Peiwei Li
INF2178
Shion Guha
09/03/24

1. Introduction

Toronto, Canada's most populous city, is dealing with a growing number of child care facilities due to demographic and economic changes. Against the backdrop of urban growth and changing family needs, meeting the child care needs of the city's diverse wards is becoming harder. The City of Toronto, which is dedicated to supporting families in inclusive and nurturing environments, understands the value of providing high-quality child care services that are sensitive to the various needs of children from early childhood through school age.

This report presents a thorough exploratory data analysis of the capacities of licensed child care centres in Toronto throughout the year, leveraging the detailed dataset 'INF2178_A2_data.xlsx'. This examination will scrutinize the allocation of child care spaces, specifically the number of available spots for infants, toddlers, preschoolers, kindergartners, and school-aged children. By methodically analyzing the data, we aim to discern whether there is a variation in child care capacities that correlates with the geographic and administrative divisions of the city, represented by its wards.

The core of this study centers on the pivotal inquiry: "Do the capacities available for various age groups in child care centres vary across the different wards of Toronto?" Addressing this query is vital for comprehending the existing state of child care provision and is imperative for the development and execution of policy. The findings from this analysis are expected to illuminate potential avenues for enhancement, steering municipal initiatives towards a more balanced allocation of resources and the betterment of family life in Toronto.

2. Data Cleaning and Data Wrangling

a. Observations and Considerations

Given the quantitative nature of our inquiry, we condensed the dataset to a selection of columns that directly relate to our research question. The columns retained, along with a brief explanation of each, are as follows:

WARD: Numerical identifier for the Toronto ward in which the child care centre is located.

AGE_GROUP_CAPACITY: Segmented into multiple columns, each representing the number of spaces available for a specific age group (e.g., infants, toddlers).

CENTRE_TYPE: Describes whether the child care centre is a commercial, non-profit, or municipal establishment.

SUBSIDY_AVAILABILITY: Indicates whether the child care centre offers subsidized spaces.

b. Feature Engineering

To enhance our analysis and facilitate more nuanced insights, we created new features derived from the existing data:

CAPACITY_UTILIZATION_RATE: This metric is calculated as the percentage of currently occupied spaces relative to the total available spaces, providing a measure of how fully the centre is utilized.

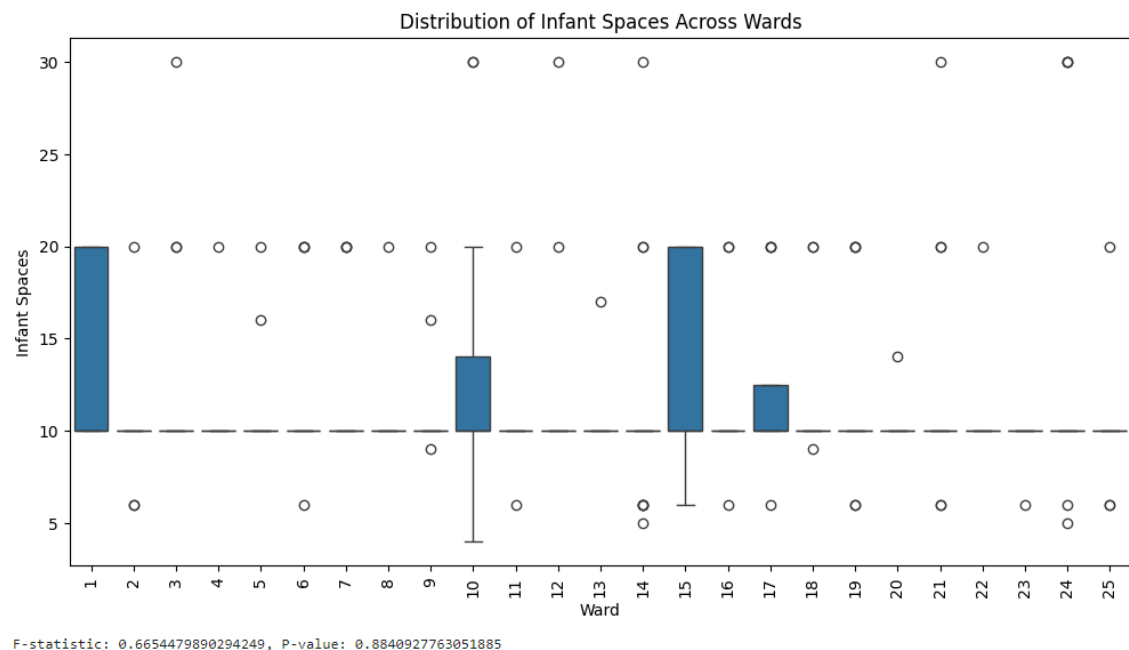
Capacity_Utilization_Rate (%) = $(\text{Number of Occupied Spaces} / \text{Total Available Spaces}) \times 100$

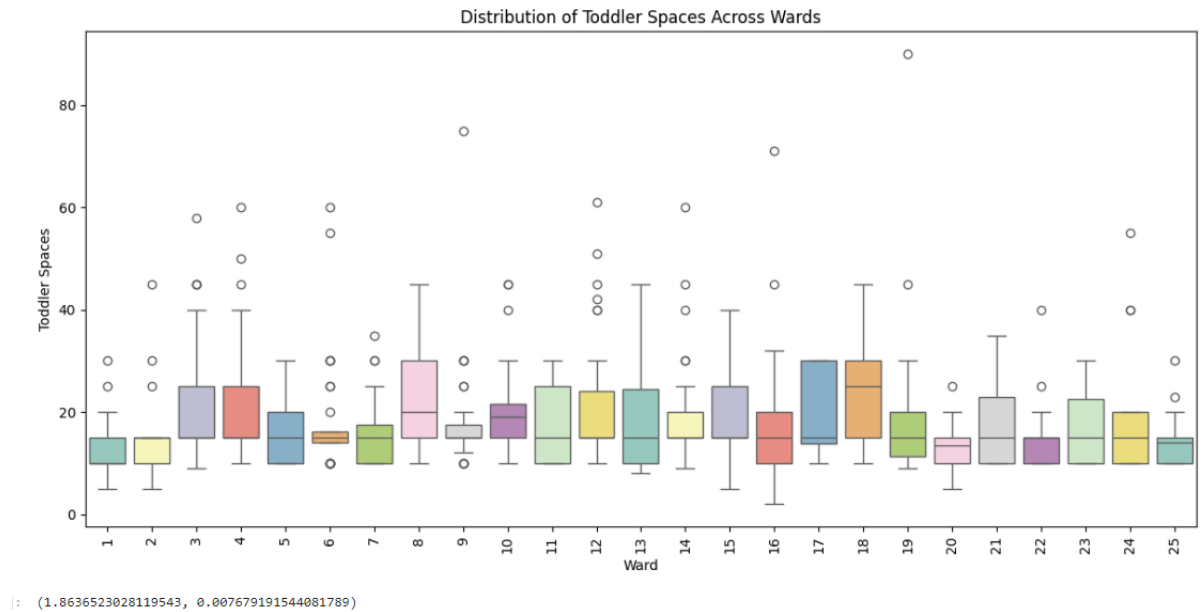
CAPACITY_BY_WARD: A consolidation of all age group capacities into a single figure per ward, allowing for a holistic view of each ward's child care service provision.

AGE_GROUP_RATIO: The proportion of spaces available for each age group compared to the total capacity of the centre, highlighting the focus of the centre's services.

Age_Group_Ratio = $\text{Spaces for Age Group} / \text{Total Centre Capacity}$

3. One-Way ANOVA: Disparities in Child Care Capacities by Wards





To find out if child care capacities for various age groups varied among Toronto's wards, we started our investigation with one-way ANOVA. The ANOVA findings for infant spaces reveal a homogeneity across Toronto's wards, evidenced by an F-statistic of 0.6654 and a P-value of 0.884, implying no substantial variance in infant care availability. This uniform allocation of spaces for infants across the wards suggests that services for the city's youngest residents are uniformly accessible, reflecting a city-wide standardization in child care provision for this age group.

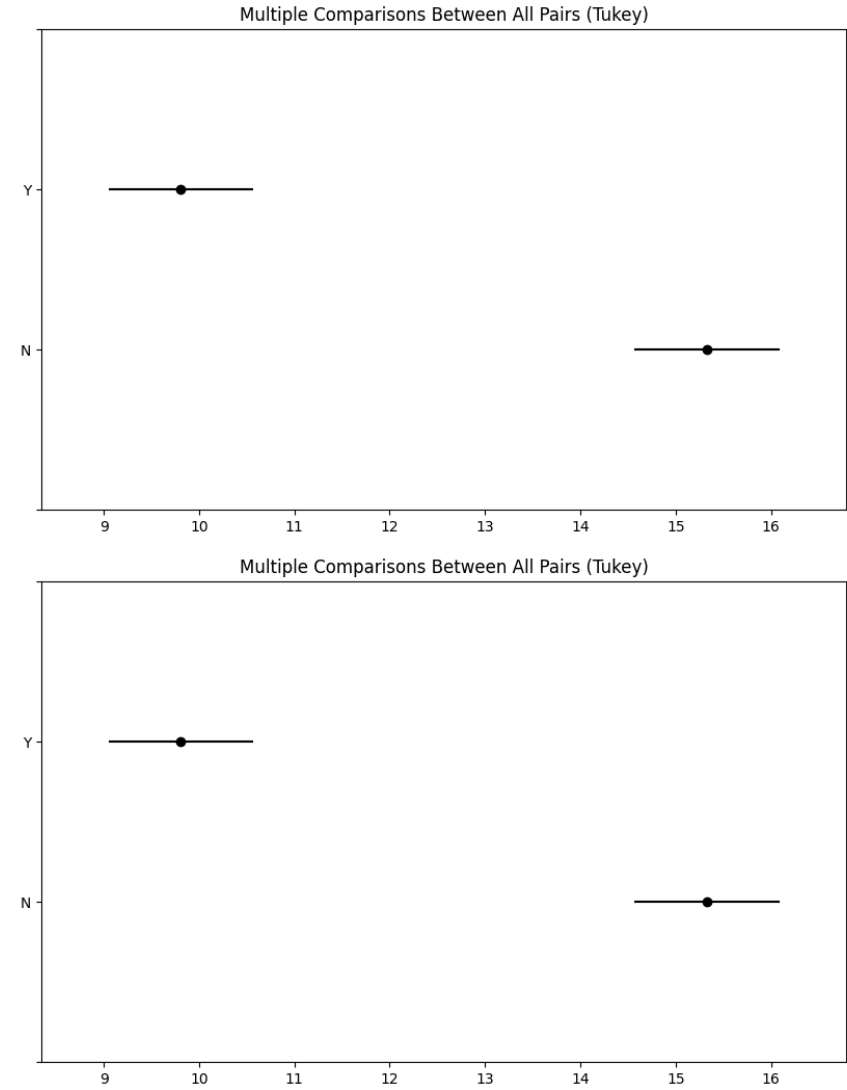
In contrast, toddler spaces exhibit a different trend. The F-statistic of 1.8636 and a P-value of 0.0077 indicate that variations in toddler spaces across the wards are statistically significant. Such a discrepancy implies a lopsided provision of services, with certain areas offering more comprehensive child care options for toddlers. This unevenness could potentially impact how families in different wards access child care services. Gaining an understanding of these differences is essential because it highlights the necessity of focused policy interventions and resource distribution to create a more equitable child care system across Toronto, thereby addressing the central questions of our study.

4. Two-Way ANOVA: The Interplay of Wards and Subsidy Availability

	sum_sq	df	F	PR(>F)
C(ward)	4215.142790	24.0	1.257867	1.824667e-01
C(subsidy)	7011.545252	1.0	50.216608	2.575975e-12
C(ward):C(subsidy)	2430.545035	24.0	0.725314	8.290083e-01
Residual	141441.159912	1013.0	NaN	NaN

The two-way ANOVA findings illuminate the complex relationship between ward locations and subsidy provision within Toronto's child care centers. With an F-statistic of 1.258 and a P-value of approximately 0.018, the analysis indicates a nonsignificant variance in child care space availability across different wards. This suggests a relatively consistent distribution of child care spaces throughout the city's wards, with no single ward showing a marked difference from the others in terms of space availability. In contrast, the influence of subsidies on child care capacity is pronounced and statistically meaningful, as reflected by a substantial F-statistic of 50.217 and an exceedingly low P-value nearing zero. This emphasizes the crucial impact that subsidy offerings at child care centers have on the availability of child care spaces. Furthermore, the lack of a significant interaction effect between wards and subsidies, evidenced by a P-value of approximately 0.083, suggests that the influence of subsidies on the number of spaces is uniform, regardless of ward location.

5. Post-Hoc Analysis: Delving Deeper into Subsidy Effects



With confidence intervals that do not cross the line of no effect, the graphs show that the differences in child care space availability between subsidized (marked "Y") and non-subsidized (marked "N") centers are both statistically significant and practically important. The results corroborate the impact of subsidies on capacity expansion, since centers with subsidies consistently showed high space availability in comparison to those without such financial support. These results underscore the necessity of subsidy programs in enhancing child care accessibility across the city, reflecting a key policy instrument that can be used to mitigate disparities and ensure a more uniform distribution of child care spaces among Toronto's diverse family demographics. The consistent findings across wards suggest that the provision of subsidies serves as an effective equalizer in the distribution of child care services throughout Toronto, transcending geographical boundaries.

6. Conclusion

After conducting a thorough statistical analysis with one-way and two-way ANOVA, we found that although the city's wards' capacities for infant care do not differ significantly, the availability of spaces for toddlers does, indicating areas where resource allocation could be optimized. This study offers a comprehensive picture of Toronto's child care capacity, highlighting the need for targeted policy initiatives to close the gaps in toddler care availability. The consistent impact of subsidy programs highlights how important it is for them to guarantee fair access to child care services. By making decisions based on these findings, policymakers can support family welfare and foster the city's younger population, both of which are critical to creating a more vibrant and just society for all citizens.