Rosa Lee

Assignment 1

Introduction:

The following data analysis provides a comprehensive examination of shelter usage trends in Toronto for the year 2021. The data analysis leverages statistical and visual exploratory data analysis (EDA) methodologies. The primary statistical tool employed in this study is the t-test, a method used to determine if there are significant differences between the means of two groups, which can be related to various categorical variables within the shelter data, such as shelter type and sector.

Data was pulled from the INF2178_A1 file, that tracked the daily occupancy and capacity of Toronto shelters for the year 2021. There are 15 categories with 50945 lines of data input.

Calculating the Shelter Occupancy Rate:

Occupancy Rate was calculated by:

'Bed_Occupancy_Rate= OCCUPIED_BEDS / CAPACITY_ACTUAL_BED

- **Continuous Variable for Comparison:** Creating a continuous variable (bed occupancy rate) was necessary for comparing and analyzing trends across different shelter types and sectors. This allowed for a quantitative comparison of occupancy rates, enabling the identification of statistically significant differences through t-tests.
- **Standardized Metric:** Using occupancy rates as a continuous variable provided a standardized metric, allowing for a more meaningful and consistent comparison of data. This helped in removing biases that may arise from absolute bed counts, as shelters vary in their capacities.

The calculation of shelter program occupancy rates would allow comparison and analysis of t-test results of Service Type and Sector in a standardized way, providing a clearer picture of the trends in shelter usage across Toronto.

Quantitative Analysis (T-tests):

The t-test results provide valuable insights into the differences in bed occupancy rates among different shelter types and sectors in Toronto. Providing potential valuable insights into potential trends and disparities in shelter occupancy rates in 2021 Toronto. The t-test results provide valuable insights into the differences in bed occupancy rates across different shelter sectors and types.

<u>Sectors:</u>

1. Mixed Adult:

- T-statistic: 20.56 | P-value: 1.24e-88

- The high t-statistic and extremely low p-value suggest a significant difference in bed occupancy rates between Emergency and Transitional programs within the Mixed Adult sector.

2. Men:

- T-statistic: 23.35 | P-value: 3.59e-112

- Similar to the Mixed Adult sector, there is a highly significant difference in bed occupancy rates between Emergency and Transitional programs for Men.

3. Women:

- T-statistic: 20.72 | P-value: 1.76e-88

- The results for Women also indicates a difference in bed occupancy rates between Emergency and Transitional programs.

4. Youth:

- T-statistic: 11.57 | P-value: 1.57e-30

- Youth shows a significant difference in bed occupancy rates between Emergency and Transitional programs, though the effect size is slightly smaller compared to other sectors.

<u>Shelter Types:</u>

1. Shelter:

- T-statistic: -0.50 | P-value: 0.61

- The p-value is relatively high, suggesting that there is no significant difference in bed occupancy rates between different programs within the Shelter.

2. 24-Hour Respite Site:

- T-statistic: 42.34 | P-value: 0.0

- The low p-value indicates a highly significant difference in bed occupancy rates within the 24-Hour Respite Site.

3. Warming Centre:

- T-statistic: -4.43 | P-value: 1.60e-05

- The negative t-statistic suggests a significant difference, with lower bed occupancy rates in Warming Centres compared to the overall average.

4. 24-Hour Women's Drop-in:

- T-statistic: -16.86 | P-value: 3.32e-51

- A highly significant difference is observed, with lower bed occupancy rates in 24-Hour Women's Drop-in programs compared to the overall average.

5. Motel/Hotel Shelter:

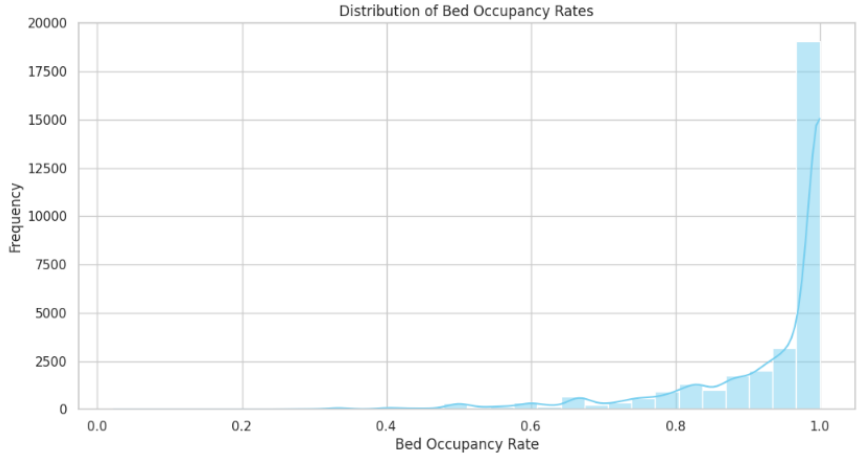
- T-statistic: -17.40 | P-value: 4.48e-40

- Like the Women's Drop-in, Motel/Hotel Shelters show significantly lower bed occupancy rates compared to the overall average.

Overall, the t-tests suggest, Emergency and Transitional programs within the Men, Women, Mixed Adult, and Youth sectors exhibit significant differences in bed occupancy rates. Also, Shelter types such as '24-Hour Respite Site', '24-Hour Women's Drop-in', and 'Motel/Hotel Shelter' also show substantial variations in bed occupancy rates compared to the overall average.

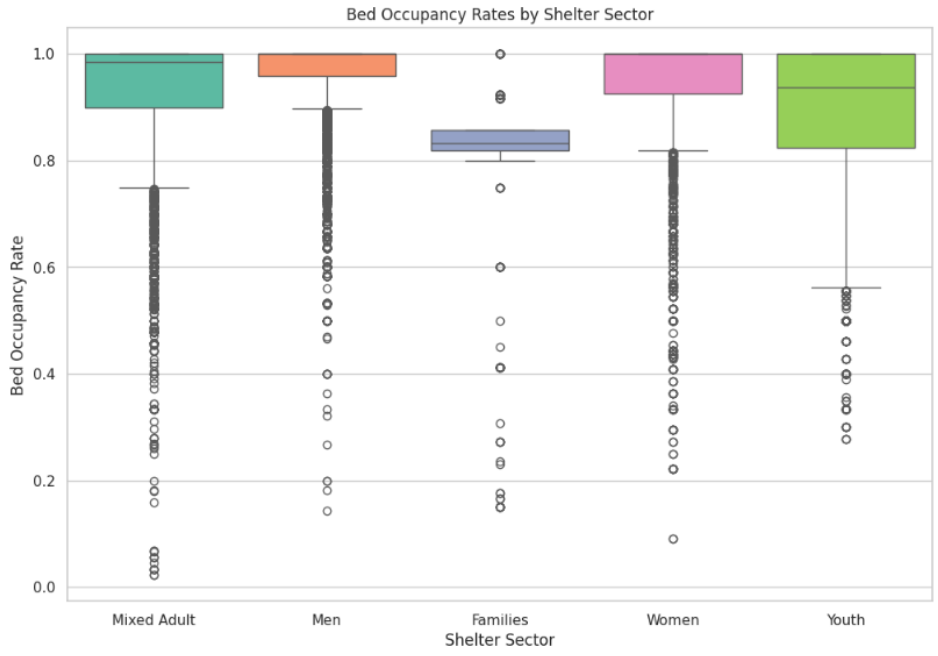
Exploratory Data Analysis (EDA):

Bed Occupancy Rate Distribution:



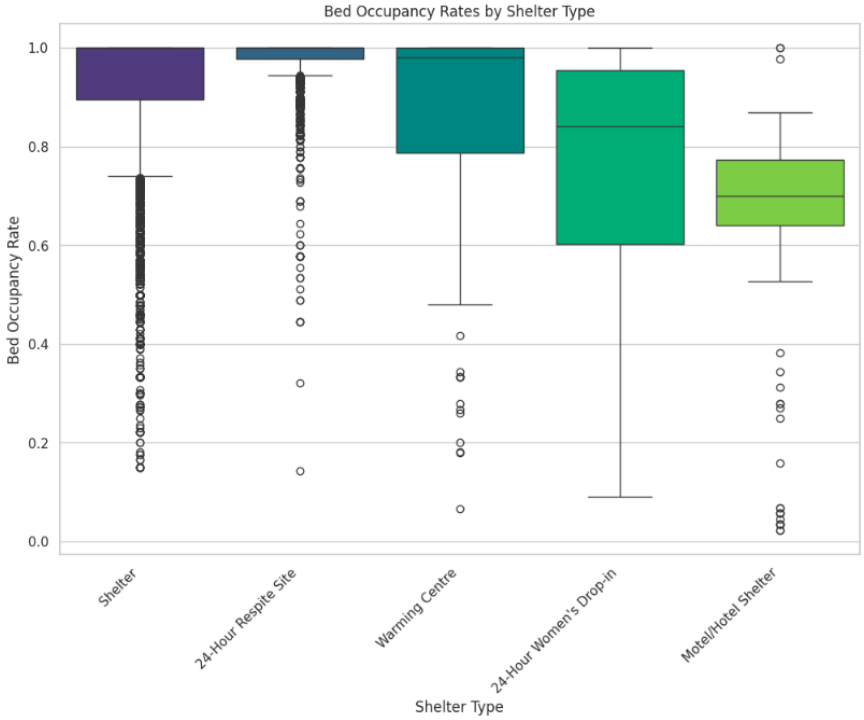
The plot displays the distribution of bed occupancy rates across all shelter programs. The histogram displays the distribution of bed occupancy rates across shelter programs. Most programs have high occupancy rates, with a significant spike at 100%, indicating many shelters operate at full capacity.

Bed Occupancy Rates by Shelter Sector:



The boxplot compares bed occupancy rates among different shelter sectors. This visualization helps identify variations in bed occupancy rates across Men, Women, Mixed Adult, and Youth sectors. Families and women's sectors have higher median occupancy rates, close to full capacity, while men's and mixed adult's sectors have a broader range of occupancy rates with lower medians, suggesting more variability in how full these shelters are.

Bed Occupancy Rates by Shelter Type:



The boxplot shows bed occupancy rates by shelter type (Shelter, 24-hour Respite Site, Warming Centre, 24-Hour Women's Drop-in, Motel/Hotel Shelter). The 24-hour respite sites and warming centres have higher median occupancy rates, indicating they are often near capacity. The motel/hotel shelters have the broadest range of occupancy, which suggests a greater variability in how full these shelters are.

The graphs collectively highlight a shelter system that is frequently operating at or near capacity. This full operational capacity is especially prevalent in sectors serving families and women. The variability in some sectors and shelter types could point to disparities in demand or available resources. From the narrative analysis made from the EDA, please see below for a list of potential research questions to explore:

- Question: What factors contribute to the consistently high demand in the Families and Women's shelter sectors?
    - Requires further analysis: Explore demographic trends, seasonal variations, or specific program features that might explain the observed patterns.
- Question: What influences the variability in bed occupancy rates in the Men's and Mixed Adult shelter sectors?
    - Requires further analysis: Investigate external factors such as seasonal trends, outreach efforts, or program design that could contribute to the observed fluctuations.

To explore these potential research questions, additional data collection and analyses could be helpful. That could be detailed demographic information, temporal trends, external factors affecting shelter demand, and a deeper examination of specific program characteristics.

Ultimately, the data suggests a Toronto shelter system with consistent operational challenges, especially in sectors serving Families and Women. While high occupancy rates indicate robust demand, the variability observed in certain sectors and shelter types raises questions about resource allocation and program effectiveness.