INF2178 A1
Chloe Jinxuan Li #1010033321

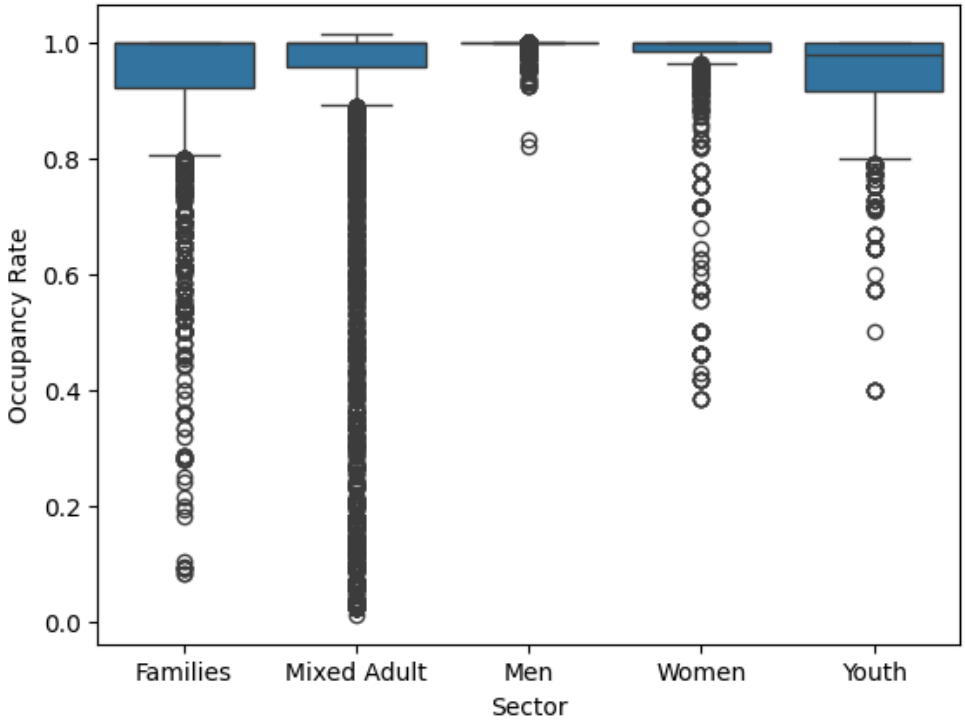
Shelter Data Analysis in Toronto

This analysis aimed to uncover the trends and patterns in the occupancy and utilization of Toronto's shelter system in 2021. This exploration was particularly focused on understanding the dynamics across different sectors and program areas, as well as assessing the impact of program models on occupancy rates. This narrative recounts the steps taken and the insights gleaned from the data visualizations and statistical analyses.

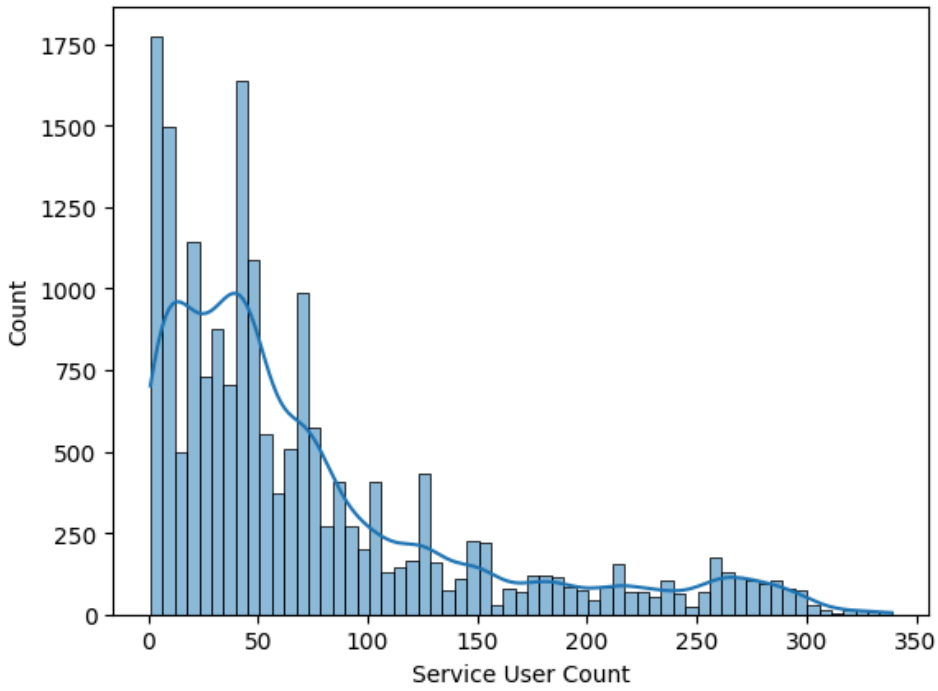
Exploratory Data Analysis

Began by calculating the room occupancy rate. This calculation quantifies how many rooms were occupied in relation to the total available rooms across the shelters.

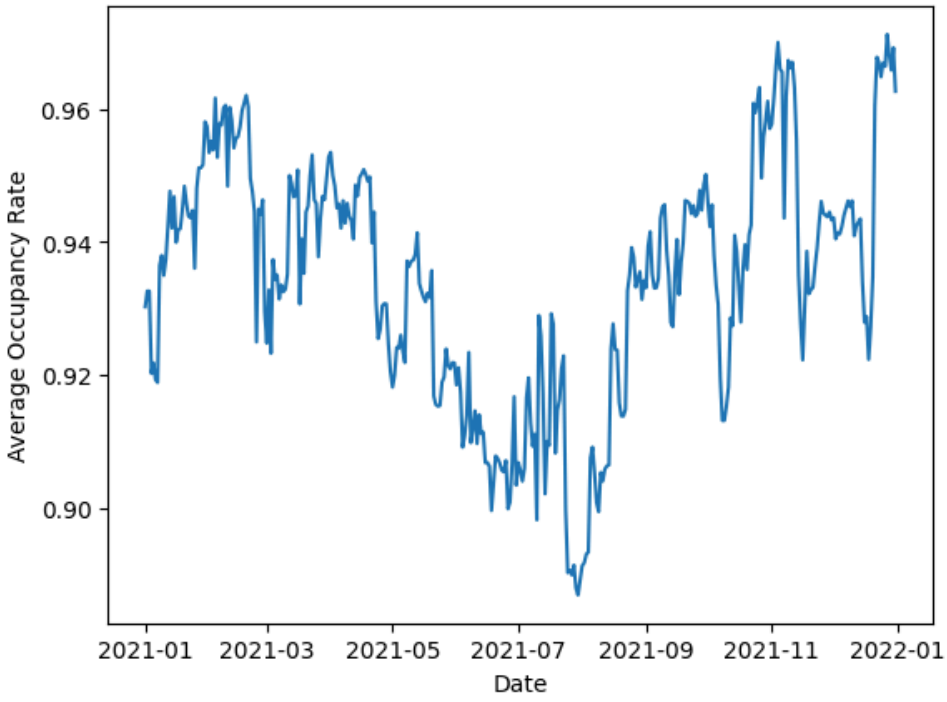
- Box Plot of Occupancy Rate by Sector:
  The boxplot of occupancy rates by sector, offering a visual symphony of medians, quartiles, and outliers. The Families and Youth sectors showed relatively high median occupancy rates, suggesting a strong and consistent demand for shelters in these groups. On the other hand, the Mixed Adult sector showed lower median occupancy but a wide range of rates, indicating a potential inconsistency in shelter usage.



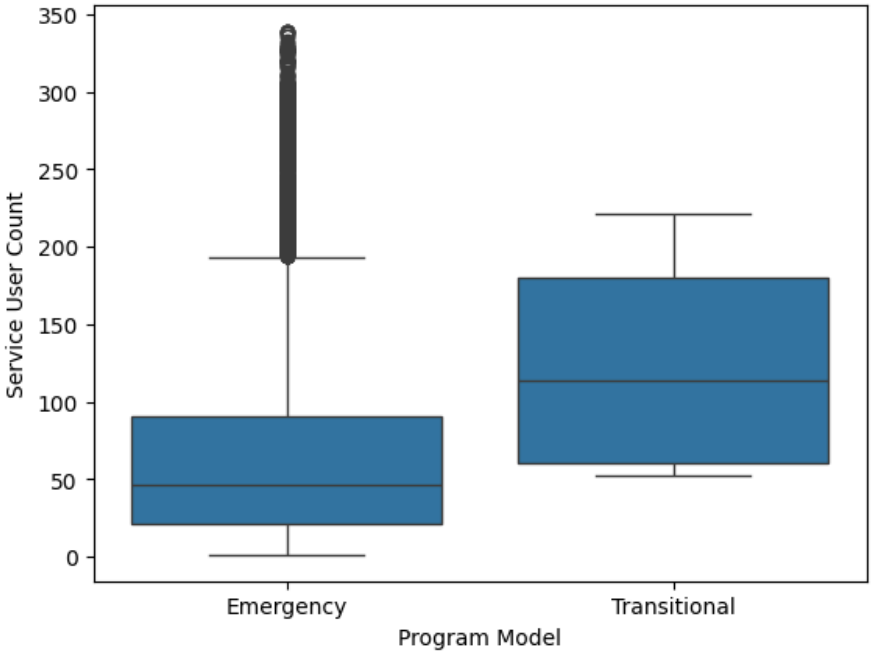
- Histogram of Service User Count:
  This is a histogram that visualizes the service user count. The clear skew towards lower counts suggested that most shelters served a smaller population, while a handful catered to larger groups—a potential indication of resource concentration in specific locations.



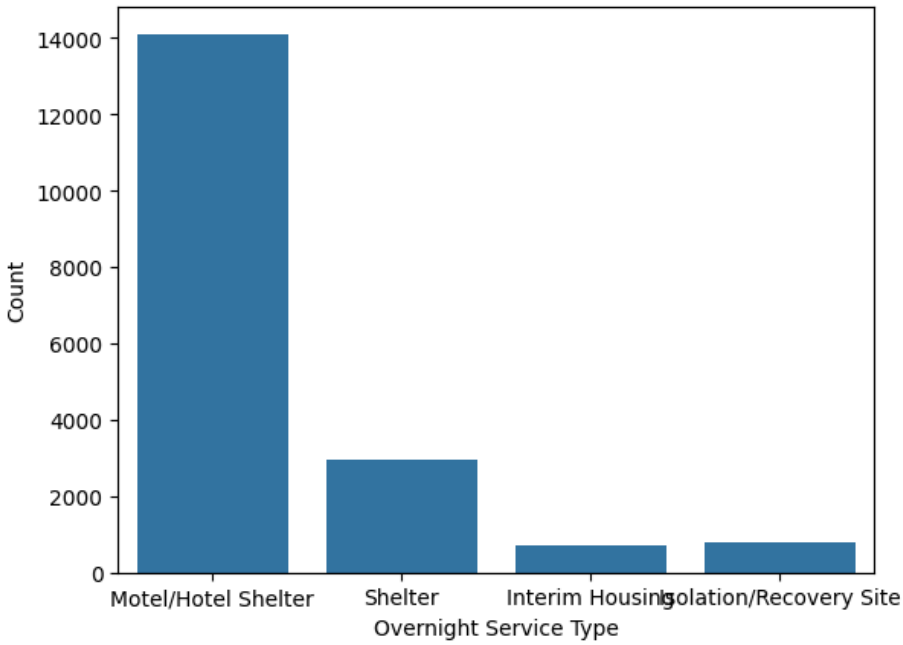
- Trend of Average Occupancy Rate Over Time:
  The average occupancy rate over time revealed through a line plot offered a narrative of its own. Significant variance was observed over the months, with distinct troughs and peaks that could be indicative of seasonal effects or policy-driven changes affecting the shelter system's usage.



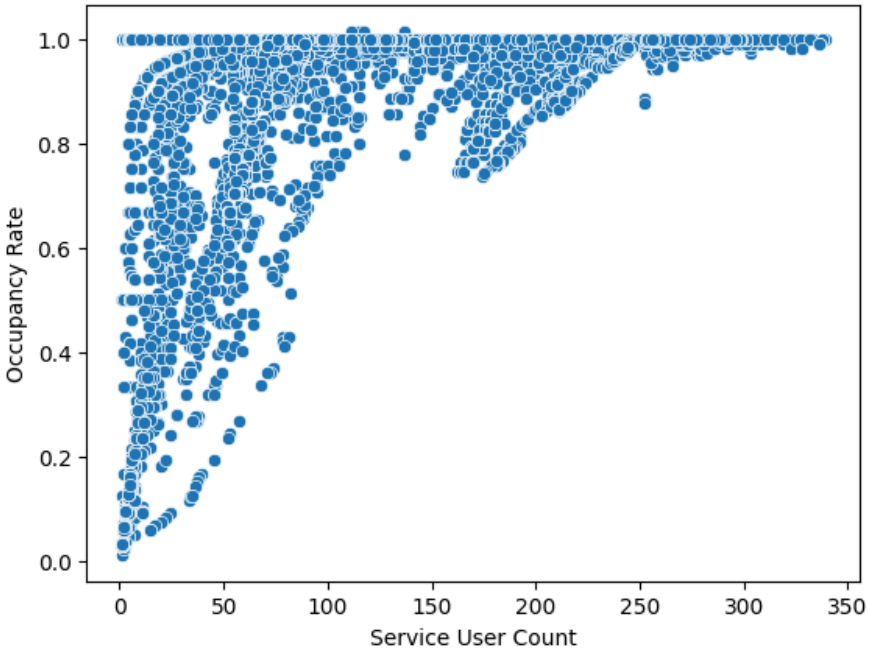
- Box Plot of Service User Count by Program Model:
  The boxplot shows differences between Emergency and Transitional programs, which displayed a wide interquartile range for Emergency programs, with outliers suggesting sporadic surges in usage, while Transitional programs showed tighter clustering, hinting at more predictable service user counts.



- Bar Plot of Record Count by Overnight Service Type:
  In the bar chart of overnight service types, the predominance of the Motel/Hotel Shelter category was striking. This finding revealed a heavy reliance on such facilities for overnight services, potentially reflecting a trend or a gap in the traditional shelter infrastructure.



- Scatter Plot of Service User Count vs. Occupancy Rate:
  The scatter plot shows a pattern where occupancy rates tend to be lower when the service user count is small. As the service user count increases, the occupancy rate appears to consolidate towards higher values. This indicates that facilities with more users generally tend to operate at or near full capacity.



Quantitative Analysis
For the overnight service types, the t-tests revealed statistically significant differences in occupancy rates for each category when compared against the overall rates within the dataset:

- Motel/Hotel Shelters showed a significantly higher occupancy rate with a T-value of 27.89 and a P-value of 7.007613270131504e-169 (near 0), indicating a much higher utilization compared to the collective average of all shelter types.

- Standard Shelters had a lower occupancy rate than average, as indicated by a negative T-value of -9.95 and a P-value of 4.578136366982899e-23 (near 0), suggesting less utilization relative to the overall shelter data.

- Interim Housing demonstrated a higher occupancy rate, but the difference was less pronounced, with a T-value of 2.70 and a P-value of 0.007, still below the 0.05 threshold for statistical significance.

- Isolation/Recovery Sites had a markedly lower occupancy rate compared to the overall average, with an extremely negative T-value of -65.48 and a P-value of 0.0, emphasizing a significant underutilization or possibly a larger capacity to handle surges in need.

When examining the room occupancy rates by program area, the results varied:

- The COVID-19 Response program area showed no significant difference from the overall occupancy rate, with a T-value of 0.79 and a P-value of 0.43, suggesting that the COVID-19 response programs' occupancy rates align with the average shelter occupancy.

- The Temporary Refugee Response program area had a significantly higher occupancy rate, with a T-value of 14.15 and a P-value of 1.7618563911229607e-41 (near 0), indicating these programs are operating at higher capacity than the average.

- The Base Shelter and Overnight Services System showed a significantly lower occupancy rate than the overall average, with a T-value of -4.74 and a P-value of 2.1635149419946284e-06 (near 0), pointing to lower utilization or possibly a deliberate policy of maintaining available capacity.

The t-test results provide a clear indication that there are notable differences in how different shelter types and program areas are utilized. Specifically, the data suggests that Motel/Hotel Shelters and Temporary Refugee Responses are experiencing high usage, whereas standard shelters and Base Shelter and Overnight Services Systems are less utilized in comparison. The Isolation/Recovery Sites' significant underutilization might reflect a capacity for emergency situations like health crises, which could see sudden increases in demand.