

Toronto Shelter Occupancy Analysis Overview

Introduction

This document provides a comprehensive overview of the data analysis undertaken on the occupancy rates within Toronto's shelter system. It details the procedures for loading, cleaning, visualizing, and performing statistical analysis on the data, and it outlines the key findings derived from the study.

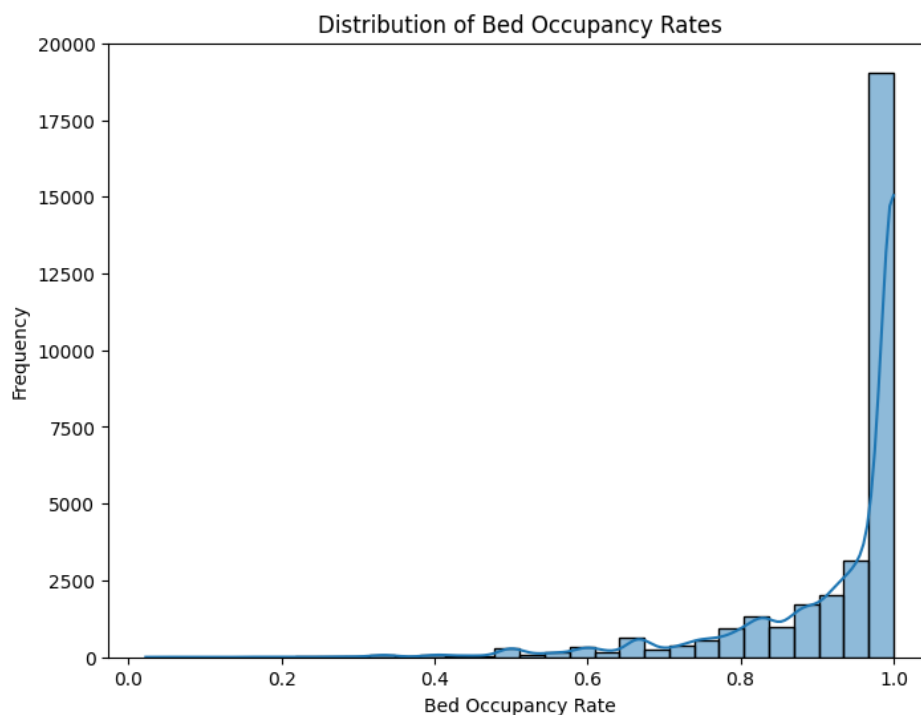
Data Loading and Cleaning

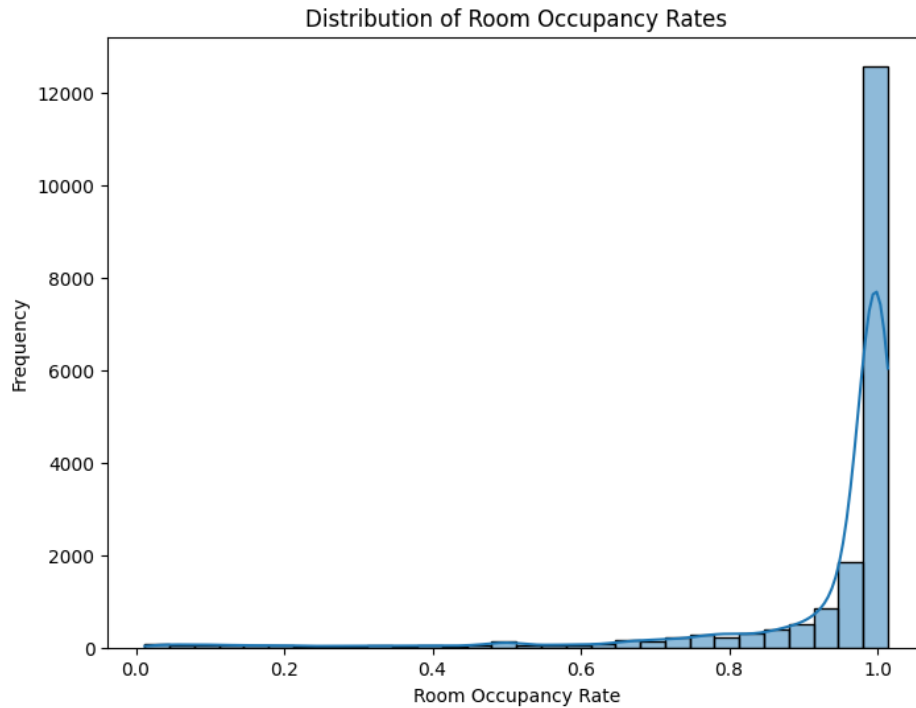
The data was loaded from an Excel file using Pandas. I calculated bed and room occupancy rates and introduced a 'TIME_PERIOD' variable. Missing values were addressed, and categorical data types were set appropriately.

Visualization

- **Histograms and density plots for occupancy rates to examine their distribution**

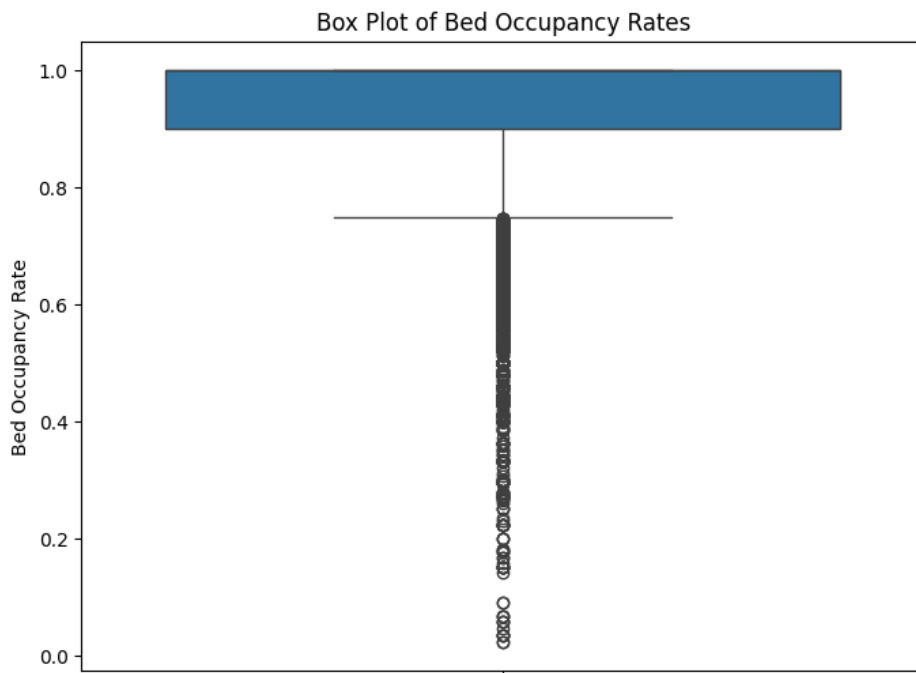
The distribution of bed and room occupancy rates, as shown in the histograms, indicated a skewed distribution, suggesting a majority of shelters operate below capacity. This raises questions about resource allocation efficiency.

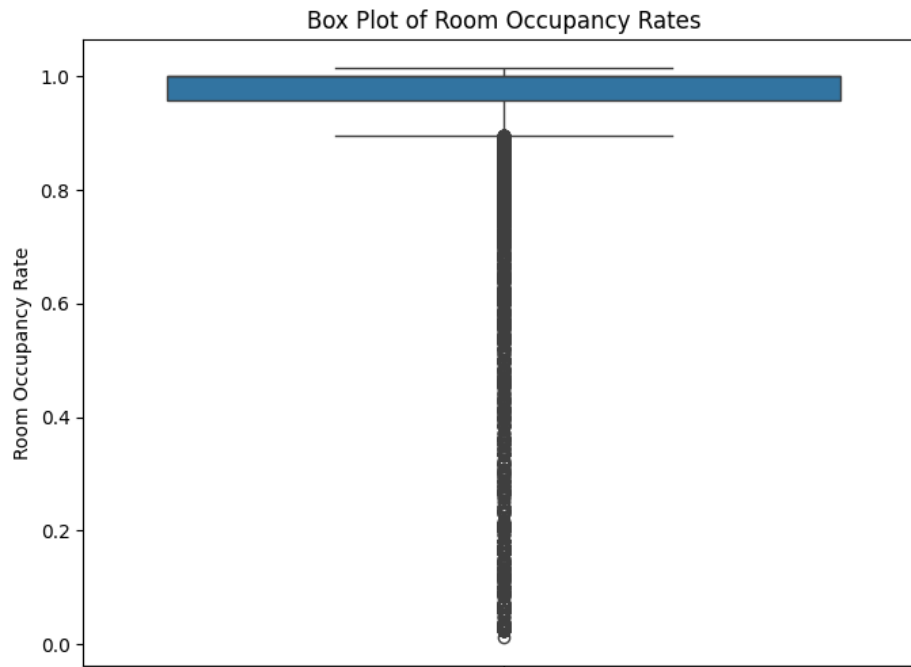




- **Box plots to visualize the spread and identify outliers**

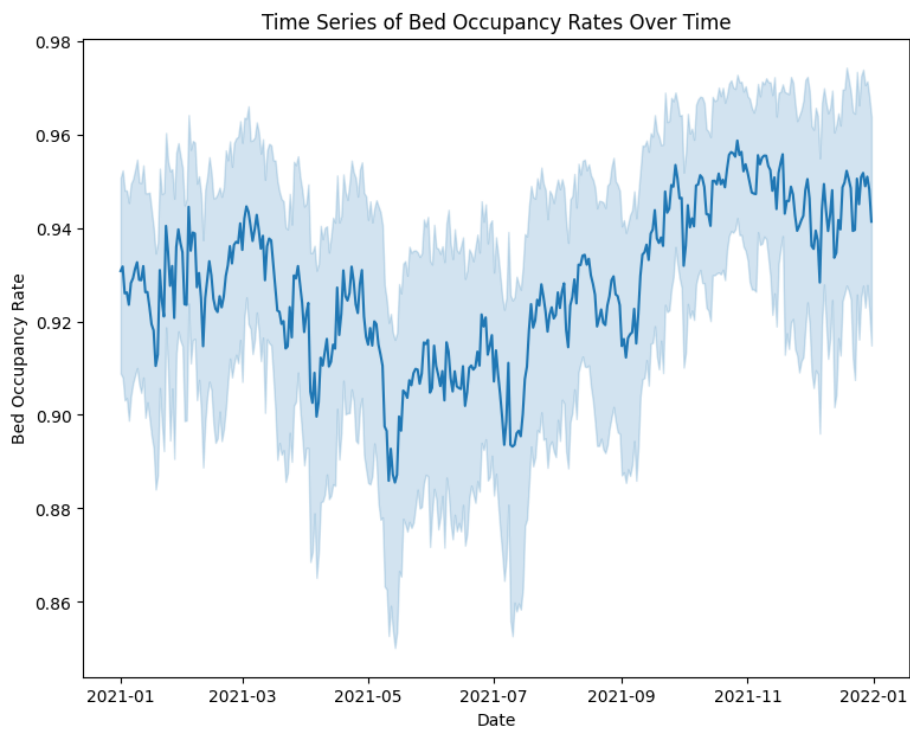
The box plots revealed a number of outliers in both bed and room occupancy rates, indicating that some shelters consistently operate at or above capacity, hinting at potential overutilization or resource constraints.





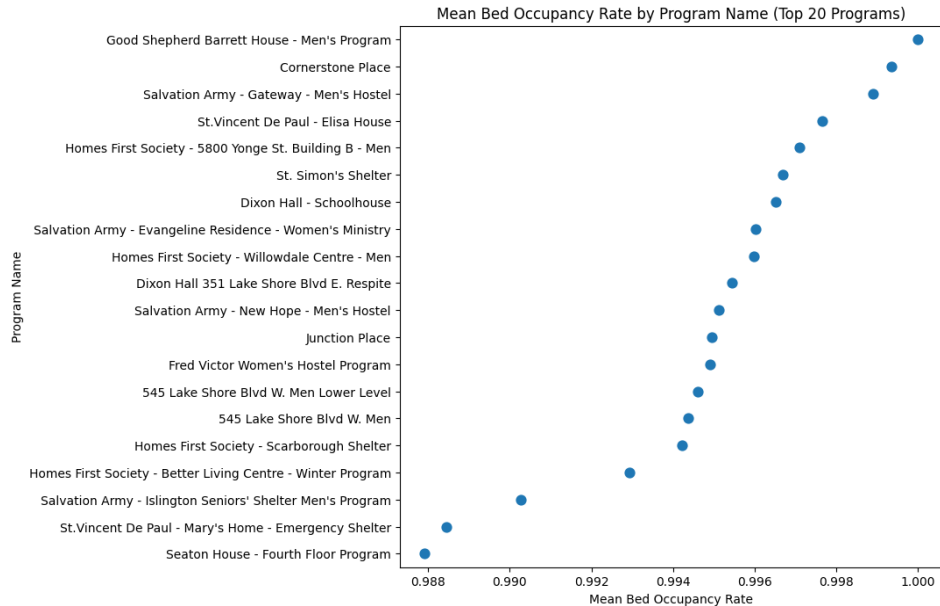
- **A time series line plot to observe trends over time**

The time series plot of bed occupancy rates over time showed seasonal trends, with peaks and troughs corresponding to winter and summer months, respectively



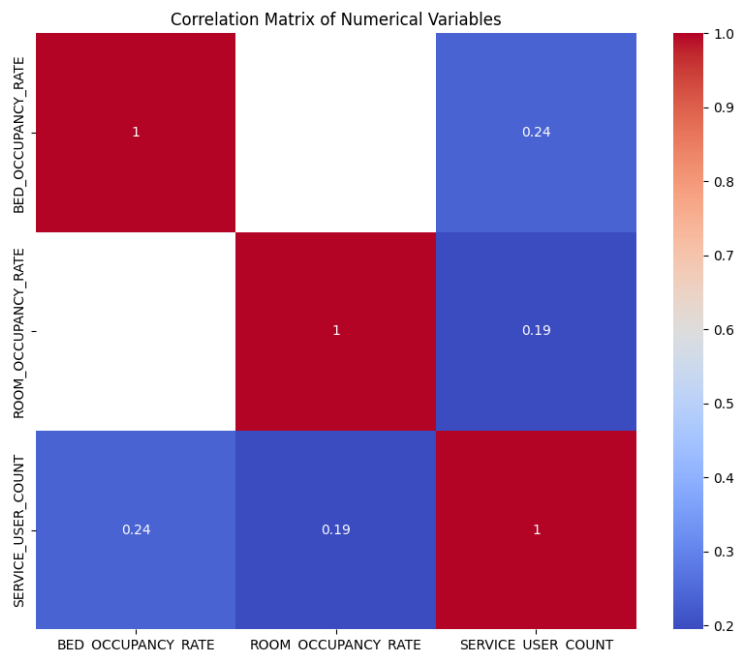
- **A dot plot to compare the mean occupancy rate among the top 20 programs**

The dot plot highlighted that the top programs, in terms of bed occupancy rate, varied greatly, which may reflect program-specific factors or local demand for services.



- **A correlation matrix to examine relationships between numerical variables**

The correlation matrix revealed a strong positive relationship between bed and room occupancy rates, while the correlation with service user count was less pronounced, suggesting other factors may influence the number of service users.



Statistical Analysis :

Our analysis involved conducting t-tests to evaluate differences in 'SERVICE_USER_COUNT' across 'CAPACITY_TYPE' and 'PROGRAM_MODEL' categories.

The statistical analysis reveals significant differences in service user counts between different capacity types and program models within the shelter system. The t-test comparing Room Based and Bed Based Capacities showed an extremely high statistic value, indicating a vast difference in occupancy rates, ***with Room Based Capacities being significantly higher on average.*** This suggests that room-based facilities might be dealing with a higher demand or a different demographic compared to bed-based ones.

Similarly, the comparison between Emergency and Transitional program models through t-tests showed a significant difference in service user counts, with Emergency programs having a higher average count. ***This could indicate that emergency shelters are in higher demand or that they have a higher turnover rate, reflecting more immediate or short-term needs of the population they serve.***

These findings suggest that resource allocation, program funding, and capacity planning need to be carefully considered to address the apparent disparities in service use and to ensure that shelters meet the varying needs of their service users efficiently.

Key Findings

1. The occupancy rates show that the majority of shelters operate below capacity, indicating potential inefficiencies in resource use and highlighting the need for targeted capacity planning.
2. The t-tests revealed significant differences between program models, suggesting that some models are more effective or more in demand, which warrants further investigation into program attributes.
3. The correlation matrix indicates a strong positive relationship between bed and room occupancy rates, which confirms that these two metrics are closely linked, while the weaker correlation with service user count suggests additional factors are at play in influencing service usage.

Conclusion :

The statistical analysis of Toronto's shelter system reveals distinct discrepancies in occupancy rates between different shelter capacities and program models. Room Based Capacity shelters are significantly more utilized than Bed Based Capacity shelters, indicating a pressing need for strategic reallocation to improve occupancy efficiency. The overwhelming demand for

Emergency programs compared to Transitional ones suggests a critical review of funding and support structures is necessary to address the immediate needs of the homeless population in Toronto. Moreover, the pronounced seasonality in shelter use requires targeted policy interventions to anticipate and manage the fluctuating demands throughout the year, ensuring no individual is without shelter during peak periods.

Usage

To replicate this analysis:

1. Run **load_and_clean_data(file_path)** with the path to your Excel data file.
2. Use **visualize_data(shelter_data)** to generate the plots.
3. Execute **compare_groups_and_summarize(data, 'COLUMN_NAME', 'NUMERIC_COLUMN')** to perform t-tests and get summary statistics.