

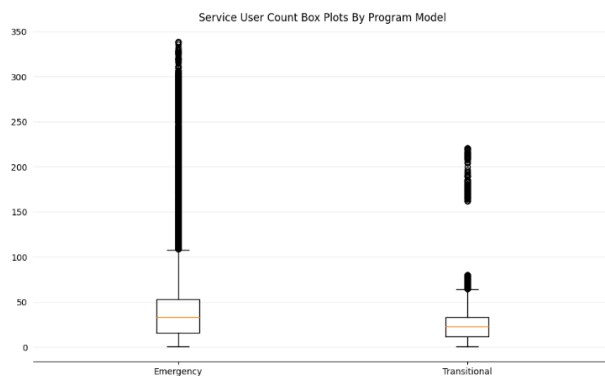
To analyze shelter usage based on the program model: 'Emergency' & 'Transitional'.

1. Null hypothesis: There is no significant difference in shelter usage between the 'Emergency' and 'Transitional' program models.
2. Alternative hypothesis: There is a significant difference in shelter usage between the 'Emergency' and 'Transitional' program models.
3. The extremely low p-values obtained from the Welch's t-test and the Two-sample t-test provide strong evidence to reject the null hypothesis in favor of the alternative hypothesis, supporting the conclusion that there is a significant difference in shelter usage between the 'Emergency' and 'Transitional' program models.

Welch's t-test result : 38.85174699254652
p-value: 0.0

Two-sample t-test result : 29.937570467283667
p-value: 3.1720139638162956e-195

4. The emergency shelter program has a higher overall usage, with a higher mean and median compared to the transitional program.
5. Interpretation of the boxplot:



- Mean and Central Tendency: The relatively low median in both the "Emergency" and "Transitional" categories implies that a substantial portion of the data lies towards

the lower end of the distribution. This suggests that most observations exhibit lower shelter usage, contributing to the central tendency.

- **Outliers and Skewed Distribution:** The presence of numerous outliers in the "Emergency" plot, extending significantly beyond the upper whisker, indicates the existence of sporadic instances with exceptionally high shelter usage. This suggests a positively skewed distribution, where most data is concentrated towards lower values, but there are instances of extreme values pulling the mean and upper quartile upwards.
- **IQR:** The larger IQR for the "Emergency" program model compared to "Transitional" implies greater variability in shelter usage within the emergency context. This increased spread could be indicative of diverse and potentially dynamic factors influencing emergency shelter utilization, contributing to a wider range of user counts.

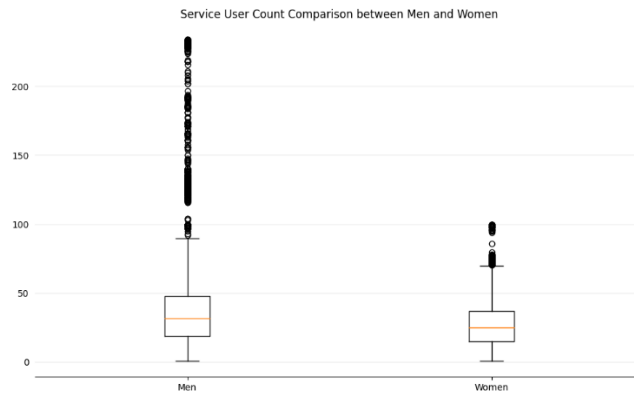
To analyze the service user count between the 'Men' and 'Women' sectors.

1. **Null Hypothesis:** There is no significant difference in the user counts between the Men and Women sectors.
2. **Alternative Hypothesis:** There is a significant difference in the user counts between the Men and Women sectors.
3. There is a statistically significant difference in the user counts between the Men and Women sectors cause the extremely low p-values obtained from both Welch's t-test and the Two-sample t-test.

Welch's t-statistic = 30.50421164151915
p-value = 1.0532742925183108e-199

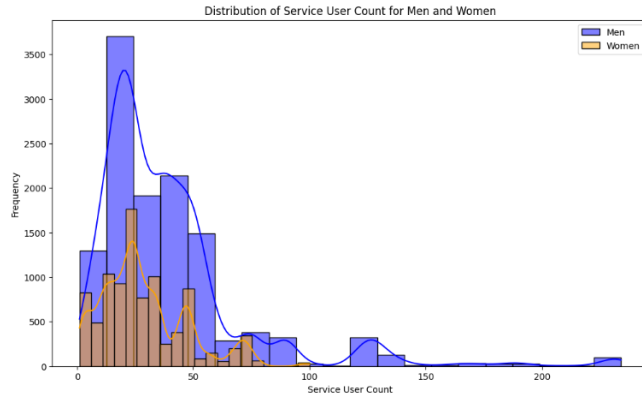
Two-sample t-test result: 28.24955042305982
p-value: 2.022346382679794e-172

4. Intepretation of the boxplot:



- Mean and Central Tendency: The median for men is positioned slightly above 50, For women, the median is slightly lower but still close to 50. This implies a central tendency around the midpoint for both genders.
- Outliers: The presence of numerous outliers for men indicates instances of significantly higher service user counts. In contrast, there are fewer outliers for women, suggesting a lower frequency of exceptionally high service user counts. This observation highlights greater variability and potential extremes in the distribution of service user counts for men.
- IQR: The IQR is broader for men than for women. This indicates greater variability in service user counts among men, with a wider range of values between the first and third quartiles. In contrast, women exhibit a narrower IQR, suggesting more consistency in service user counts.
- the central tendency for both men and women's service user counts is near 50. This alignment with the median indicates that, on average, the middle values for service user counts are comparable between genders.

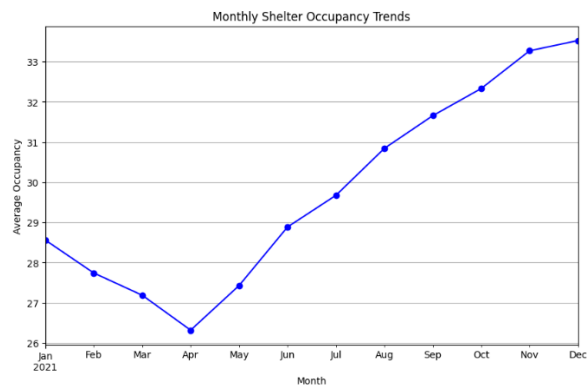
5. Intepretation of the histogram:



- Both distributions have a right-skewed shape, indicating that there are higher frequencies of lower service user counts and fewer frequencies as the count increases.

To analyze shelter occupancy trends on a monthly basis over time.

1. Interpretation of the line chart:

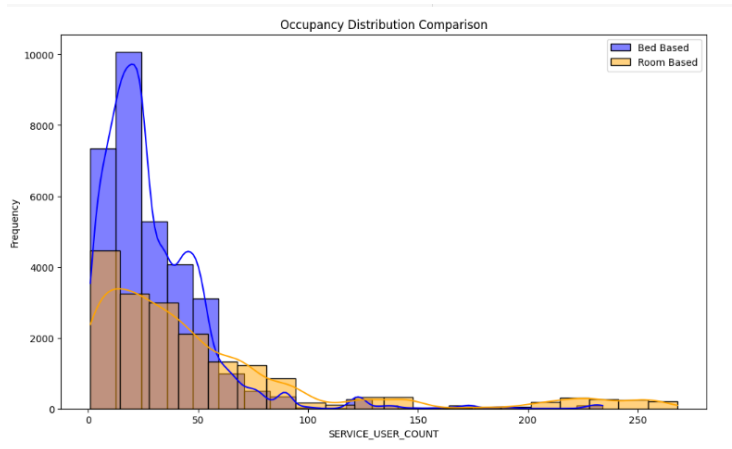


- From January to December, the occupancy of shelters experienced a declining trend from January to April, reaching its lowest point in April. Subsequently, from April to December, there was a significant surge in occupancy.

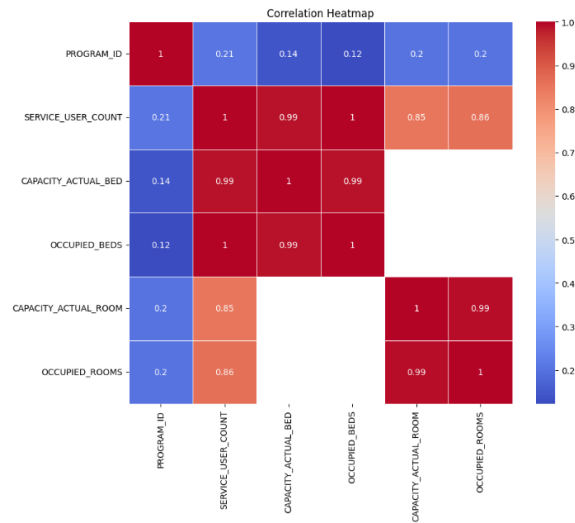
To analyze occupancy between different capacity type: 'Room Based' and 'Bed Based'.

- Null hypothesis: There is no significant difference in the mean values of capacity between "Room Based Capacity" and "Bed Based Capacity."

2. Alternative hypothesis: There is a significant difference in the mean values of capacity between "Room Based Capacity" and "Bed Based Capacity."
3. The extremely low p-values obtained from both Welch's t-test and the Two-sample t-test provide strong evidence against the null hypothesis, indicating that there is a statistically significant difference in the mean values of capacity between "Room Based Capacity" and "Bed Based Capacity."
4. Interpretation of the histogram:



- Both distributions have a right-skewed shape, The bed-based distribution appears to have higher frequencies for lower service user counts compared to the room-based distribution. Additionally, the room-based distribution seems to have a longer tail, suggesting that there are instances with higher service user counts, although they are less frequent.
5. Interpretation of the heatmap:



- 'SERVICE_USER_COUNT' and 'OCCUPIED_ROOMS': These two variables show a high correlation of 0.86, suggesting a strong relationship between the number of service users and the number of occupied rooms.
- 'SERVICE_USER_COUNT', 'CAPACITY_ACTUAL_BED', and 'OCCUPIED_BEDS': These three variables have very high correlations with each other, as indicated by the red squares with values of 0.99. This suggests that as the number of service users increases, the actual bed capacity and the number of occupied beds tend to increase proportionally.