

For this database I would like to look at the utilization of individual shelters because emergency shelters have a high turnover rate and need to wait in line every day, so there is a high probability that homeless people who miss a good spot in the morning will choose to go to another shelter.

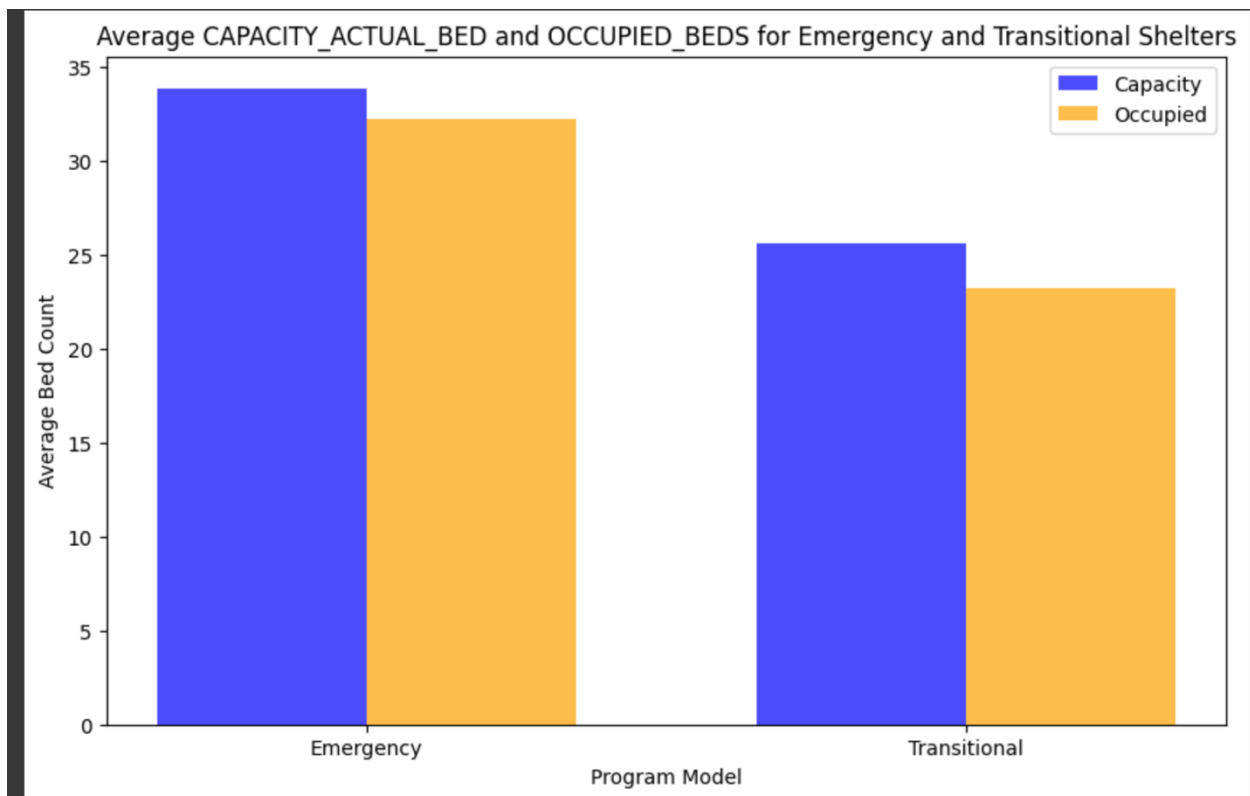
So it would be necessary to see if emergency shelters have a lower bed utilization rate than Transitional shelters.

For the Service user count

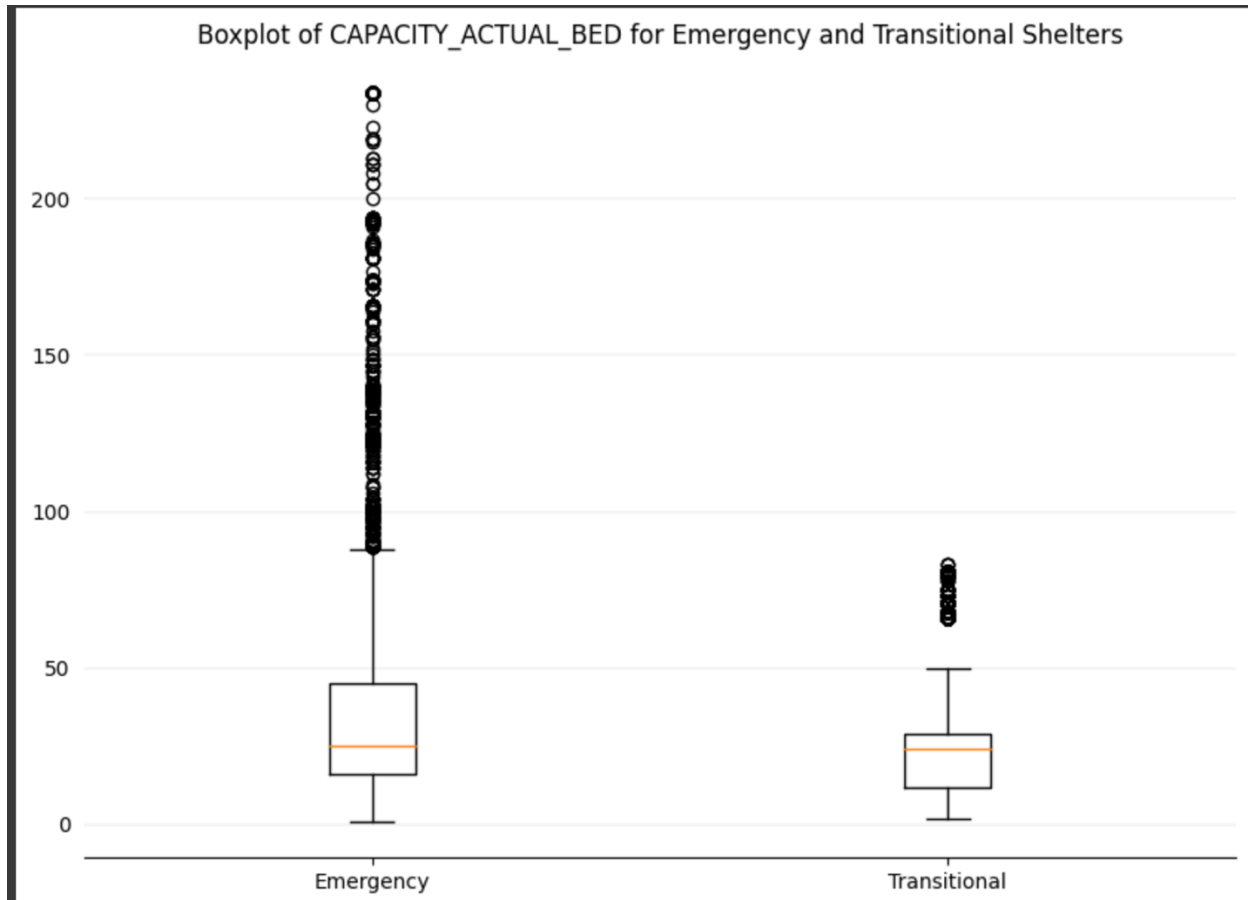
The average number of users of shelters is 45.73, while the median number of users is only 28, indicating that small shelters still predominate and that there are a small number of shelters with a particularly high capacity. We can see the max number is 339 which should be counted as extreme value.

And for Occupied beds and CAPACITY_ACTUAL_BED

The values for actual and occupied beds in the shelters are very close to each other, with an average difference of less than 2, and each shelter appears to be close to full capacity. Excluding the possibility that the homeless left on their own in the middle of the night, it may be that the shelters had to reserve a bed for a homeless person who might freeze to death during the night. This suggests that more shelters may need to be built to meet the large demand.



For this diagram we can see that for both Transitional and Emergency Shelters, the utilization rates are all very high that almost no free bed



We can clearly see there are more and large extreme values of the emergency shelter. Maybe it is because Transitional shelters' homeless are more stable and not very like to have large amount of people suddenly to live in. But emergency shelters are different, the turnover of people is very high, there may be one day suddenly a lot of homeless people need to move in.

Two-sample t-test for CAPACITY_ACTUAL_BED:

T-statistic: 24.425704744840953

P-value: 1.3885176166874997e-130

Two-sample t-test for OCCUPIED_BEDS:

T-statistic: 27.414488367392458

P-value: 1.3637601442103477e-163

In both cases, the small p-values strongly indicate that there is a significant difference between the means of "CAPACITY_ACTUAL_BED" and "OCCUPIED_BEDS" for the Emergency group and

the Transitional group. The results suggest that the observed differences are not likely due to random chance, supporting the hypothesis that there is a true difference in means.

Now lets see if the utilization rate are different for two program method

Two-sample t-test for utilization_rate:

T-statistic: 38.780694714817365

P-value: 0.0

The mean of emergency_utilization_rate is 0.943501 with standard deviation 0.115901

The mean of transitional_utilization_rate is 0.885194 with standard deviation 0.129907

So the emergency shelter utilization rate is actually better than transitional shelter and observed differences is significant.