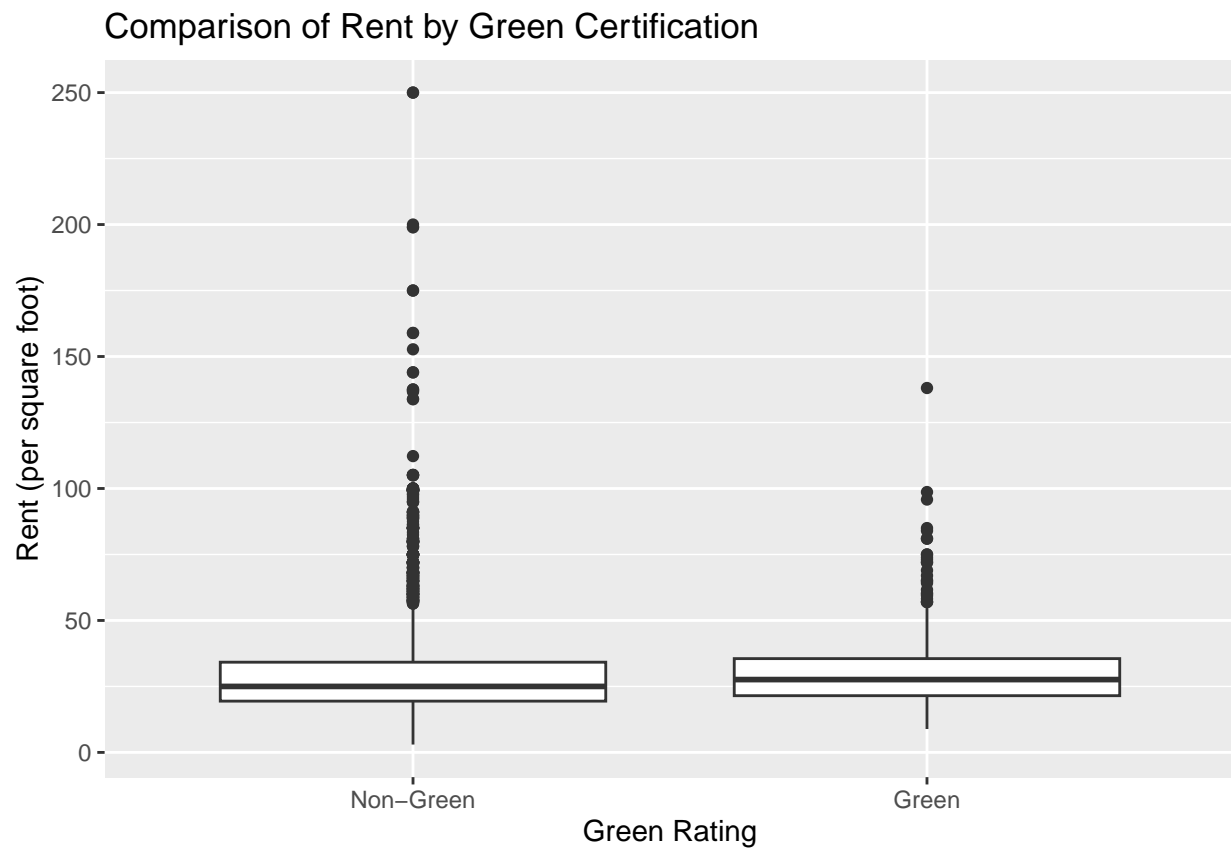


Visual story telling part 1: green buildings

1. Data Cleaning and Initial Exploration

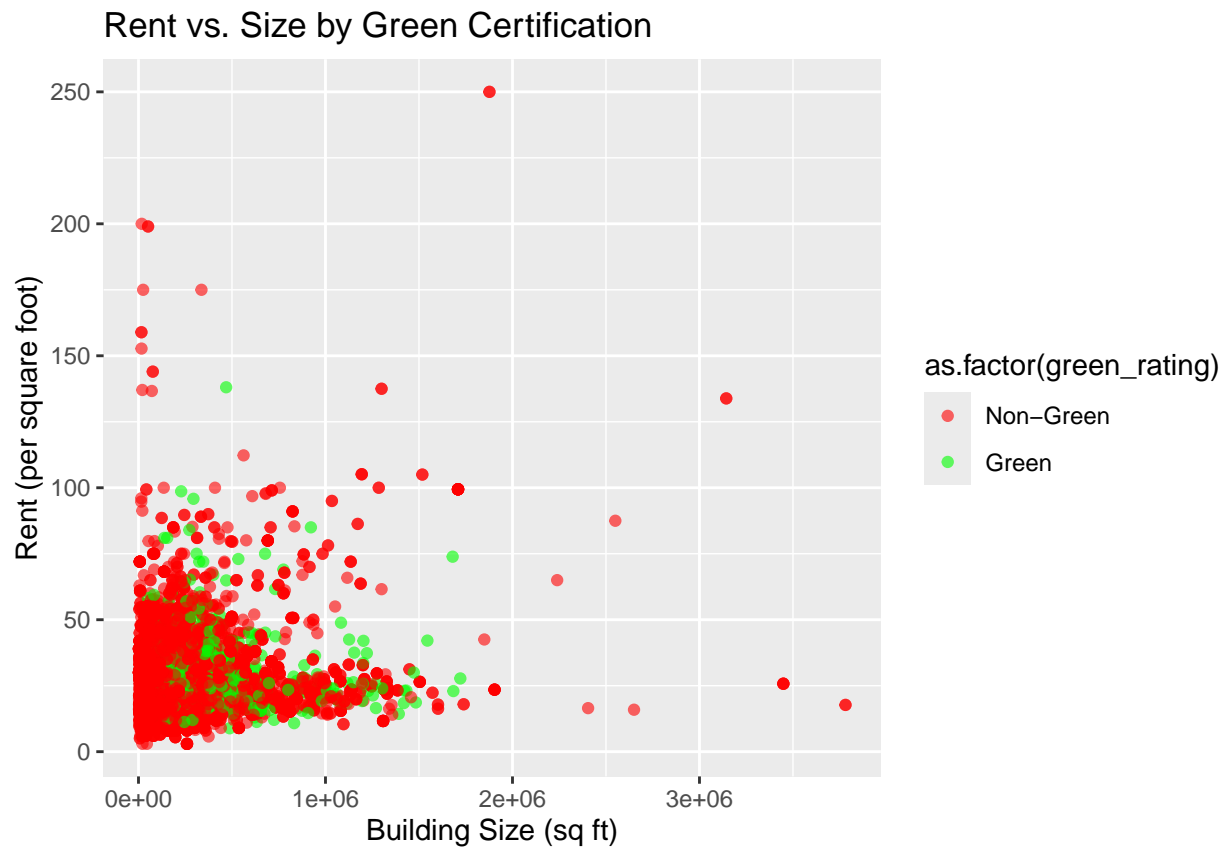
Similar to the Excel guru's analysis, I also filtered out buildings with an occupancy rate less than 10%.

2. Compare Median Rent between Green and Non-Green Buildings



The median rent in dollars per square foot per calendar year for non-green buildings is 25 dollars, while the median rent for green buildings is 27.60. This aligns with the initial analysis that green buildings indeed bring a higher rent.

3. Examine Potential Confounding Variables



Rent vs. Size by Green Certification:

The scatter plot shows that both green and non-green buildings have a wide range of sizes, but green buildings tend to have slightly higher rents across different sizes. However, size only doesn't fully account for the rent differences.



Rent by Building Class and Green Certification:

The boxplot indicates that green buildings, particularly those in Class A (1), generally have higher rents. This suggests that building class might be a confounder, with Class A buildings being more likely to be green-certified and commanding higher rents.

4. Visualize the Effect of Confounders

```
## 'geom_smooth()' using formula = 'y ~ x'
```



Rent vs. Cluster Rent by Green Certification:

The scatter plot shows that rents tend to correlate with the average rent in the building's cluster, regardless of green certification. However, green buildings generally have higher rents, even within the same cluster.

Conclusion

- The initial analysis by the Excel guru focused on the median rent difference without adjusting for possible confounders like class, size, or cluster location.
- The plots suggest that green buildings are more likely to be high-end (Class A) and located in areas with higher overall rents, which could explain some of the rent differences.
- Green buildings seem to command higher rents however, which supports the initial guru's conclusion but with an asterisk - part of the premium may be due to these confounding factors.