Green Buildings

2024-08-18

```
library(ISLR2)
library(plotly)
```

```
## Loading required package: ggplot2
```

```
##
## Attaching package: 'plotly'
```

```
## The following object is masked from 'package:ggplot2':
##
##     last_plot
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
## The following object is masked from 'package:graphics':
##
##     layout
```

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v lubridate 1.9.3     v tibble    3.2.1
## v purrr     1.0.2     v tidyr     1.3.1
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks plotly::filter(), stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(viridis)
```

```
## Loading required package: viridisLite
```

```
library(broom)
library(knitr)
library(shiny)
```

Visual Story Telling Part 1: Green Buildings

```
gb = read.csv('/Users/krummelha/Downloads/greenbuildings.csv')
names(gb)
```

```
##  [1] "CS_PropertyID"     "cluster"           "size"
##  [4] "empl_gr"           "Rent"              "leasing_rate"
##  [7] "stories"           "age"               "renovated"
## [10] "class_a"           "class_b"           "LEED"
## [13] "Energystar"        "green_rating"      "net"
## [16] "amenities"         "cd_total_07"       "hd_total07"
## [19] "total_dd_07"       "Precipitation"     "Gas_Costs"
## [22] "Electricity_Costs" "cluster_rent"
```

```
gb%>%glimpse
```

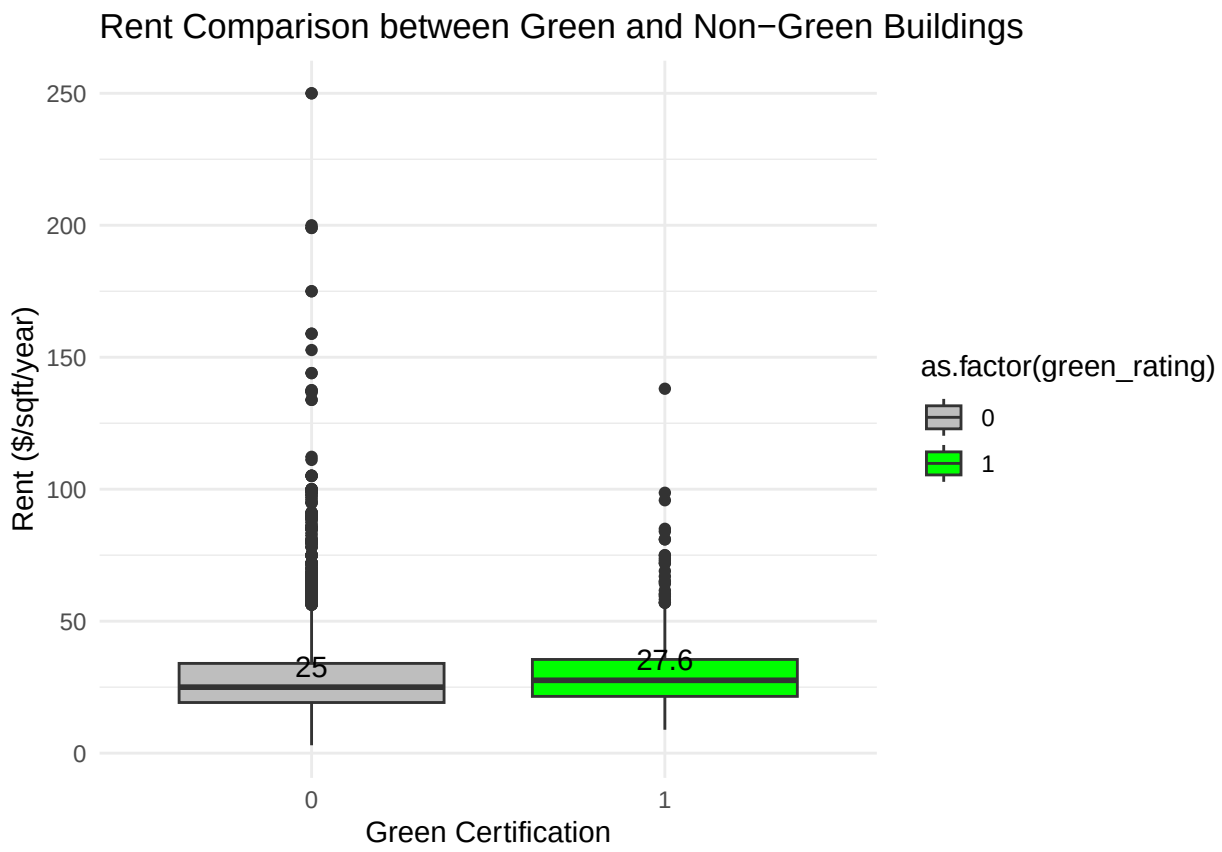
```
## Rows: 7,894
## Columns: 23
## $ CS_PropertyID     <int> 379105, 122151, 379839, 94614, 379285, 94765, 236739~
## $ cluster           <int> 1, 1, 1, 1, 1, 1, 6, 6, 6, 6, 6, 6, 6, 6, 6, 8, 8~
## $ size              <int> 260300, 67861, 164848, 93372, 174307, 231633, 210038~
## $ empl_gr           <dbl> 2.22, 2.22, 2.22, 2.22, 2.22, 2.22, 4.01, 4.01, 4.01~
## $ Rent              <dbl> 38.56, 28.57, 33.31, 35.00, 40.69, 43.16, 12.50, 14.~
## $ leasing_rate      <dbl> 91.39, 87.14, 88.94, 97.04, 96.58, 92.74, 94.33, 91.~
## $ stories           <int> 14, 5, 13, 13, 16, 14, 11, 15, 31, 21, 11, 15, 15, 3~
## $ age               <int> 16, 27, 36, 46, 5, 20, 38, 24, 34, 36, 32, 25, 26, 2~
## $ renovated         <int> 0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0~
## $ class_a           <int> 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0~
## $ class_b           <int> 0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0~
## $ LEED              <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ Energystar        <int> 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0~
## $ green_rating      <int> 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0~
## $ net               <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ amenities         <int> 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0~
## $ cd_total_07       <int> 4988, 4988, 4988, 4988, 4988, 4988, 2746, 2746, 2746~
## $ hd_total07        <int> 58, 58, 58, 58, 58, 58, 1670, 1670, 1670, 1670, 1670~
## $ total_dd_07       <int> 5046, 5046, 5046, 5046, 5046, 5046, 4416, 4416, 4416~
## $ Precipitation     <dbl> 42.57, 42.57, 42.57, 42.57, 42.57, 42.57, 25.55, 25.~
## $ Gas_Costs         <dbl> 0.01370000, 0.01373149, 0.01373149, 0.01373149, 0.01~
## $ Electricity_Costs <dbl> 0.02900000, 0.02904455, 0.02904455, 0.02904455, 0.02~
## $ cluster_rent      <dbl> 36.78, 36.78, 36.78, 36.78, 36.78, 36.78, 17.50, 17.~
```

```
#important variables
# size, empl.gr, Rent charged to tenants, leasing rate: measure of occupancy,
# LEED, energystar: different kinds of green certifications, ElectricityCosts
```

```
ggplot(gb, aes(x = as.factor(green_rating), y = Rent, fill = as.factor(green_rating))) +
  geom_boxplot() +
  stat_summary(fun = median, geom = "text", aes(label = round(..y.., 2)),
               position = position_dodge(width = 0.75), vjust = -0.5, color = "black") +
  labs(x = "Green Certification", y = "Rent ($/sqft/year)",
```

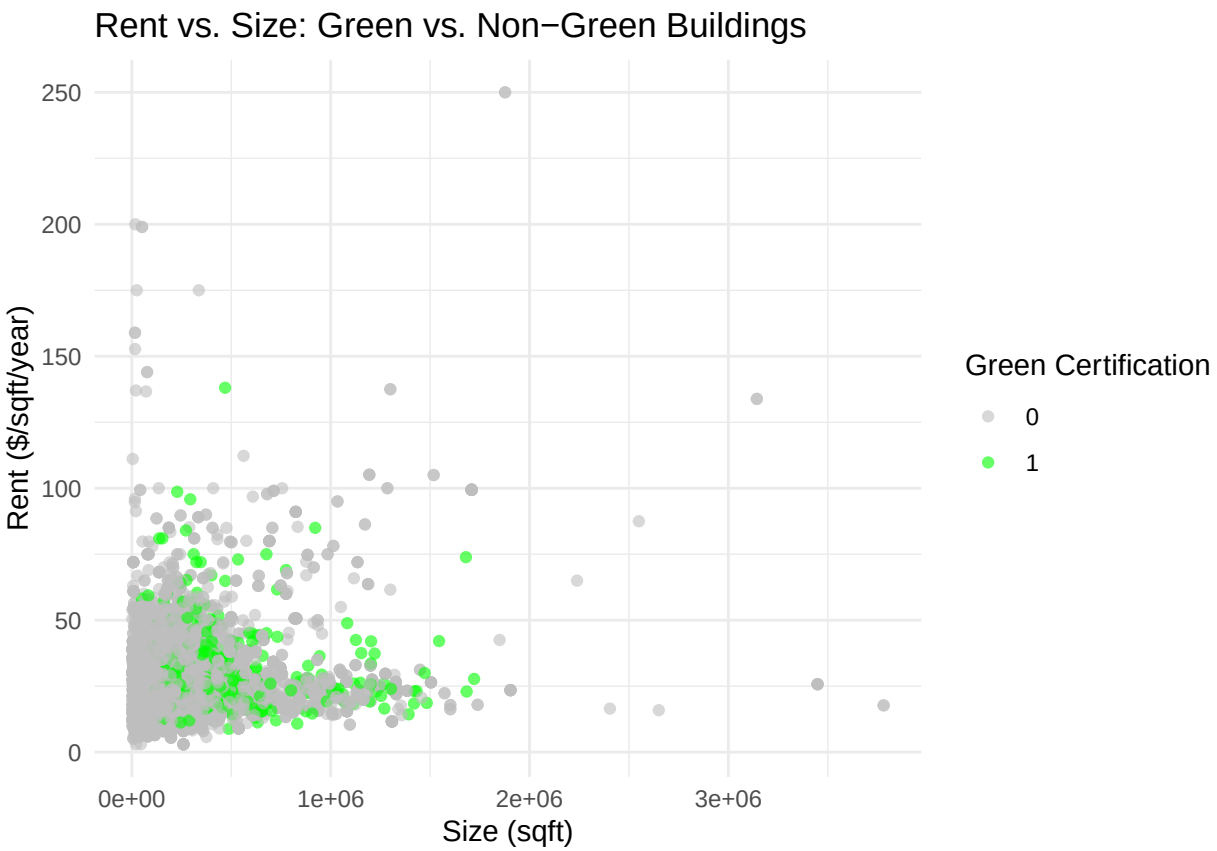
```
       title = "Rent Comparison between Green and Non-Green Buildings") +
  scale_fill_manual(values = c("gray", "green")) +
  theme_minimal()
```

```
## Warning: The dot-dot notation (‘..y..‘) was deprecated in ggplot2 3.4.0.
## i Please use ‘after_stat(y)‘ instead.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```



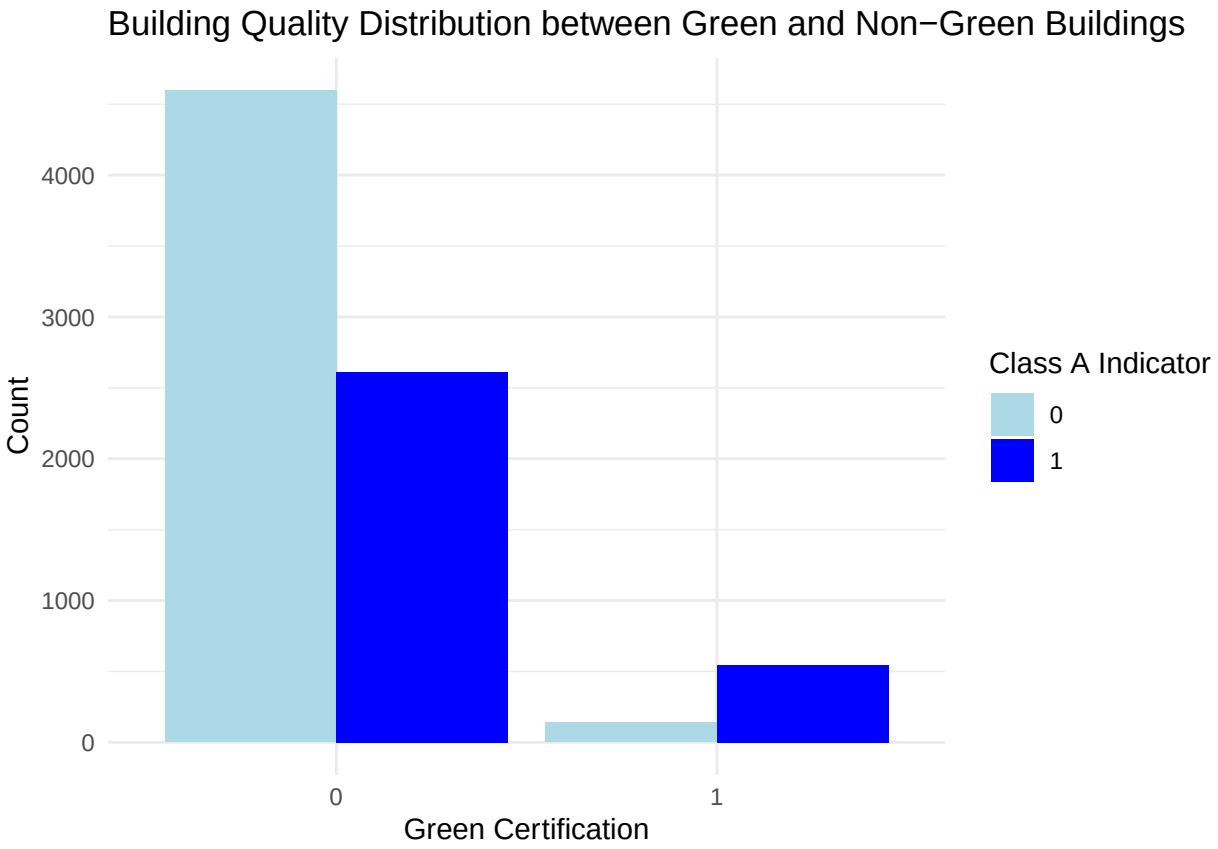
The above graph verifies that the staff member was indeed correct that the median rent for green buildings is $2.6 higher than it is for non-green buildings. There are definitely some high outliers for non-green buildings, but as a whole, green buildings are accruing higher rents than non-green buildings.

```
# Scatter plot of Rent vs Size
ggplot(gb, aes(x = size, y = Rent, color = as.factor(green_rating))) +
  geom_point(alpha = 0.6) +
  labs(x = "Size (sqft)", y = "Rent ($/sqft/year)",
       title = "Rent vs. Size: Green vs. Non-Green Buildings") +
  scale_color_manual(values = c("gray", "green"), name = "Green Certification") +
  theme_minimal()
```



This graph is helpful in that it illustrates that the size of the buildings does not have a huge affect on the rent per square foot per year of a building whether it is a green building or a non-green building. This is helpful to demonstrate because the developer can be assured that extra space is not required to draw in attracted buyers but rather a different quality, the quality of green buildings, for example. I intend to demonstrate this in the following graph.

```
# Bar plot for Building Quality Distribution
ggplot(gb, aes(x = as.factor(green_rating), fill = as.factor(class_a))) +
  geom_bar(position = "dodge") +
  labs(x = "Green Certification", y = "Count",
       title = "Building Quality Distribution between Green and Non-Green Buildings") +
  scale_fill_manual(values = c("lightblue", "blue", "darkblue"), name = "Class A Indicator") +
  theme_minimal()
```



The above graph demonstrates that green buildings have a Class A reputation in comparison to non-green buildings. Though there are much fewer green buildings as obvious in this depiction, the data shows they are of esteemed quality, which is a great reason to invest and become a part of a developing community that not only offers a long term financial benefit for the company, but also a PR and culture boost.

Final Conclusion and Recommendation

Our recommendation would be to follow through with this green building development based on the data available to us from the dataset. Building a green building would not only offer the company financial compensation as I previously stated, but it would also enhance company culture due to the executive team's social awareness. Moreover, the company could boast front-end buy-in into a green-focused culture that currently engulfs the present generations of the labor force, Millenials and Gen Zers.