Team Number: 079

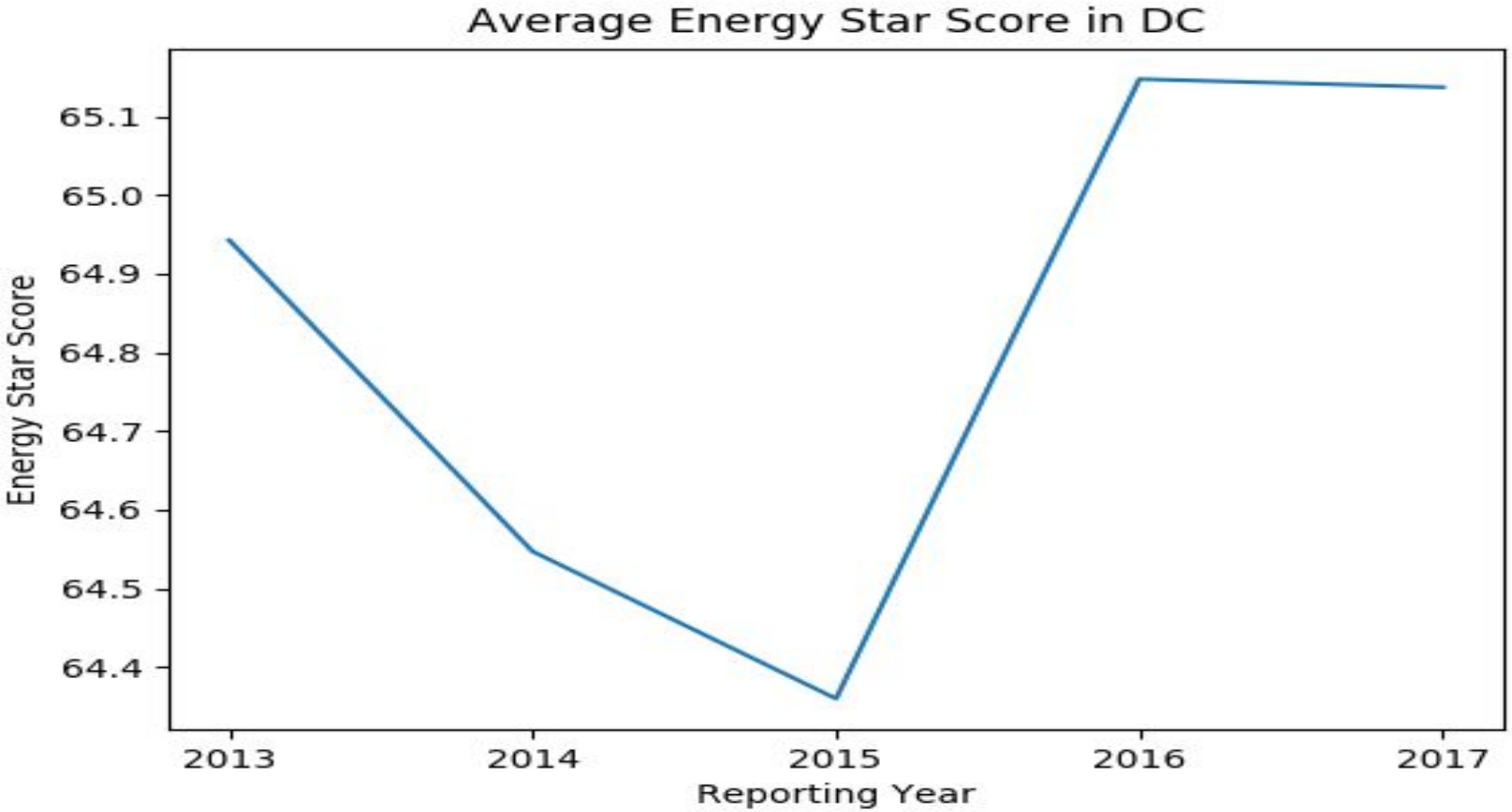
Dataset: DOEE Energy Star Benchmarking Database

The Data

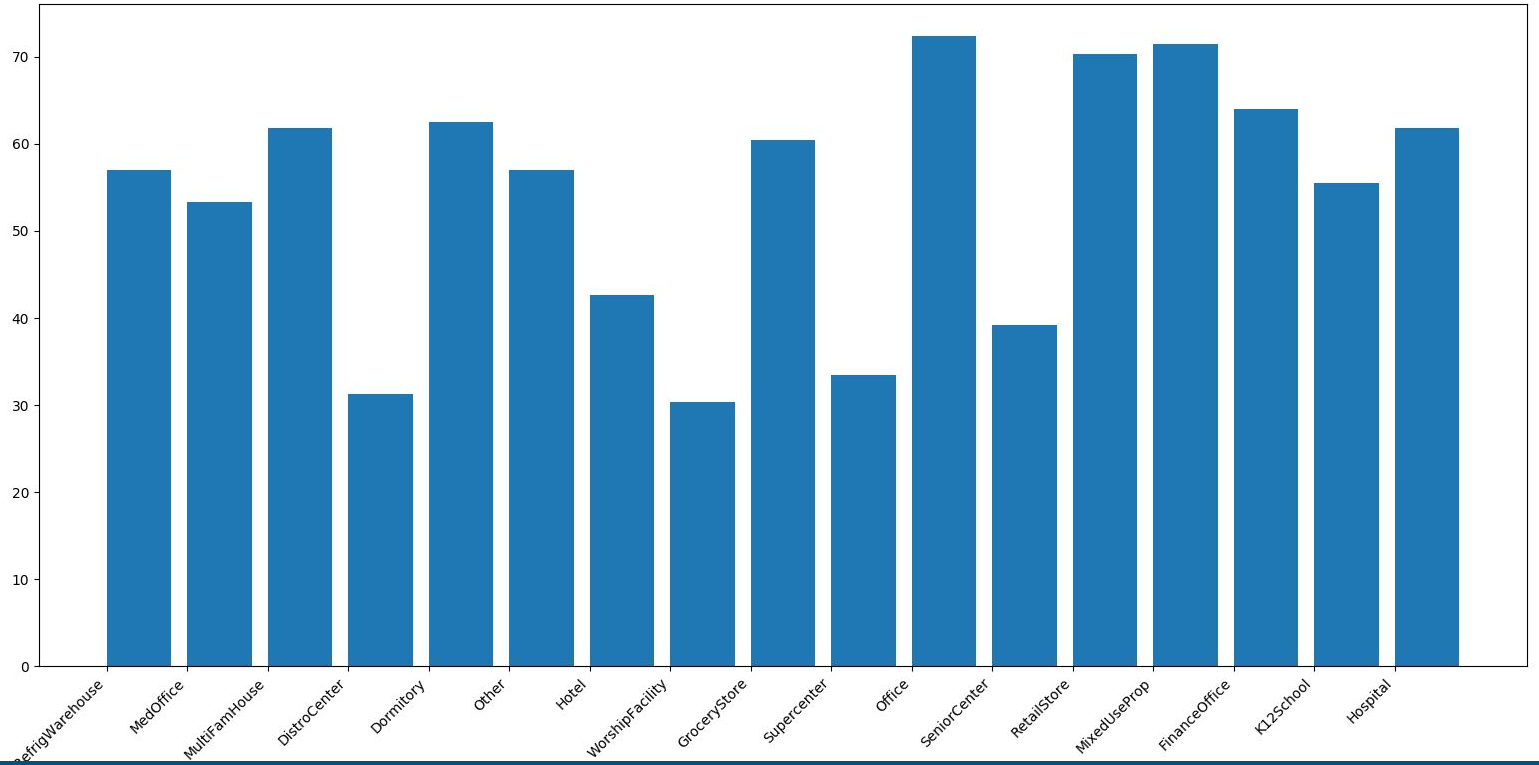
- Our dataset contained a list of values and fields linked to particular buildings in the DC metropolitan area. These fields included Greenhouse Gas (GHG) emissions, site and source energy usage, as well as the efficiency score calculated by energy star for each building.
- The data was cleaned in advance, although there were outliers that we had to remove from the dataset.

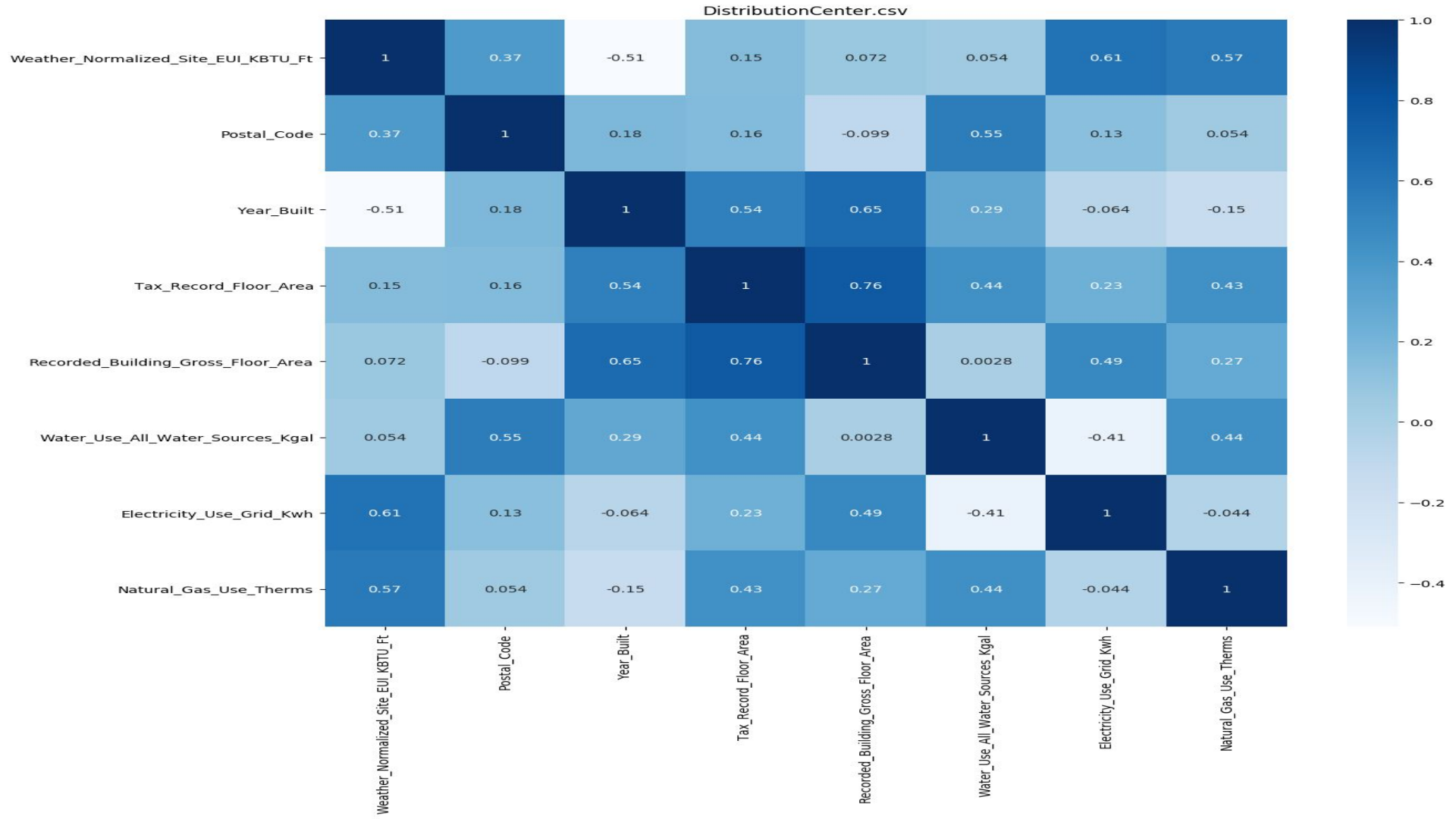
Central Questions and Areas of Inquiry

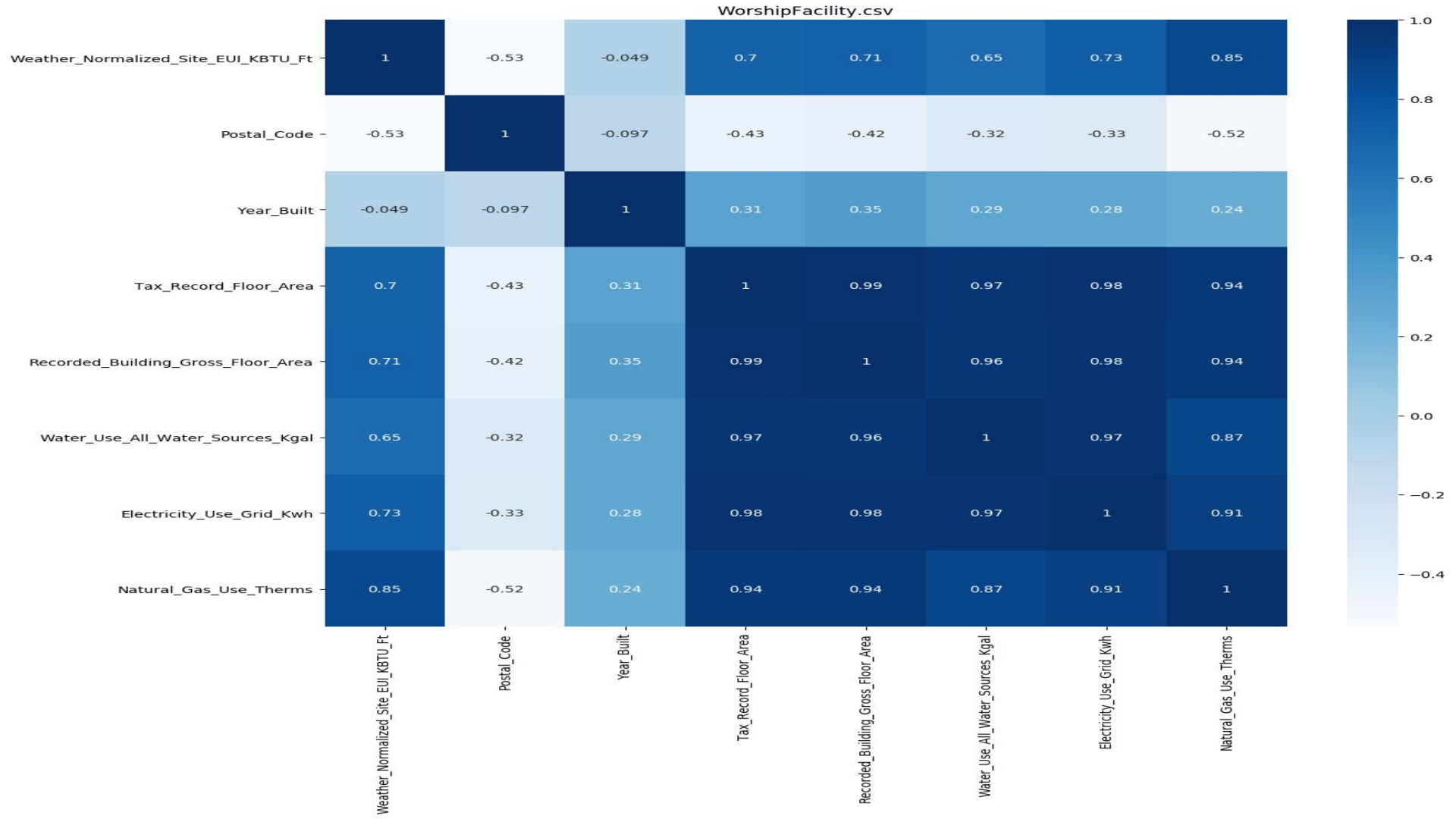
1. Which factors most significantly impact the energy usage in DC?
2. Is there a correlation between energy efficiency and carbon emission?
3. Find energy consumption patterns (increase or decrease) between 2013-2017, for each zip code and building types.
4. Identify the factor/ratio between source energy and site energy.

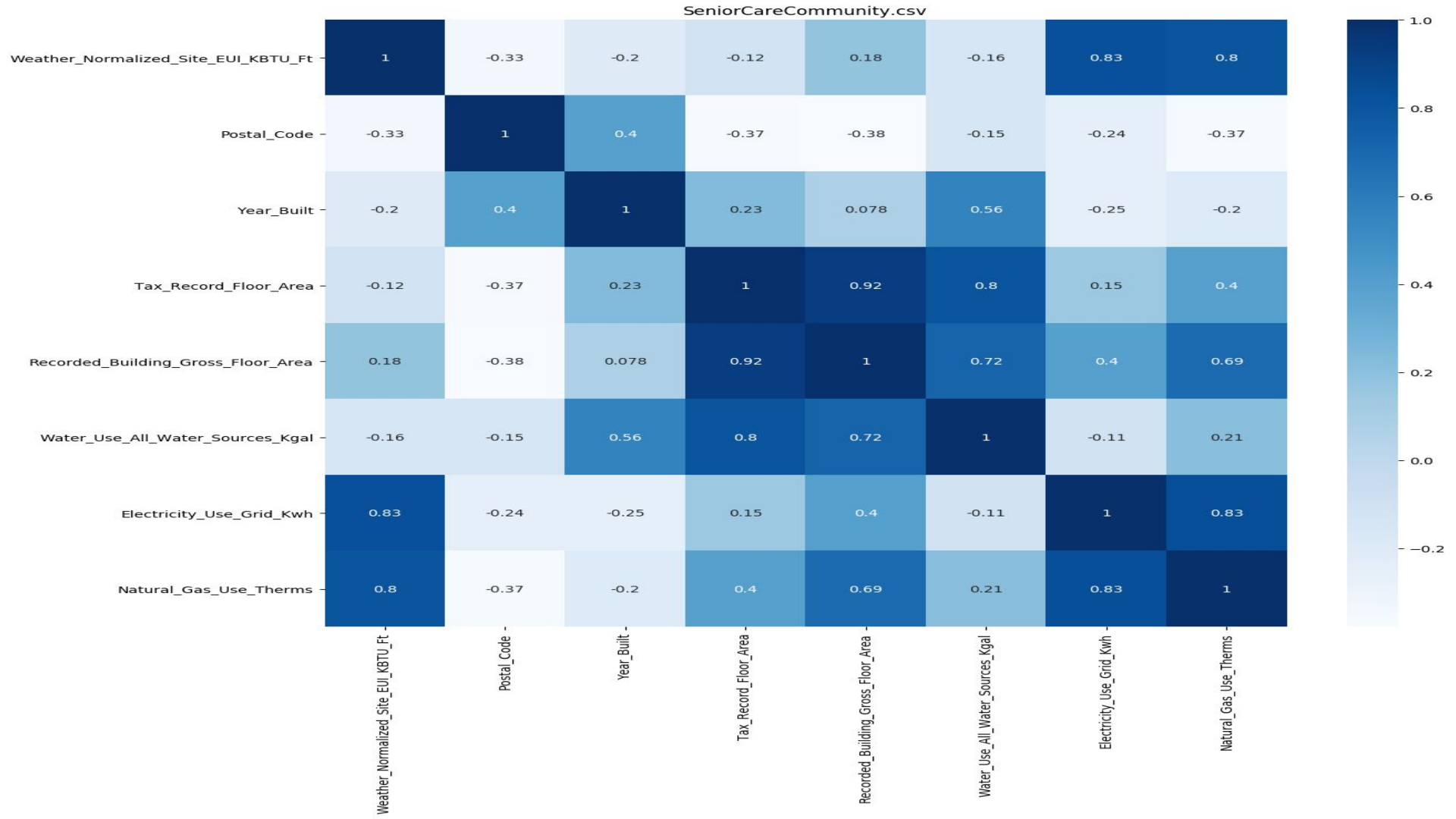

Average Energy Star Score in DC
Energy Star Score
65.1
65.0
64.9
64.8
64.7
64.6
64.5
64.4
2013
2014
2015
2016
2017
Reporting Year

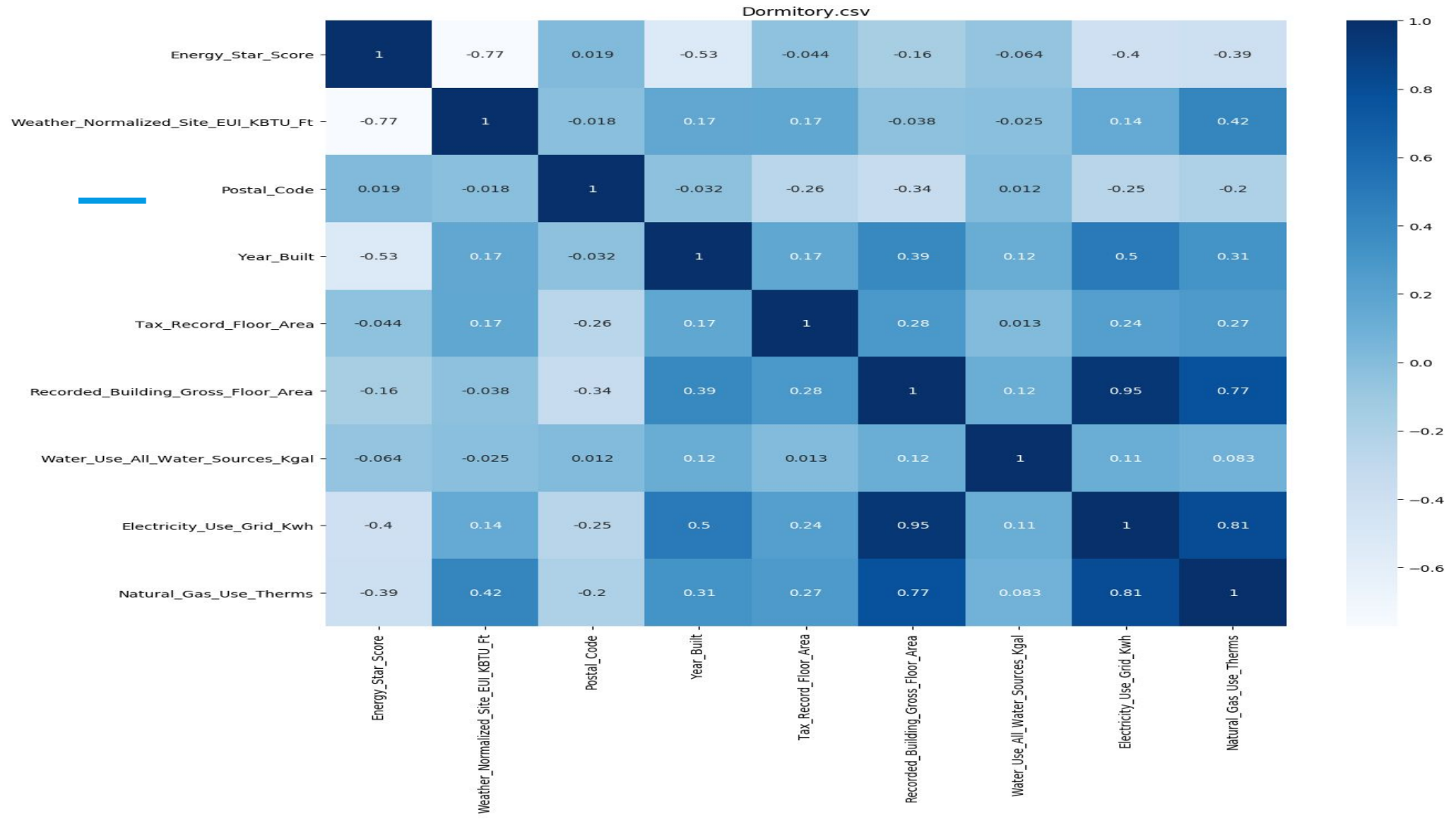
E-Star Score vs. Building Type

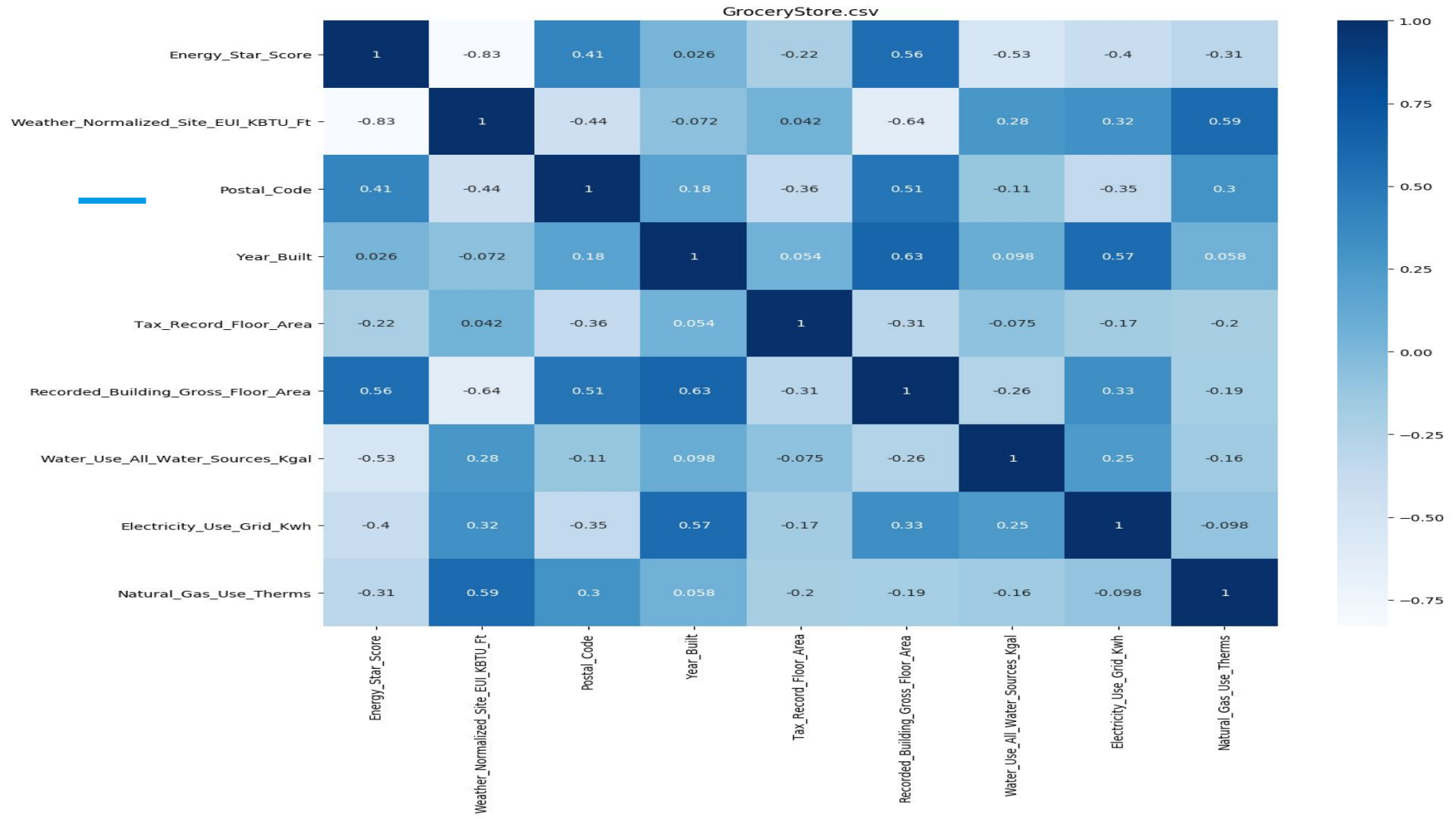


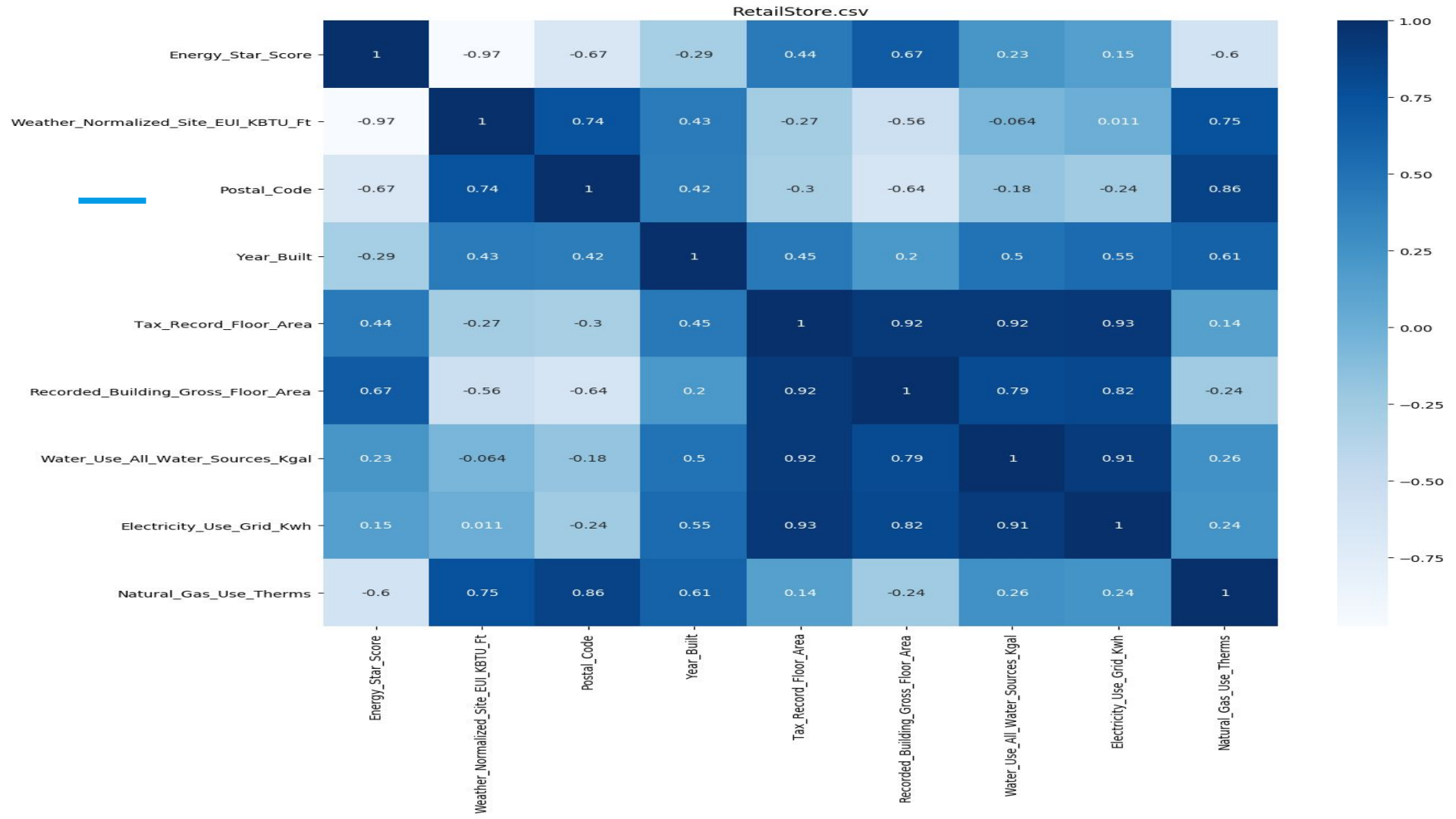

DistributionCenter.csv
Weather_Normalized_Site_EUI_KBTU_Ft
Postal_Code
Year_Built
Tax_Record_Floor_Area
Recorded_Building_Gross_Floor_Area
Water_Use_All_Water_Sources_Kgal
Electricity_Use_Grid_Kwh
Natural_Gas_Use_Therms
1 0.37 -0.51 0.15 0.072 0.054 0.61 0.57
0.37 1 0.18 0.16 -0.099 0.55 0.13 0.054
-0.51 0.18 1 0.54 0.65 0.29 -0.064 -0.15
0.15 0.16 0.54 1 0.76 0.44 0.23 0.43
0.072 -0.099 0.65 0.76 1 0.0028 0.49 0.27
0.054 0.55 0.29 0.44 0.0028 1 -0.41 0.44
0.61 0.13 -0.064 0.23 0.49 -0.41 1 -0.044
0.57 0.054 -0.15 0.43 0.27 0.44 -0.044 1
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4

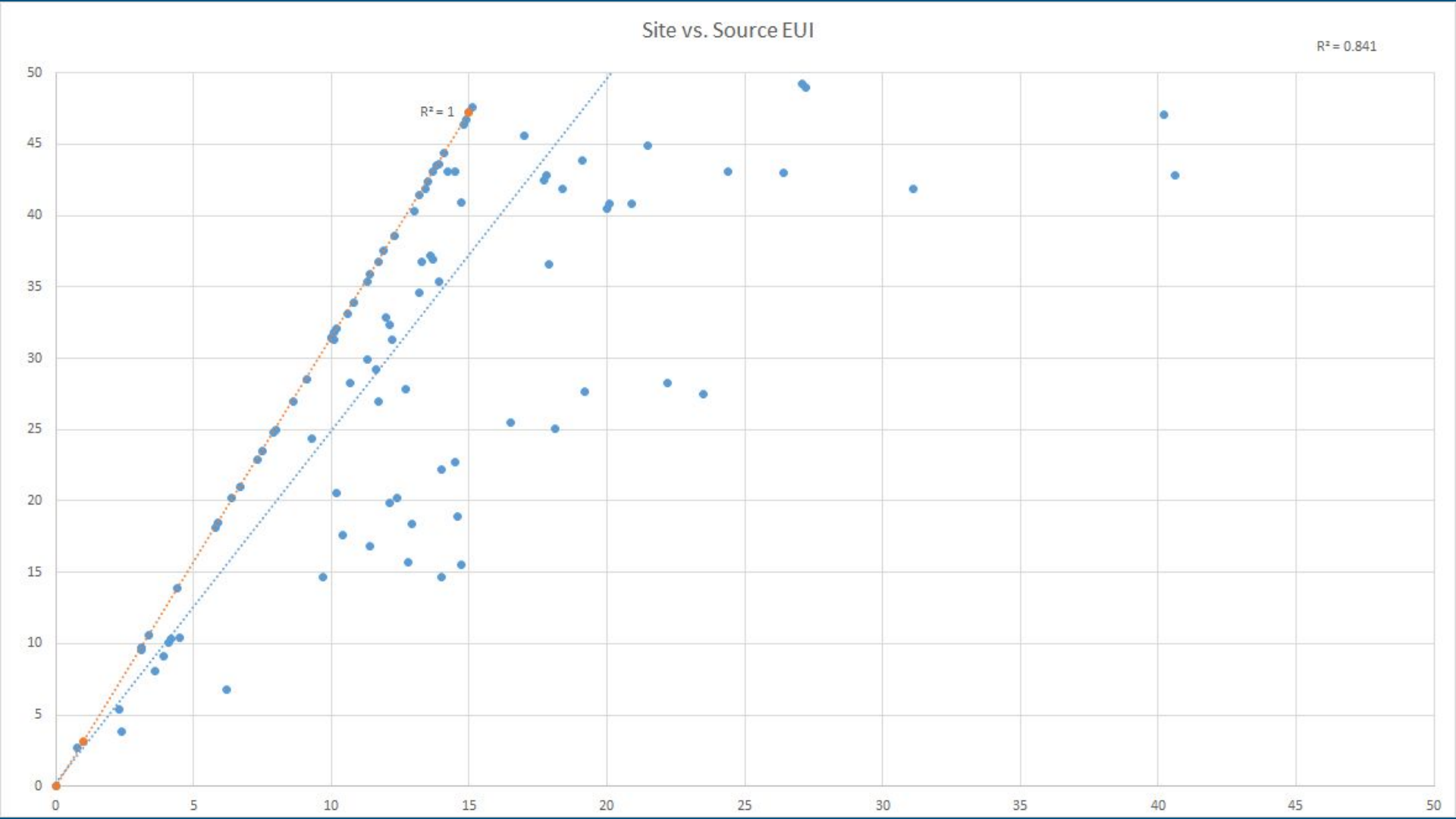

WorshipFacility.csv
Weather_Normalized_Site_EUI_KBTU_Ft
Postal_Code
Year_Built
Tax_Record_Floor_Area
Recorded_Building_Gross_Floor_Area
Water_Use_All_Water_Sources_Kgal
Electricity_Use_Grid_Kwh
Natural_Gas_Use_Therms
1 -0.53 -0.049 0.7 0.71 0.65 0.73 0.85
-0.53 1 -0.097 -0.43 -0.42 -0.32 -0.33 -0.52
-0.049 -0.097 1 0.31 0.35 0.29 0.28 0.24
0.7 -0.43 0.31 1 0.99 0.97 0.98 0.94
0.71 -0.42 0.35 0.99 1 0.96 0.98 0.94
0.65 -0.32 0.29 0.97 0.96 1 0.97 0.87
0.73 -0.33 0.28 0.98 0.98 0.97 1 0.91
0.85 -0.52 0.24 0.94 0.94 0.87 0.91 1
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4


SeniorCareCommunity.csv
Weather_Normalized_Site_EUI_KBTU_Ft
Postal_Code
Year_Built
Tax_Record_Floor_Area
Recorded_Building_Gross_Floor_Area
Water_Use_All_Water_Sources_Kgal
Electricity_Use_Grid_Kwh
Natural_Gas_Use_Therms
1 -0.33 -0.2 -0.12 0.18 -0.16 0.83 0.8
-0.33 1 0.4 -0.37 -0.38 -0.15 -0.24 -0.37
-0.2 0.4 1 0.23 0.078 0.56 -0.25 -0.2
-0.12 -0.37 0.23 1 0.92 0.8 0.15 0.4
0.18 -0.38 0.078 0.92 1 0.72 0.4 0.69
-0.16 -0.15 0.56 0.8 0.72 1 -0.11 0.21
0.83 -0.24 -0.25 0.15 0.4 -0.11 1 0.83
0.8 -0.37 -0.2 0.4 0.69 0.21 0.83 1
1.0
0.8
0.6
0.4
0.2
0.0
−0.2


Dormitory.csv
Energy_Star_Score
Weather_Normalized_Site_EUI_KBTU_Ft
Postal_Code
Year_Built
Tax_Record_Floor_Area
Recorded_Building_Gross_Floor_Area
Water_Use_All_Water_Sources_Kgal
Electricity_Use_Grid_Kwh
Natural_Gas_Use_Therms
1 -0.77 0.019 -0.53 -0.044 -0.16 -0.064 -0.4 -0.39
-0.77 1 -0.018 0.17 0.17 -0.038 -0.025 0.14 0.42
0.019 -0.018 1 -0.032 -0.26 -0.34 0.012 -0.25 -0.2
-0.53 0.17 -0.032 1 0.17 0.39 0.12 0.5 0.31
-0.044 0.17 -0.26 0.17 1 0.28 0.013 0.24 0.27
-0.16 -0.038 -0.34 0.39 0.28 1 0.12 0.95 0.77
-0.064 -0.025 0.012 0.12 0.013 0.12 1 0.11 0.083
-0.4 0.14 -0.25 0.5 0.24 0.95 0.11 1 0.81
-0.39 0.42 -0.2 0.31 0.27 0.77 0.083 0.81 1
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6


GroceryStore.csv
Energy_Star_Score
Weather_Normalized_Site_EUI_KBTU_Ft
Postal_Code
Year_Built
Tax_Record_Floor_Area
Recorded_Building_Gross_Floor_Area
Water_Use_All_Water_Sources_Kgal
Electricity_Use_Grid_Kwh
Natural_Gas_Use_Therms
1 -0.83 0.41 0.026 -0.22 0.56 -0.53 -0.4 -0.31
-0.83 1 -0.44 -0.072 0.042 -0.64 0.28 0.32 0.59
0.41 -0.44 1 0.18 -0.36 0.51 -0.11 -0.35 0.3
0.026 -0.072 0.18 1 0.054 0.63 0.098 0.57 0.058
-0.22 0.042 -0.36 0.054 1 -0.31 -0.075 -0.17 -0.2
0.56 -0.64 0.51 0.63 -0.31 1 -0.26 0.33 -0.19
-0.53 0.28 -0.11 0.098 -0.075 -0.26 1 0.25 -0.16
-0.4 0.32 -0.35 0.57 -0.17 0.33 0.25 1 -0.098
-0.31 0.59 0.3 0.058 -0.2 -0.19 -0.16 -0.098 1
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75


RetailStore.csv
Energy_Star_Score
Weather_Normalized_Site_EUI_KBTU_Ft
Postal_Code
Year_Built
Tax_Record_Floor_Area
Recorded_Building_Gross_Floor_Area
Water_Use_All_Water_Sources_Kgal
Electricity_Use_Grid_Kwh
Natural_Gas_Use_Therms
1 -0.97 -0.67 -0.29 0.44 0.67 0.23 0.15 -0.6
-0.97 1 0.74 0.43 -0.27 -0.56 -0.064 0.011 0.75
-0.67 0.74 1 0.42 -0.3 -0.64 -0.18 -0.24 0.86
-0.29 0.43 0.42 1 0.45 0.2 0.5 0.55 0.61
0.44 -0.27 -0.3 0.45 1 0.92 0.92 0.93 0.14
0.67 -0.56 -0.64 0.2 0.92 1 0.79 0.82 -0.24
0.23 -0.064 -0.18 0.5 0.92 0.79 1 0.91 0.26
0.15 0.011 -0.24 0.55 0.93 0.82 0.91 1 0.24
-0.6 0.75 0.86 0.61 0.14 -0.24 0.26 0.24 1
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75


Site vs. Source EUI
R² = 0.841
R² = 1
50
45
40
35
30
25
20
15
10
5
0
0
5
10
15
20
25
30
35
40
45
50

Recommendations

- An assessment of the feasibility of incentivizing building owners to install solar panels or find other ways of generating power locally.
- Further study on the possibility of improving energy efficiency in distro centers, senior centers and retail stores.
    - Establish causal relationships between correlated factors

Assumptions

- That the massive outlier could be safely excluded from the dataset.
- That 2% of energy star values being 1 and 4% being 100 was within the norm.
- That self reported data was accurate (The value of property square footage varied)

Issues with the Data Set

- The set came with little information on how data was collected for each field.
- No context on how sampling was determined (super centers)
- There were several large outliers that skewed the data until they were removed.
- Only two Supercenters were provided

Further Questions

- Source of the data fields (e.g. EPA reported vs. self reported)
- Whether geographical data can be correlated to efficiency metrics