Data Analysis of Building Energy Benchmarking Data

GitHub repository link - https://github.com/davidfakolujo/HW-04.git

Introduction

The Building Energy Benchmarking Dataset provides information on the energy and greenhouse gas emission performance information for properties owned and operated by the City of Calgary.

Regular expressions were used to extract numeric values from text-based numeric columns, standardize postal codes, as well as to clean and extract meaningful text from Property Names and Addresses.

An analysis was done on the dataset and the following insights were derived.

Key trends in energy consumption and efficiency

- **Histogram of Site Energy Use**
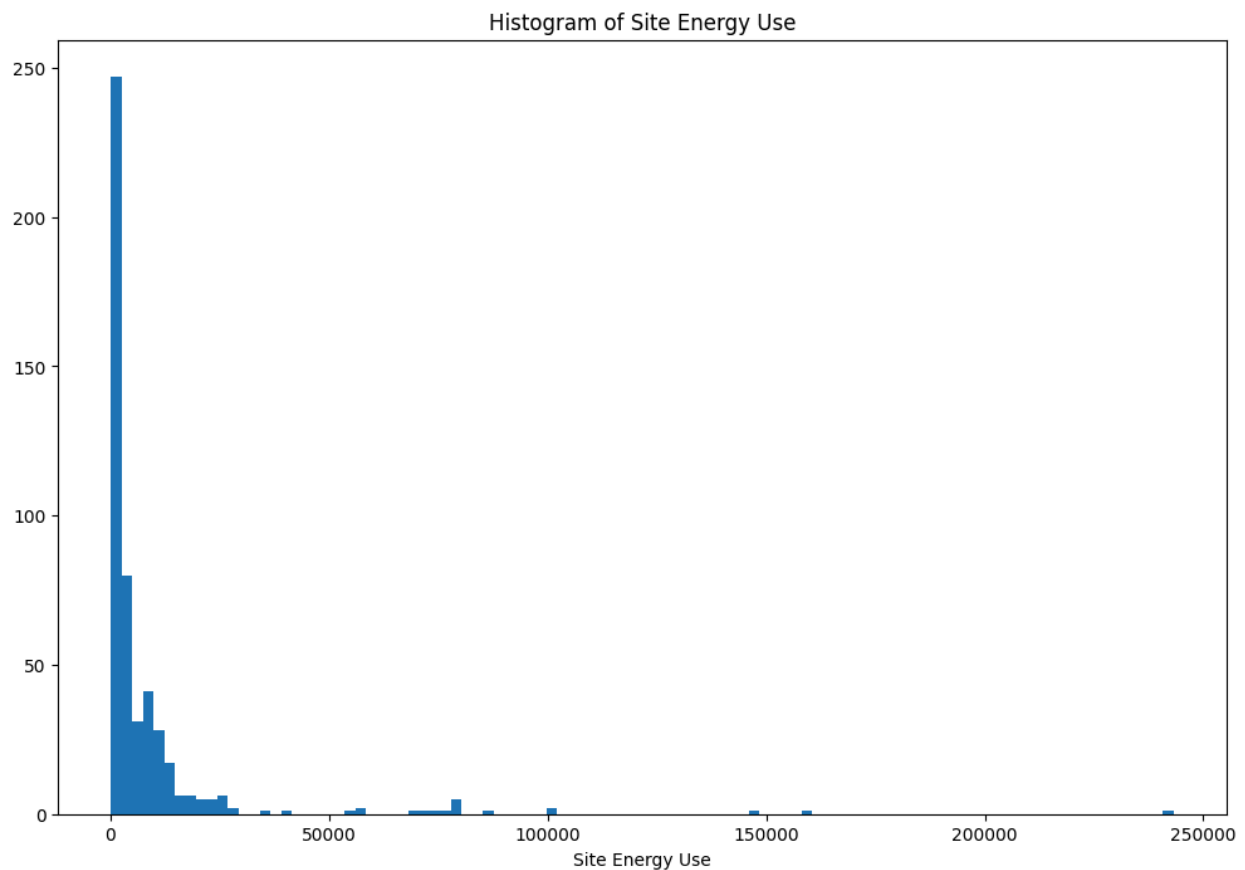
The plot above shows a right skewed histogram, which indicates that most of the buildings had energy consumptions between 0 GJ and about 25,000 GJ, while very few buildings had energy consumptions being above 25,000 GJ.

- **Scatter plot of efficiency against Site Energy Use**

Scatter Plot of Efficiency Against Site Energy Use
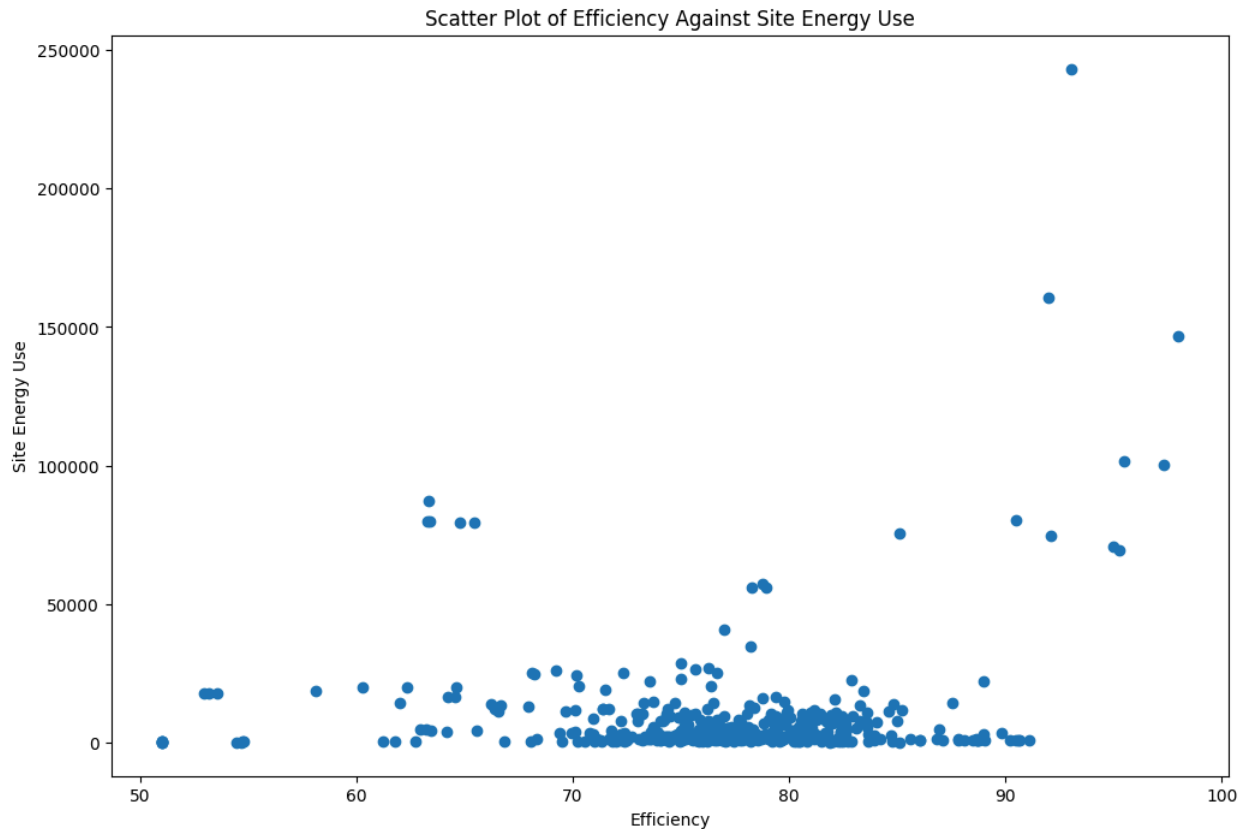
The scatter plot above generally shows that most of the energy consumption was less than 25,000 GJ, while most of the efficiencies was between 70% and 85% for consumptions less than 5,0000 GJ, with few exceptions having less than 70% or more than 85%. The very high consumptions (above 50,000 GJ) generally had higher efficiencies of 90% and above, with some exceptions having efficiencies between 60% and 90%.

The plot suggests the buildings with higher energy consumption have better methods of consuming energy such as using renewable energy sources which reduces emissions.

Seasonal and Property Type Variations

- **Yearly trend of total and direct Greenhouse Gas (GHG) emissions**
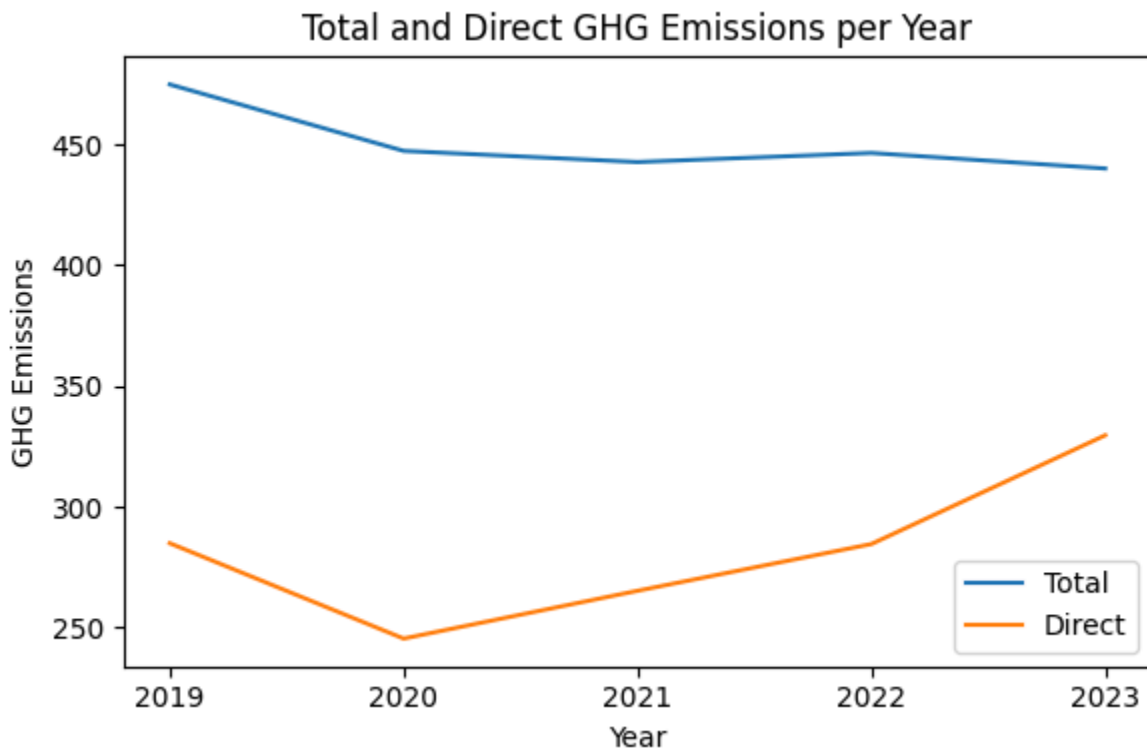
For total GHG emissions, the plot above shows that 2019 had the highest average GHG emission, with a gradual reduction through the years. For direct GHG emissions, the average GHG emission initially reduced between 2019 and 2020. However, the average GHG emission then went back up gradually through the remaining years. This indicates that while the direct GHG emissions rose from 2020 to 2023, the indirect GHG emissions must have reduced in those years, since there was also a reduction in total GHG emissions.

- **Bar chart of average Site Energy Use by Property Type**

Average Site Energy Use per Property Type
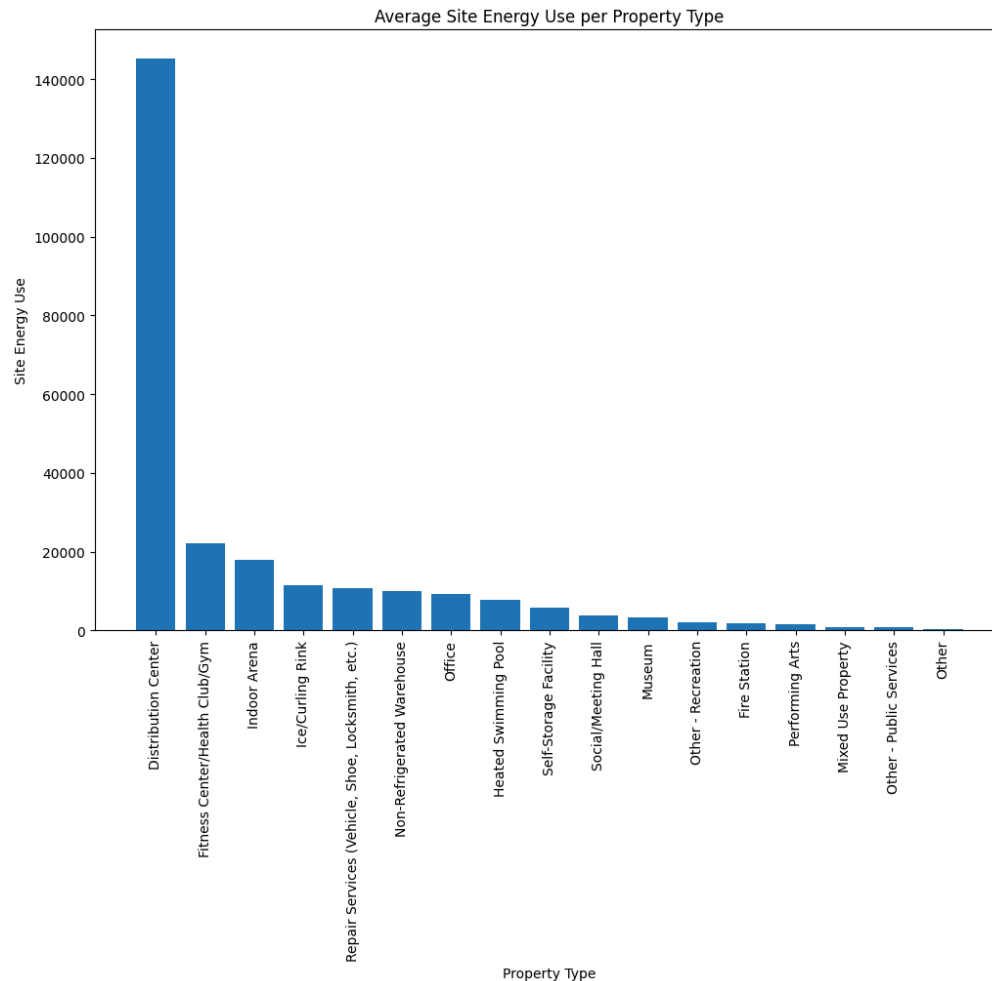
The bar chart above shows that distribution centers had the highest average consumption of energy. This average consumption was about 7 times more than that of fitness centers/health clubs/gyms which had the second highest average energy consumption. The lowest average energy consumptions came from properties such as fire stations, performing arts, mixed use properties, and others.

Recommendations for improving energy efficiency and reducing emissions.

- Buildings with low efficiencies should explore ways better ways of consuming energy and reducing losses, including renewable energy sources
- Buildings need to work on reducing their direct GHG emissions. Since there has already been a reduction in indirect GHG emissions, reducing direct GHG emissions would lead to much further reductions in total GHG emissions.