

Digital Scholars Data Science Discovery - “You and Data Science” Paper

Global warming is constantly becoming a greater and greater worldwide crisis. It impacts climate all around the world through human abuse of fossil fuels and the uncontrollably increasing amounts of carbon dioxide emissions. However, nowhere is this issue more conspicuous than in large, densely-populated, bustling cities. Factories import and burn coal, petroleum, and natural gas; endless concrete buildings through thermal decomposition; cars, trucks, and planes belch out foul smog into the skies - even through this small glance into the city's landscape, it's obvious that the environment of the megalopolis fuels the development and perpetuates the impact of emissions on climate like no other place on Earth. However, there remains a means to fix the urban climate issues that we've caused. As awareness towards climate change has risen, huge companies, contractors, city lawmakers, and even the populace have pushed to make energy greener. By moving to a more sustainable energy source, carbon dioxide emissions decrease, thereby helping cities thrive. Chicago, the USA's third largest city, has faced much criticism for its terrible environment, dating back even to the early 1900s. However, as time progressed and technology became available and less expensive, sustainability has seen a significant increase over the years in Chicago. Nevertheless, due to the inequality in the regions that see these progressions and the increasing energy needs of the ever-growing city, it's obvious that a lot of work still needs to be done.

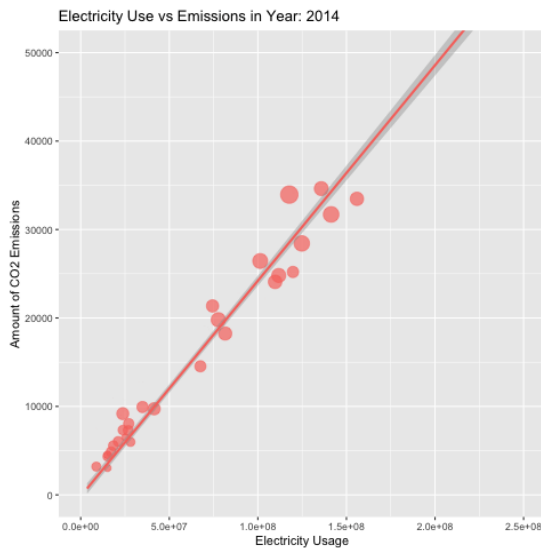
This issue piques my interest every time I hear about it. I care a lot about the environment of the city in which I've spent the entirety of my childhood. Moreover, as someone from the suburbs, I never experienced the full extent of city pollution until I started spending much more time in urban areas, in Chicago and around the world. Resultantly, this data is really important to

me, because I hope that with my analysis of these issues, I can shine a light on problems like these and help solve them. When I first encountered this dataset, I knew that Chicago was working on solving problems like these, but there was still a lot to do. My goal in this analysis project, therefore, was to find exactly what still needs to be done in improving the environment in Chicago urban areas, in terms of inequality and energy efficiency. The data that I used in this analysis was the [Chicago Energy Benchmarking](#) data, found through the Chicago Data Portal, which details the energy emissions, electricity usage, and efficiency of the biggest buildings in Chicago.

In my data analysis method, I hoped to utilize many of the skills that I had learned in my Data Science class. I first cleaned my data, filtering out unnecessary information and keeping the data that I needed. This helped me keep my data analysis easy-to-use, which was really helpful in keeping my code succinct. Then, I created several graphs to help me identify patterns in my data. This included histograms, scatter plots, regression plots, and even geographical maps to show the locations, sizes, emissions, electrical usages, and efficiencies of the biggest buildings. Because I wanted to focus on the progression of energy usage and efficiency throughout the years, I relied on animated graphs to show development over time. In addition, my geographical maps showed the locations of the most energy-efficient buildings in Chicago, which really helped me visually process the disparity between the sustainability of buildings in various regions of the city. Finally, I conducted various statistical tests (such as constructing confidence intervals, P-testing, and chi-square testing for independence) to find out more about the relationships between parts of my data.

From my data analysis, I made a lot of interesting and unexpectedly surprising discoveries. Firstly, I discovered that the regions of greatest correlation were between electricity

usage, total carbon dioxide emissions, and energy efficiency. A chi-square test helped me reject



the possibility of independence between electricity

usage and carbon dioxide emissions, and an animated

scatter plot (click [here](#) to view it if it doesn't play on

the PDF) further helped to emphasize this strong

correlation. But more specifically, my data analysis

helped me conclude that the buildings with the

greatest energy efficiency over *all* types of buildings -

or, in other words, the ones that were able to make the

most out of their electricity usage, with respect to the

size of the building in question - were the ones with the least carbon dioxide emissions per

square foot. This was extremely important, because it proves that the buildings with the greatest

efficiency were the ones with the least relative carbon dioxide emissions, or in other words, the

ones with the greenest energy sources. Over time, in fact, I found that the buildings of Chicago

tended to turn to less carbon-emitting energy sources, but still used constant amounts of energy,

proving that they were adopting more sustainable energy sources for better efficiency. My

hypothesis test for the average energy efficiency

in Chicago also proved that energy efficiency has

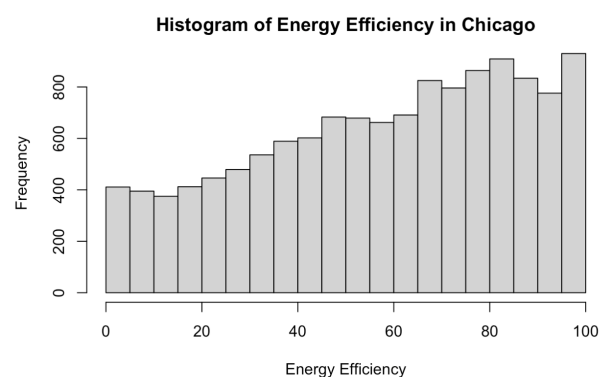
definitely seen a significant rise over the years.

In addition, I also conducted an analysis on the

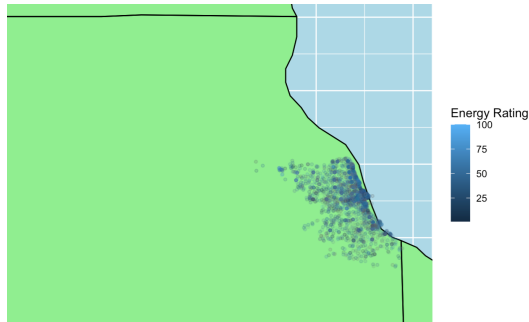
spread of the energy efficiency of the buildings

of Chicago, which showed that although

nowadays, Chicago is above the average energy efficiency rating in the US, it's not above the



average by very much, and still has a long way to go in becoming more energy-efficient. In fact, the 95% confidence interval for energy efficiency rating in Chicago, from a scale of 0 to 100, was 57.77 to 58.72 (with the median being 50); although this is better than average, it's not that



great. Finally, I conducted a geographical analysis of the regions in Chicago with the most energy-efficient and reduced-emission buildings, and found my most important result: the distribution of energy-efficient structures in Chicago is not very equal. The heart of the city - the tourist-frequented areas by the river and the

lakefront - enjoy much higher values of energy efficiency and the least amount of air pollution; however, the more densely populated and economically burdened areas of Southern Chicago were still frequented by rampant air-pollution and energy efficiency loss, which could be attributed to energy mismanagement or poor funding to be able to develop more efficient (and more sustainable) energy.

In conclusion, I found from my data that Chicago's sustainable energy sector is growing as corporations take advantage of the growing efficiency of sources of green energy. However, the developing city is still dealing with issues of regional inequality, hindered access to renewables, and a still weak energy efficiency rating. Nevertheless, over the years, Chicago's green energy usage (and resultantly, its energy efficiency) is bound to increase, and with proper judicial management, we can work together to solve problems of energy inequality and bring more green energy to Chicago. Now that our advancements in green energy have resulted in the indisputable fact that the link between energy efficiency and energy sustainability has been

growing stronger and stronger, we can expect to see green energy rise in popularity throughout the city.

This analysis was very eye-opening for me, and helped me learn lots of things about the energy usage, efficiency, and inequalities that Chicago experiences. It also helped me learn a lot about the power and efficacy of green energy sources. This analysis has inspired me to conduct more data analyses on related data (for example, I'd like to check out other solutions to urban climate issues, such as green roofs). This analysis has definitely encouraged me to go and discover new things, including researching the reasons for the energy inequality in Chicago - I want to find out why some regions experience less energy efficiency as other regions. I hope that my analysis of this data was eye-opening, and helped shine some light on the true sources of energy efficiency; and hopefully, my findings can help support some serious action involving bringing sustainable and efficient energy to more regions of Chicago and the nation.