Andy Carson
DATA 608
12/21/2018

Final Project: Energy Usage

Introduction

The environmental effects of global warming are having and will continue to have a negative impact on people around the world, and in particular, on the poorest and least capable of adjusting to the change through the use of resources. In addition, the higher cost of cleaner alternative energy will be more burdensome to the poorest people and countries. This is a serious matter of social justice.

Reducing energy usage (when the burning of fossil fuels for energy production to sustain energy consumption is contributing to global warming via CO2 emissions) is one way of combating global warming and addressing this matter of environmental and social concern. Hence, it is important for businesses and homeowners to understand what factors increase energy usage and what might be done to reduce energy usage and thereby reduce their contribution to global warming.

The task of this project is to visually demonstrate and understand the prominent factors that lead to increased energy usage at the local level. To motivate this effort, I first produce visualizations that demonstrate the relationships between global energy usage, CO2 emissions, and global average temperatures. Having demonstrated the strong correlations among these three, I next produce visualizations that show how home energy usage relates to various factors responsible for my home energy usage increasing and decreasing, including: local temperature, time of day, day of week, month, and (inferred) appliances/devices in use.

Visualizations

The visualizations referred to are located here:

- https://carsonar.shinyapps.io/energyusage/

Data Sources

The data sources are:

- Home energy usage data from my own house from September 1st, 2017 through September 30, 2018. After having exported from the device (reference: http://eyedro.com/how-to-export-to-excel-from-myeyedro/), the data is now stored in my GitHub repo:
    - https://github.com/anrcarson/CUNY-MSDA/tree/master/DATA608/FinalProject
- Global yearly average temperatures:
    - https://climate.nasa.gov/vital-signs/global-temperature/
    - https://climate.nasa.gov/system/internal_resources/details/original/647_Global_Temperature_Data_File.txt
- Global energy/fuel usage:

    - https://www.eia.gov/totalenergy/data/browser/index.php?tbl=T01.01#/?f=A&start=1949&end=2017&charted=4-6-7-14
- Global CO2 emissions:
    - http://cdiac.ess-dive.lbl.gov/trends/emis/tre_glob_2014.html
    - https://www.epa.gov/ghgemissions/global-greenhouse-gas-emissions-data
- Local weather history to Seattle, WA:
    - https://www.wunderground.com/history/monthly/us/wa/seattle-boeing/KBFI/date/2017-8

All the above data has been exported for use in the visualizations and is located here:

- https://github.com/anrcarson/CUNY-MSDA/tree/master/DATA608/FinalProject

Parameters of the Data

The data sources above combine into two different data sets with the following parameters:

- Global
    - Energy Production and Consumption
        - Nuclear, exports, imports, stock change/other, fuels, renewable, totals
        - Unit: Quadrillion BTU
    - CO2 emissions
        - gas fuel, liquid fuel, solid fuel, cement production, gas flaring, per capita
        - Unit: million metric tons of Carbon
    - Global temperature change
        - the change in global surface temperature relative to 1951-1980 average temperatures (mean relative temperature), Lowess smoothed mean relative temperature
        - Unit: degrees Celsius
    - Year
- Home
    - Calendar
        - year, month, date, hour, day of the week
    - Watt hours
        - total and average; port A, port B, and total
        - Unit: Watts
    - Daily Temperature
        - max, min, average, difference from 65 degrees
        - Unit: degrees Fahrenheit
    - Daily dewpoint
        - max, min, average
        - Unit: degrees Fahrenheit
    - Humidity
        - max, min
        - Unit: percentage %

- Windspeed
    - max, min
    - Unit: mph
- Pressure
    - max, min
    - Unit: Hg
- Precipitation
    - average
    - Unit: inches

Strengths and Weaknesses of the Data

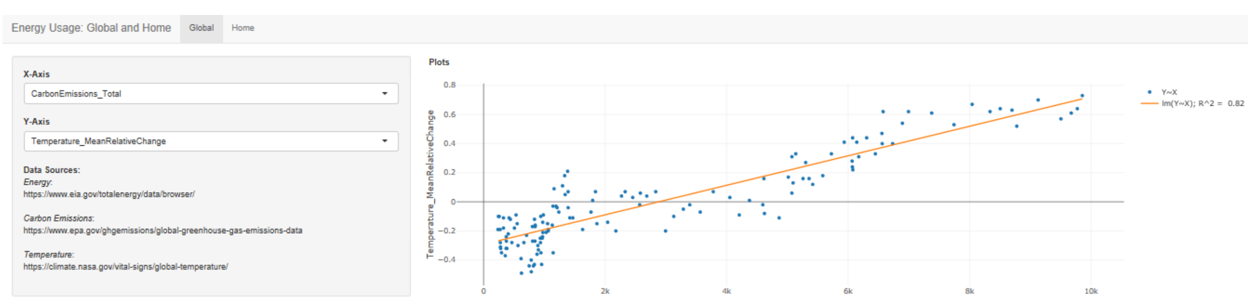
The global data is primarily limited by the years available. While the carbon emissions data dates to 1751, the global temperature data is not available until 1880. The energy data isn't available until 1949. Consequently, comparison among these variables is limited by the variable with the fewest number of data points. However, while limited in the number of years of data available, the global data does have useful breakouts by type of carbon emissions and type of energy production and consumption. This allows for a more improved analysis to determine precisely which kinds of carbon emissions and which kinds of energy production and consumption are most associated with global temperature increases.

The home energy data is extremely detailed and there are lots of data points. The weather data is complete and could only be improved by having the data at an hourly or minute level of granularity. The biggest limitation is in the energy usage data. While I have the totals for given points in time, I cannot directly observe which appliances are using the energy. For example, I have no data showing the energy usage breakouts for the refrigerator, the heaters/AC, the hot water tank, or the stove/oven. I can infer which appliances are in use based on the day and time of the energy usage and my knowledge of our family's home appliance and electrical usage, but I cannot directly and conclusively show where the energy is being used the most.

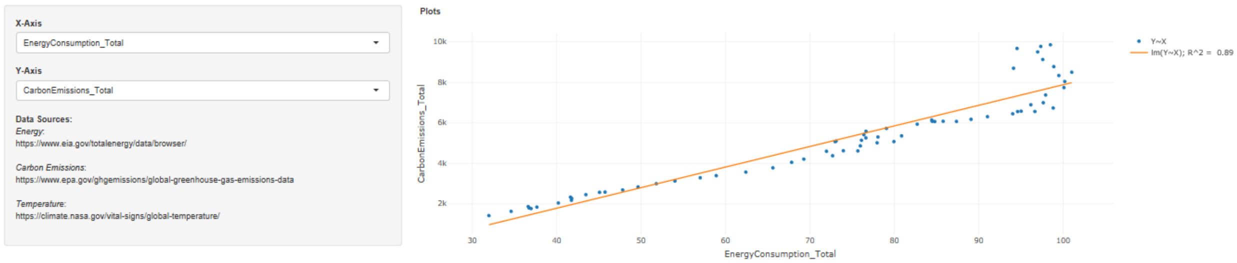
Analysis: Visual

Global

The data shows that there is a strong correlation between carbon emissions and increasing average global temperatures (in my visualization, a simple linear model returns an R^2 of 0.82). This supports the widely accepted view that carbon emissions are the cause of increasing average global temperatures.

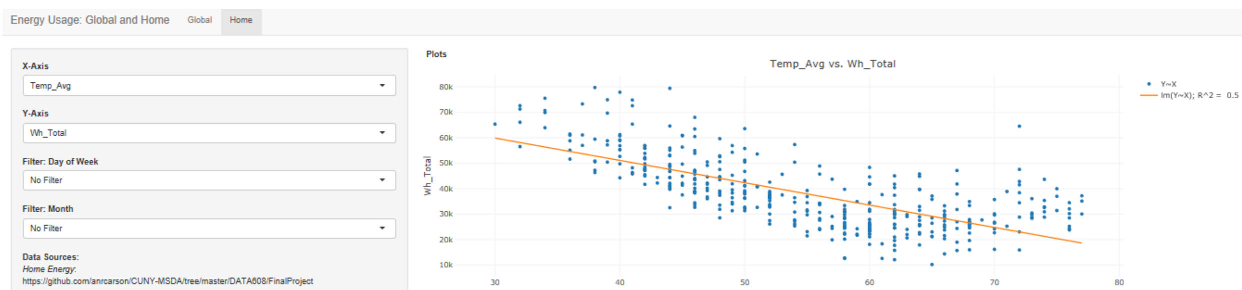


Those carbon emissions are (presumably) the direct result of energy production and consumption (my simple linear model returns an R^2 of 0.89). Energy production largely consists of burning fossil fuels which produces CO2 emissions.

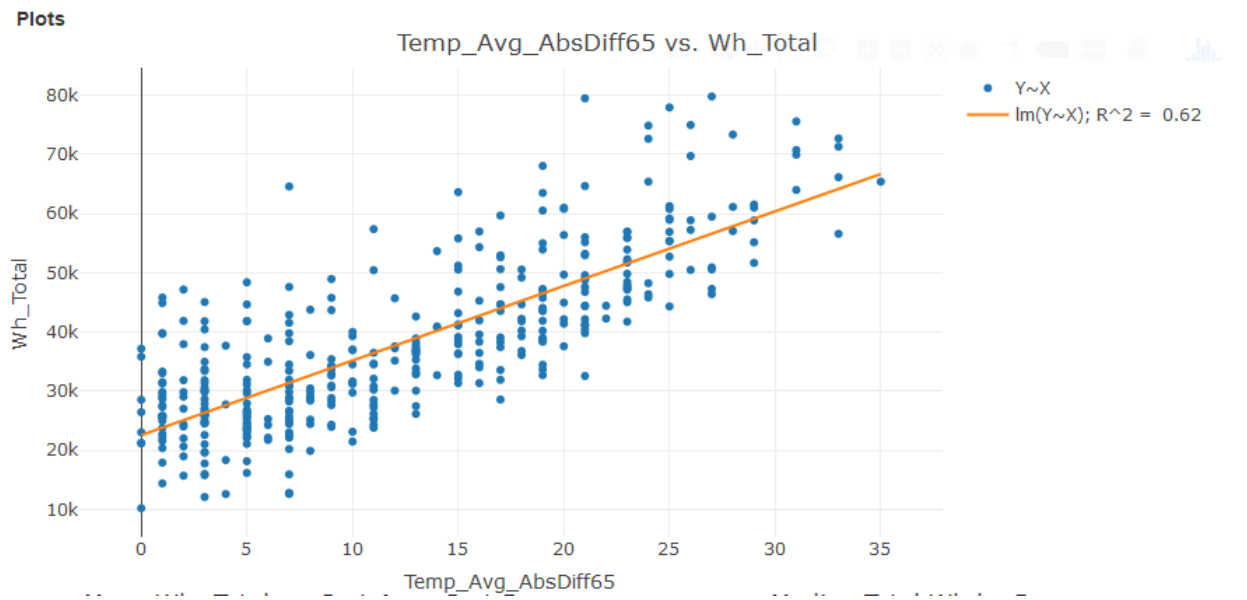


Home

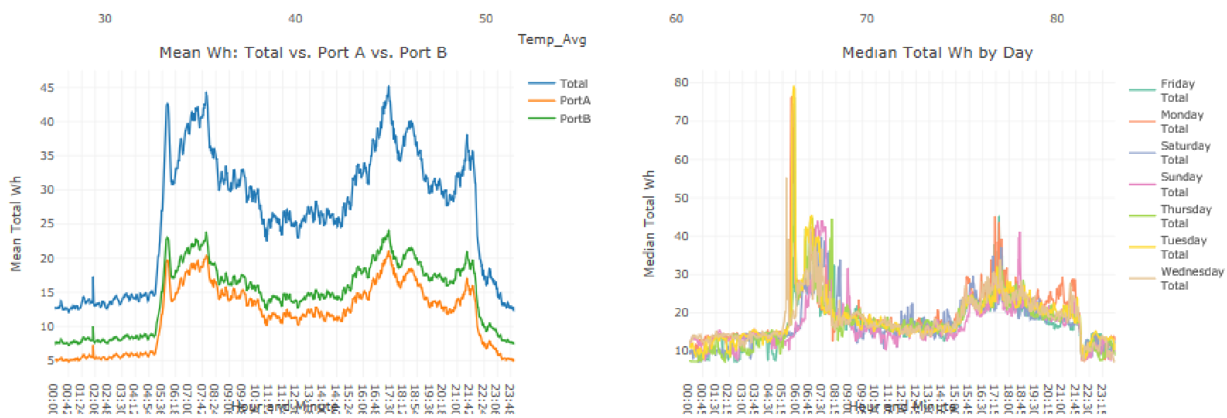
To understand energy usage at the local home or business level, I used my own home energy usage over the course of a year and related that information to the local weather, which I assumed to be among the major factors associated with energy usage. Based on the visualization below, I can immediately see that there is a rough negative linear relationship between temperature and energy consumption (lower temperature = higher energy usage; R^2 = 0.5).



However, knowing that I have air conditioning, I can see that energy usage goes up a bit when the temperature is above 70 degrees and the air conditioning is turned on. Thus, a main energy consumer in my home are heaters and air conditioning, where increasing energy usage is associated with more extreme (hot or cold) temperatures outside. This can be more easily shown comparing the absolute temperature difference between the daily temperature and 65 degrees with total energy usage. A simple linear model returns an adjusted R^2 of 0.62.



Another source of energy consumption can be inferred from these charts:



The highest peaks are in the morning, followed by a lull midday with some lesser peaks around dinner time. The peaks correspond to water usage, more specifically, hot water usage, and kitchen appliance usage. The morning peaks correspond to daily showers while the dinner time peak corresponds to hand washing dishes, bath-time for the kids, running the dishwasher, and on certain days, doing laundry. The stove and oven are also primarily used at dinner time.

Analysis: Regression

Global

When predicting global mean relative change in temperature, a linear model that considered the following variables returned a model with an adjusted R^2 of 0.88:

- CarbonEmissions_GasFuel
- CarbonEmissions_LiquidFuel
- CarbonEmissions_SolidFuel
- CarbonEmissions_CementProd

- CarbonEmissions_GasFlaring
- EnergyConsumption_Nuclear
- EnergyProd_Nuclear
- EnergyStockChangeOther
- EnergyConsumption_Fuels
- EnergyProd_Fuels
- EnergyConsumption_Renewable
- EnergyProd_Renewable

Using StepAIC to automatically reduce the variables produced a model with an adjusted R^2 of 0.89, and contained the following variables, each of which was statistically significant at 0.05 or lower:

- CarbonEmissions_LiquidFuel
- CarbonEmissions_CementProd
- CarbonEmissions_GasFlaring
- EnergyConsumption_Renewable

Liquid fuel and cement production emissions were positively correlated with global mean temperature relative change, while gas flaring and energy consumption from renewables were negatively correlated.

A model that considered only energy consumption and production in predicting global mean temperature relative change yielded a model with an adjusted R^2 of 0.87. The variables that remained in the model after step AIC were:

- EnergyConsumption_Nuclear
- EnergyStockChangeOther
- EnergyConsumption_Fuels
- EnergyProd_Renewable

Energy consumption from nuclear sources was positively correlated, along with energy production from renewables. Stock change and consumption from fuels were negatively correlated. A model only considering total energy consumption and production yielded a model with an adjusted R^2 of 0.70, where after StepAIC, only the consumption variable remained. It positively correlated with global mean temperature relative change. However, due to missing energy data prior to 1950, the small number of data points should make us cautious in taking this model too seriously.

A model that only considered carbon emissions in predicting global mean temperature relative change yielded a model with an adjusted R^2 value of 0.85. After StepAIC, each of the following variables was statistically significant at 0.001:

- CarbonEmissions_GasFuel
- CarbonEmissions_LiquidFuel
- CarbonEmissions_CementProd
- CarbonEmissions_GasFlaring

Emissions from gas fuel and gas flaring were negatively correlated, while emissions from liquid fuel and cement production were positive correlated with mean temperature relative change. A model only considering total emissions produced a model an adjusted R^2 of 0.82, where total carbon emissions were positively correlated.

To summarize, in comparing energy production and consumption to carbon emissions as factors in predicting global mean temperature relative change, it appears that carbon emissions are more statistically significant. This should not surprise us as most energy production and consumption produces CO_2 emissions and thus energy production and consumption will be highly correlated with increasing global mean relative temperatures. However, energy production and consumption are presumably the remote causes of global warming, while carbon emissions are presumably the proximate causes of global warming. Thus, it should not surprise us that carbon emissions, being more directly related, should be more statistically significant in predicting to global mean relative temperature.

Home

When predicting total Watt hours (Wh_Total) for a given day at my house, a linear model consisting of the following variables returned a model with an adjusted R^2 of 0.68:

- DayOfWeek
- Month
- Temp_Max
- Temp_Min
- Temp_Avg
- Temp_Avg_AbsDiff65
- DewPoint_Avg
- DewPoint_Max
- DewPoint_Min
- Humidity_Min
- Humidty_Max
- WindSpeed_Max
- WindSpeed_Min
- Pressure_Max
- Pressure_Min
- Precip_Avg

. Most variables were not significant, although the following were:

- DayOfWeekMonday - positively correlated and added much more than other days. This is our primary laundry day.
- Months: August, December, January, February – the hottest month and the coldest months in which the AC or heat is working the hardest. All were positively correlated and had much higher magnitudes than other months.
- Temperature: while the min and average were significant, the most important variable was Temp_Avg_AbsDiff65, that is, the absolute difference between the day’s average temperature and 65 degrees. I had determined 65 degrees to be the temperature at which the AC/heater had the lowest energy usage, so any deviation above (AC) or below (heat) would be drawing more energy.

After having binarized the character variables, I ran StepAIC on the above model to reduce the variables to those that were most significant. A model with an adjusted R^2 of 0.69 resulted. Most significant were:

- Temp_Avg_AbsDiff65 – positively correlated
- DayOfWeek_Monday – positively correlated and adding a large amount in magnitude.
- Months: September, October, March, April, May: negatively correlated, accounting for the milder weather and reduced need for heating and cooling (and thus, less need for energy usage).

Finally, to remove any variables highly correlated with each other (e.g., Temperature average vs temperature min vs. temperature max), I selected only those variables that were most significant in each grouping (temperature, windspeed, pressure, dewpoint, humidity). The result was a model whose variables were all statistically significant at 0.1 and most were significant at less than 0.001. With an adjusted R^2 of 0.68, the model contained the following variables:

- Intercept: 31,972 watts per day (positive correlation)
- Temp_Avg_AbsDiff65: 738 watts added per degree away from 65 degrees
- Dewpoint_Avg: -395 watts added per degree
- Humidity_Max: 197 watts added per humidity percent
- Days:
    - DayOfWeek_Monday: 5637 watts added
    - DayOfWeek_Wednesday: 1995 watts added
    - DayOfWeek_Saturday: 2125 watts added
- Months
    - Sep-Nov: negatively correlated, with October contributing -9131 watts to the model.
    - March-Jun: negatively correlated, with March and May contributing -8366 and -8026 watts respectively.

Consequently, I see these two major factors in predicting energy usage at my home:

- Temperature/ weather (including dewpoint, humidity, and month)
- Hot water usage (inferred from day of week):
    - Monday – primary laundry day
    - Wednesday; Saturday – primary bath days for kids

Conclusion

The impact of increased global temperatures is a matter of environmental concern and of social justice. Consequently, one should care about energy production and consumption in relationship to global warming, for if one can reduce energy production and consumption (at least where such production and consumption comes from CO2 emission producing processes), one should be able to reduce one's impact on global warming.

What can a responsible average person do to reduce global warming? A straightforward thing to do is to reduce one's energy consumption, thus reducing the need to produce that energy by burning fossil fuels that produce CO2 emissions that contribute to global warming. How can one reduce one's energy consumption? One must first understand where that energy consumption is largely coming from and what factors play into increasing or decreasing energy consumption. Once understood, one can determine how best to reduce one's energy consumption by targeting these contributing factors.

Understanding these factors is what I have attempted to do in this project. While a more detailed analysis could be undertaken to more specifically understand the causes of energy consumption in my home, it seems that the primary sources of energy consumptions come from heating/cooling the air inside the house (related to the temperature outside) and heating the water (related to hot water usage inferred from day and time of energy usage).

Having identified these factors, these can be targeted for reducing energy consumption in my home through:

- Better insulation, which would counteract the need for heating and cooling the home through heaters and air conditioners. This would make the energy usage less dependent on the outside temperature.
- More efficient appliances, especially, a hot water tank, that use less energy to produce the same result (e.g., hot water, a heated stove and oven).

With these relatively simple home improvements, energy consumption could be reduced, thus lowering the need to produce that energy by burning fossil fuels which produce CO_2 emissions that contribute to global warming. While such measures would have no noticeable global impact if only done by myself, such measures would likely have an impact if implemented by most homeowners and business owners in the world such that a significant reduction in energy consumption (and hence production) was achieved.