## Introduction

As part of the EPA, our objective in this report is to analyze the energy generation data to find opportunities for optimization in the allocation of resources, trends in energy consumption, types of fuels used in the overall energy mix, propose regulations on different technologies and identify opportunities for emission reductions.

## Dataset

We have been provided with data for power plants in the US for 2010. The Key information provided by this data are:

a) Power Generated
b) Fuel type used
c) Amount of heat input required
d) Emissions (CO2 equiv. emissions, SO2 emissions and NOX emissions)

During our analysis, we found the data to be of high quality with very few errors or omissions. The data provided was for 5393 power plants across the United States. This helped us understand the energy mix used in various states as well. Although the data provided 38 different variables for the power plants, we needed to calculate the following additional values for our analysis:

a) Plant Capacity Factor: This gives us insight on the portion of the plant's total capacity that is being utilized over a year.
b) Plant Nominal Heat Rate (Heat Rate): This tells us how efficiently a plant converts input energy to electricity.
c) Annual Output Emission Rate for different pollutants: It is difficult to compare two different plants in terms of which one is more polluting as the higher generation also leads to higher emissions (due to more fuel being burnt). We circumvent this problem by dividing the total annual emissions by the net generation of the plant over the year.

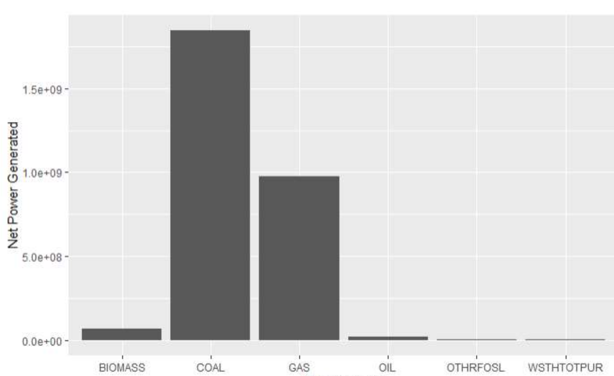
Finally, the data provides us information on plants that use combustion to generate electricity, non-combustion plants as well as partial combustion plants. This complicates our analysis as we need to compare efficiencies as well as emissions across different plants. So, in order to narrow our research, we first see the portion of energy produced through combustion plants. We see from our data that 71% of total US electricity is generated through combustion plants. So, we will remove the non-combustion plants from our dataset analysis.

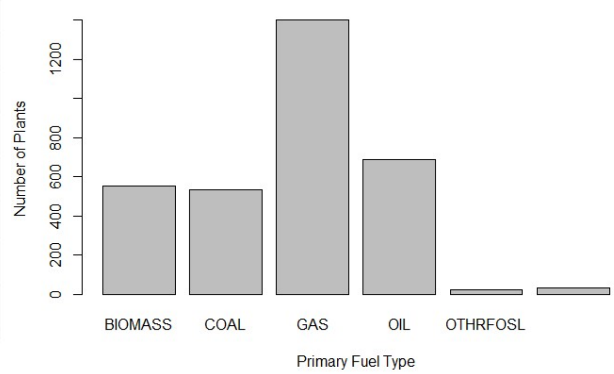
Next, we see that among the combustion plants, only 2.9% of the energy is produced through partial combustion plants. Therefore, we exclude them from our analysis as well.

## Analysis

Now, that we have prepared our data for analysis, we will start with our analysis. Our data provides information on the Primary fuel used in each of the power plants under the FuelCat variable. There are 11 such primary fuels that contribute to the US total combustion power mix: Coal, Oil, Gas, Biomass, Other Fossil Fuels, and Other Unknown/Purchased/Waste Heat.

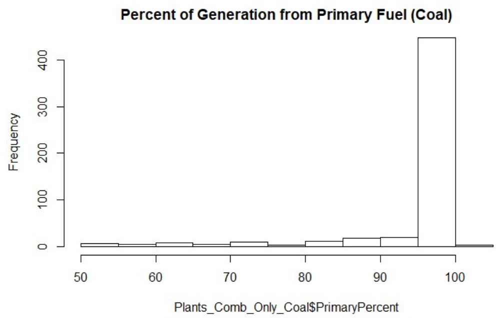
The distribution in this energy mix is as follows:

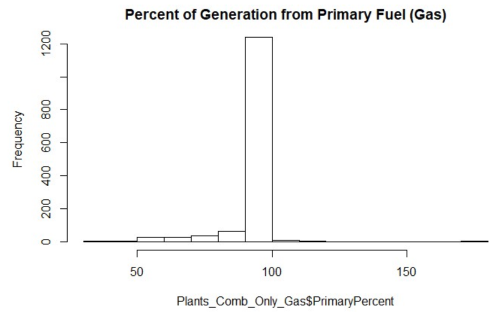


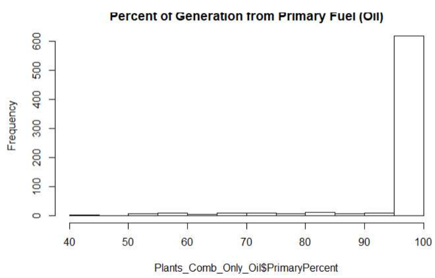



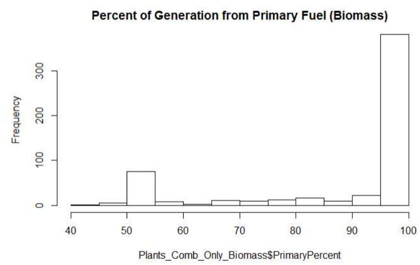

As we can see, the Primary fuels of BioMass, Coal, Gas and Oil make up a large percentage of the total combustion plants. Also, most of the power in the US is generated through Coal power plants, followed by Gas plants. However, if we look at the total number of power plants, the number of Gas plants is more than any other type. In fact, it seems that the coal power plants provide the lion's share of our energy mix while accounting for a fifth (533) of the total number of combustion plants (3233).

Now, each plant is indicated by a primary fuel source, but there are still smaller portions of generation associated with other fuels. We will now perform an analysis to show that the primary fuel source is a large fraction for large percentage of plants.



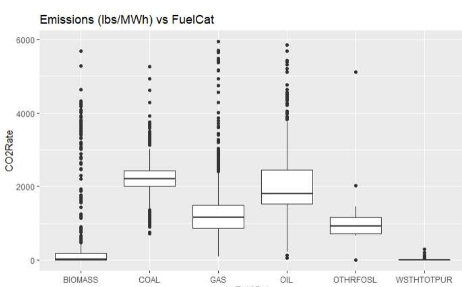



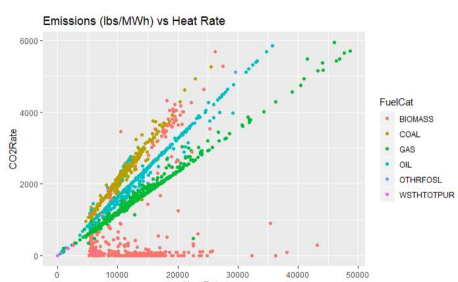



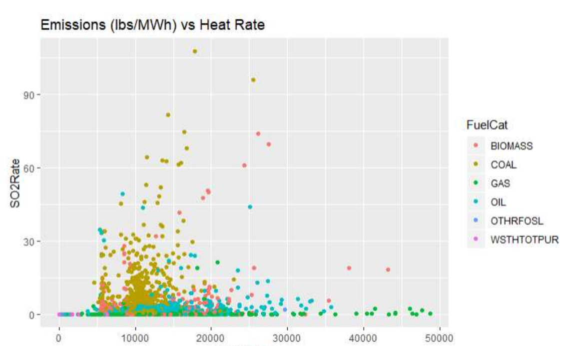



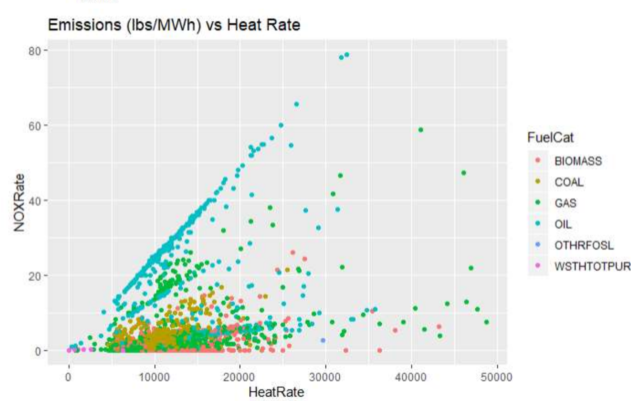

Based on the above, we can see that the majority (close to 100%) of power is indeed produced by the primary fuel. It thus is reasonable to study the data by groups of primary fuel category, keeping the small fraction of mixed fuel sources in mind.

Now, we will characterize emissions and performance of plants. We start this by analyzing emissions by fuel type to understand if there is any correlation. Let us compare the emissions rate vs the heat rate of plants. This way we can compare different primary power sources in terms of emissions.



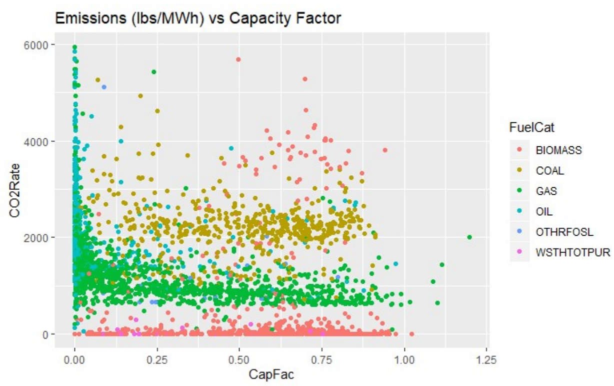



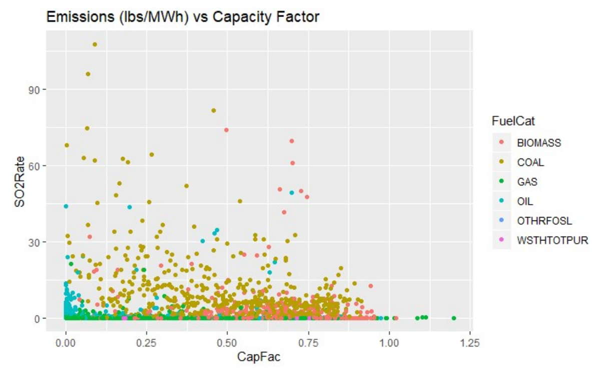



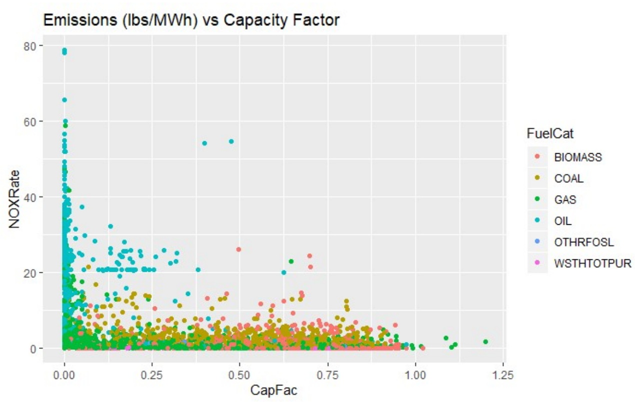
Based on the above plot, we can see that as the heat rate increases, the emissions also increase. This is probably because we are inputting more heat to get the same Net Generation, so we are losing more heat to waste. From this, we can say that low efficiency leads to higher emissions.

Another key observation is that for a given heat rate, mean emissions rate (per MWh generated) is highest for the coal plants as compared to other fuel types (except in case of Oil plants, where NOX is highest). The outliers present an interesting insight as well – there seem to be many outliers for the Biomass sector, so some standardization may be required on these power generators.

Let us compare the emissions rate vs the capacity factor of plants. This way we can compare different primary power sources in terms of emissions.



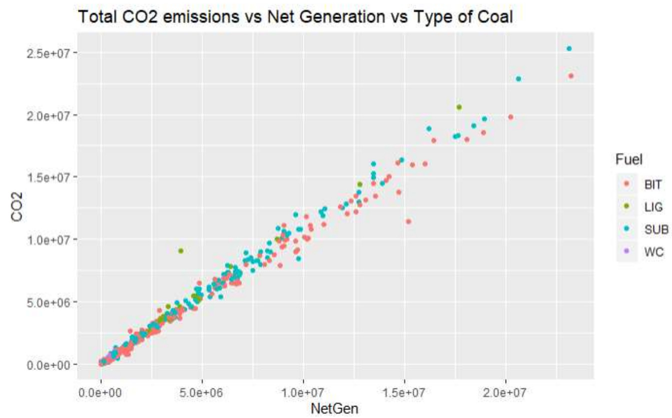



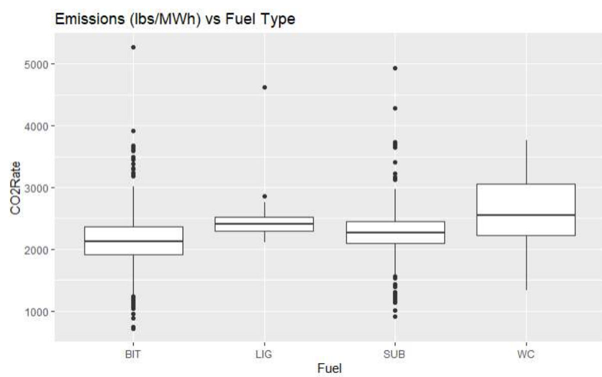

Based on the above plot, we can see that as the emissions generally remain the same with increasing capacity factors. This is probably because we are higher capacity factor plants are also high net generation plants, so the efficiency gained is cancelled out by the sheer mass of emissions. We also see that for the same capacity factor, Biomass tends to have higher emissions than coal, which has higher emissions than Gas.

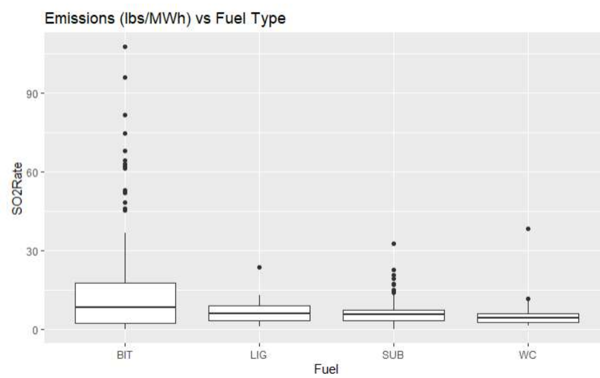
Let us now look at specific fuel categories and try to find other drivers for emissions

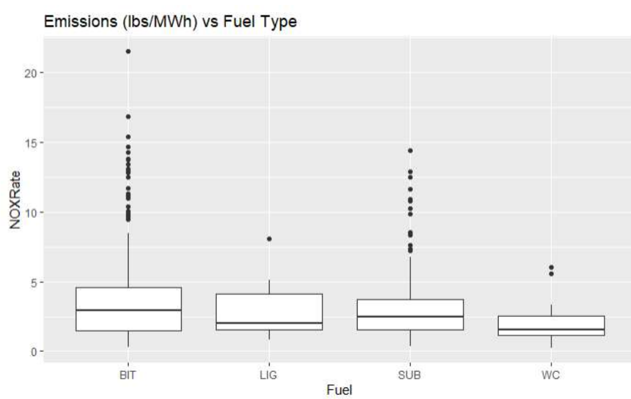
1. Coal Plants




An important insight from this plot is that Lignit Coal seems to have the highest CO2 emissions as compare to Bitumen. However, once again this could be because there are more Lignit coal plants. So, We need to divide the emissions with the energy generated and plot that for different types for plants. However, we see a few outlier plants that are way out of scale. We should probably look into these plants. After checking the dataset, we see that same outlier is way out of scale for all types of emissions. So, we will disregard this outlier to see any relation between types of coal used in a plant vs emissions.

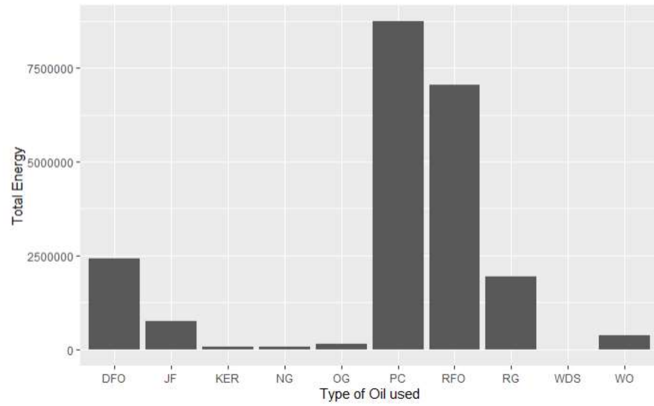


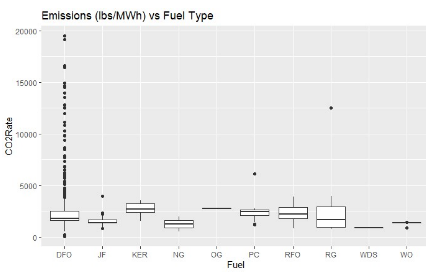

So, now that we look at the emissions per energy generated, we see that Bituminous coal plants have more SO2 and NOX emissions per unit of energy generated than Lignit coal, while having less CO2 emissions than Lignite. So, type of coal is definitely a factor in the emissions. This also represents a trade-off in CO2 emission s vs. SO2 and NOX emissions in determining the type of coal to be used.

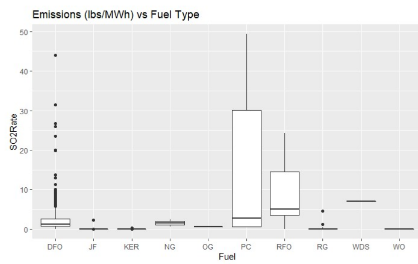
2. Gas Plants - Almost all gas plants used Natural Gas as their fuel, so unlike the coal plants, we do not need to analyze the distribution of emissions based on type of Gas. However, we do see that total CO2 emitted increases with increase in power generated by a plant, which is to be expected.

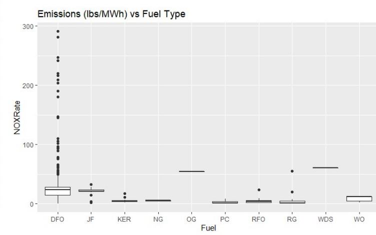



3. Oil -

An important insight from these plots is that there seems to be a significant spread of fuel (Diesel Fuel Oil (DFO), Petroleum Coke(PC) and Residual Fuel Oil / Petrol (RFO) and Regfinery Gas (RG)) used among various plants contributing to total power mix. So, we will analyze if there is some significance in the CO2 emissions rate per type of fuel.

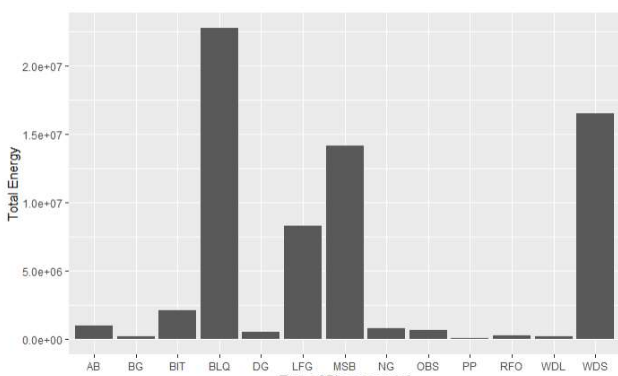


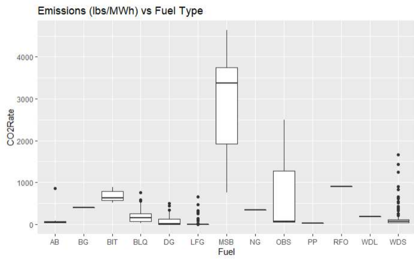



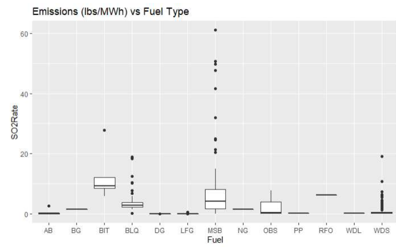

From the above plots, we see that there is a huge variety of generators run by Diesel oil that have vastly different emission patterns. As they form a substantial portion of the energy mix, we should target this category of fuels through future policy decisions.

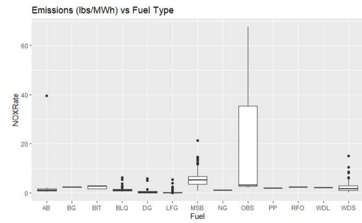
4. Lastly, we check the above for Biomass fuels as well:




We see that the majority of Biomass energy mix is by 4 types of Biomass - Black Liquor, Landfill Gas, Municipal Solid Waste Biomass and Wood. Out of these Landfill Gas and Municipal Solid waste generate the most CO2.




From the above, we can clearly see that MSB is the biggest culprit in high CO2, SO2 and NOX emissions.

Conclusion

In conclusion, we see that the main factors that we should target to optimize the energy mix as well as reduce emissions are:

a) Switch from Coal and Oil generation to Gas (more capacity factor and less CO2 equiv., SO2 and NOX emissions)
b) Bituminous coal plants have more SO2 and NOX emissions per unit of energy generated than Lignit coal, while having less CO2 emissions than Lignite. So, type of coal is definitely a factor in the emissions. This also represents a trade-off in CO2 emission s vs. SO2 and NOX emissions in determining the type of coal to be used.
c) There is a huge variety of generators run by Diesel oil that have vastly different emission patterns. As they form a substantial portion of the energy mix, we should target this category of fuels through future policy decisions.
d) We need to reduce the use of Municipal Solid waste as a Biomass fuel as it is has a very high rate of emissions compared to other biomass fuels.