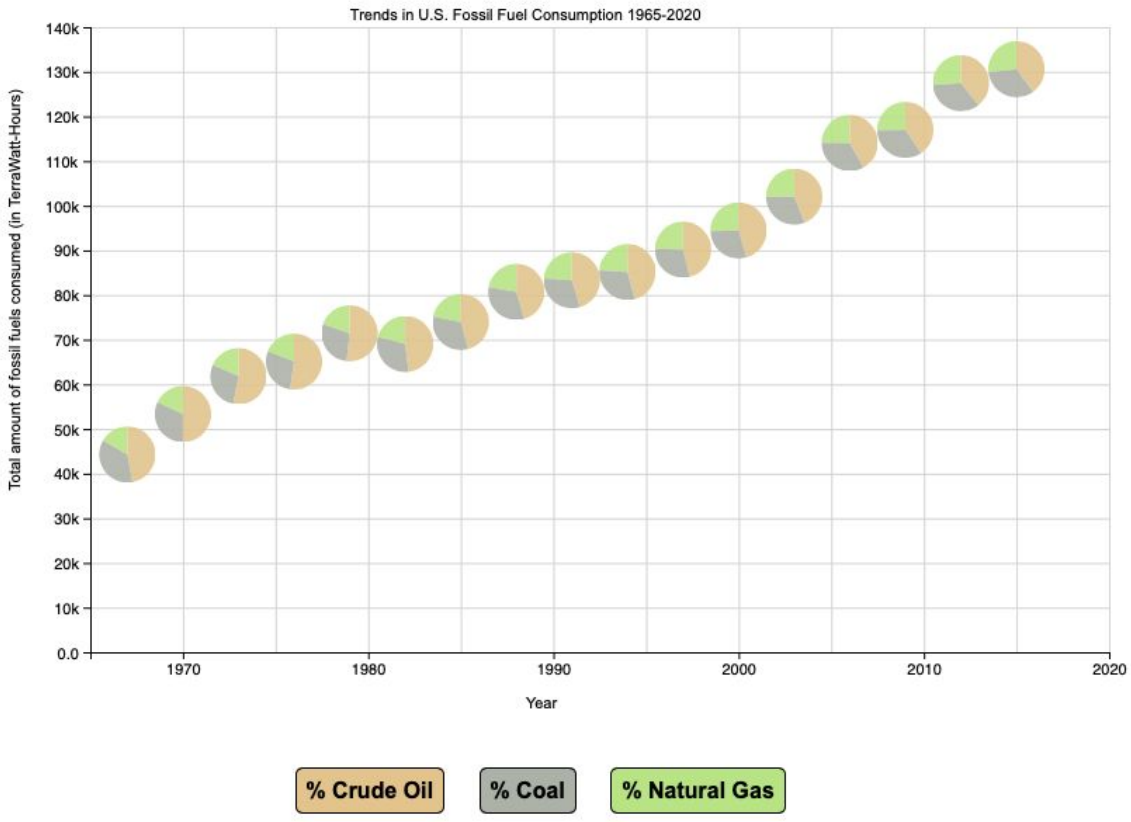
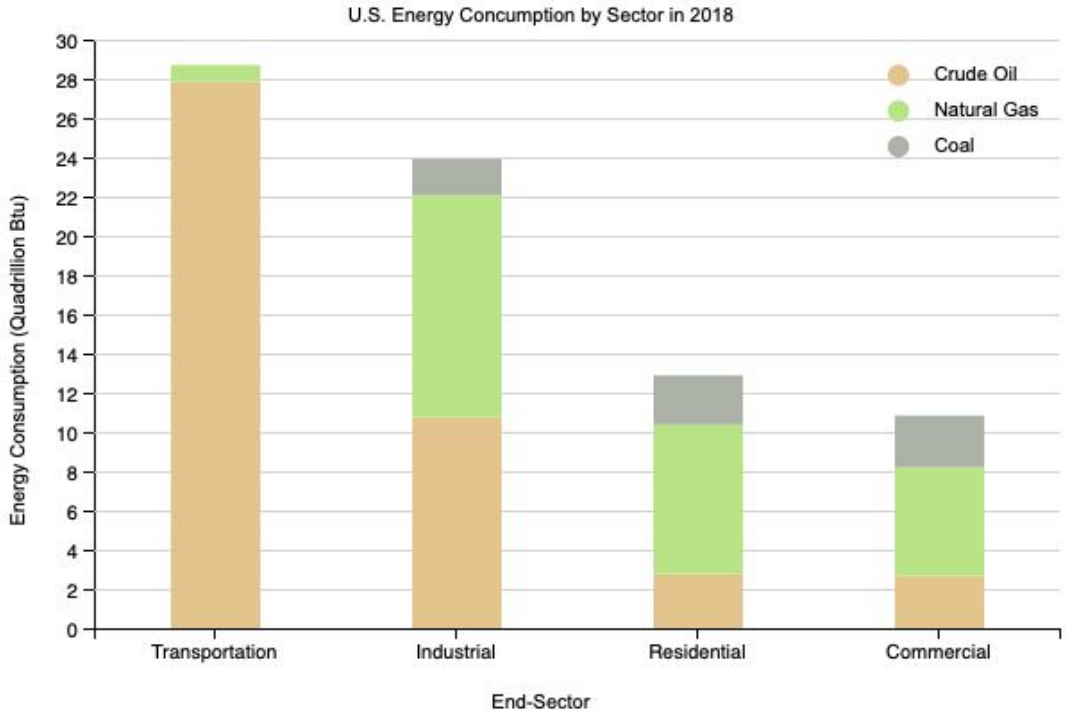
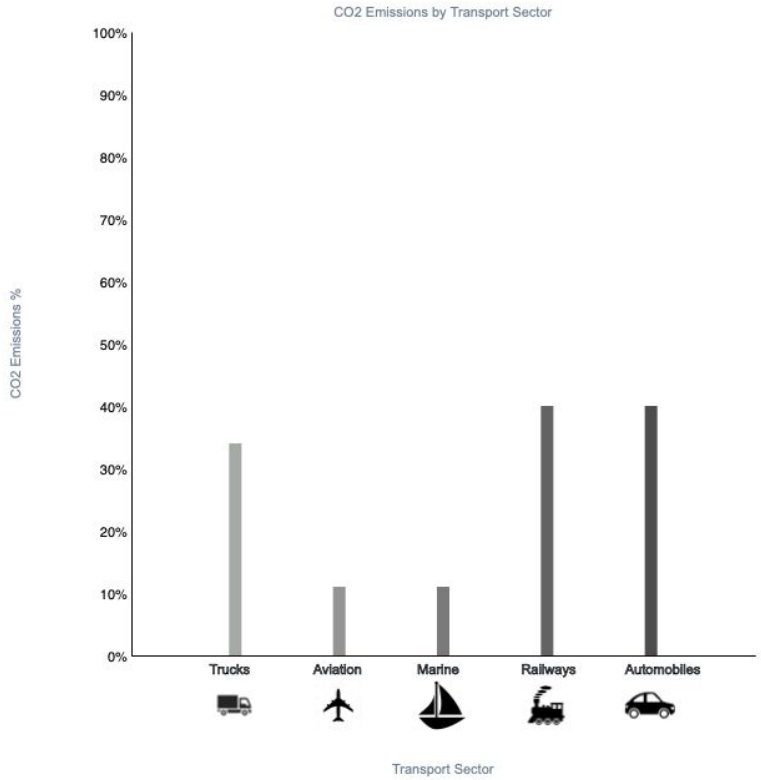
Group: Chris (co253), Katie (kvs46), Kartikay (kj295)
CS3300 - Project 1
3/3/20

Project 1 Report

Our final visualization consists of three separate plots focusing on fossil fuel consumption and greenhouse gas emissions. It can be found in the *index.html* file in the root of our repository at https://github.coecis.cornell.edu/kj295/INFO-3300-Project-1. Images of each of our plots are below.




U.S. Energy Concumption by Sector in 2018
Energy Consumption (Quadrillion Btu)
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0
Crude Oil
Natural Gas
Coal
Transportation
Industrial
Residential
Commercial
End-Sector


CO2 Emissions by Transport Sector
CO2 Emissions %
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Trucks
Aviation
Marine
Railways
Automobiles
Transport Sector

Description of Data

For the first plot in our visualization, the data was obtained from https://ourworldindata.org/fossil-fuels. This data is about fossil fuel consumption, specifically how the total consumption amount has changed over the past 50 years, and how what the percentages are from what resources this total is made up. This dataset is in *data/global_consumption.csv* and is the sole source of data for the first plot. I filtered this data solely by only using values from 1965 though 2017 and only picking every third point, so that the pie-chart points don't overlap. This filtering is all done in Javascript.

While our first plot gives viewers an insight into the issue of increased fossil fuel consumption in the United States. But for a lot of individuals it's mysterious what that is actually going toward. So to make the insights more relevant for the majority of people, I found data relating the consumption of energy sources to the end-sectors those are contributing to. This data is from the recent year of 2018, which is two years ago but I concluded would probably have indiscenerable differences from the current year. I found this dataset at https://www.eia.gov/totalenergy/data/monthly/.

They provide a lot of data on the breakdowns of every year but I was only concerned with fossil fuel, coal, and natural gas, as well as the standard end-sectors of Transportation, Industrial, Commercial and Residential. They had a table and infographic that compiled this data into usage percentages here: https://www.eia.gov/energyexplained/us-energy-facts/ . The infographic was a little convoluted but I took the information from the table and was able to hardcode that dataset into our index.html file to make the U.S. Energy Consumption 2018 graphic. The only editing that needed to be done was converting the percentage of energy that goes from coal to the electricity power grid to an end-sector like residential to be just a transfer of energy from coal to the residential sector. This required simple math of converting the percentages into Quadrillion Btu and computing new percentages.

For the third plot showing carbon dioxide emissions vs the several industries in the transportation sector, we got the data from the following source: https://transportgeography.org/?page_id=15778. This source shows how in percentages CO2 emissions is varying among marine, airlines, railways, automobiles and truck industries. Filtering this data was very simple as we could just create y-scale that varied from 0% to 100% which represented the CO2 emission percentages and then on the x cale plot out for each industry how much CO2 emissions were made. Originally the idea was to have percentages for all greenhouse gases however there weren't that many reliable sources on certain greenhouse gas emissions such as methane in certain industries so we decided to stick with CO2 emissions as the data is widespread and very reliable.

Design Rationale

We decided initially to break out visualization into multiple plots for two reasons -- 1) because it would allow each of us to focus on one significant aspect, but 2) primarily, it would allow us to convey enough detail that would otherwise be hard to fit into a single plot. There's a lot of data about this topic and it would have made for a crowded visual.

For the first plot, I (Chris) wanted to show not only how fossil fuel usage has increased, but how it's broken up by source. I thought that using small pie charts instead of just circles was a great way to do this. This way, the reader can take a step back and look at the position of each pie, or, if they'd like, they can focus on how each pie chart is split up. I think these two separate groupings of visual channels were a great way to convey both datasets without overwhelming the reader. I chose the lighter colors for the pie chart solely because they are easy to look at, easy to distinguish and follow conventional representations of fossil fuels. I used the d3 pie() and arc() functions -- as scales essentially -- to make these charts out of the data. I used linear scales for x/y positioning of the charts themselves. The marks in this plot are the pie charts themselves, and visual channels include the area of pie slices, color of pie slices, and the aligned vertical/horizontal position of each pie chart.

For the second plot, I focused on building off the information portrayed in the first graphic. Therefore I continued the same color scheme to represent the same three fossil fuels of crude oil, natural gas and coal. I represented the breakdown of fuel consumption per end-sector with a bar graph because lines are a visual channel that is easily comparable to viewers. That way, someone looking at the visualization will quickly be able to comprehend what sector consumes the most, and how much of each fossil fuel it consumes. I added only horizontal lines in light grey so that nothing was too distracting in the background, while still providing the ability to get a clear picture into exactly what the consumption for a sector was. The labeling of the axis follows that of the first and third graph to maintain consistency. The marks of this graph are the line elements making up the graph. The channels include the varying aligned length and horizontally aligned position of the bars.

For the third plot (i.e the plot of CO_2 emissions vs transport sector industries) I (Kartikay) used a bar chart to distinctly show how CO_2 emissions are varying amongst the different industries. The marks hence were the bars (the rectangular svg elements) and the visual channels employed were the color hue difference in grey and black. CO_2 is a pollutive chemical and so grey is often associated with pollution so I decided to vary that color channel. In order to vary it I actually went with how impactful the industry is to greenhouse gas emissions. Although aviation and marine and truck are definitely causing global pollution the most important stakeholders are in fact railways and automobiles. This is actually because a vast amount of the world population actually functions with railways and automobiles (cars/buses). Daily many people commute from place to place using a train or a car compared to a ship or a plane or even a truck (which is used to carry cargo). Hence by darkening those

bars I wanted the viewer's attention to be drawn to it as it stands out more against the white background compared to the lighter grey options of marine, planes and trucks. I used a linear scale as we were dealing with percentages and I tried to keep the plot simple so it is a fast and easy visualization for viewers to view through. If I considered too many other non important gases like NO2 as bars too then the viewer's attention would be diverted from the main greenhouse gas which is CO2 (sometimes simple is better).

Visualization *Story*

The focus of our visualization was to provide some detail and scientific context to the ongoing discussion of greenhouse gas emissions and fossil fuel consumption. We all hoped to use our designs to convey just *how* reliant we are on fossil fuels and just *how* much greenhouse emissions were resulting from our fossil fuel use. The first plot not only conveys how our reliance on fossil fuels has increased over the past 50 or so years, but how much of each of the subcategories of fossil fuels we use also. In this plot, it was surprising how the relative percentages of each subcategory hasn't seriously changed over the years.

The second plot dives deeper by providing a comprehensive view of how that energy is used. We can see that Transportation and Industry are by far the highest consumers. The results can be surprising because one might typically think the modes of production would be the highest energy consumers. But instead, Transportation comes out on top. This demonstrates how far our demand for travel and convenience might affect the environment. A potential viewer would learn the best ways to cut back on energy consumption would be through those top users.

In the case of the third plot, we dive deeper into the sector with the greatest usage and learn that Railways and Automobiles lead the CO2 pollution race which makes sense as they are the most used civil transport. What really was interesting was that Trucks also had a big impact in CO2 pollution as it was the next big contributor from the plot. This could convey to the viewer that a lot of cargo transport takes place such as Amazon deliveries, and other item deliveries around the world which also contributes to a lot of pollution. Aviation and Marine don't populate much C02 as they are less frequently used.

Work Distribution

We basically broke the work up strictly by each putting together a different plot for our final visualization. Each of our plots has a different specific focus, but they all aid in conveying the theme of increasing fossil fuel consumption and greenhouse gas emissions. Specifically:

- **Chris** worked on the first plot, taking the total amount of fossil fuel consumption over the years and breaking it down by category. I collected my own data, and sketched/coded this visualization entirely on my own. I spent roughly 20 hours in total on this project, with collecting and formatting the data taking the longest. Though, it did take some time learning how to make pie charts with d3.
- **Kartikay** worked on the third plot, C02 emissions vs the industry within the transport sector. I originally tried collecting data for other gases as well but they were unreliable and barely available so I then made the visualization focusing on C02 gas emissions which is the main greenhouse gas and the plot took me the weekend to make.
- **Katie** worked on the second plot, U.S. Fossil Fuel Consumption broken down by sector. I spent a considerable amount of time searching for data that would connect sets that display how much is being used and the emissions resulting from that to where that energy is going. From there I worked in putting together a visualization that would tie all of our ideas together and display the three of them in a format that would get our visualization story across well.