Executive summary

# Introduction

## Background and Motivation

This visualization aims to provide information about global energy production and consumption for the general public as well as the energy domain experts. The reason for our choice of energy topic is due to the fact that although energy has brought a wide range of benefits to our daily social lives, its manufacture and consumption have led to a catastrophic environmental crisis. According to the International Energy Agency (IEA), the power sector accounted for nearly two-thirds of global emissions growth in 2018 (World-nuclear 2021).

In order to assist society in moving towards a more sustainable future, my team will create visualizations about the energy use profile around the world. The infographics not only enable individuals to understand the energy consumption profile, but it also helps them to recognize the severe environmental consequences caused by power use. This in turn makes way for appropriate actions to be undertaken to reduce the amount of energy consumption.

## Project Objectives

The primary objective of this report is to demonstrate the following energy data to the audiences:

- The amount of global energy used and how it was distributed in different sectors in the community in 2019.

- The total energy consumption by different regions around the world in the 10-year period from 2010 to 2019.

- The amount of CO2 emissions generated from different types of energy use (e.g. coal, oil and gas combustion) in the 5-year period from 1960 to 2020.

These visualizations are beneficial for the scientists as well as people interested in energy domains and its impact on the environment. As a result, they could answer the following questions respectively to the above topics:

- How is the energy used in the community?

- How different countries around the world make use of the energy?

- What is the environmental impact of different types of energy ?

## Project Schedule

To finish this project which includes the interactive website and Process Book, both of us have agreed on the schedule below:  
Week 3-6:

Decide what kind of graph should appear on the website and start to sketch the design idea.

Week 6-7:

Collecting Data from different sources

Deciding which datasets are appropriate for the visualization.

Start writing the Process Book.

Week 8:

Finishing the Data process part in process book

Week 9:

Start writing code for Visualization

Week 10:

# Data

## Data source

### Energy demand per sector

For the visualization of energy demand by sector, we use the dataset from IEA (International Energy Agency) (Figure 1.). This dataset contain data on the analytic information of energy demand for each sector and each type of products from 1971 to 2019. It is an analysis of information and trade-offs of different technology and policy choices, thereby providing quantitative data of energy demand for different sectors (IEA 2019).

Graphical user interface, application, table, Excel

Description automatically generated

Figure 1. Dataset from IEA

Below is the table of attributes that appear in this data set

|  |  |  |
| --- | --- | --- |
| Attributes | Note | Type |
| Country | Name of countries that involved in this statistic | Nominal |
| Product | Name of energy types | Nominal |
| Flow | The sector that each energy types distributed to | Nominal |
| Years | The year included in study period | Interval |
| Energy consumption | Amount of energy use measured in Pj | ratio |

Because we intend to visualise this data using the Sankey library. As a result, we must convert this data into a format that is compatible with this library.

### Primary energy consumption around the world

### In order to provide the infographics about the primary energy consumption in different countries around the world, the following dataset is going to be examined:

### - Primary energy consumption, 2020, Our World in Data:

### <https://ourworldindata.org/explorers/energy?facet=none&country=USA~GBR~CHN~OWID_WRL~IND~BRA~ZAF&Total+or+Breakdown=Total&Energy+or+Electricity=Primary+energy&Metric=Annual+consumption>

### 

### Figure 2. A part of dataset of primary energy consumption in different countries around the world

### This dataset is the table displaying the total amount of primary energy consumption (Twh) in each country in each single year. It consists of the following attributes:

|  |  |  |
| --- | --- | --- |
| Attributes | Description | Type |
| Entity | This attribute refers to different countries around the world | Nominal |
| Code | Each country is encoded by a code | Nominal |
| Year | The financial years that are reported and examined in the dataset | Interval |
| Primary energy consumption | The total amount of energy used in each country in the specified year. The unit measured is in terawatt-hour | Ratio |

### 

### - GeoJSON file for the location of all nations on the world:

### <https://github.com/johan/world.geo.json/blob/master/countries.geo.json>

### This json file will enable us to create the choropleth for the visualization of primary energy consumption in different countries.

### 

### Co2 emissions from different types of energy

For the visualization of the amount of Co2 emissions from different types of energy, we have decided to use 3 datasets from Statista. Those three were published by the same project which is Global Carbon Project (Global Carbon Atlas) and it contain data about CO2 emission from different types of energy (Coal, Gas Combustion and Oil) from 1960 to 2020:

* <https://www.statista.com/statistics/1198050/carbon-dioxide-emissions-from-coal-use-in-select-countries/> (CO2 emissions from coal)

Graphical user interface

Description automatically generated with medium confidence  
Figure 3. Dataset about CO2 emissions from Coal.

* <https://www.statista.com/statistics/1198084/carbon-dioxide-emissions-from-gas-combustion-in-select-countries/> (CO2 emissions from gas combustion)

A picture containing graphical user interface

Description automatically generated

Figure 4. Dataset about CO2 emissions from Gas Combustion.

* <https://www.statista.com/statistics/1198124/carbon-dioxide-emissions-from-oil-use-in-select-countries/> (CO2 emissions from oil use)

Graphical user interface, application, table

Description automatically generated

Figure 4. Dataset about CO2 emissions from Oil use.

Those three excel files provide us with tables of CO2 emissions from various types of energy in different countries such as Japan and the United States. Besides, the CO2 emissions were measured in million metric tonnes. And the attributes included in those two tables are shown in the table below:

|  |  |  |
| --- | --- | --- |
| Attributes | Note | Type |
| Year | The year included in the study of CO2 emissions from the mentioned type of energy | Interval |
| Country’s name | The CO2 emissions of that country measured in million tons | Ratio |

To obtain the correct data, we have to combine the information from those three tables, as we did in the first visualization. Then we need to remove some countries that do not appear in all three datasets.

## Data processing

### Energy demand per sector

As previously said, the goal of this sankey chart (flow chart) is to visualise energy demand in each industry in 2019. As a result, we must filter out the columns from the other year that are no longer required. Furthermore, because the library we're using requires a certain json file format. In particular, in the json file, we must contain objects including Links and Nodes. Nodes will comprise the names of various attributes to which energy flows. The source and target nodes, as well as the value of energy consumption allocated between them, will be included in the Links object. After filtering and transforming data we will have a JSON file below.

Graphical user interface, text

Description automatically generated

Figure 5. Data after transforming

Text

Description automatically generated

Figure 6. Data after transforming

### Primary energy consumption around the world

For reducing the data size when loaded into D3, the “Code” attribute which does not contribute to the data semantics, is deleted. Besides, the “Entity” column is renamed to “Country” to make it more interpretable for audiences. Below is the resulting dataset after applying those filters:

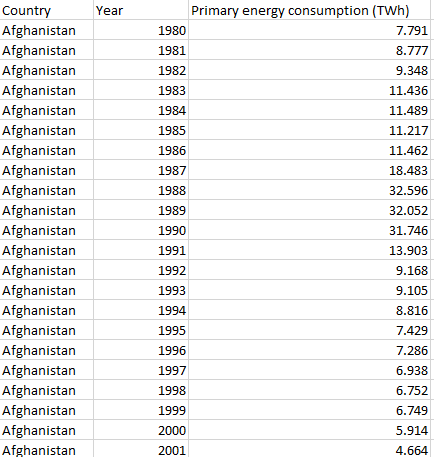


Figure 7: Primary energy consumption dataset after filtering

Furthermore, some “country” records are not actually a country. For example, there are records for “Africa” and “Asia Pacific”, which should be treated as a continent or region rather than a country.





Figure 6: Some “not-country” records in the dataset

Those records are then removed out of the dataset because it does not match with the chosen Geojson file as well as provide unnecessary information. Besides, some countries in the dataset do not match with the properties’ names in the Geojson file. For instance, the US is marked as “United States” in the dataset, whilst the json file marks it as “United States of America”. Thus, we have to change the property’s name of the US in json file to make it properly corresponds to its record in the dataset.

The dataset is intended to be visualized as a choropleth map based on different year choices. To be specific, the primary energy consumption in all countries (with available data) will be displayed in a choropleth according to the selected year. Thus, we have to group the data by year in D3 in order to make the processed data more organized.

var groupByYear = d3.group(data, function(d) {

return d.Year;

});

Figure 8: Code to group the dataset by year

### Co2 emissions from different types of energy

To extract the data needed for this visualization, we must integrate the three files about different energy kinds (Coal, Gas Combustion and Oil). To do this we have to use functions offered by both Excel and D3.

Except the datasets for Coal and Gas Combustion that contain the data of 5-year period from 1960 to 2020, the other one (dataset for Oil use) has data for each year from 1960 to 2020. Therefore, we have to filter unnecessary years in the dataset of Oil. To filter the rows related to unnecessary years, we can use the filter function in excel and choose appropriate years (Figure 9.).

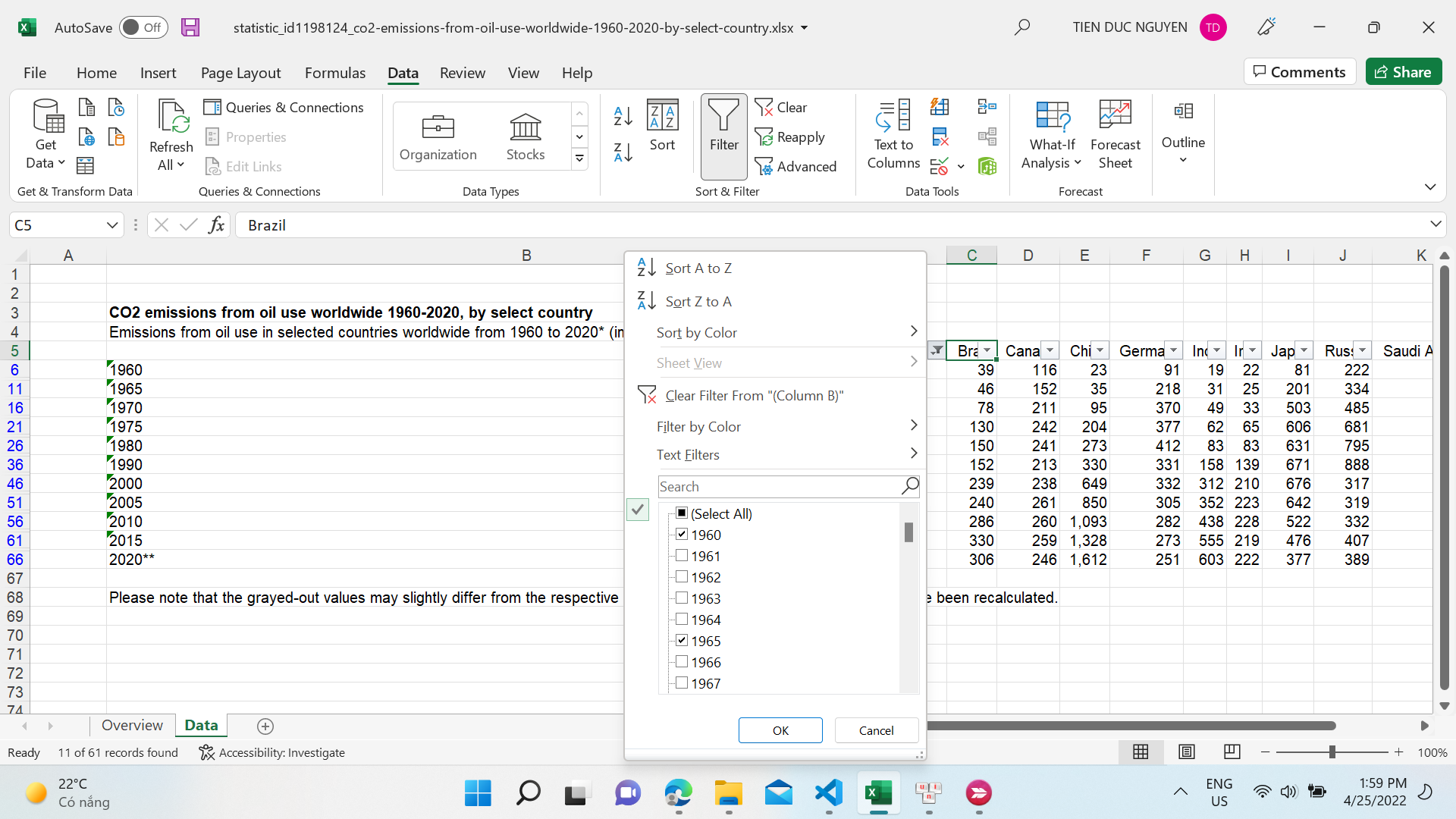


Figure 9.

However, not all of the countries will be represented on the graph. Because a number of countries did not have data in all three datasets, we must select countries that appear in all datasets, which include Japan, Germany, the United States, and China. After that, we can combine the datasets, and the result is a csv file like the one below:

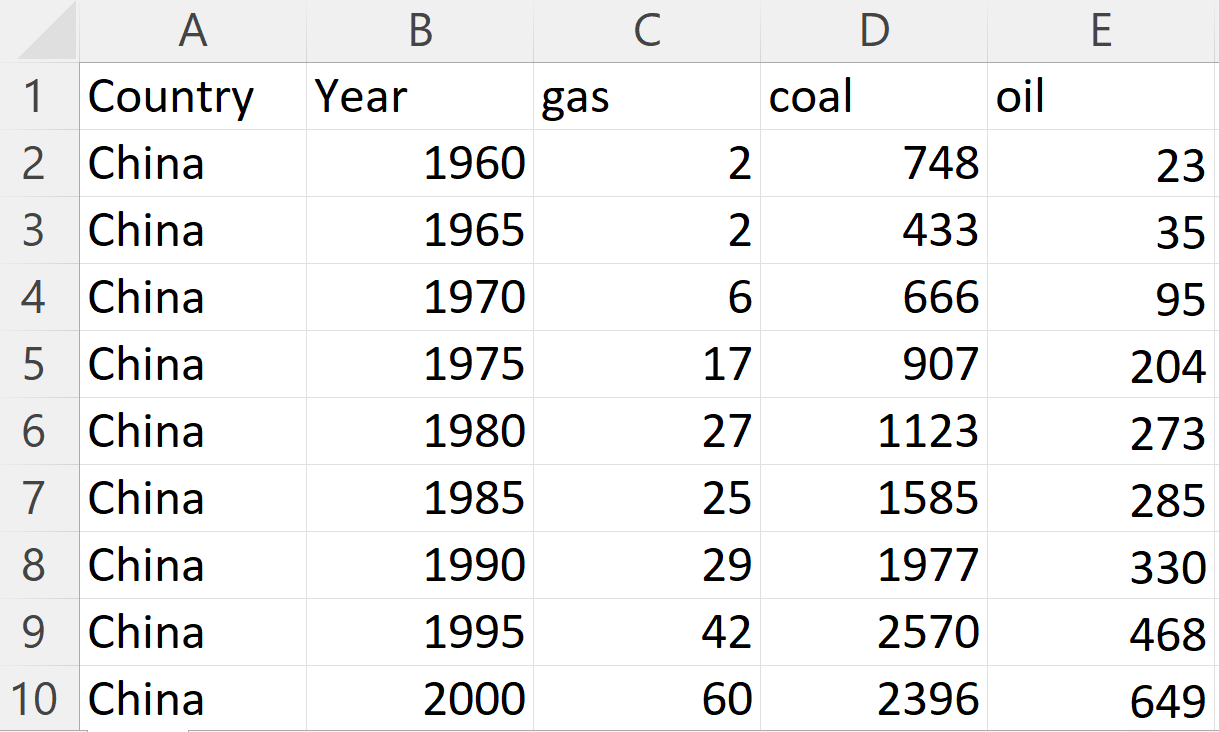


Figure 10.

Figure CSV file for the third Visualization.

The primary objective of this chart is to generate a line chart with each line representing the variance of CO2 emissions in each country over a 5-year period. Thus, we need to use the “Group” function in D3 to generate an organized array with the key element being the name of a country.

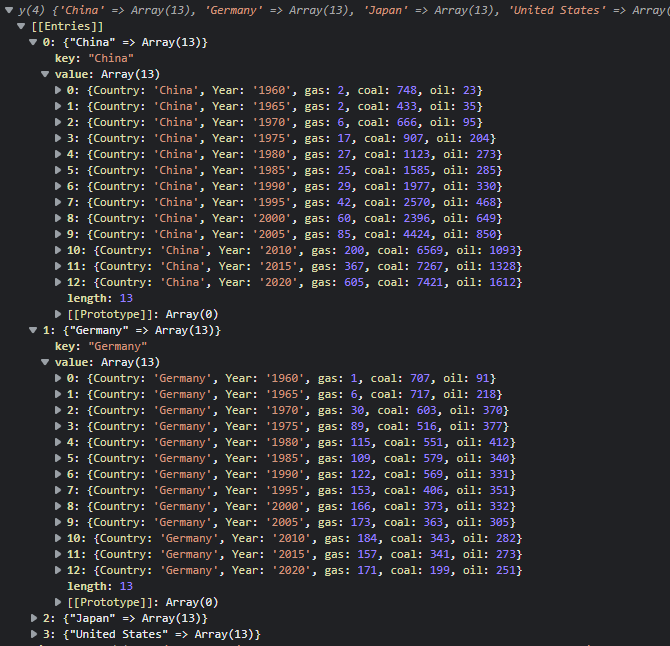


Figure 11.

Figure. Data after grouping

# Requirement

## Must-have Features

There are several features that are considered vital for the purposes of this visualisation because without them, the project could not be progressed appropriately:

1. A chart that demonstrates the demand of multiple sources of energy over the world in 2019 in different sectors. There must be colour legends to help the audiences figuring out how the data in each sector is represented.
2. A choropleth that displays the total amount of primary energy consumption around the world in a 10-year period starting from 2010 to 2020. The colour saturation must be used in order to encode the amount of energy consumption, with the higher saturation indicating the higher energy consumption.
3. A line chart indicates the CO2 emissions of different types of energy in four different countries namely The United States of America, Germany, Japan and China. As a result, four different colour hues are utilised to illustrate the volatility of the CO2 emissions in each country.

## Optional features

# Visualization Design

## Energy demand per sector

### Initial Idea

Initially, we intend to create multiple pie charts to represent the energy demand of various types of energy for each sector (Industry, Residential, Commercial and public services, Transport and Other). As a result, there will be buttons that allow users to select which sector they are interested in. Besides, different types of energy will be encoded by different colors (hues). And the larger the slices, the higher the value of energy demand for a specific energy type.

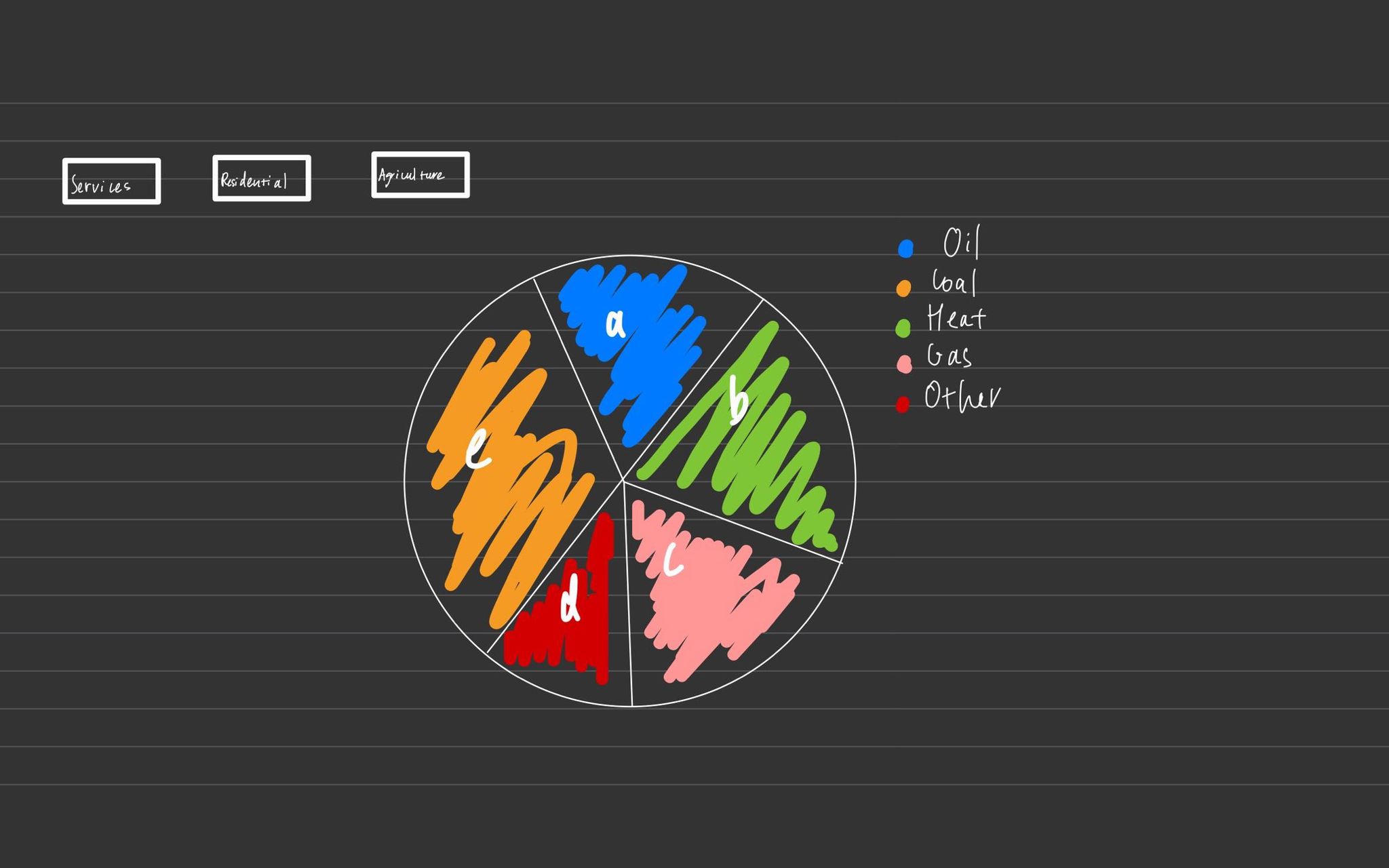


Figure 12. The sketch for the first design

Even though this approach has numerous advantages such as it allows the audience to immediately observe a comparison between each sort of energy and perform an immediate analysis or grasp information (Rajasekhar 2022). However there are several issues with this visualization:

* When slices are close in size, it’s difficult to tell which is bigger.
* Because the human eyes isn’t good at ascribing quantitative value to two-dimentional space. Therefore, pie chart are hard for people to read. (Storytelling) (need to paraphrase).
* Since this approach requires to create 5 more pie charts correspond to 5 different sectors, which is time consuming and repetitive

In second attempt, we decided to go with the bar chart. With this method, the amount of energy consumption will be encoded by line mark and vertical spatial position channel. Which means the larger the amount of energy consumption, the higher the bar is.

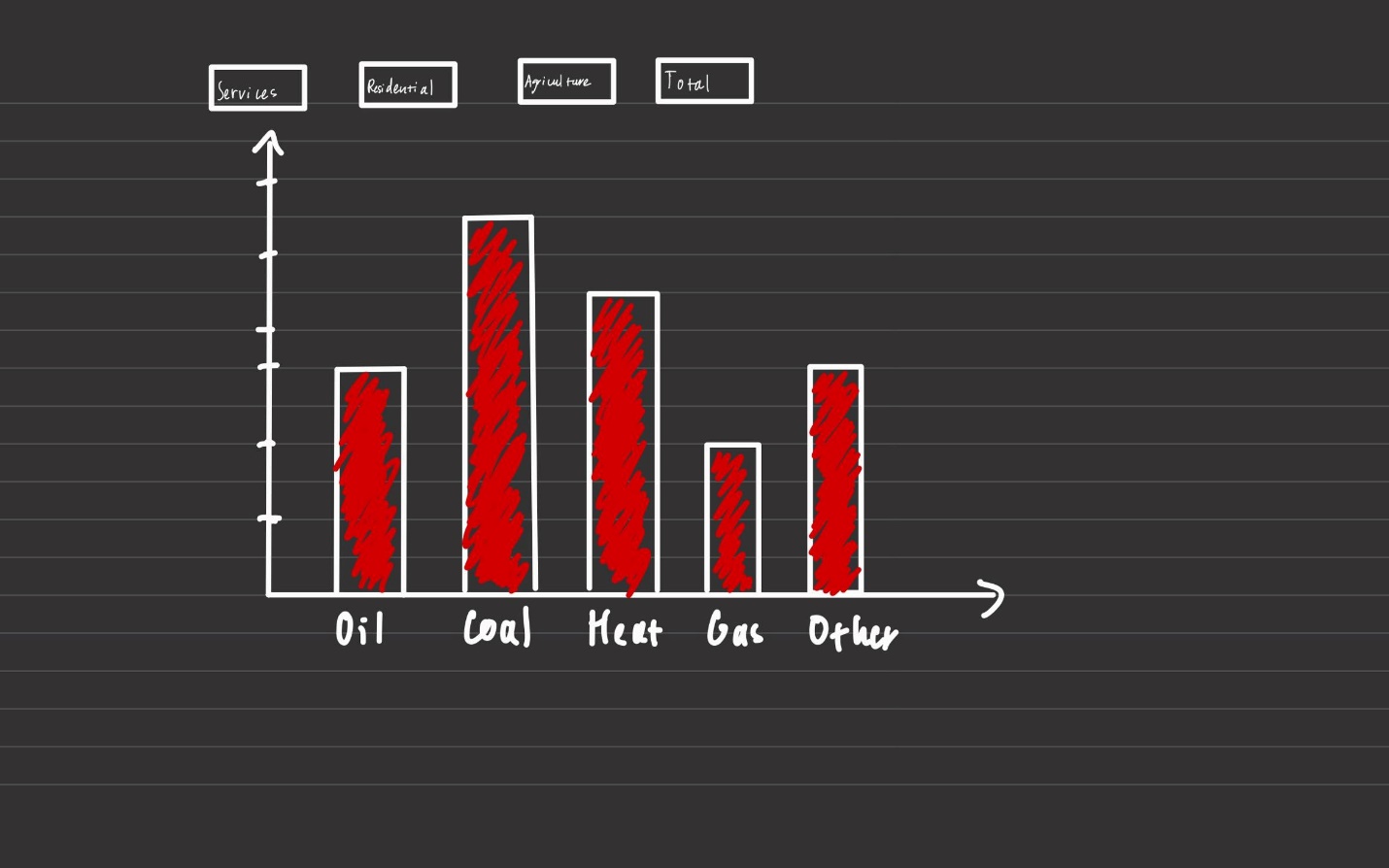


Figure 13. The sketch of second design.

In addition we also want to implement a stacked bar chart. With this method each sector will be encoded by different colors (hue). Besides, each sector will be encoded by different colors (hue)

### Final

## Energy consumption per region

### Initial Idea

### Final

## Co2 emissions from different types of energy

### Initial Idea

### Final

# Conclusion

References

[world.geo.json/countries.geo.json at master · johan/world.geo.json (github.com)](https://github.com/johan/world.geo.json/blob/master/countries.geo.json) (GeoJson)

[Energy Technology Perspectives 2017 - Data product - IEA](https://www.iea.org/data-and-statistics/data-product/energy-technology-perspectives-2017-2) (First Chart)

[Carbon Dioxide Emissions From Electricity - World Nuclear Association (world-nuclear.org)](https://www.world-nuclear.org/information-library/energy-and-the-environment/carbon-dioxide-emissions-from-electricity.aspx?itid=lk_inline_enhanced-template#:~:text=According%20to%20the%20International%20Energy%20Agency%20%28IEA%29%2C%20the,greenhouse%20gases.%20For%20more%20information%20see%20pages%20on%3A)

[Advantages and disadvantages of pie charts (excelr.com)](https://www.excelr.com/advantages-and-disadvantages-of-pie-charts/#:~:text=Advantages%20of%20a%20Pie%20Chart%201%20A%20simple,manipulate%20pieces%20of%20data%20in%20the%20pie%20chart.) (pie chart)

Appendix