Swinburne University of Technology

COS30045 Data Visualization

DATA VISUALIZATION

ENERGY CONSUMPTION AND PRODUCTION

LA2_G08

Group members: Hoang Chau Anh Pham – 103533839
Trung Kien Nguyen – 103506132

Tutorial: Monday 12.30 - 2.30pm, Semester 1 2022
Word count: 4075 words
Visualization: https://kiennguyen2403.github.io/datavisualization/

Table of Contents

Introduction

1.1. Background and motivation

1.1.1. Background

Energy is essential to our modern life, especially when it is connected to everything we do. The energy system is undergoing the greatest transformation since the 1960s. As for Australia, these changes are driven by numerous factors such as society, economy, and technology. Customer needs are also changing as increased choices are available, and people desire to control their choice of freedom as well as optimize their usage and spending.

Most electricity in Australia is produced, imported, exported, and transported sufficiently and promptly to supply consumers' demand in real-time. There is an interconnected electricity system known as the National Electricity Market (NEM), one of the greatest retailers, however, this wholesale market only supplies energy to the East Coast and Southern states. What about Western Australia and Northern Territory? These regions have their own electricity systems. The energy supply and conversion in Australia are therefore especially interesting as each state has its own partner energy supply organization and separate regulatory arrangements. The data itself is then an ideal source to analyze and do research, as well as training or applying it to visualization.

Up to now, a substantial portion of Australian needs have been met by fossil fuels, and this low-cost resource is contributed for about 70% of Australia’s domestic electricity and underpin some of the cheapest electricity in the world.

Australia can completely rely on self-generated energy. The total output of all electric power generating facilities is 243 billion kWh, accounting for 106 percent of total requirements. The remainder of self-produced energy would either be exported or goes unused. Along with pure consumption, production, imports, and exports all play significant roles. Natural gas and crude oil are often used as energy sources.

These energy units have contributed to Australia's complicated and organized electrical infrastructure, and without the coal boom, the country would not have much of the complex electrical framework we rely on in our homes, railways, and cities. The economic stability has also been strengthened by coal exports, which would have hampered our social, cultural, and community growth.

1.1.2. Motivation

As Swinburne is committed to helping individuals and society work towards a more sustainable future, we are here to help. Rather than considering it as a project, we see itself as a research and information design pathway through which we learn ourselves the university's ambition and approach.

By visualizing the statistical energy data, our website intended to show the public how the physical unit of energy has been conversed and transmitted, as well as how much energy is consumed in each period. The user who approaches the websites should also acquire some knowledge about the flow of energy or be able to find the information they need on the energy usage in their regions. Moreover, we hope this presentation will help somebody with their ideas for data visualization or even a source of reference.

We consolidate statistical data and resources from multiple authorized sources and make them available in easy approach forms such as graphs and charts. It provides users with several options to choose from, to deeply understand the data or use the visualization in the way they like.

1.2. Project objective

This project is about to answer several questions that underlying the usage of energy and the energy production process:

- How much energy a state consumed?
- Which state consume the largest amount of energy, and which consume the least?
- Is the energy demand increased?
- How much energy is generated to supply users demand?

Our ambitions at the completion of the project are:

- Learn and have a better understanding about how the energy is conversed and transmitted.
- To be aware of energy shortage and energy market.
- Find an efficient way to introduce the data and analyze the growth pattern.

The project has well informed us the insight regarding the energy production and consumption. Approaching the data and comparing the values each year wake me up on how much the energy market has changed, as well as how fast are we growing the country and the economy.

1.3. Project Schedule

The tasks are planned and scheduled on Monday every week at the meeting. The testing and checking processes occur on Friday. Each team member’s timetable does not have many similarities; therefore, we can only meet twice a week. Members must work independently and have the responsibility to finish their tasks before Friday. Any reasons for the absence and unfinished tasks must be announced 2 days before the checking day. The schedule and tasks are displayed in Google Sheets.

Data

2.1. Data source

2.1.1. Energy consumption

Australian Government
Department of Industry, Innovation and Science
Office of the Chief Economist

Table D1

Australian energy consumption, by fuel, energy units

	Consumption of fuels														Production of derived fuels					
	Black coal	Brown coal	Coke	Coal by-products	Liquid/gas biofuels	Wood, woodwaste	Bagasse	Refinery input	Petroleum products	Natural gas	Town gas	Electricity	Solar energy		Coke	Coal by-products	Petroleum products a	Town gas	Thermal electricity	Total energy consumption b
	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ		PJ	PJ	PJ	PJ	PJ	PJ
1999-00	1,387.5	672.5	98.8	80.0	na	108.2	105.7	1,672.8	1,778.6	924.3	3.3	756.8	2.6		105.9	80.4	1,728.5	6.1	703.9	4,971.0
2000-01	1,418.6	666.0	83.9	77.6	na	109.4	98.5	1,702.0	1,740.5	961.7	3.3	776.4	2.6		94.9	69.7	1,741.0	5.1	739.9	5,011.8
2001-02	1,457.1	673.4	81.7	76.8	na	95.0	91.7	1,674.8	1,760.3	976.4	3.4	786.4	2.7		92.9	68.3	1,685.1	5.1	749.7	5,097.0
2002-03	1,428.0	699.1	71.2	72.6	11.2	105.3	95.1	1,629.5	1,800.6	942.1	5.2	799.1	2.8		89.1	66.3	1,628.2	5.1	737.5	5,138.7
2003-04	1,510.2	714.6	76.7	76.0	10.6	103.2	101.1	1,531.0	1,906.7	974.0	5.2	826.7	2.6		95.5	63.7	1,627.1	5.1	765.6	5,284.8
2004-05	1,554.4	711.7	73.7	72.8	9.2	100.4	108.3	1,543.7	1,969.4	996.4	5.2	822.1	2.6		95.0	65.3	1,644.0	5.1	763.5	5,399.2
2005-06	1,576.4	725.1	72.6	75.9	9.9	98.9	109.1	1,409.2	1,998.9	1,000.9	5.1	837.2	2.4		90.8	65.4	1,438.1	5.0	774.0	5,547.0
2006-07	1,601.9	721.9	73.3	73.8	10.7	98.7	110.8	1,505.7	2,027.0	1,131.9	5.0	874.3	6.0		90.7	63.5	1,542.3	4.7	813.3	5,726.1
2007-08	1,581.2	725.1	74.6	74.8	15.2	98.7	110.8	1,464.6	2,088.4	1,168.7	3.8	874.5	6.7		90.7	64.9	1,564.8	3.7	820.6	5,742.5
2008-09	1,601.4	749.7	63.8	48.0	20.1	88.5	65.4	1,480.7	2,071.0	1,266.6	0.5	892.7	8.2		67.5	44.7	1,544.6	0.4	841.3	5,858.2
2009-10	1,489.0	742.0	75.6	59.3	22.9	86.7	93.3	1,439.3	2,101.2	1,269.7	0.4	908.7	10.5		85.5	55.6	1,473.2	0.3	839.7	5,844.4
2010-11	1,403.4	728.2	79.0	70.5	26.0	83.9	83.3	1,510.3	2,188.2	1,284.2	0.5	912.3	11.7		91.4	66.0	1,488.6	0.2	825.0	5,910.4
2011-12	1,353.1	739.3	58.8	56.0	27.6	82.0	84.9	1,486.6	2,269.6	1,345.3	0.2	903.1	12.4		80.6	51.6	1,514.9	0.2	817.9	5,953.8
2012-13	1,318.2	645.7	56.8	50.1	26.4	82.7	94.9	1,443.1	2,346.6	1,379.8	0.1	899.1	13.1		86.8	49.8	1,501.0		790.8	5,928.3
2013-14	1,248.4	627.8	55.4	47.9	29.3	84.5	90.2	1,336.7	2,362.1	1,412.5	0.1	893.9	13.2		80.3	47.6	1,441.6		773.2	5,859.5
2014-15	1,253.4	669.1	57.0	46.5	30.5	83.8	102.9	1,130.8	2,376.0	1,430.9	0.1	908.5	14.8		71.9	46.2	1,269.3		797.2	5,919.6

Notes:
Totals may not add due to rounding.
Data concerning refinery input, brown coal briquettes, ethylene, hydrogen and other energy commodities have not been shown in the table for reasons of confidentiality, but are included in totals where appropriate.
a Production may exceed refinery input as some petroleum products are produced from other petroleum products.
b Total energy consumption is the total quantity (in energy units) of primary and derived fuels consumed less the quantity of derived fuels produced.
na Not available.

AUS | NSW | VIC | QLD | WA | SA | TAS | NT

Ready Accessibility: Investigate

Figure 1. Australian energy consumption, by fuel, energy units

The visualizations make use of data from the Department of Industry, Science, Energy and Resources, which is quite clean and detailed (see Figure 1). Because there is no noise or inconsistency in the data, the data cleaning process is forgone. This numerical dataset contains some information about consumption across variety energy units and some other in production.

2.1.2. Energy Production

O56 3402.5

Australian Government
Department of Industry, Innovation and Science
Office of the Chief Economist

Table J

Australian energy production and trade, energy units

	1973-74	1974-75	1975-76	1976-77	1977-78	1978-79	1979-80	1980-81	1981-82	1982-83	1983-84	1984-85	1985-86	1986-87	1987-88	1988-89	1989-90	1990-91	1991-92	1992-93	1993-94
	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ	PJ
Imports																					
Coke																					
Crude oil and ORF	405.2	393.6	375.5	391.5	434.0	402.8	435.9	443.1	482.2	455.9	331.0	282.3	239.4	298.9	370.6	466.6	449.0	518.2	593.3	751.6	785.4
LPG								0.0	0.0	0.0	0.2	0.1	0.0	1.0	1.1	1.0	2.2	0.9	1.2	2.9	4.1
Automotive gasoline	16.3	14.0	30.1	31.7	26.1	24.4	16.8	14.4	13.7	19.0	11.6	20.2	17.3	43.6	31.1	47.8	54.5	20.0	7.0	13.2	13.9
Aviation gasoline	2.7	3.0	2.6	3.2	2.4	2.8	1.3	1.3	0.4	0.2	0.0			0.2		0.2	0.3				
Aviation turbine fuel	3.8	3.5	3.7	4.8	3.7	7.1	6.6	5.5	3.9	3.9	2.3	3.5	6.1	8.1	6.3	7.2	8.6	3.8	3.8	1.3	7.0
Lighting kerosene	3.4	3.5	2.5	1.0	0.3	0.3															
Power kerosene	0.3	0.2										0.0				0.7	1.4	1.3	0.5		
ADO	20.1	18.4	14.5	18.9	20.3	15.8	23.8	24.5	20.1	18.0	12.4	26.2	27.6	39.2	27.3	31.2	36.6	17.1	14.7	25.9	28.5
IDF	1.1	1.5	7.4	8.3	16.5	3.1	2.4	0.5	0.4	1.1	0.2	0.2	0.9	0.5	1.9	1.5	3.2	0.7	0.3	1.2	1.0
Fuel oil	116.1	105.3	86.8	91.4	81.8	101.5	108.3	84.7	62.5	48.3	58.0	45.1	44.6	48.1	39.8	43.8	45.5	40.8	36.9	45.9	38.5
Lubes and greases	2.2	2.7	2.3	1.1	2.0	2.2	2.8	2.6	2.0	1.8	1.0	1.7	1.4	1.7	1.7	1.9	2.1	0.9	1.3	1.1	1.7
Bitumen					0.2	0.2	0.2	0.1	0.4	0.2	0.1	0.4	0.7	0.6	0.4	0.2	2.9	0.3	0.2	0.1	0.5
Other petroleum products	7.4	7.7	4.6	1.8	5.2	6.4	14.6	11.9	10.1	9.2	11.6	10.0	7.0	9.4	15.2	17.6	19.3	21.1	17.6	17.1	17.6
Natural gas																					
Total	**578.6**	**553.4**	**530.0**	**553.7**	**592.5**	**566.6**	**612.7**	**588.6**	**595.7**	**557.6**	**428.4**	**389.7**	**345.0**	**451.3**	**495.4**	**619.7**	**625.6**	**625.1**	**676.8**	**860.3**	**898.2**
Exports																					
Black coal	836.9	950.3	892.7	1,038.4	1,111.0	1,118.7	1,258.6	1,377.2	1,371.4	1,608.4	1,913.9	2,398.0	2,558.7	2,718.2	2,910.4	2,834.4	2,965.6	3,217.0	3,502.2	3,660.6	3,668.1
Coke **b**	3.8	11.4	4.9	5.1	4.0	4.0	3.3	0.6	0.2	0.1	0.4	0.3	5.9	4.6	22.1	26.5	15.5	22.6	19.6	16.2	14.3
Briquettes	0.1	0.0	0.0	1.0	0.9	0.6	0.6	0.9	1.0	1.0	1.2	1.0	1.4	1.2	1.7	0.5	0.7	1.0	1.8	1.3	2.2
Crude oil and ORF					7.8	13.7	4.7	3.2	1.6	2.3	39.1	215.3	186.9	211.0	238.8	181.9	266.5	326.4	332.0	373.6	352.9
LPG	53.0	53.4	52.1	60.2	76.5	80.9	73.8	68.6	70.0	62.3	75.6	69.4	78.9	70.9	63.6	57.7	52.6	40.0	41.6	39.3	34.2
Automotive gasoline	15.5	8.6	5.3	7.3	9.8	11.7	10.7	9.2	11.7	17.6	20.2	11.7	13.6	8.6	12.3	8.4	7.2	10.8	23.7	23.2	30.5
Aviation gasoline	0.7	0.8	0.5	0.7	0.8	0.7	0.3	0.6	1.4	1.9	2.4	2.7	2.3	2.2	2.7	1.9	2.8	2.1	5.2	2.3	2.0
Aviation turbine fuel	13.9	13.1	10.3	10.0	12.0	11.6	7.7	10.2	9.9	9.9	14.1	13.8	12.1	11.2	14.6	18.4	19.9	11.8	9.1	14.4	14.7

Aus

Ready

Figure 2. Australian energy production and trade, energy units

The data in Figure 2 is used in Production visualization – Sankey Diagram. As it is clearly shown, the imports and exports amount of energy in different categories is collected throughout a long period. The variables and values are already clean but for some missing data, and hence the cleaning process is performed.

2.1.3. States and cities boundaries


```
var statesData = {
    "type": "FeatureCollection",
    "features": [
        {
            "type": "Feature",
            "id": 0,
            "properties": {
                "STATE_CODE": "1",
                "STATE_NAME": "New South Wales",
                "consumption": 69397.8
            },
            "geometry": {
                "type": "MultiPolygon",
                "coordinates": [
                    [ [ [ 159.105420, -31.563994 ], [ 159.097750, -31.564275 ], [ 159.099634, -31.573372 ], [ 159.094217, -31.570970 ], [ 159.
                    074599, -31.597285 ], [ 159.068740, -31.595697 ], [ 159.077238, -31.543533 ], [ 159.056995, -31.519772 ], [ 159.038463, -31.
                    522425 ], [ 159.037960, -31.512156 ], [ 159.064657, -31.509666 ], [ 159.063790, -31.517379 ], [ 159.081178, -31.526607 ], [
                    159.081913, -31.539767 ], [ 159.094678, -31.545036 ], [ 159.105420, -31.563994 ] ] ], [ [ [ 151.145550, -33.824079 ], [ 151.
                    143844, -33.829174 ], [ 151.169325, -33.839309 ], [ 151.180206, -33.836165 ], [ 151.174277, -33.843006 ], [ 151.143088, -33.
                    835565 ], [ 151.135902, -33.836469 ], [ 151.148783, -33.839386 ], [ 151.133837, -33.844837 ], [ 151.112863, -33.829724 ], [
                    151.108625, -33.835555 ], [ 151.092828, -33.819609 ], [ 151.070345, -33.816445 ], [ 151.051336, -33.823662 ], [ 151.084007,
                    -33.822080 ], [ 151.075192, -33.832683 ], [ 151.080873, -33.845750 ], [ 151.083228, -33.827897 ], [ 151.093817, -33.824443 ]
                    , [ 151.089778, -33.833044 ], [ 151.098921, -33.831641 ], [ 151.093602, -33.839764 ], [ 151.102215, -33.838102 ], [ 151.
                    099653, -33.844209 ], [ 151.104144, -33.834387 ], [ 151.120767, -33.841788 ], [ 151.113783, -33.853971 ], [ 151.122305, -33.
                    856994 ], [ 151.119070, -33.864711 ], [ 151.128397, -33.860961 ], [ 151.122165, -33.848934 ], [ 151.126983, -33.844195 ], [
                    151.142064, -33.845845 ], [ 151.142273, -33.853358 ], [ 151.149526, -33.846908 ], [ 151.144434, -33.843531 ], [ 151.154043,
                    -33.842070 ], [ 151.164849, -33.855249 ], [ 151.144520, -33.863162 ], [ 151.149559, -33.866131 ], [ 151.141168, -33.869844 ]
                    , [ 151.150870, -33.872265 ], [ 151.185490, -33.846008 ], [ 151.190538, -33.852375 ], [ 151.184547, -33.853645 ], [ 151.
```


Figure 3. State and Territories boundaries. Link: data.gov.au

The map on consumption site is based on data from Australian Government released in August 2021. The data shows the network and links between regions, which also includes names, features, coordinates of specific states or cities. The data is up to date and filtered out the noise so there is nothing to do with the data cleaning, but to make use of it, we have to update the amount of energy used in each area and merge this piece of data in to the same Geojson file.

2.2. Data processing

2.2.1. Consumption Visualization

Data_Visualization > Sankey_Attempt > example > {} Test.json > {} 13 > Total energy consumption b


```
{
    "Time": "1973-74",
    "Black coal": 662.9,
    "Brown coal": 262.6,
    "Coke": 139.8,
    "Coal by-products": 128.9,
    "Liquid/gas biofuels": "na",
    "Wood, woodwaste": 92.5,
    "Bagasse": 56.4,
    "Refinery input": "1,230.1",
    "Petroleum products": "1,348.9",
    "Natural gas": 172.5,
    "Town gas": 20.6,
    "Electricity": 250.5,
    "Solar energy": 0.1,
    "": "",
    "Coke_1": 144.6,
    "Coal by-products_1": 129.4,
    "Petroleum products a": "1,245.8",
    "Town gas_1": 20.6,
    "Thermal electricity": 203.1,
    "_1": "",
    "Total energy consumption b": "2,615.2"
},
{
```


Data_Visualization > Sankey_Attempt > example > output_2.txt


```
{
    "Time": "1973-74",
    "Black coal": 662.9,
    "Brown coal": 262.6,
    "Coke": 139.8,
    "Coal by-products": 128.9,
    "Liquid/gas biofuels": "na",
    "Wood, woodwaste": 92.5,
    "Bagasse": 56.4,
    "Refinery input": 1230.1,
    "Petroleum products": 1348.9,
    "Natural gas": 172.5,
    "Town gas": 20.6,
    "Electricity": 250.5,
    "Solar energy": 0.1,
    "": "",
    "Coke_1": 144.6,
    "Coal by-products_1": 129.4,
    "Petroleum products a": 1245.8,
    "Town gas_1": 20.6,
    "Thermal electricity": 203.1,
    "_1": "",
    "Total energy consumption b": 2615.2
},
{
```


Figure 4. Convert the string of number separated by a comma to a whole number.

The cleaning process is challenging because we had to make adjustments to a Geojson file with dictionary variables. When the amount exceeds 1000, the Geojson file generated will force the value to be a string of numbers separated by a comma. We obtained the ideal values that can be used in Geojson syntax by utilizing a Python program that reads from line to line and then writes to another text file.

The code for converting value from string to a whole number can be refer to as follows:

```python
import json

with open('Test.json', 'r+') as f:
    data = json.load(f)
    f = open("output_2.txt", "w")
    for objects in data:
        f.write("{\n")
        for key, value in objects.items():
            input = objects[key]
            output = ''
            if key == "Time" or value == "na":
                f.write('"%s": "%s",\n' % (key, value))
            elif type(value) == int or type(value) == float:
                f.write('"%s": %s,\n' % (key, value))
            elif value == "":
                f.write('"%s": "%s",\n' % (key, value))
            elif key != "Time" or value != "na" and type(value) == str:
                for i in range(len(input)):
                    if input[i] != ',':
                        output += input[i]
                f.write('"%s": %s,\n' % (key, output))
        f.write("},\n")
    f.close()
    print(f'Processed')
```

Figure 5. String to number converter.

However, when visualize the data on the line chart, we recognize that the lines are close to together since certain categories of energy units have similar patterns and the same amount of consumption. In this situation, we decided to aggregate the energy units using the Autosum function in Excel into relevant categories as follows:

1. Fossil resources: Black coal, Brown coal, Coke, Coal by-products, Liquid/gas biofuels, Refinery input, Petroleum products.
2. Clean: Wood - Woodwaste, Bagasse, Electricity, Natural gas, Town gas, Solar energy.

Type of data is also summarized in the following table:

Attribute	Description	Type	Subtype
Consumption of fuels	This attribute contains multiple fuels type and is expressed as the main source of energy in Australia.	Categorical	Nominal
Black coal	Contain the consumption in Petajoule (PJ). 1 PJ = 31.6 million m^3 of natural gas or 278 million kilowatt hours of electricity.	Numeric	Continuous
Brown coal			
Coke			
Coal by-products			
Liquid/gas biofuels			
Wood, woodwaste			
Bagasse			
Refinery input			
Petroleum products			
Natural gas			
Town gas			
Electricity			
Solar energy			
Year	This attribute represents the period of time	Numeric	Continuous

2.2.2. Production Visualization

As we use the data across years, the whole data set from Figure 2 in the consumption section was extracted into a separate Excel file. The file is then saved as CSV to be input into a Python code, which generates command line as exact syntax using for Geojson.

Figure 6. Python command release processor.

Along with transferring the csv syntax to Geojson syntax, the code also removes the comma in the values equal or greater than 1000, so that it will not raise an error in Geojson. The data transformation looks like this:

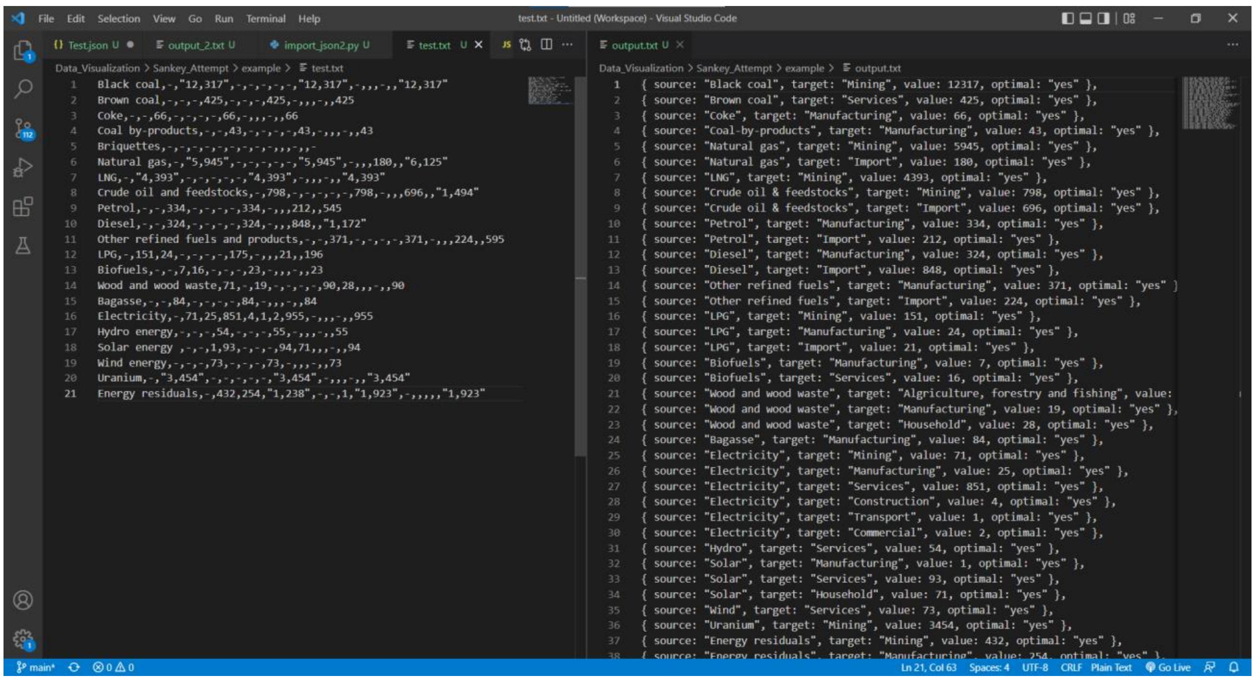

Figure 7. From CSV to Geojson.

As the data is numeric and subtype is numerous, it is easy to transform and make some statistical calculation.

Visualization Feature

3.1. Must-Have Features

We determined some characteristics that should be included in the visualization before beginning the design process to lay the groundwork for the visualization. There are some elements that are required, as well as some that would improve the visualization.

3.1.1. Energy consumption

As it is aimed to visualize the data of energy consumption by states and territory, there should be a graph that ideally present both the data and the position of each state. The data should also be compared between years so there should be a feature that can change the graph from one year to another.

- The chart used here is the bubble chart on a map. The bubble size will present how much amount of energy used while the bubble position will be placed at the states' coordinate.
- Choropleth map to describe the quantitative amount of energy consumption of a region when zoom in
- A time bar that updates the data of the graph and map from the year to years.
- A line chart to represent the total consumption of Australia
- A bar chart to describe the consumption of each state by energy units.
- Animated transitions and effects when hover or click on the line and bar chart elements.
- Zoom in and out effect when changing from bubble map to choropleth map.
- Color scale to emphasize the amount of energy usage.

3.1.2. Energy production

The physical units will be converted to numerous types of energy (e.g., thermal, electricity, chemical, ...) so that it should be display as a flow with different levels corresponding to the stages of units converted. In other words, if we use coal to generate thermal energy, coal will lie at the first level and thermal energy is placed at the second one.

- By considering those issues, the decision was made on creating a Sankey diagram. The linking between levels has its size to emphasize how much amount of a unit was used to generate the energy at the next stage.
- Animated transition when hovering on an energy unit
- A table or list to control the link when click on.

3.2. Optional Features

3.2.1. Energy consumption

Pie chart

An optional feature that we are not able to deploy was a pie chart to describe the portion of specific energy units contributed to the fossil and clean resources as indicated in *Data*

Processing. The numerical data is cumulated to have the fossil resources attribute, so that a pie chart would be favorable in clarifying the smaller units that made up the new attribute.

- Color hue to depicts the energy usage. The largest contribution will be referred to the most extreme color, and the least contribution will be assigned the least extreme.
- Animated transition when hovering on an energy unit ((i.e., data will appear when hovering on a slice of pie)

3.2.2. Energy production

Table of value

It is inconvenient to hover on each attribute to see its value. The table should be attached with the list of categories so that it presents the value and control the link in the flow the production.

Sources	**4,589,768**		**$36.58**
Solar (Rooftop)	82,797	1.8%	$48.53
Solar (Utility)	28,337	0.6%	$65.15
Wind	148,460	3.2%	$57.73
Hydro	273,630	6.0%	$49.74
Battery (Discharging)	413	0.009%	$141.75
Gas (Waste Coal Mine)	507	0.01%	$137.97
Gas (Reciprocating)	5,921	0.1%	$78.61
Gas (OCGT)	70,237	1.5%	$90.12
Gas (CCGT)	220,439	4.8%	$49.43
Gas (Steam)	76,965	1.7%	$43.78
Distillate	1,211	0.03%	$302.78
Bioenergy (Biomass)	3,303	0.07%	$62.41

Figure 8. Table of attributes and values. Source: OpenNEM (opennem.org.au, n.d.)

However, in such a short amount of time, we consider it as redundant since displaying the value of attributes not giving any knowledge or interest to user.

Visualization design

4.1. General ideas

Instead of making a single all-encompassing chart for our visualization, we chose to make a more 'conventional' scroll-down webpage with multiple charts displaying different data. This would allow me to give an indication to users while showing the charts with the explanation.

There are 2 sections to visualize, one is the consumption, and the other is production. As the subjects are too general, we have broken it down to the smaller categories. For the consumption, it will go by each state, city and industry while the production is the energy's type.

The secondary chart as line and bar chart is added on one side so that when user clicking on a region, the data is displayed right next to the original map.

4.2. Sketches

Below are some sketches from a brainstorm we did for our visualization, based on the requirements outlined above. Bar charts, pie charts, Sankey diagram, and choropleth were considered as visualization components because the visualization will compare the data.

4.2.1. Consumption Visualization

The layout of the webpage is presented as the Figure 10, as we want to offer the categorical and numerical data for a given region at the same time. To describe distinct types of data, the line chart and bar chart are taken place one after the other.

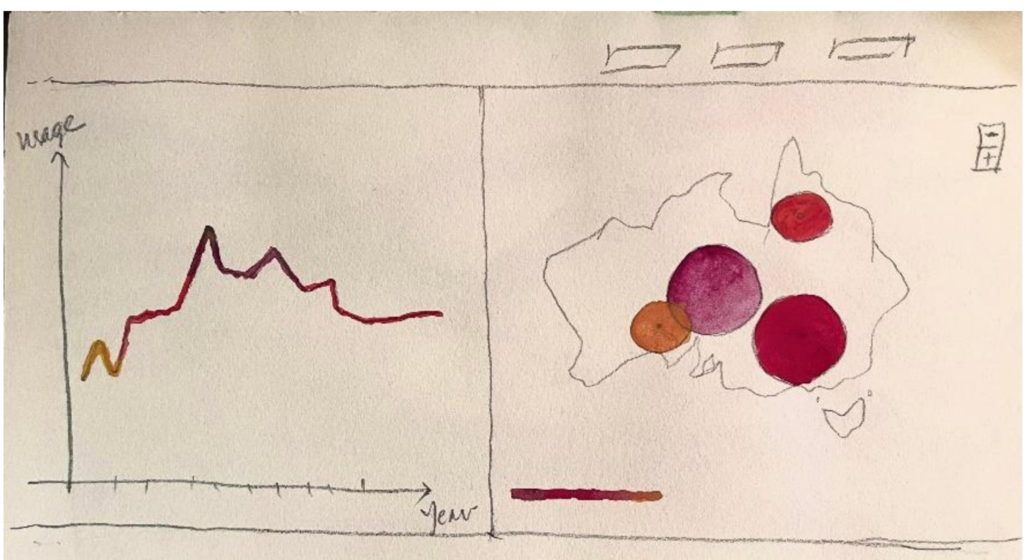

Figure 9. Low fidelity of consumption webpage.

The line chart shown a quite clear concept, which is chosen to present the total energy consumption of Australia region across years while the bubble map is more complex, displaying both the location and the size of data. The bubble map is then designed as follows:

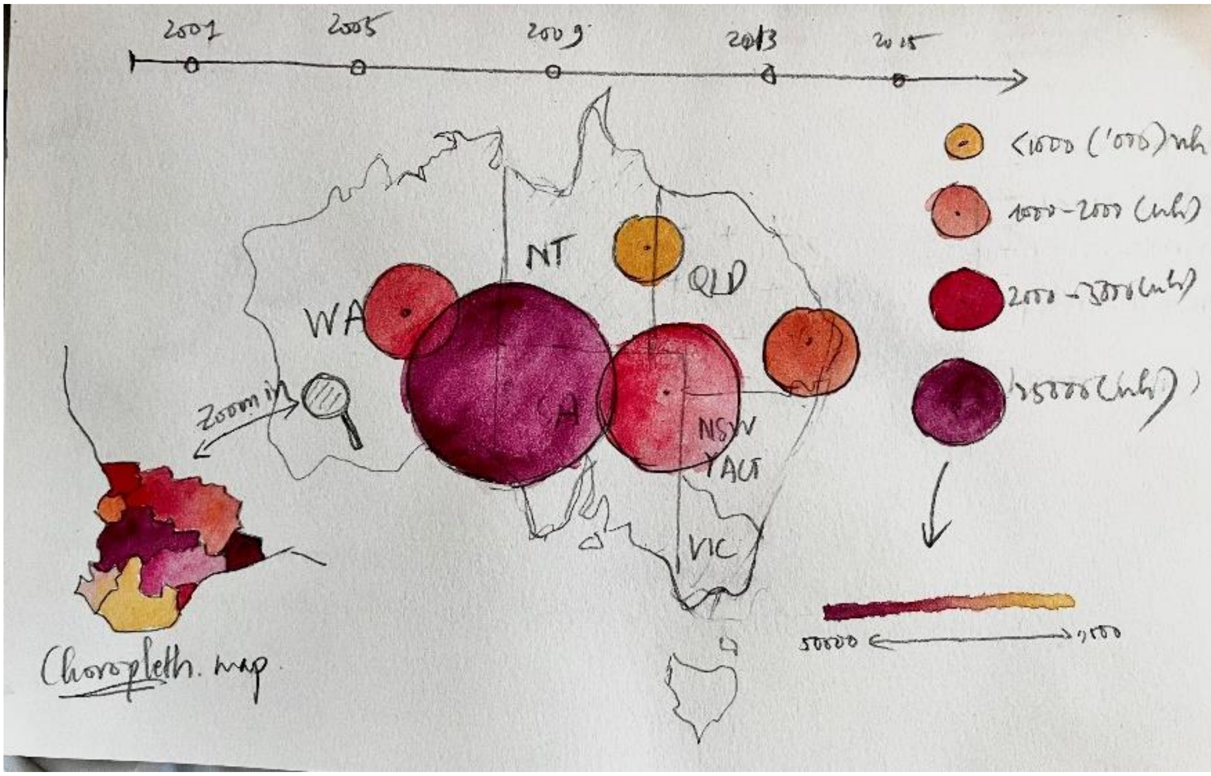

Figure 10. Low fidelity bubble map in details.

By sketching the visualization at the first stage, we realize some of the noteworthy points:

Plus points

- Easily check the energy usage details of each region
- Can compare the total energy between states and territories and between years.

Limit

- More text description is required to provide information about which graph is displaying which data.
- Cannot directly identify the pattern throughout a period since color hue is utilized to presents the amount of energy used. It is difficult to figure out how much extreme a color is, and the values therefore cannot be easily identified as more or less significant.

Features to Implement

- Transition when bubbles are clicked.
- Data appears when hovering on the lines and bars of charts.

- Switch view from Australia regions to state and territories regions.
- Update the data when users click on a year in the time bar.

Visual Encoding

- Color hue and saturation: coordinate the amount of energy consumption.

4.2.2. Production Visualization

The layout of the webpage is presented as the Figure 12, as we want to offer the categorical and numerical data for a given region at the same time. On the left-hand side, the Sankey diagram represents the link and the flow of resources while the other hand side contains a table of data and pie chart. The table only represent the raw data in the dataset, and pie chart will describe the percentage of contributed energy units of a production.

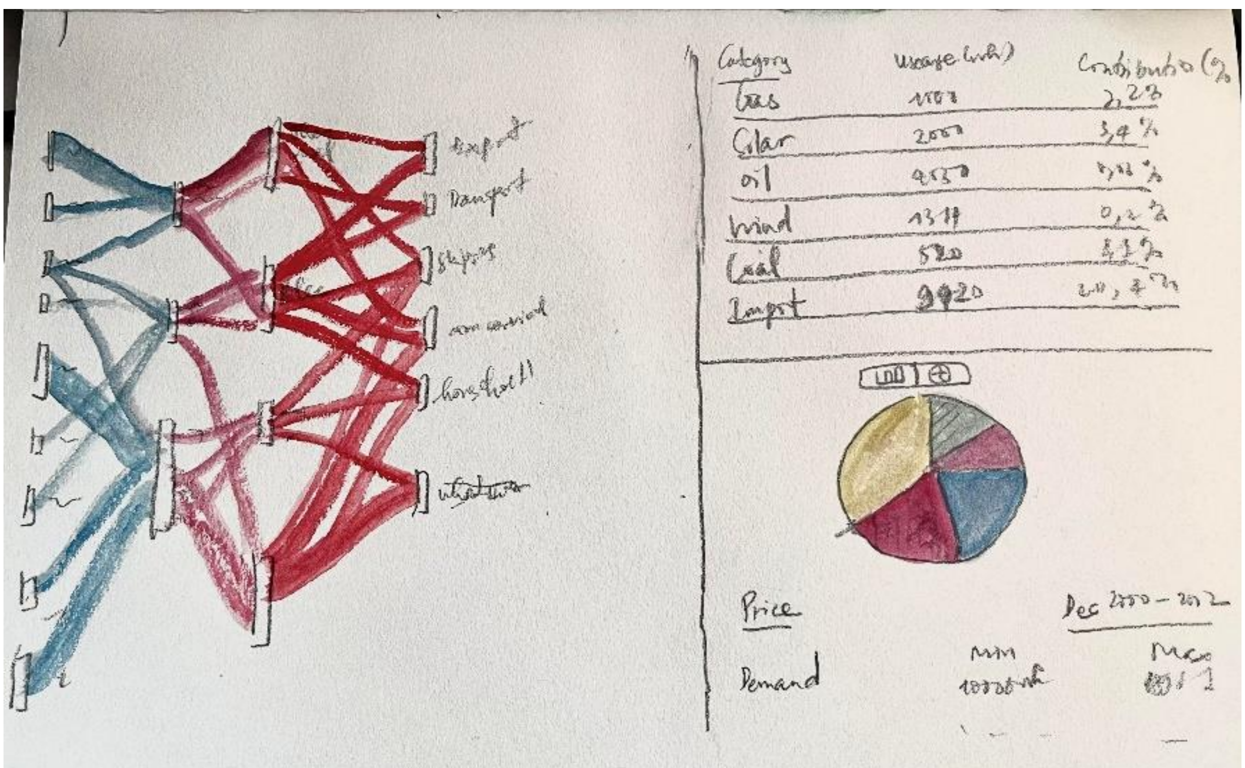

Figure 11. Low fidelity of production webpage.

As Sankey diagram mainly depicts the flows and links, the general look of Sankey diagram is designed as below:

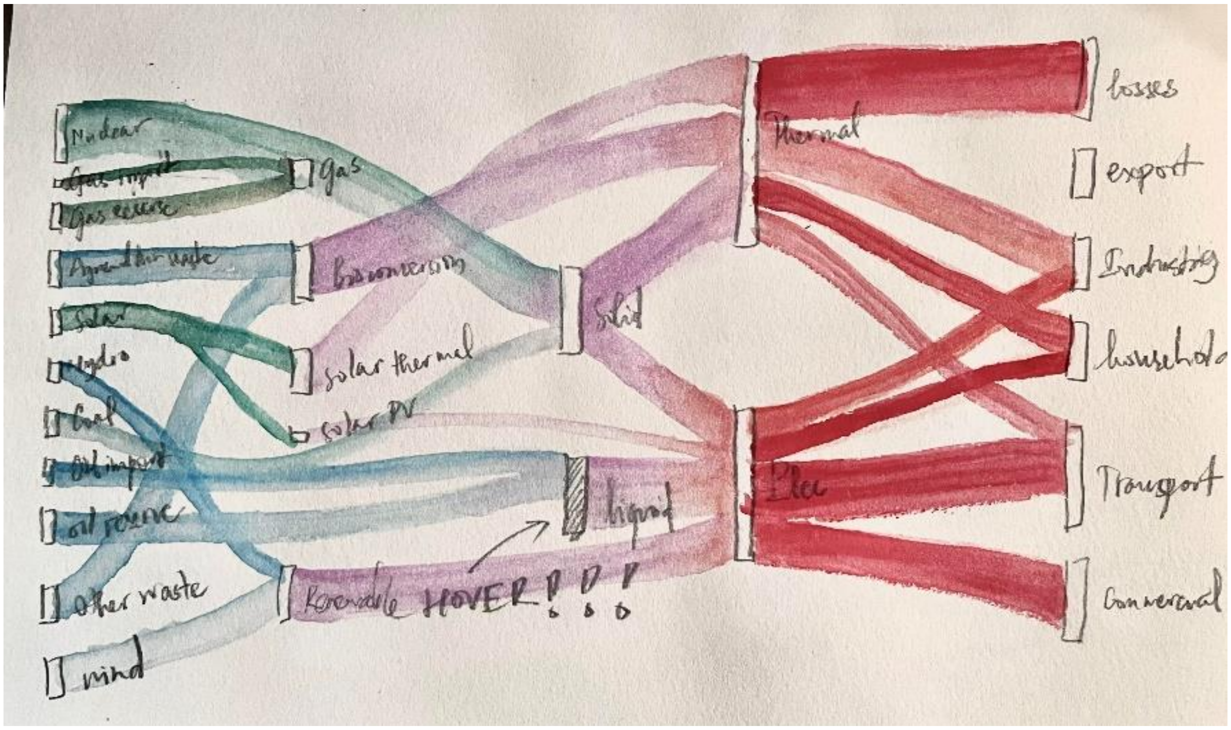

Figure 12. Low fidelity of Sankey diagram for energy flows.

The analysis can be found as below:

Plus points

- Identify what energy units is contributed to what energy production type and process with an ease (i.e., energy flows).
- Easily figuring out how much the production is changed across years.
- Audience can find out the energy produced by hovering on specific energy units.

Limit

- Large datasets cause the visualization to be overwhelmed.
- Cannot directly identify the pattern throughout a period since the link between energy units is more prioritized.

Features to Implement

- Transition when list of categories is clicked.
- Data appears when hovering on the energy units
- Update the data when users click on a year in the time bar.

Visual Encoding

- Color hue and saturation: To emphasize the linking.
- Width of the line: Demonstrate the size of energy unit is used (i.e., the contribution of one unit to the production of one energy).

Design Evolution

After developing the website with sketches, we had a good vision of what we wanted our visualizations to look like.

Bubble map

For the first webpage, we implement the bubble map before inserting the line chart on the side. The iteration of this map is first done with a random color scheme, which is not provide the visual effect we expected or meet the ground standard that we have indicated. However, having such a similar version from a low-fidelity prototype is a commendable effort.

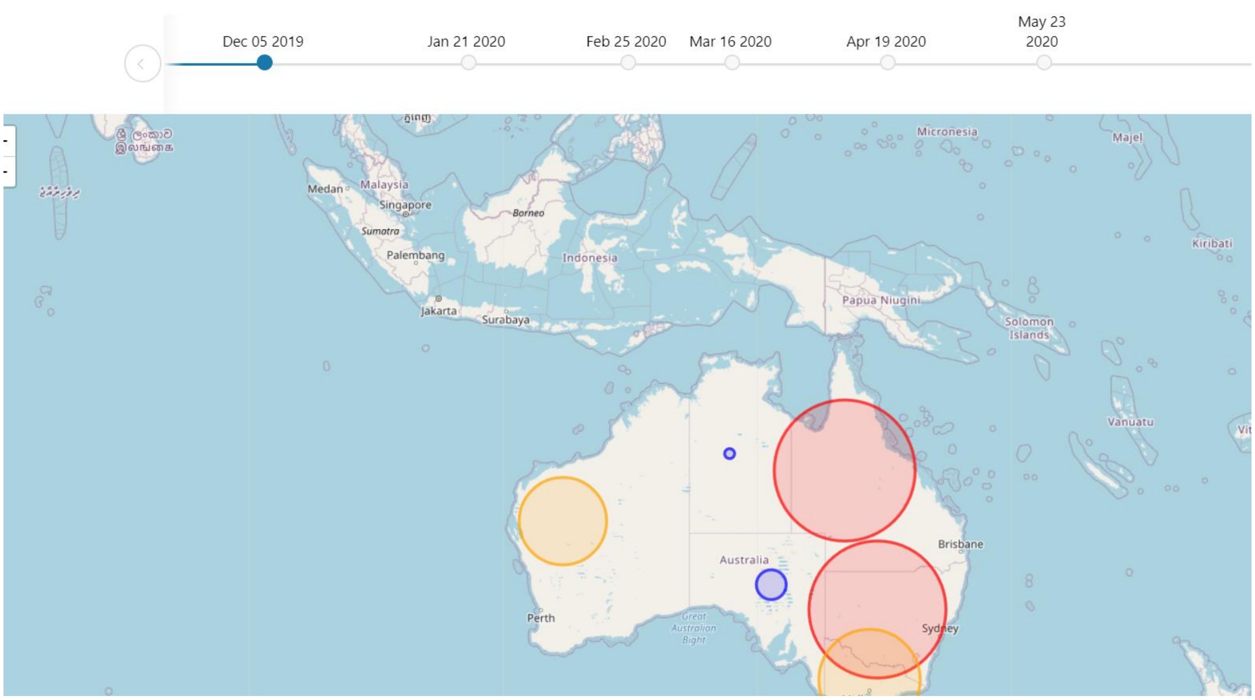

Figure 13. The first version of bubble map.

For the legends, we initially specified the size of the bubbles to represent power consumption, and it would look like this:

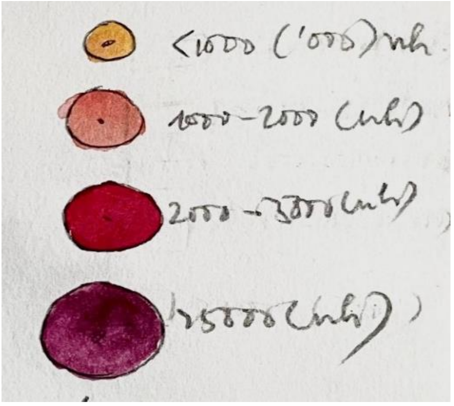

Figure 14. The legends for bubble map - first idea.

But soon we felt that it is no longer satisfied our purpose and by adding color, the visualization is more dramatic and vivid. So instead of just using the circle size, we changed that notation to a color bar:

Figure 15. The legends for bubble map – final idea.

The final iteration of the map after rescaling the color and adding the legends looks like:

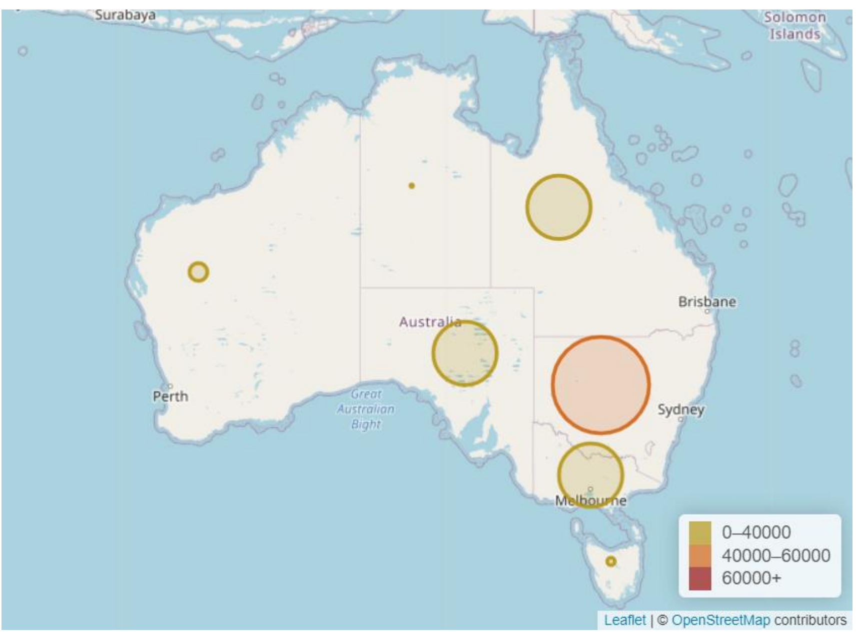

Figure 16. Bubble map.

Line chart

The creation of line chart is straightforward since it is the most commonly used graph. The first look is not optimal because each line has too many points, which confuses the users. Each line also has a varied tip form, resulting in inconsistency. Although the chart displays the data correctly, it is shown that the y-axis is not appropriately scaled and represent values.

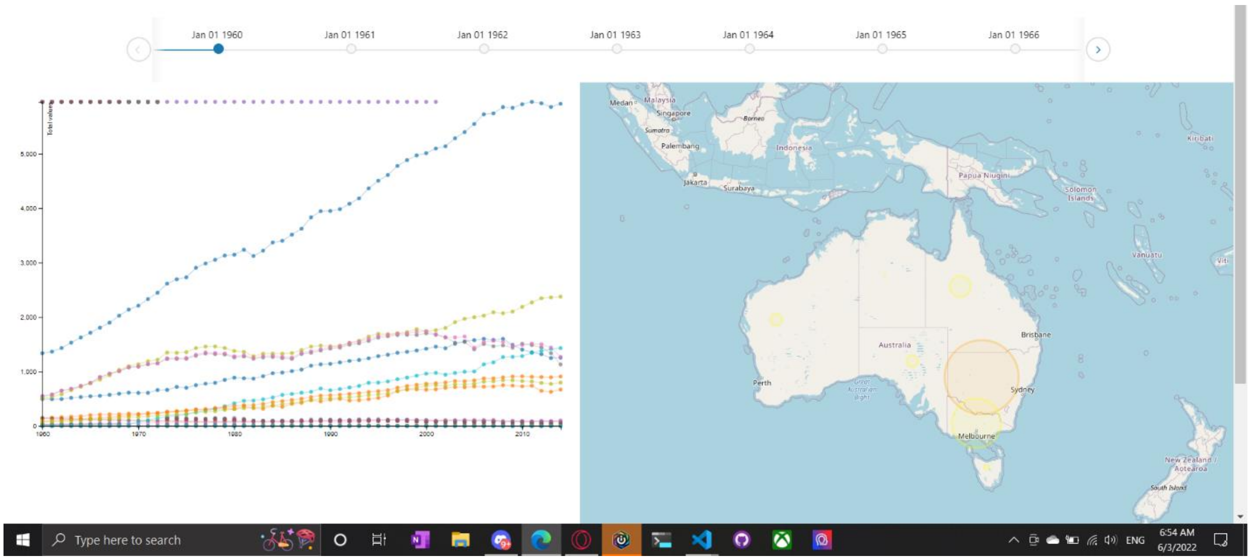

Figure 17. Line chart – First attempt.

At this point, we opted to create points for the year indicated on the x-axis only, with dots utilized as the shape. The property of the line is visible at the first glance, not only when mouse in the critical region.

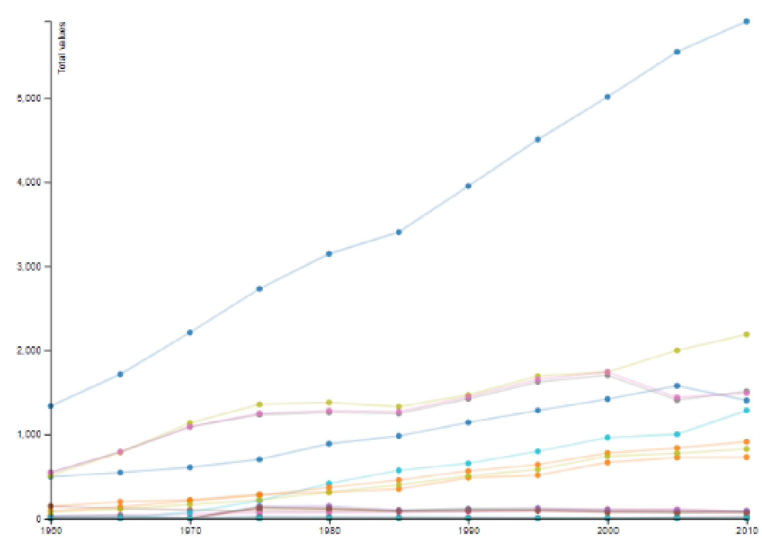

Figure 18. Line chart – After amendment.

Bar chart

For the bar chart, we initially started with all the energy units appear in the dataset. Figure 20 is the first chart created, and it contains minor text scaling errors. The bar chart does not immediately display any data and does not include any interaction or animation.

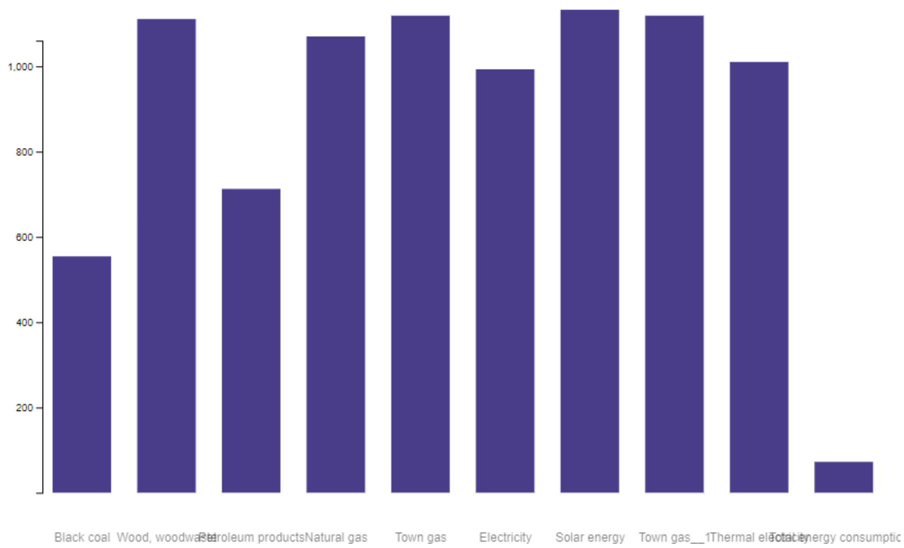

Figure 19. Bar chart – First attempt.

After that, we added tooltips that, when hovered over, show the value it represents. When the user hovers over a bar, it highlights that piece and brings the focus to that specific region of the chart, which is an extra filter tool. Figure 20 depicts the final iteration, while Figure 21 depicts the tooltip and opacity features.

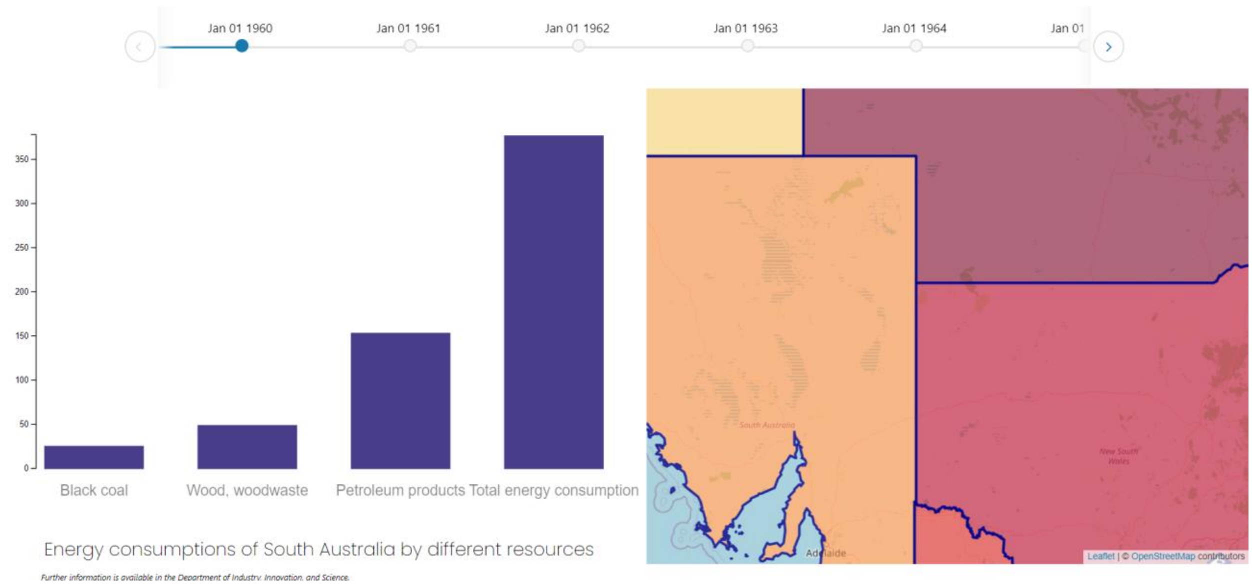

Figure 20. Bar chart – Final attempt.

Figure 21. Tooltips for bar chart.

Sankey diagram

The energy production had been visualized on a Sankey diagram, which is often used for linking and flows. The energy units are added as nodes and there is a link between different levels of energy in the production process. When hovered on one node, that node and the link from the node will be emphasized while the other is lightened. The first version of this diagram is solely controlled by a dropdown, which shows the year from 2015 – 2020.

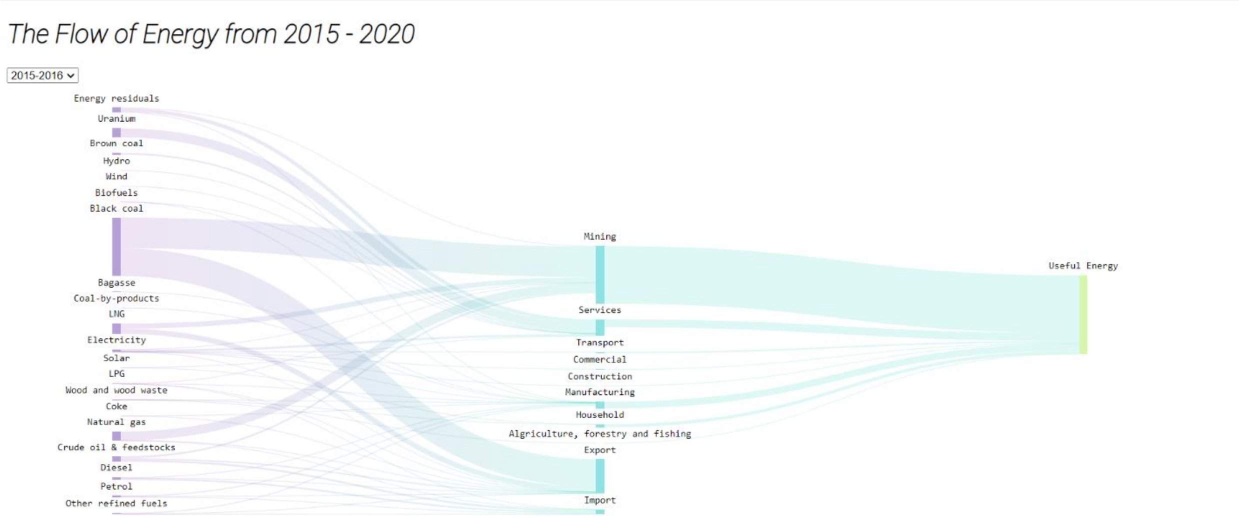

Figure 22. First version of Sankey Diagram.

As some of the nodes have a small width, which is difficult to mouse on, we have created a list of categories so that audience can control the energy units and link at ease. The drop down is also changed to time bar to maintain the consistency between websites.

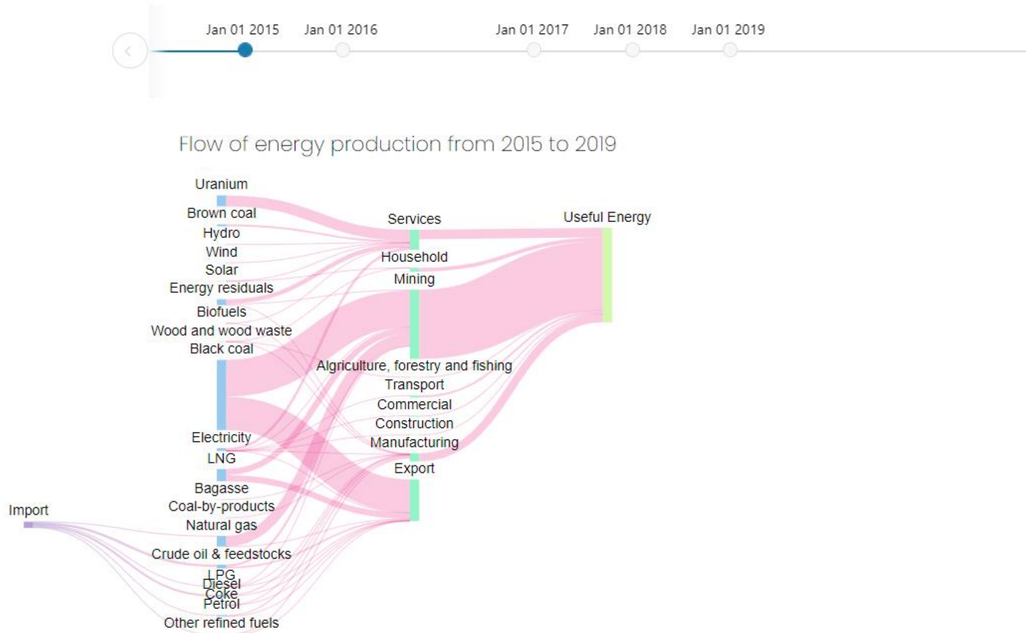

Figure 23. Final version of Sankey Diagram.

Usability Testing – Validation

After creating a website to visualize energy data, we perform some formative evaluations to validate the interface. Summative usability testing is being carried out in order to increase the accuracy of task completion, rates, and timings, as well as to discover underlying flaws in the visualization.

The summative examination is held in two days and is limited to six participants. Despite the number of participants in the assessment process is relatively small, the case study and data gathered adequately supply the measure and development stages.

5.1. Demographic data of participants

All participants were aged between 18-34 and have interacted with variety websites, especially social media, and magazine websites. Most people prefer to approach the data from diagrams and graph while some minors want to see the statistical results. Participants were asked if they had used some of the websites and applications for creating data visualization as Google Chart and Chart.js before, but the most used tool turns out to be Excel, which does not meet our expectation. The further details on participants characteristics can be found as follows:

Data Visualization Demographic Questionnaire (Responses)

Timestamp	ID. Participant ID (we will tell you what to put for this Question)	1. Which of the following includes your age?	2. What is your gender	3. What types of website have you interacted with?	4. Which way(s) do you prefer when approaching the data?	5. How confident are you in design (e.g., create charts and diagrams, make a presentation, formatting or drawing ...)?	6. Have you used any of the following websites/ applications?	7. How satisfied you are when using the application(s) listed in previous questions?	8. For what purpose do you use the website(s) or application(s) listed above?	9. How fascinated are you when you hear about energy issues?	10. Did you experience any of the following problem(s) when searching for energy or electricity data on a website?
3/11/2022 0:51:4(	01	18 - 24	Male	E-commerce websites (e.g., Amazon, Alibaba, Ebay, ...), Social media websites (e.g., Facebook, Instagram, ...)	See diagrams or graphs	4	Microsoft Excel	5	Exercise or assignment	3	It takes time to understand the data as it is too lengthy and statistically reported (e.g., the data is presented by table, ...)
3/11/2022 1:28:4:	02	18 - 24	Male	Magazine websites (e.g., Forbes, National Geographic, Time, Cosmopolitan ...), Blogs, Social media websites (e.g., Facebook, Instagram, ...)	See diagrams or graphs, View statistical results	3	Microsoft Excel	2	Exercise or assignment	4	Cannot find the data in a specific region, It takes time to understand the data as it is too lengthy and statistically reported (e.g., the data is presented by table, ...)
3/11/2022 2:46:21	03	18 - 24	Female	Magazine websites (e.g., Forbes, National Geographic, Time, Cosmopolitan ...), E-commerce websites (e.g., Amazon, Alibaba, Ebay, ...), Blogs, Social media websites (e.g., Facebook, Instagram, ...)	See diagrams or graphs	3	Microsoft Excel	2	Exercise or assignment	3	Cannot find the data in a specific region, It takes time to understand the data as it is too lengthy and statistically reported (e.g., the data is presented by table, ...), Unreliable data source
/12/2022 23:30:3	04	25 - 34	Female	Magazine websites (e.g., Forbes, National Geographic, Time, Cosmopolitan ...), E-commerce websites (e.g., Amazon, Alibaba, Ebay, ...), Blogs, Portfolio websites, Social media websites (e.g., Facebook, Instagram, ...)	See diagrams or graphs	5	Google Charts, Infogram, Chartblocks, Microsoft Excel, IBM SPSS Statistics	4	Exercise or assignment	3	Cannot find the data in a specific region, It takes time to understand the data as it is too lengthy and statistically reported (e.g., the data is presented by table, ...), Unreliable data source
/12/2022 23:39:1	5	18 - 24	Female	Magazine websites (e.g., Forbes, National Geographic, Time, Cosmopolitan ...), E-commerce websites (e.g., Amazon, Alibaba, Ebay, ...), Portfolio websites, Social media websites (e.g., Facebook, Instagram, ...)	See diagrams or graphs	5	Google Charts, Chart.js, Microsoft Power BI, Microsoft Excel	5	Exercise or assignment	5	It takes time to understand the data as it is too lengthy and statistically reported (e.g., the data is presented by table, ...)

Figure 24. Demographic Questionnaire's response.

5.2. Evaluation Tasks

There are five tasks in total, all of which focus on determining whether or not the audience can locate the data on the site. The activity requires participants to not only experience and engage with the website, but also to respond to the questions linked to the tasks.

Data Visualization Evaluation Tasks (Responses)

Timestamp	1. Participant ID (we will tell you what to put for this Question)	The region consume the largest amount of energy is:	The region consume the least amount of energy is:	Task 1 - Please rate the difficulty is task:	Petroleum Products in 2000:	Petroleum Products in 2010:	Task 2 - Please rate the difficulty is task:	Wood and woodwaste consumption of Melbourne in Jan 1 2003:	Task 3 - Please rate the difficulty is task:	The amount of Black coal is:	Task 4 - Please rate the difficulty is task:	Energy portion(s) that contributed to Service category is/are:	Task 5 - Please rate the difficulty is task:
6/11/2022 1:07:4	01	Melbourne	Perth	3	1741	1488.6	2	1500	5	1000	3	NA	5
6/11/2022 1:52:1	02	Sydney	South Australia	3	1741	2188.2	4	1430	3	23158	2	Uranium, Brown Coal, Energy Residuals, Electricity	2
6/11/2022 3:03:5	03	Queensland	Northern Terrirory	2	1741	2188.2	1	N/A	2	23158	1	Uranium, Brown Coal, Hydro, wind, solar, energy residuals, Biofuels, Electricity	2
6/11/2022 23:34:	04	New South Wales	Northern Territory	1	1741	2188.1	3	30	3	23158	3	Uranium, Brown Coal, Hydro, wind, solar, energy residuals, Biofuels, Electricity	2
6/11/2022 23:58:	05	NSW	Northern territory	1	1741	2188.2	2	30	1	23158	1	Uranium, Brown Coal, Hydro, wind, solar, energy residuals, Biofuel	2
6/11/2022 0:20:4	06	New south wales	Northern territory	1	1741	2188.2	2	30	1	23158	1	Uranium, Brown Coal, Hydro, wind, solar, energy residuals, Biofuels, electricity	2

Figure 25. Evaluation Tasks' response.

5.3. Reflection

By hosting the usability evaluation and taking notes, we found out some features need to be improved.

*Tasks.txt - Notepad

```
- Identify which region consume the largest amount of energy resources and which is the least.
- Find the total energy consumption of Thermal Energy in 2000 and fill the value in the form.
- In Western Australia, read the energy consumtion value of Natural Gas in Jan 1 2003
- The amount of Black coal is used as production resource
- Find out the energy portions that contributed to the 'Services' category and fill it in the form.


Feedback:

- The line chart needs to be separated a bit, when hovering over a line, the data should be displayed all the time, the size of the
- The bar chart should be visible because so many people can't find it
- scale the word bar chart bigger please
- hover on bar chart must have value
- to both onlick and hover, so that if the user clicks, it still shows the value (sankey)
- After clicking on the categories on the table, the link part is not relevant should be greyed out, and the link should be darker,
```

Ln 15, Col 124 | 100% | Windows (CRLF) | UTF-8

Figure 26. Note taken in the validation session.

Line chart

As illustrated in Figure 19, several of the lines are entangled. We only see problems when participants try to hover on a yellow line but are confused by the pink one. We then opted to lower the number of categories because the spread of data was not suitable for rescaling the chart. The data is then aggregated into two categories and is processed as indicated in the *Data Processing*. The aggregation was done in Excel with the following outcomes:

			Consumption of fuels													
	Clean	Fossil	Black coal	Brown coal	Coke	Coal by-products	Liquid/gas biofuels	Wood, woodwaste	Bagasse	Refinery input	Petroleum products	Natural gas	Town gas	Electricity	Solar energy	energy consumpti
1960-61	289.6	1696.4	495.3	144.4	0.0	0.0	0.0	145.3	27.1	546.7	510.0		27.9	89.3	0.0	1,336.6
1961-62	290.1	1758.8	492.3	152.8	0.0	0.0	0.0	139.6	28.2	576.4	537.3		27.9	94.4	0.0	1,365.8
1962-63	303.8	1887.2	494.3	162.2	0.0	0.0	0.0	134.3	36.5	648.8	581.9	0.1	28.3	104.6	0.0	1,432.2
1963-64	310.6	2027.3	517.0	170.9	0.0	0.0	0.0	129.4	34.8	682.6	656.8	0.1	29.8	116.5	0.0	1,531.4
1964-65	325.8	2164.9	531.5	178.4	0.0	0.0	0.0	124.2	44.1	735.1	719.9	0.1	29.2	128.2	0.0	1,625.3
1965-66	329.5	2320.6	545.9	198.8	0.0	0.0	0.0	119.7	41.4	791.2	784.7	0.2	30.7	137.5	0.0	1,713.1
1966-67	346.1	2512.3	554.3	208.6	0.0	0.0	0.0	115.2	49.4	896.3	853.1	0.2	31.8	149.5	0.0	1,805.8
1967-78	354	2680.2	572.0	214.9	0.0	0.0	0.0	111.0	49.4	961.0	932.3	0.2	32.9	160.5	0.0	1,898.9
1968-69	372.6	2846.0	592.5	216.8	0.0	0.0	0.0	106.6	53.3	1,007.5	1,029.2	2.1	35.0	175.6	0.0	2,025.9
1969-70	371.6	3017.6	612.5	229.0	0.0	0.0	0.0	102.6	45.7	1,077.5	1,098.6	29.5	0.0	193.8	0.0	2,137.6
1970-71	432.2	3051.8	607.2	219.5	0.0	0.0	0.0	98.9	50.2	1,091.0	1,134.1	74.4	0.0	208.7	0.0	2,210.3
1971-72	472.9	3173.7	609.2	238.6	0.0	0.0	0.0	95.7	56.1	1,136.8	1,189.1	102.2	0.0	218.9	0.0	2,331.2
1972-73	525.2	3270.8	660.5	240.8	0.0	0.0	0.0	92.3	55.5	1,151.9	1,217.6	144.1	0.0	233.3	0.0	2,447.8
1973-74	592.6	3773.2	662.9	262.6	139.8	128.9	0.0	92.5	56.4	1,230.1	1,348.9	172.5	20.6	250.5	0.1	2,615.2
1974-75	625.3	3845.0	717.6	270.8	149.0	131.3	0.0	89.9	59.7	1,231.9	1,344.4	189.2	20.0	266.4	0.1	2,694.8
1975-76	657.7	3837.6	702.2	286.8	135.6	123.4	0.0	87.4	63.0	1,233.7	1,355.9	211.2	19.3	276.6	0.2	2,730.8
1976-77	724	4025.8	749.7	304.0	132.9	118.7	0.0	81.9	68.2	1,289.8	1,430.7	256.2	18.4	299.0	0.3	2,905.9
1977-78	761.1	4111.2	776.8	297.4	130.3	116.4	0.0	80.9	68.3	1,331.1	1,459.2	283.0	17.2	311.3	0.4	2,985.1
1978-79	804.9	4155.5	792.7	312.9	142.1	126.4	0.0	82.0	59.9	1,320.4	1,461.0	314.9	16.5	331.0	0.6	3,053.0
1979-80	872	4168.0	846.5	321.0	137.2	118.0	0.0	83.6	60.4	1,312.8	1,432.5	362.6	16.4	348.2	0.8	3,131.4
1980-81	954.5	4070.9	887.6	312.0	116.4	114.7	0.0	84.1	68.5	1,261.7	1,378.5	416.0	16.4	368.5	1.0	3,146.3
1981-82	1018	4104.0	876.7	357.7	125.1	107.8	0.0	84.1	74.1	1,276.1	1,360.6	462.0	14.6	381.9	1.3	3,237.6
1982-83	1023.2	3897.6	869.6	329.4	96.9	85.0	0.0	85.1	71.9	1,235.6	1,281.1	466.2	14.8	383.6	1.6	3,122.7
1983-84	1067.7	4000.2	914.4	316.5	95.4	84.2	0.0	85.4	69.1	1,260.9	1,328.8	490.0	15.5	405.9	1.8	3,221.2

Figure 27. Excel output.

The line chart is then turned into this look:

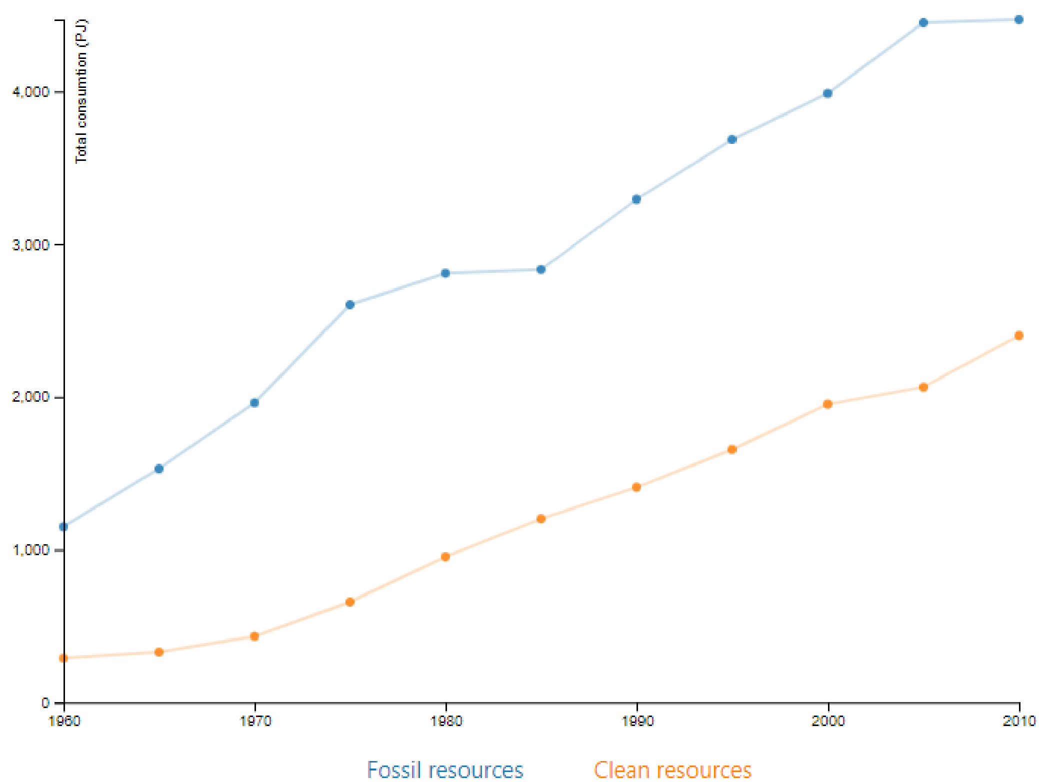

Figure 28. Line chart after validation.

Sankey Diagram

Participants appear to be overwhelmed by the list of energy units for the Sankey Diagram. They struggle to discover the attributes specified in the task description because the variable was at the end of the list. Recognizing the problem, we created a filter and a sort function so that the attributes could be sorted out or the list might appear in an order.

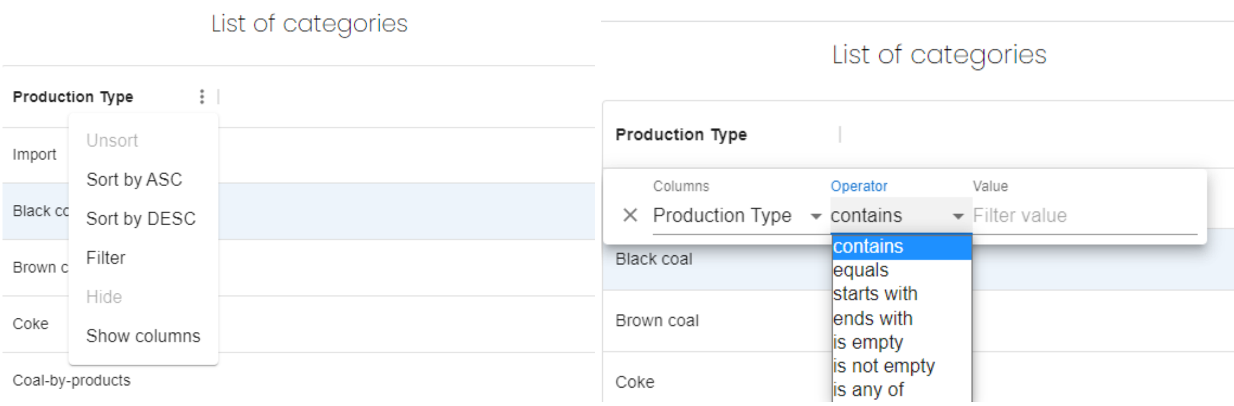

Figure 29. Filter and sort features.

Final Visualization

Finally, in this section, we will give screenshots of the diagrams in the final visualization that have not yet been presented. Many of the final visuals make use of code from the D3 Graph Gallery and unit material.

6.1. Consumption Visualization

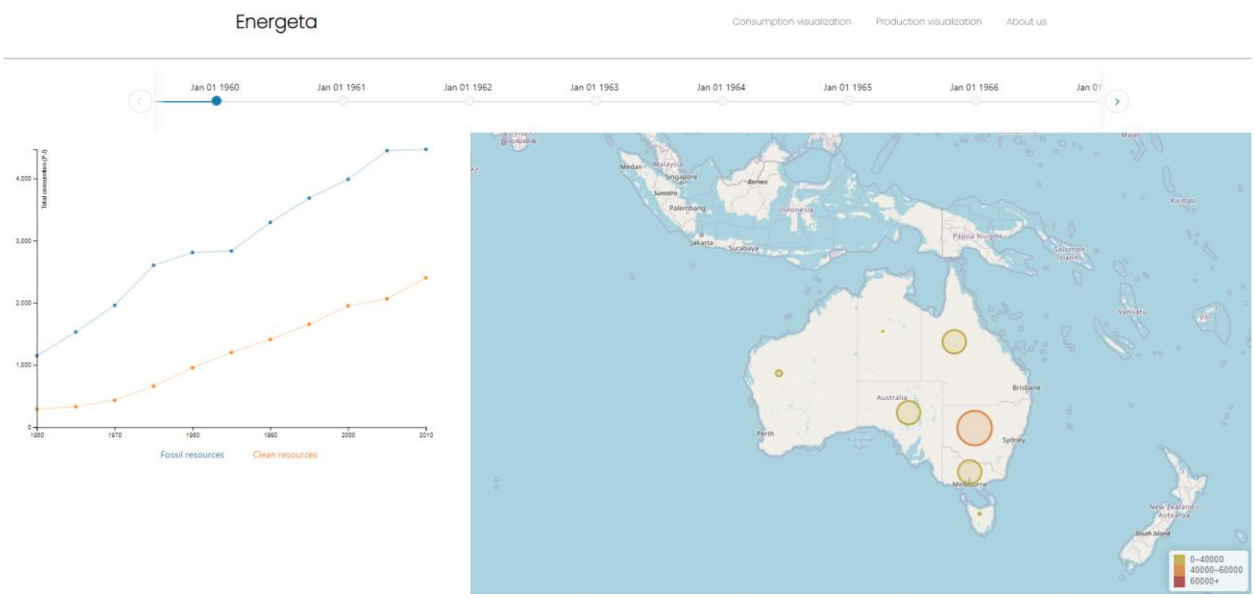

Figure 30. Energy consumption webpage.

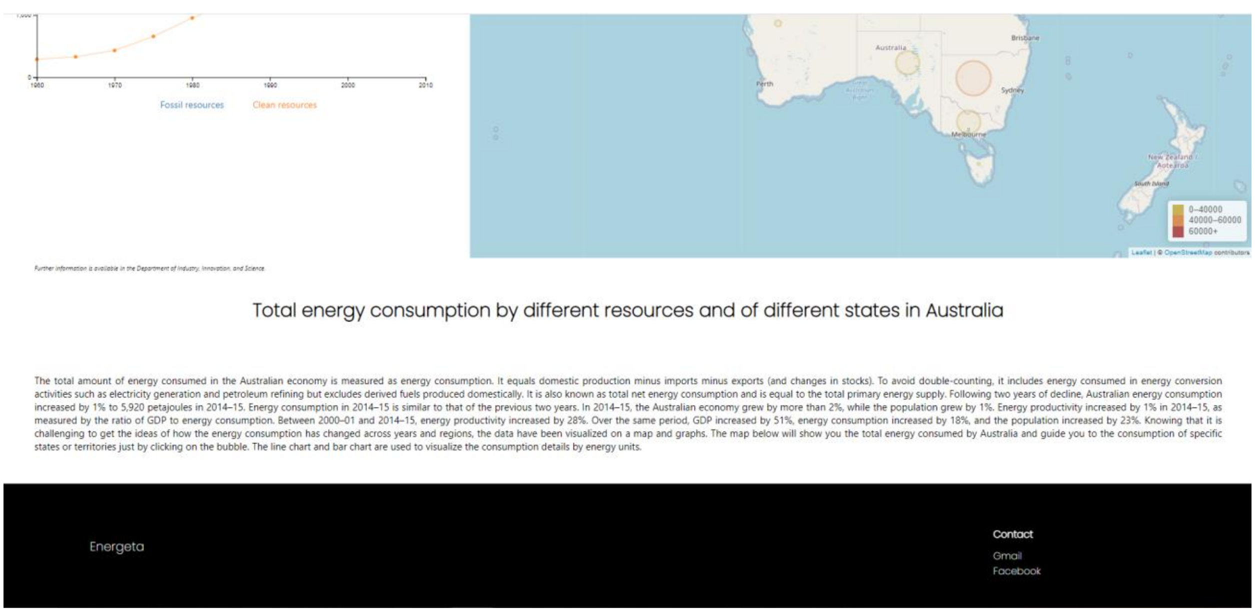

Figure 31. Energy consumption webpage.

6.2. Production visualization

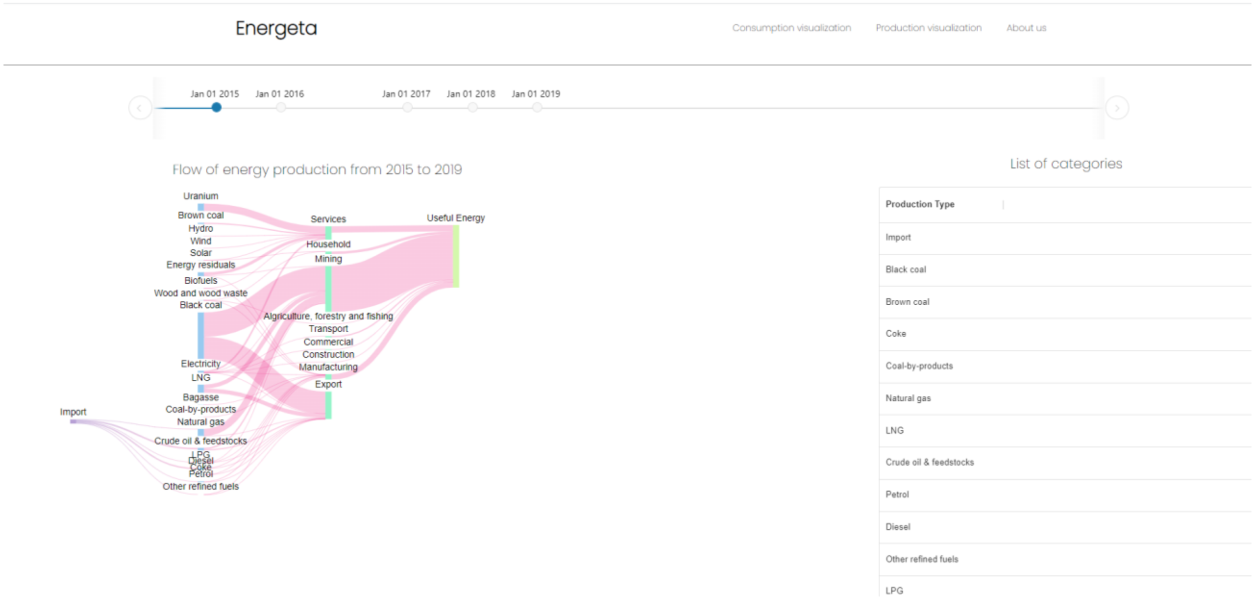

Figure 32. Energy production webpage.

Conclusion

After finishing our project, we learnt a lot about our topic and believe we accomplished all of the goals we set for our group. Overall, we learned how importance the data visualization process is and what should be done during the data transformation. The project has assisted us in gaining insight into their massive volumes of data, recognizing new patterns and errors, making sense of these patterns, and paying attention to regions that represent red flags or progress.

In addition, we discovered that being prepared is essential while moderating a usability test. Although we do not have the opportunity to conduct the testing session on a wider scale and to invite some primary and secondary users, the data gathered accurately describes what we are lacking and what needs to be corrected. This approach gave us a better understanding of how to do an evaluation.

These process, in turn, drives us forwards.

Reference

1. OpenNEM
opennem.org.au. (n.d.). *OpenNEM: NEM*. [online] Available at:
https://opennem.org.au/energy/nem/?range=all&interval=1y

2. Government data
data.gov.au. (n.d.). Datasets - Page 1. [online] Available at:
https://data.gov.au/data/dataset?tags=Boundary [Accessed 13 Jun. 2022].

3. Graphs
observablehq.com. (2020). Gallery. [online] Available at: https://observablehq.com/@d3/gallery

Validation Raw Data

1. Data Visualization - Evaluation Explanatory Statement
2. Data Visualization Demographic Questionnaire
3. Data Visualization Evaluation Tasks
4. Data Visualization Post Study Questionnaire 1 (Response)
5. Data Visualization Post Study Questionnaire 2 SUS (Response)