

Drexel University
College of Computing and Informatics

INFO 250 Information Visualizations

Individual Project 1A

Prof. Dr. Chaomei Chen

Lixiao Yang
ly364@drexel.edu

January 2022

Table of Contents

1.	Data Set Chosen and Visualizations	2
1.1	Data Set Chosen	2
1.2	Four Visualizations.....	2
2.	Analysis and Discussions.....	5
2.1	Matrix Plot.....	5
2.2	Sunburst Diagram.....	5
2.3	Linear Dendrogram	6
2.4	Sankey Diagram	6
3.	Recommendations.....	7

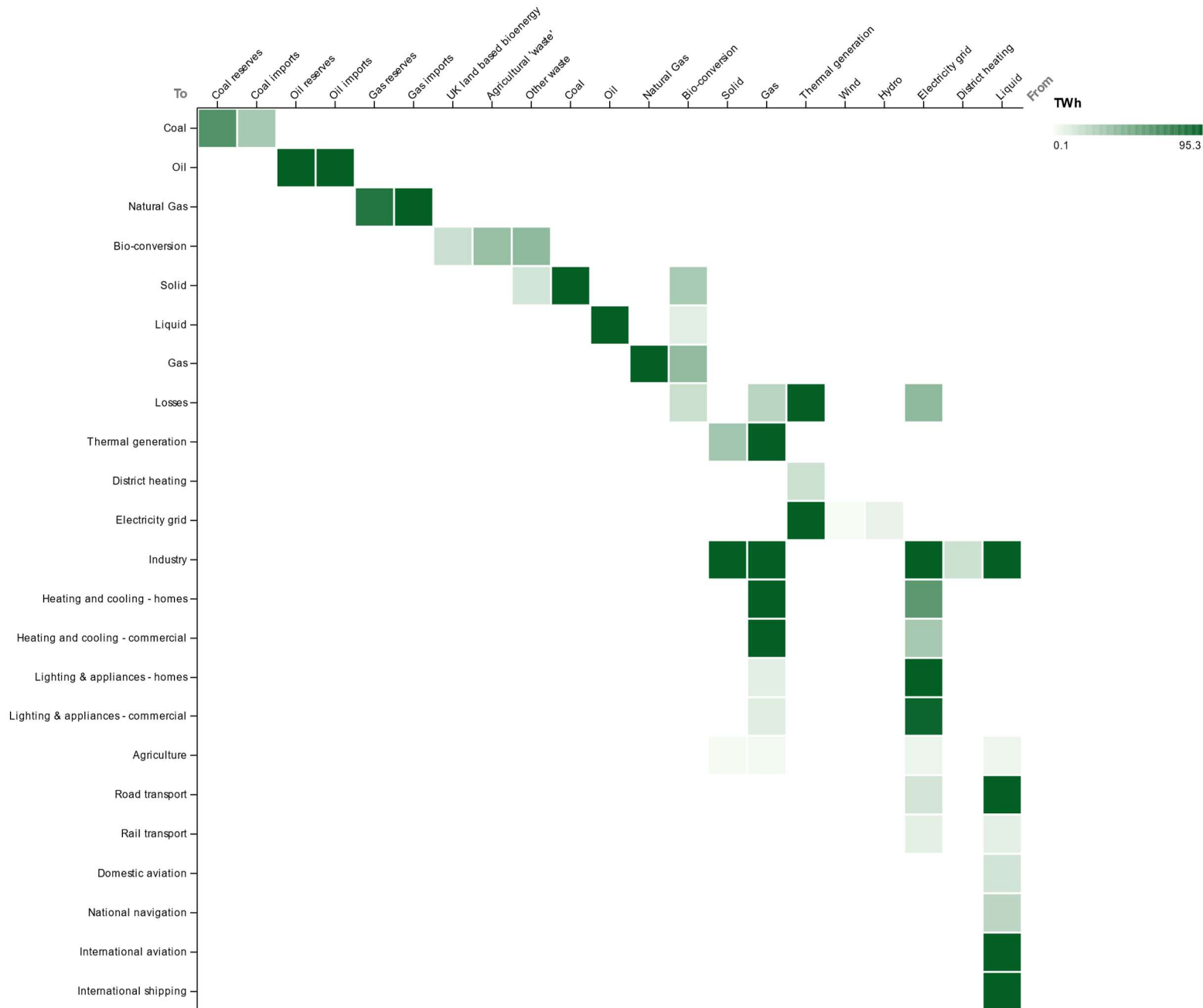
1. Data Set Chosen and Visualizations

1.1 Data Set Chosen

The data set chosen for this project is “Energy flows in UK (2050)” from the source: gov.uk.

1.2 Four Visualizations

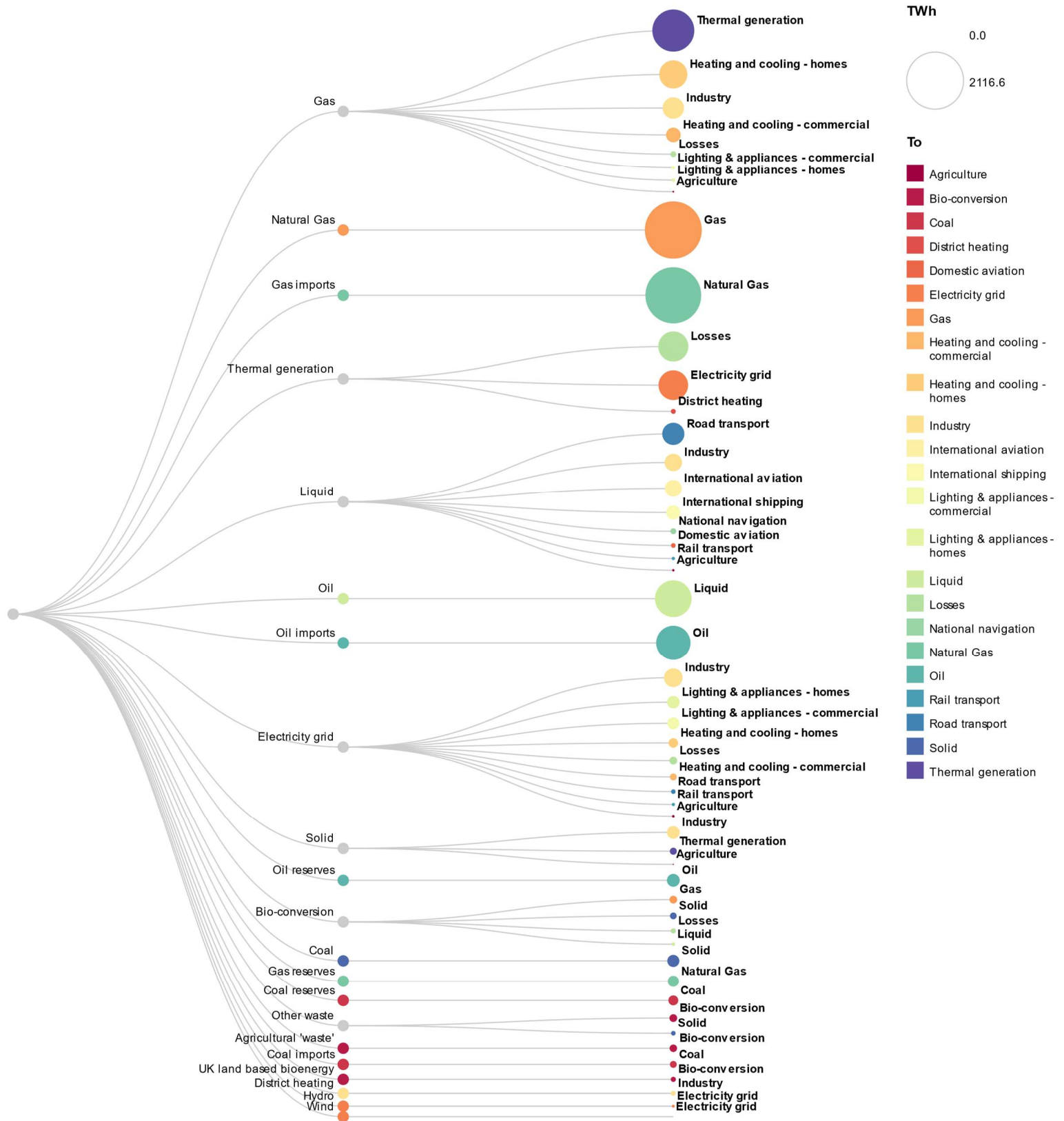
● Matrix Plot



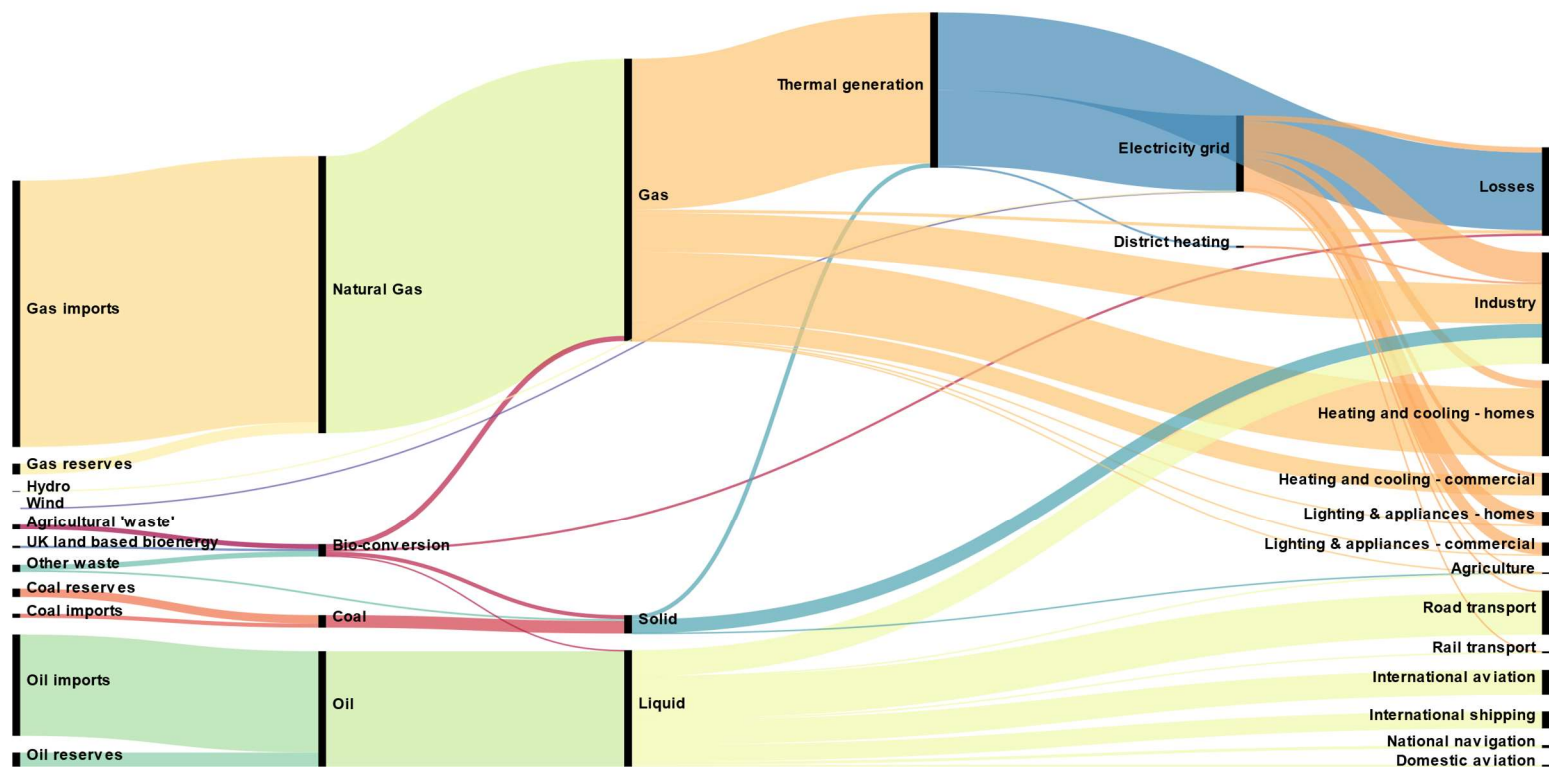
- **Sunburst Diagram**



● Linear Dendrogram



● Sankey Diagram



2. Analysis and Discussions

2.1 Matrix Plot

This matrix plot is created by using the source and destination of energy as the two axes of the matrix, and each intersection is the source of energy in the column and the destination of energy in the row, and the color shades of the intersection represent the size of the energy quantity.

- This matrix shows the **saliency and relationship patterns** of energy flows, with most of the energy sources concentrated in some of the destinations which shows **a certain degree of covariation** among energy flows.
- **Strengths:**
 - 1) Easy to observe the relative size for each type of energy flow.
 - 2) Easy to compare the correlation between different energy flow groups.
- **Weaknesses:**
 - 1) There are too many indexes on the axes which caused much blank space and reduced the efficiency of the plot display.
 - 2) It is hard to compare the proportion of different sources of energy flows.

2.2 Sunburst Diagram

This sunburst diagram shows the source of energy and the destination of energy from inside to outside in the form of double circles, and the colors in the diagram are labeled by the source of energy.

- This diagram shows the **proportion and categorization patterns** for different energy sources and reveals the **relative magnitude** of separate flows of energy.
- **Strengths:**
 - 1) The proportion of energy sources is obvious in arc lengths of different colors.
 - 2) Each energy source is easy to distinguish from others.
- **Weaknesses:**
 - 1) Hard to compare and observe the proportion of volume within different energy destinations.
 - 2) The large difference in quantity between different groups leads to the existence of overlapping or blocking of some energy flows.

2.3 Linear Dendrogram

This linear dendrogram shows the details of the energy flow in each group in relation to the magnitude of the flow within the group, with the different colors marking the direction of the different energy flows.

- This dendrogram shows the **categorization and contrast patterns** of energy flows, providing contrast within every energy source group and with every energy destination group.
- **Strengths:**
 - 1) Display every group clearly without overlapping or blocking.
 - 2) The proportion within each energy group is easy to compare.
 - 3) The total amount for different is easy to observe by colors.
- **Weaknesses:**
 - 1) Can not display the volume of different energy sources.
 - 2) Hard to observe the correlations between different energy sources.

2.4 Sankey Diagram

This Sankey diagram represents the flow of all energy sources in the form of lines while including the flow of intermediate processes in the iterative operations. Different colors are used to increase the differentiation, while the width of the lines indicates the relative size of the flows.

- This diagram shows the **relationship, trend and categorization patterns** of energy flows, detailly showing the interrelationship between energy groups and demonstrating the more implicit information about the energy flow.
- **Strengths:**
 - 1) Shows the middle layers in the whole energy flow process.
 - 2) Easiest to understand the energy flow process with clear information.
 - 3) No overlapping or blocking for category information.
 - 4) Easy to compare volume within different data groups.
- **Weaknesses:**
 - 1) Hard to compare the definite value for different energy flows.
 - 2) Some of the energy flows are not obvious enough to observe.
 - 3) Color matching can be more efficient.

3. Recommendations

The best visualization recommended for this data set is the **Sankey Diagram**.

The justifications are listed below:

1. The diagram presents the abstract data in an intuitive and meaningful graphical way.
2. Compared with other diagrams, it is the only one that can show the middle layers in energy flows (implicit information in the data set).
3. This diagram is easier to reveal insightful patterns and boost the storytelling process for energy flows.
4. It losses the minimum information during transforming raw data into the diagram.
5. It conveys a clear overview of the data as one of the tasks by data type taxonomy.