Jameson D. Brothers: jbrothers@clarku.edu

Ben M. Seidenberg: bseidenberg@clarku.edu

William J. Setlik: wsetlik@clarku.edu

Christopher K. Siems: csiems@clarku.edu

Affordable and Clean Energy Exploration

Idea

Our project will be an exploration of the data surrounding the United Nations' Sustainable Development Goal #7, ensuring access to affordable, reliable, sustainable, and modern energy for all. We chose to work with this topic due to its significant importance and its wealth of existing data, including data generated and open-sourced by the United Nations itself. Energy is one of the most important aspects of modern life, which is reflected in the United Nations' inclusion of goals related to energy in this initiative, and yet there is so much work left to do to make energy more available, reliable, and clean. Studying the current energy paradigm is necessary for completing the United Nations' goals, and that study is what we intend to do.

Plan

Data Collection - Our data, as of now, consists of transcripts from past United Nations General Assembly debates, international energy statistics, and international greenhouse gas emission statistics. All of these were downloaded from Kaggle and are suitable for in-depth analysis of statistical data and natural language.

Data Processing - The data processing step will include natural language processing of the United Nations General Assembly debates to focus the data on energy discussions. The other two datasets will need to be cleaned by filtering out missing values and narrowing the data to just the key features that we end up exploring, since the data contains a wide variety of points.

Analysis - Our analysis will start with a sentiment analysis of the United Nations General Assembly debates to gauge opinions on clean energy. We will compare sentiment and mentions of clean energy to notable events in the data time periods to see if there is correlation. Greenhouse gas emissions can be compared to clean energy levels and climate goals to see the effect of the UN's efforts. We can cluster countries into groups based on energy amounts and explore trends between Gross Domestic Product and energy availability, for example.

Visualization - We will focus on displaying the trends found in our analysis using Matplotlib. This will include graphs of the trends in the data shown with the relevant statistical distributions and graph types, scatterplots to show how data was clustered based on its features, and word clouds derived from the United Nations General Assembly debates.