# International Energy Statistics

Sai Kiran Gontyala, Yoshitha Gattu
Department of Information Systems, California State University
Los Angeles

**Abstract:** The Energy Statistics Database contains comprehensive energy statistics on the production, trade, conversion and final consumption of primary and secondary; conventional and non-conventional; and new and renewable sources of energy. The International Energy Statistics dataset has covered all about the growth of energy as an international trend. This dataset was published by the United Nations Statistics Division on the UNData site. The United Nations Statistics Division is committed to the advancement of the global statistical system. They facilitate the coordination of international statistical activities and support the functioning of the United Nations Statistical Commission as the apex entity of the global statistical system. This dataset contains data of energy trade and production from 1990-2014.

Data information:
URL: https://www.kaggle.com/unitednations/international-energy-statistics#all_energy_statistics.csv
Dataset size: 59.7 MB
Account: Elastic Cloud
Kibana version: 7.2
Memory size: 180 GB
CPU Speed: 2.195 GHz
Storage: 682 GB

## 1. Introduction

Based on the list of data approved by our instructor, we did some research which lead to exclusively decided data which would be used for this project. We then manipulate and filter the datasets via the following steps:

- International Energy Statistics; data size is 59.7 MB.
- Cleaning down the information to have a detailed analysis of the production and consumption of various sources of energy.
- From the dataset, we then sorted out the amount of energy produced and consumed, to see which locations are consuming the highest energy and also to better understand about the major source of energy.
- The tools used were ElasticSearch and Kibana.

## 2. Importing and Discovering the dataset

### 2.1 Tools and data processing

- We extracted our Energy Statistics data from the corresponding website in a .csv format: International Energy - Global trade and production 1990-2014 Statistics from Kaggle.com.


**Figure 1: Data Processing**

- We filtered out the dataset and directly uploaded the file in '.csv' format into the ElasticSearch server using the import file button on the homepage.
- Then, we verified the dataset imported using a Get command.
- We created ElasticSearch queries to discover the data.
- Lastly, we used Kibana - Version 7.2 to reproduce the selected data in the form of information by generating the appropriate graphs, maps and charts.

### 2.2 Verifying the Dataset

- Once we understood the flow of how the data processing takes place in the ElasticSearch server, we imported and then checked if the data is uploaded.
- We verified the data using the command ' /_cat/indices?v ' in the GET field from the API console of the Deployment tab.

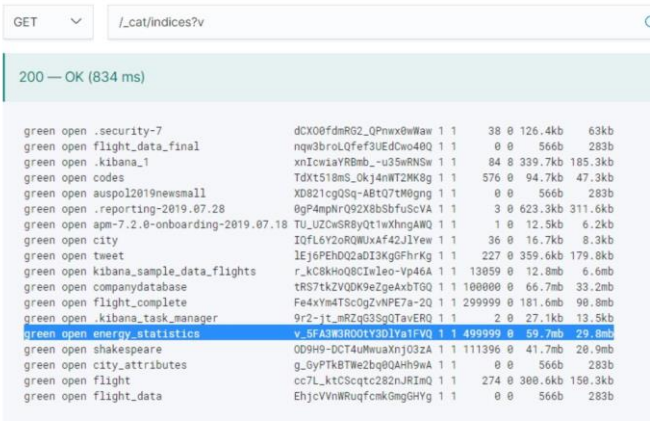


**Figure 2: Verifying the data**

## 2.3 **Defining the Index Patterns**

- Index patterns help Kibana to know which Elasticsearch indices we want to explore.
- We created an index pattern for the Energy statistics data set, which has an index named ‘energy_statistics’.

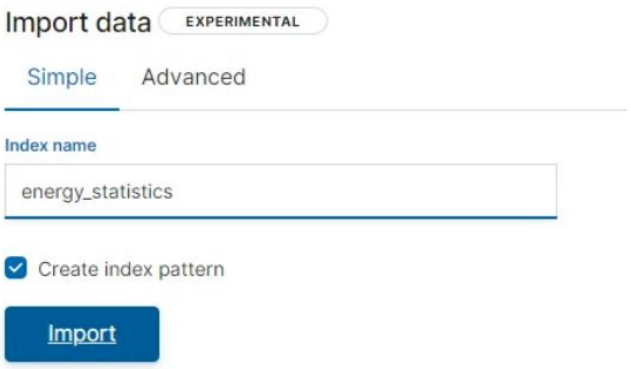


**Figure 3: Creating Index Pattern**

## 2.4 **Discovering the Dataset**

- Using the Discover application, we entered an Elasticsearch query to search our data and filter the results.
- In the search field, we entered the string ‘country_or_area : “United States” and year>2013’ and selected Update/Refresh.

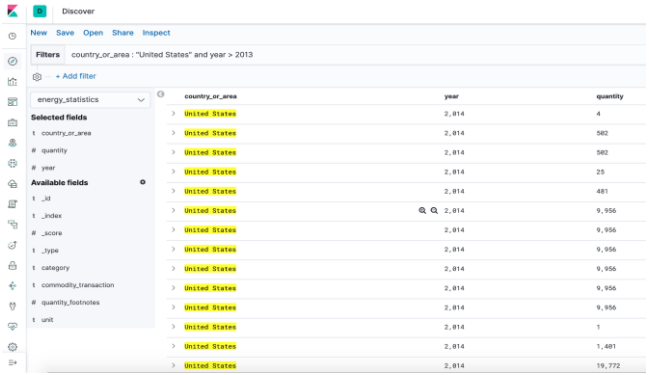


**Figure 4: Discovering the data**

## 3. Visualization Using Kibana

In the Visualize application, we can shape our data using a variety of charts, tables, and maps. We created the following visualizations with the help of our dataset.

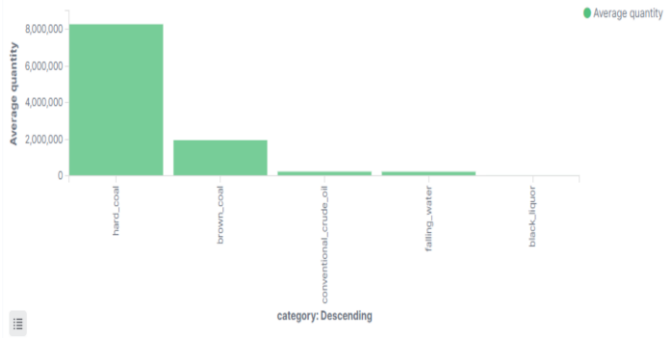


**Figure 5: Average Quantity of Energy Consumed**

Figure 5 is a Vertical Bar graph which displays the average quantity of energy consumed by each energy source. By analyzing the above bar chart, one can see the Hard_coal is utilized the most and black_liquor is consumed the least.

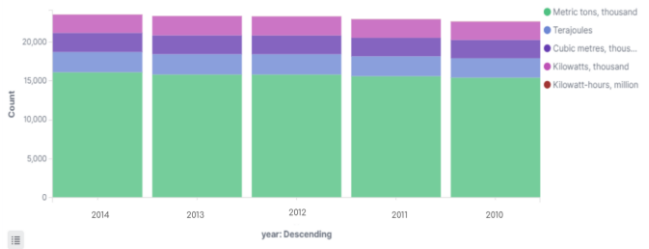


**Figure 6: Units of measurement of Energy**

Figure 6 shows the units of measurement of energy used. In the graph are different units for energy: Metric tons, Cubic Meters, Tera Joules, Kilowatt and Kilowatt-hour which signify the energy usage of various sources of energy such as

Electricity which is measured in Kilowatt-hours while energy usage of Natural gas is Cubic meters for various years.

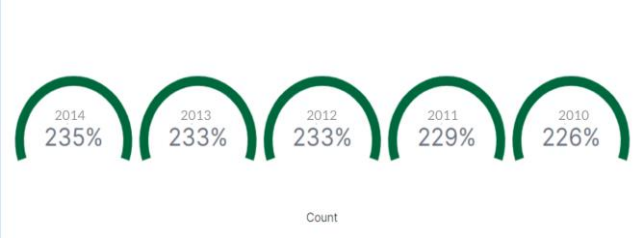


**Figure 7: Percentage of Fuel Consumed from 2010-2014**

Figure 7 shows the percentage of Fuel consumed from the year 2010 to the year 2014. From the above graph, we notice that the fuel consumption is increasing each year with 235% being the highest in 2014 (data available till 2014)*.


**Figure 8: Highest Energy Consumption Country**

Figure 8 is a tag cloud visual representation which visualizes the information pertaining to the country which consumes the highest energy. According to the data we found out that Germany consumes the highest amount of energy followed by the United States and Poland.

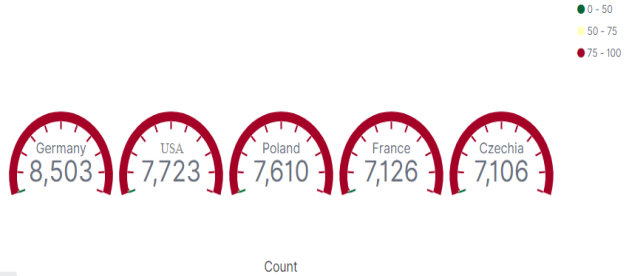


**Figure 9: Fuel Consumption in Each Country**

Figure 9 demonstrates the percentage of the Fuel consumed to produced in each country. The graph shows that all countries utilize more than 75% of their energy produced respectively where Germany tops the list of countries producing the highest amount of energy.

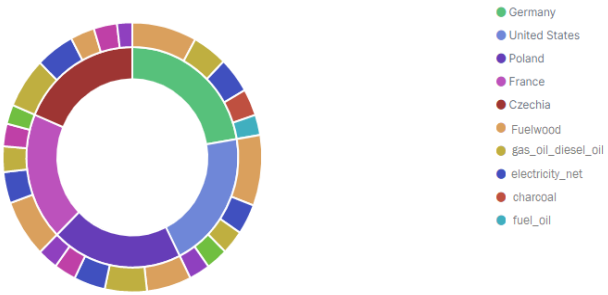


**Figure 10: Major Source of Energy in each Country**

Figure 10 illustrates the Major source of energy used in each country. The analysis shows that fuelwood is widely used in Germany, the United States, France, Poland where as gas_oil_diesel_oil is commonly used in Czechia.


**Figure 11: Major Source of Energy Per Year**

Figure 11 gives the Major Source of energy produced and consumed from the year 2005-2014. We observe that Fuelwood is playing a vital role through all the years and it is highly produced and utilized.

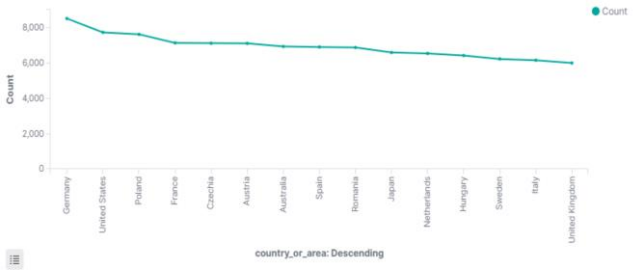


**Figure 12: Countries and Their Energy Consumption**

Figure 12 is a line graph that illustrates the amount of energy consumed by various countries. The line graph represents that Germany consumes more

than 8000 Metric Tons of energy followed by The United States and United kingdom being the last.

`

| year: Descending | country_or_area: Descending | Max quantity |
|---|---|---|
| 2014 | United States | 4,340 |
| 2013 | United States | 4,306 |
| 2012 | United States | 4,291 |
| 2011 | Germany | 3,885 |
| 2010 | Germany | 3,626 |
| 2009 | Germany | 3,223 |
| 2008 | Germany | 2,989 |
| 2007 | Germany | 2,817 |
| 2006 | Germany | 2,446 |
| 2005 | Germany | 2,126 |

**Figure 13: Countries which produced highest energy in each year**

Figure 13 exhibits the countries which produced the highest amount of energy in each year from 2005-2014. The chart illustrates that The United States produced highest energy in 2014,2013 and 2012 where as, Germany produced highest energy from 2005-2011.

## 4. Summary

- Successful application of the many tools learned in class: ElasticSearch and Kibana to use and visualize the data from any dataset.
- By creating and running individual queries in the discovery tab, we were able to run through some specific information about the Global Energy Production and Consumption.
- After visualizing the data using Kibana, we discovered that the Energy Statistics of a nation plays a major role in a country's economy.

## 5. Github URL

https://github.com/sgontya/InternationalEnergyStatistics

## 6. Reference

https://www.kaggle.com/unitednations/international-energy-statistics#all_energy_statistics.csv