

**Energy Generation of US For Last 3 Decades**

**Data Visualisation - Exploratory Data Analysis**

**B9DA106**

**Terri Hoare**

BHARATH MELAPPILLY RAJAGOPAL

STUDIENT ID: 10376807

DATE: 18-12-2018

Table of Contents

[Introduction 2](#_Toc532935201)

[2. Visualization with Tableau and R 3](#_Toc532935202)

[2.1 Most generated energy 4](#_Toc532935203)

[2.2 Most energy generated state 5](#_Toc532935204)

[2.3 Trend of energy generation 6](#_Toc532935205)

[2.4 Trend of Producer type in generating energy 7](#_Toc532935206)

[2.5 Animations showing energy generation patterns 8](#_Toc532935207)

[2.6 Energy generated by states in 2017 9](#_Toc532935208)

[2.7 Nuclear generation 10](#_Toc532935209)

[3. Conclusion 11](#_Toc532935210)

# Introduction

Data visualization is that the graphical illustration of information and data. By using visual components like charts, graphs, and maps, data visualization tools give an easier way see and perceive trends, outliers, and patterns information.

In the world of huge information, data visualization tools and technologies are essential to investigate huge amounts of knowledge and build data-driven selections.

As the part of our assignment we have worked on visualizing the data set on “Energy Generation of US for the Last 3 Decades” which had 5 attributes

* Year
* State
* Producer Type
* Energy Source
* Generation

The data set contained 49550 rows , the year attribute represents the years from 1990 to 2017, State represents the various energy producing state in US, Producer type are the companies who generates energy , Energy source represents the type of energy resources generated during the particular period and generation gives the energy generated in megawatt hours

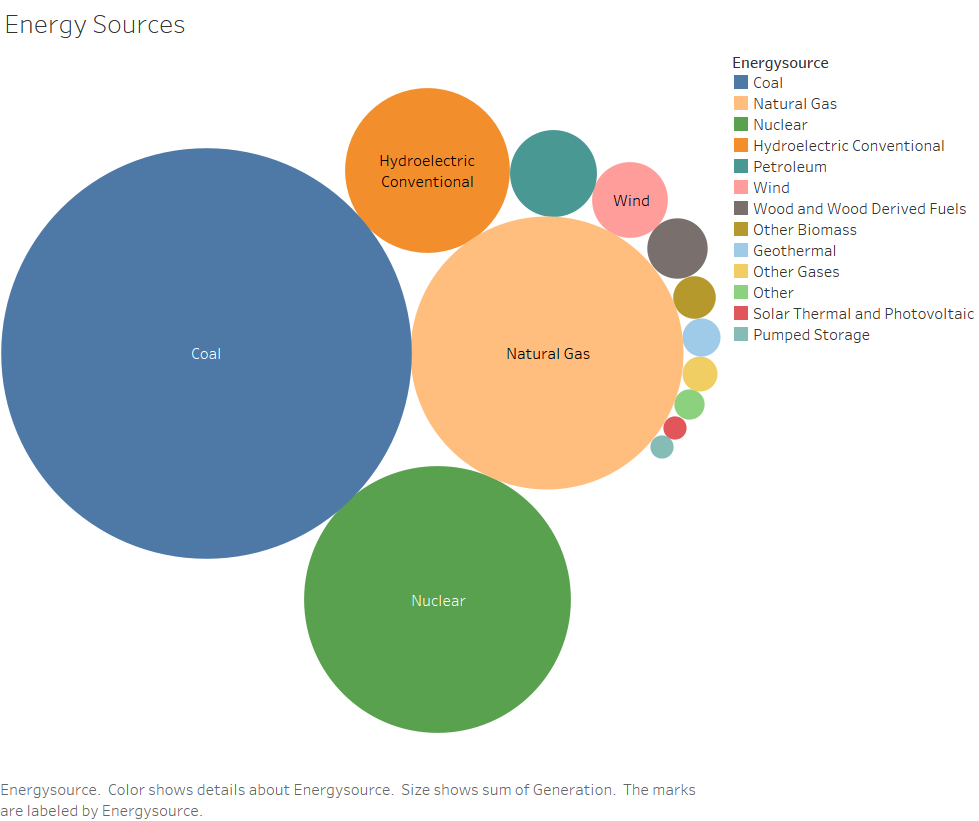
We took the data set from <https://www.eia.gov/electricity/data/state/> and it was last updated on 2017. With this data set we are trying to visualize the county’s energy production in various criteria.

# 2. Visualization with Tableau and R

Tableau is the widely used and friendly data visualization tool. Therefore, its technology is there to support complicated computations, data mixing and dash boarding for the aim of making lovely visualizations that deliver insights that can't simply be derived from looking at a computer program. It’s climbed to the highest of the information visualisation heap thanks to its dedication to the current purpose. Using drag-n-drop functionalities of Tableau, the user will produce a very interactive visual among minutes. The interface will handle endless variations whereas additionally limiting you from making charts that are against data visualisation best practices. There are many various forms of visualization choices obtainable in Tableau that enhances the user expertise. Tableau will handle countless rows of data with ease. different types of visualization will be created with an oversized quantity of data without impacting the performance of the dashboards.

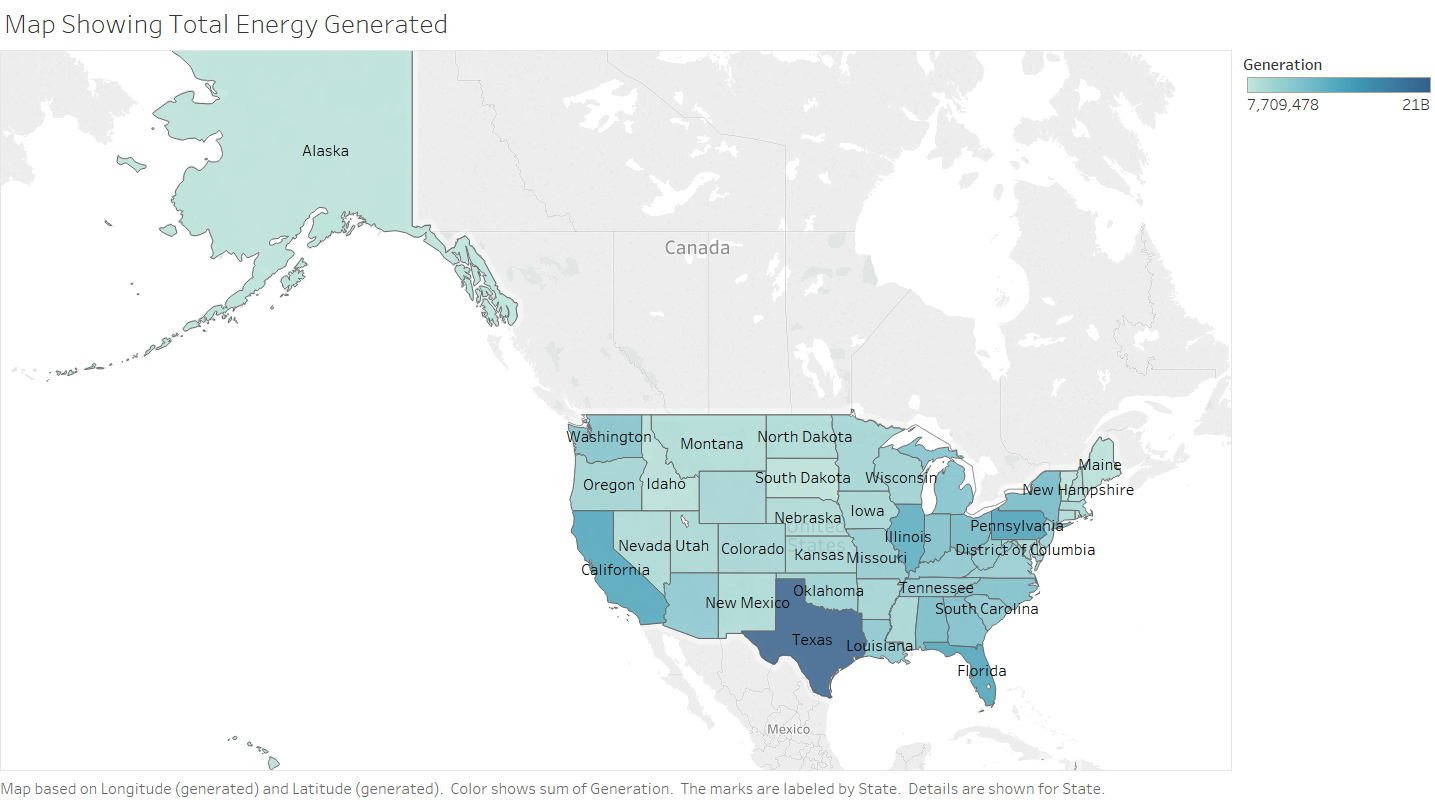
R is a language than a software in contrast, that means we have to write our script and code to generate a plot in R or we have to find some libraries that does for you .Using ggplot, plotly etc we can plot beautiful graphs in R but that might be time consuming and hectic

## 2.1 Most generated energy



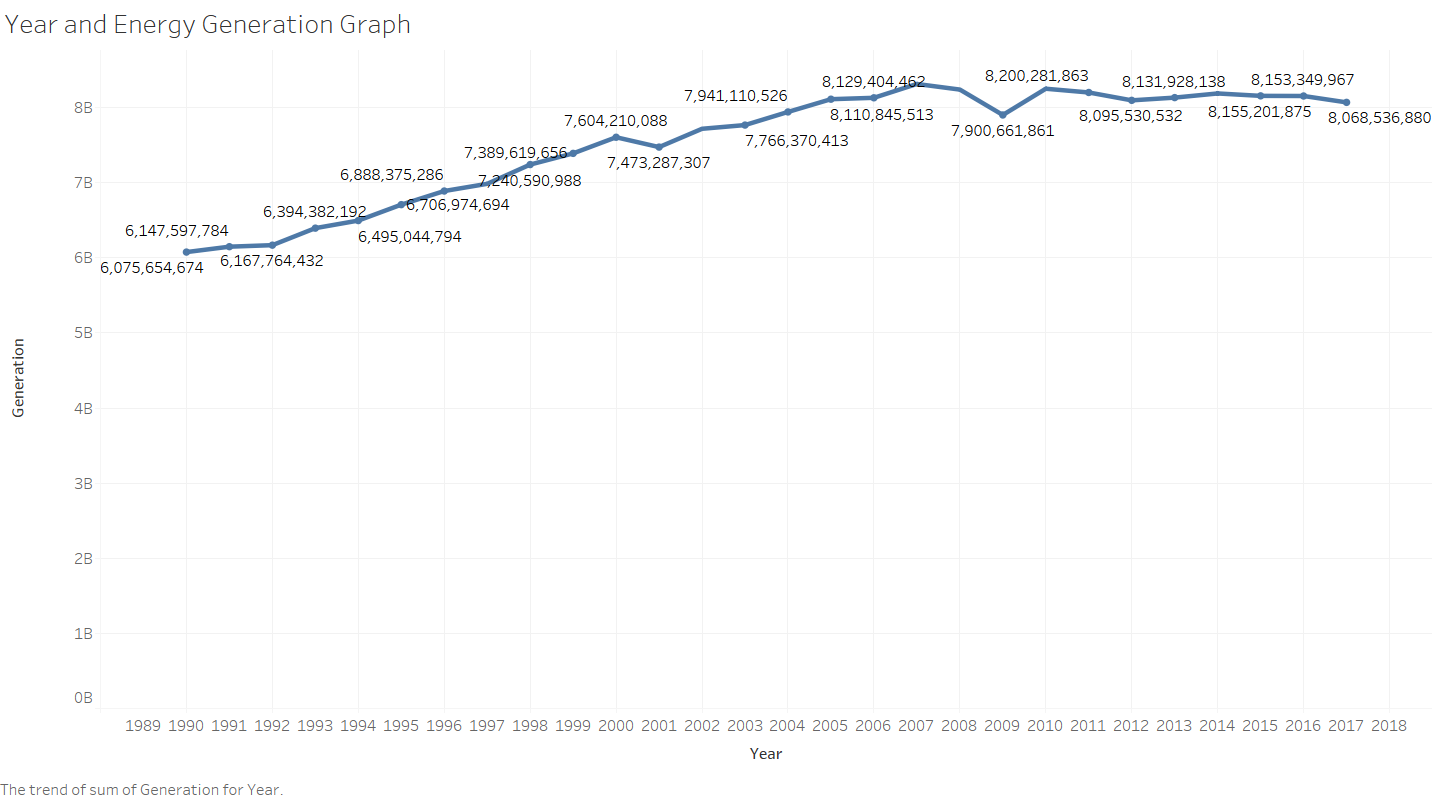
The above Bubble chart depicts the types of energy sources generated during the last 3 decades .As we can see from the above graph that Coal is the most used energy source for energy generation during the last 3 decades. Natural gas serves the second place and Nuclear energy comes in third place. The Hydroelectric conventional energy source was used less than all the above three energy sources.

## 2.2 Most energy generated state



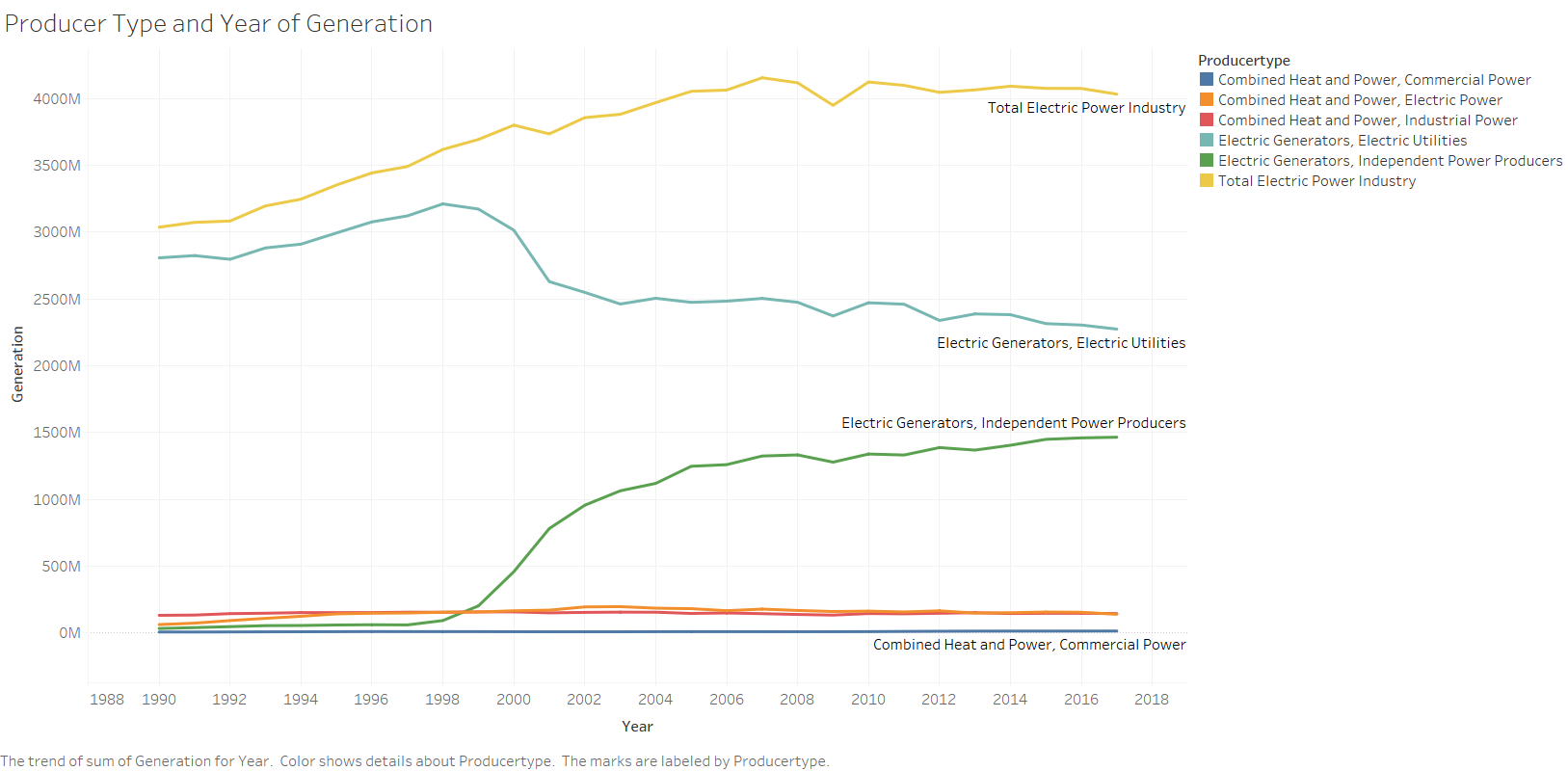
The above map shows the total energy generated by each state of US during the year 1990- 2017. We can see that the state of Texas generated the most amount of energy in following period of time, Pennsylvania accounts the second most energy generated state during the period of time where Vermont generated the least. Shades of blue were used to depict the amount of energy generation in which dark shade for the most generated states and lighter shades for least generated states.

## 2.3 Trend of energy generation



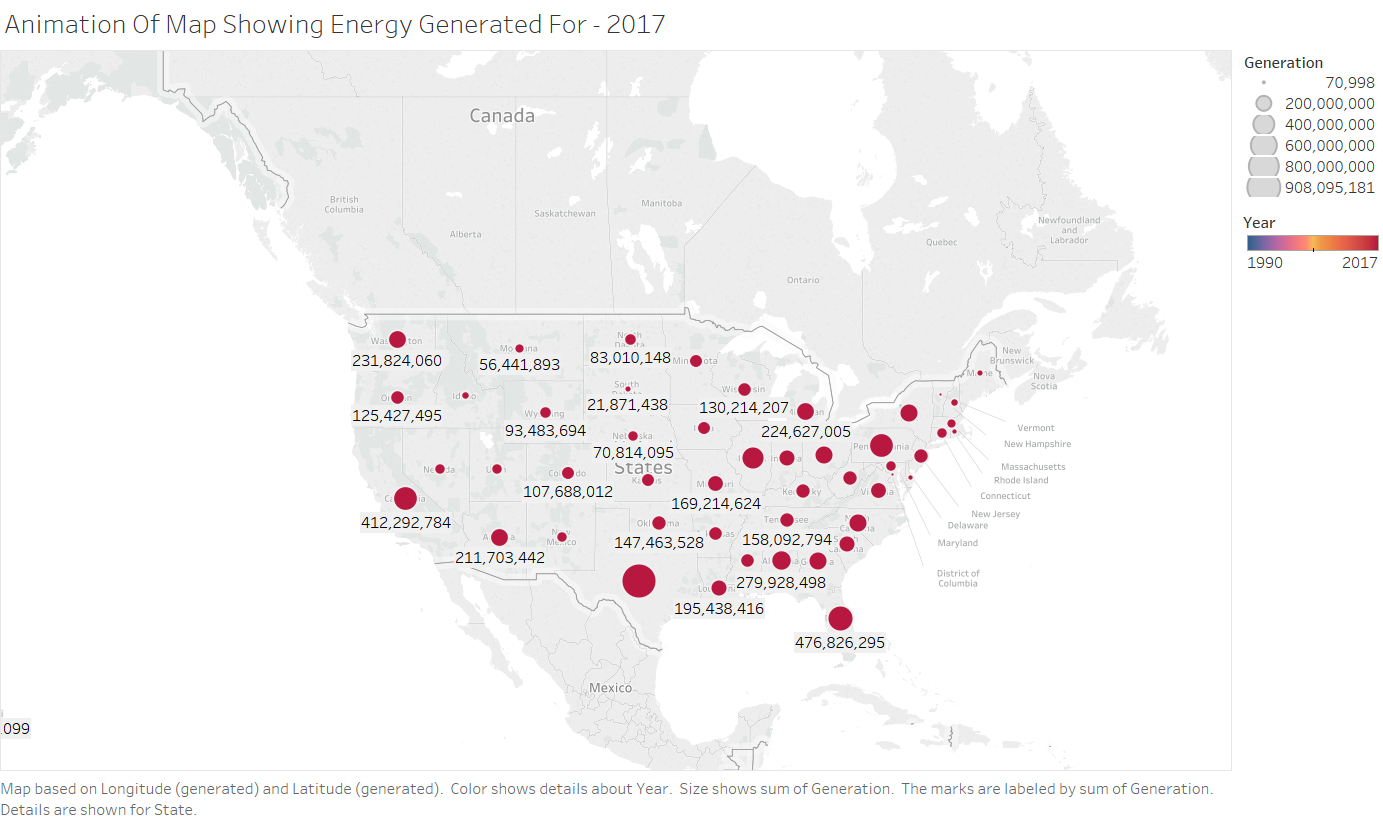
The above line graph shows the year and the total amount of energy generated each year from 1990- 2007. The total generated energy shows a gradual increase throughout the period till 2010 even though it had a small decrease in the year 2001 and 2009. 2010 was the year in which most energy was generated and after then it had a small decline and reached a lower amount in 2017 than 2010.The graph also illustrates that the gradual hype in the energy generation during the years might be the impact of population growth and industrialization that lead to an over need of energy.

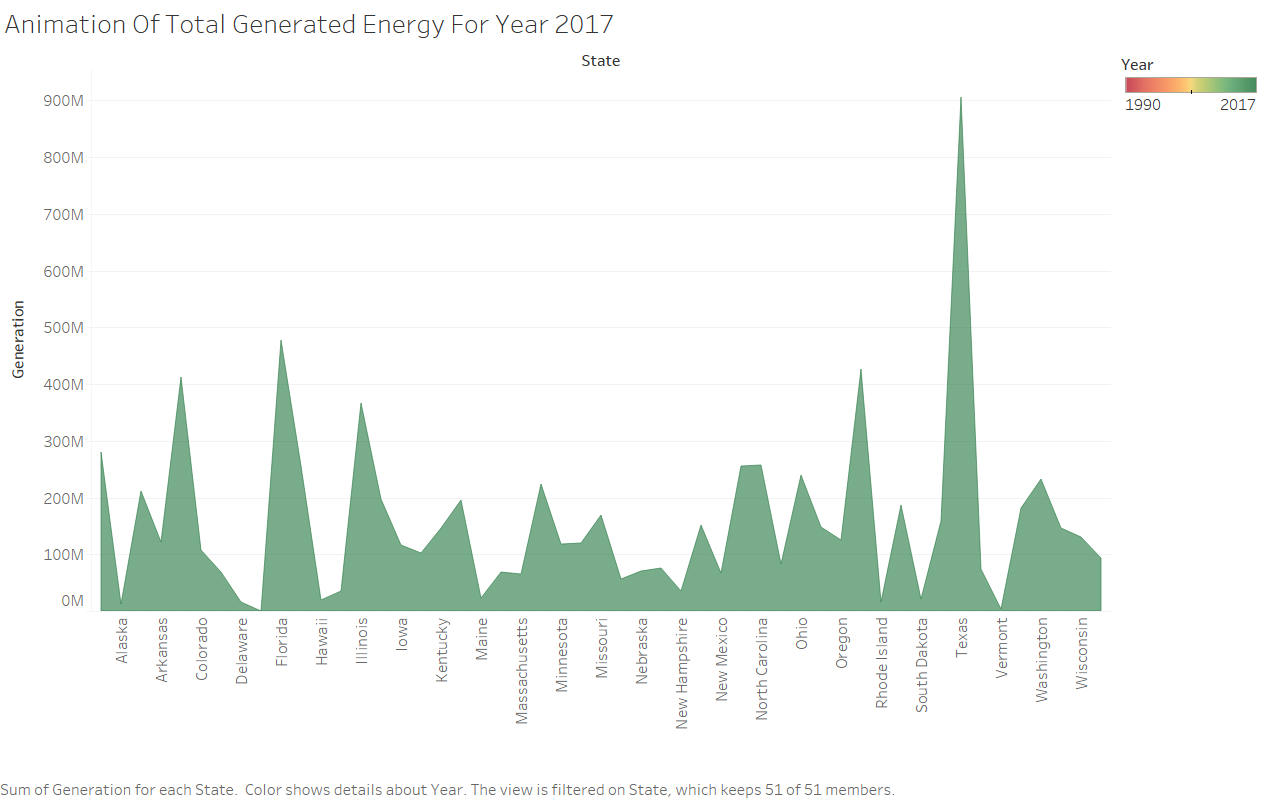
## 2.4 Trend of Producer type in generating energy



The above line graph depicts the energy generated by each company during the year 1990- 2017. Commercial power is the company who produced the least amount of energy during the time period where Total Electric Power Industry produced the most. Total Electric Power Industry always showed an increasing behaviour throughout the period. The energy generation by Commercial power is almost steady through the period whereas Combined Heat and Power doesn’t show any hype during the period and it was accounted the least producer type in producing energy.

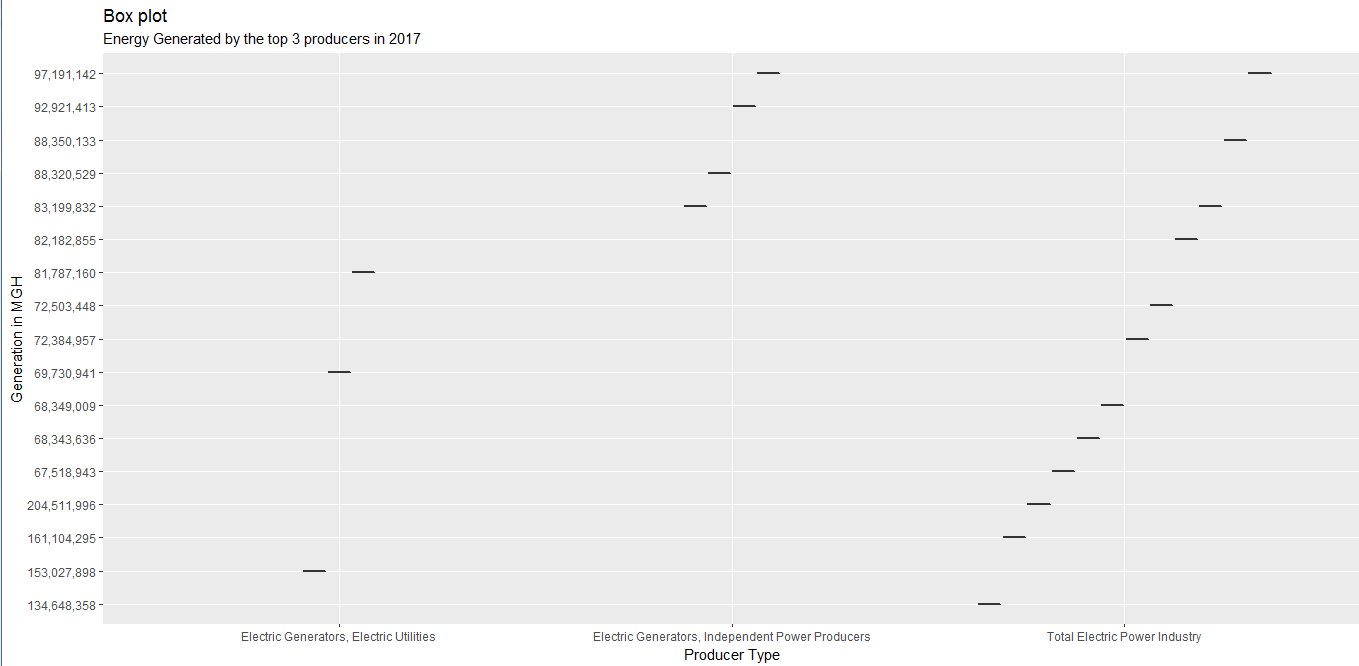
## 2.5 Animations showing energy generation patterns





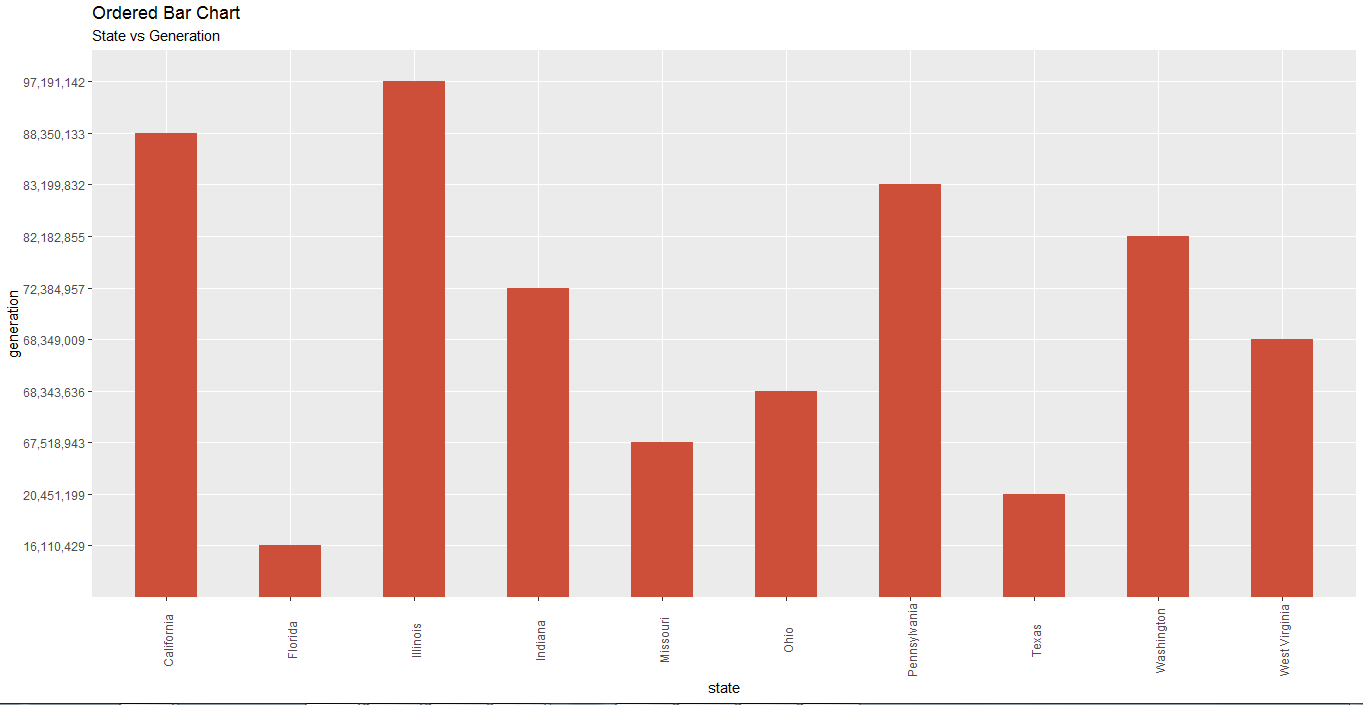
The above map and time flow depicts the animation of energy generated during the years in various places. As it is an animation a picture doesn’t makes any sense that we include our dashboard with the file sop that animation can be viewed properly

## 2.6 Energy generated by states in 2017



The above box plot represents the hits of energy generation or the count of energy generated area by the top 3 producer type in the US. As you can see from the graph that Total Electric Power Industry has generated energy sources in 12 places in 2017 whereas Independent Power Producers generated energy sources in 4 while Electric Utilities generated in 3 , the graph also depicts that the Total Electric Power Industry generated the maximum energy generation

## 2.7 Nuclear generation



The above bar graph depicts the production of energy using nuclear plant in different states of US , as from the graph we can see that Illinois accounts the highest in producing energy with nuclear energy and Florida is the least

# 3. Conclusion

In this assessment we have found that raw and complex data can’t be used to expresses what is in it but, by using a BI tool we can express what we want to tell or what the original data is. Every filed will have a lot of data from various area of working, so that these data can be used for future. Every employee or people who related to the company or the institution will not have technical knowledge, so that we need something like very user understandable method to explain the data. That is where the role of BI tools plays. By using BI tools like Tableau and R Studio, we have uncovered new plotting techniques, use of data with relevance Time, Multi-attributes on single axis etc. As we go through the visualization it gives us a story like sequence which does not break at any instances, it connects one with other like a dot. As it gives a better visualization it attracts the audience interest and attraction for longer and moreover it visualizes the data like a story.

From this assignment we have learned that:

* Visualization allows us to get faster reports, analysis with ease of understanding.
* As visualization is easily understandable every employee can be satisfied on presentation.
* No advanced programming is needed to play with the data in BI tools so, anyone can work on these tools so that company or institution can reduce cost.
* Manual visualizations can create errors by mistakes but comparatively the efficiency is high for tools.