Project 1

Alejandro Zamudio

2022-09-22

Dataset Information

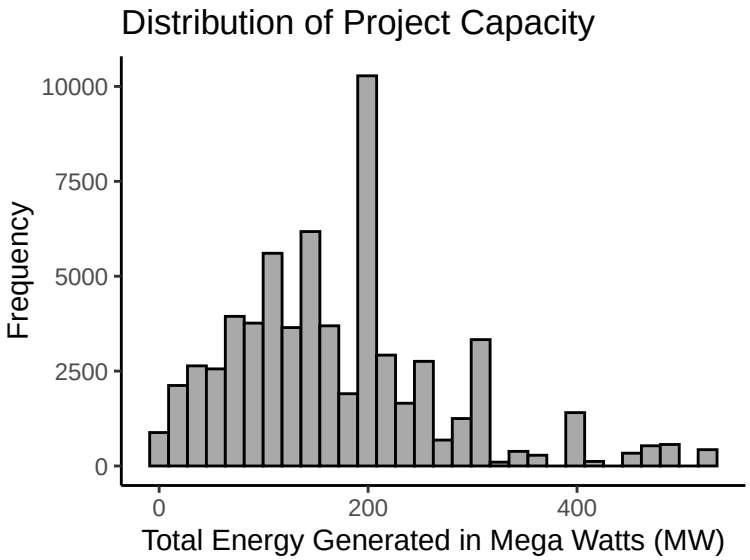
The Data in **Project1_wind_turbines.csv** provides information regarding the locations and specifications of Utility Scale Wind Turbines located in the United States. Utility Scale Turbines are wind powered turbines that are meant to generate and supply energy into grids; much larger than those that power a single home or business. Through the work of multiple organizations to collect the information, the dataset includes the location and year the particular project was begun, specific numbers and measurements relating to specifications of the design, and the number of turbines within each project and how much energy is able to be generated.
The Particular variable of interest for this investigation will be the "Project Capacity". This variable refers to the capacity of power that can generated by all the turbines in the project as a whole rather than a single turbine. In this investigation, we will be looking at this particular variable and comparing its relationship to the Rotor Diameter of each individual Turbine, the Capacity of the individual Turbines, the Longitudinal Location of each project.

Picture from https://www.rawpixel.com/search/wind%20energy?page=1&sort=curated&topic_group=_topics

Project Capacity Background

To begin, we will first look at variable "Project Capacity" and analyze the distribution of the different capacities within the dataset. Again, Project Capacity refers to the total energy generated by all the turbines within the each project.

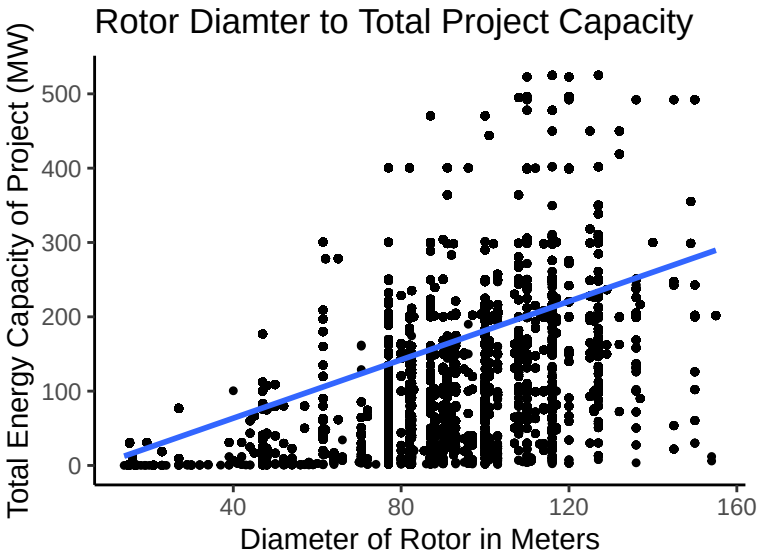
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



When Looking at the graph, it is visible that not many of the projects within the dataset generate a total capacity over 400 megawatts. There is a skew to the right of the graph, indicating that there are a good amount that surpass 400 Megawatts. While the highest total capacity generated is 525.02 Megawatts, the mean of 171.04 Megawatts and median of 160 Megawatts indicate that most of the projects generate energy just under 200 Megawatts.

Rotor Diameter

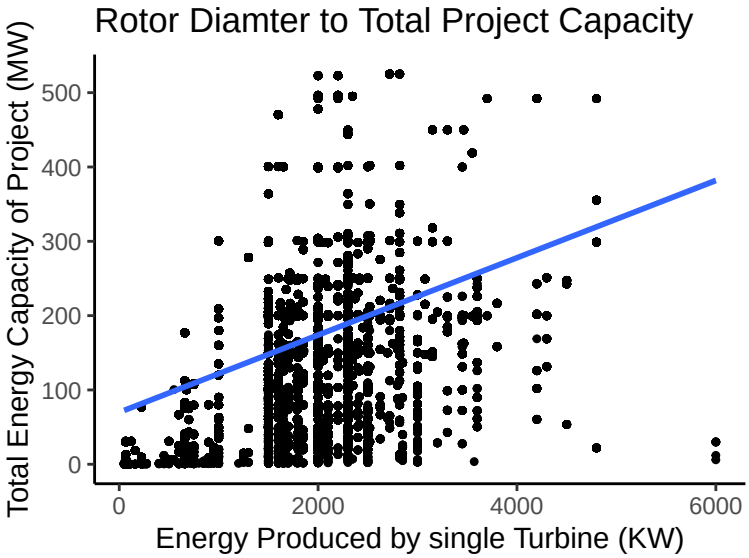
The First Variable we will compare Project Capacity with is "Rotor Diameter". The variable is a measure of the diameter of the Rotors on the turbines measured in Meters. The Diameters do vary in size with the highest diameter being 155 meters and the smallest being 14 meters. Through this first comparison, we will aim to see if higher or lower rotor diameters correlate to more capacity generated by the Project as a whole.



Through the graph, a noticeable trend can be made out of that show that projects with higher total capacities tend to have turbines with larger diameters of the rotors. This can be backed up when looking at the correlation of the graph, which is 0.45. The correlation indicates that the graph has a relatively strong relationship and even though outliers do exist, the general idea can be made that Larger diameters of rotors tends to be apart of porjects that produce higher project capacities.

Turbine Capacity

The next variable that we will compare the Project Capacity to is the individual "Turbine Capacity". The variable refers to the energy generated by a single turbine in Kilowatts. To compare, we will look at the relationship between the energy a single turbine produces and its relationship to the total energy produced by the entire project.



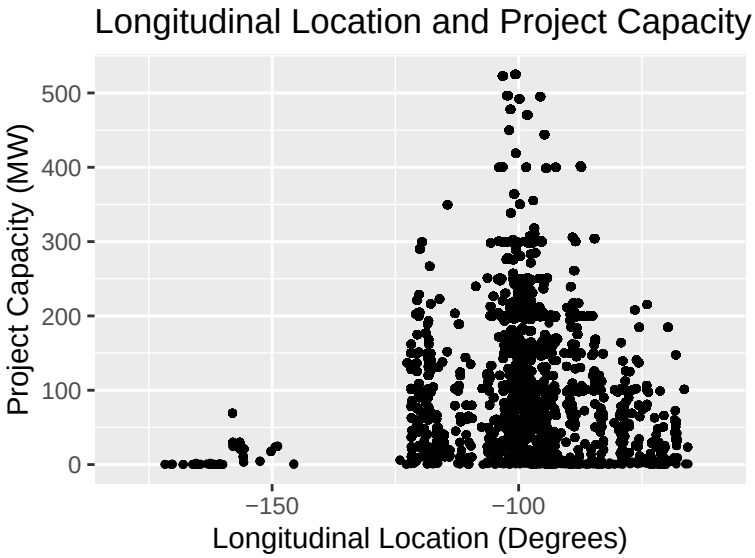
Through the graph above, we can partially make out a pattern between the two variables as the graph indicates a relative positive correlation. We see instances where Single Turbines producing lower energy are part of the projects that produce higher project capacities than others. With the correlation being 0.35, there

is a very weak relationship between the two. This suggest that a single turbine's capacity may not correlate to the capacity the entire project generates.

Longitudinal Location

The last variable we will look at in comparison to the Project Capacity is the Longitudinal Location of the Project. Longitudinal Location refers to the East to West coordinate position of a place on the surface of the Earth. For Reference, the most Eastern Point of the United States lies at 66°57 W (Located in Maine), and the most Western Point lies at 172°27 W (Located in Hawaii). The dataset contains these values as negative numbers, which simply refer to the "West" coordinate. The comparison between the two variables will look to see what locations have Turbine projects that generate more total Energy.

It is important to point out that within the data, there is an error between what is considered Latitude and Longitude. In terms of the code, I utilized the latitude Variable, however they are the values that represent Longitude as values in the -90's of latitude would point you to a location near the Southern Pole.



Through the Graph above, we observe that most of the Turbine projects lie between -125 and -75 degrees longitude (Which would range from most of the mainland United States). A good amount of the projects are centered within the -100 degrees longitude, which is backed up with the mean of the data being -99.69 and the median of the data being -99.33. This points to many of the projects being within the Midwest and Southwest Region of the United States. The most important of this, however, is that many of the higher energy producing projects are centered within this longitude range.

Photo above is included for reference to longitudinal coordinates

Conclusion

Through the investigation done to compare the variables with the Project Capacity, we can overall determine that much of the higher energy producing Turbine projects are centered towards the Mid and Southwest as a whole. The rotor diameters correlate to the total amount of capacity being produced, however the specific capacities of the Turbines themselves may not correlate to higher project capacities. This may suggest that some projects might have more turbines within the project (A variable we did not investigate and is present in the data set). The investigation also opens up to the question of what state most of these higher producing projects are located in. While we can assume from the investigation that most tend to lie within a particular region, we would need to go further at looking at the state or correlating latitude coordinate with the longitude.

For more information on the Dataset and the collected information of Utility Scale windmills in the United States, visit https://www.sciencebase.gov/catalog/item/57bdfd8fe4b03fd6b7df5ff9.

Dataset imported from https://think.cs.vt.edu/corgis/csv/wind_turbines/.