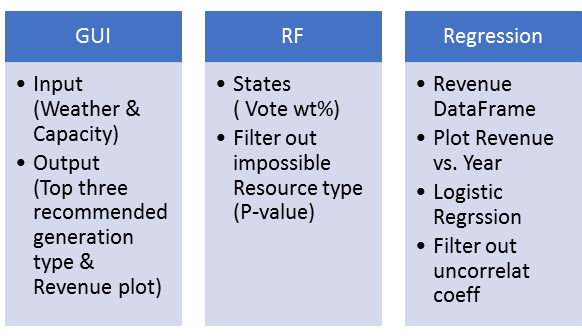
**EASE Workflow**

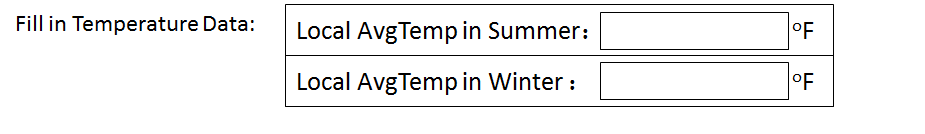
**Target Users:**

Industry users who want to build a self-sustained electricity power plant for their own factories for money saving purpose



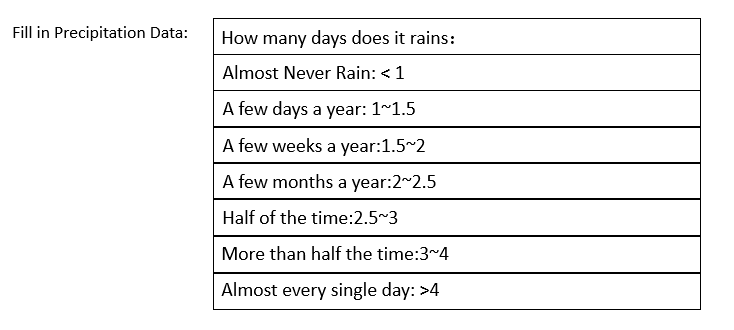
**GUI:**

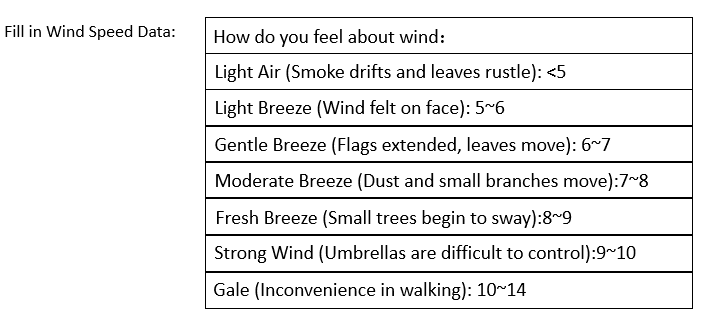
**Input weather data**



\*Note:

Eg. The summer temperature in California is about 75 ℉, suitable to wear short-sleeved T-shirt; the winter temperature in New York is about 30 ℉, suitable to wear coat.

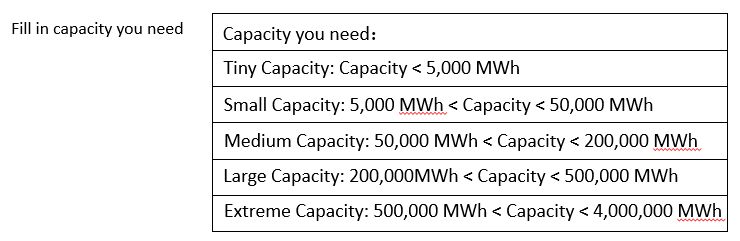




Reference Source: <https://www.unc.edu/~rowlett/units/scales/beaufort.html>

**Input electricity generation capacity**:

* Drop-down Menu Format



\*Note:

How to define the capacity you need?

The average annual electricity consumption for one household in U.S. is 10 MWh. So a small capacity plant can provide electricity for a city block; and an extreme capacity plant can provide electricity for the total Bellevue

The largest plant capacity in 2015 in U.S. is about 3500000 MWh a year.

* Scroll Format

\*Note:

How to define the capacity you need?

A small electricity plant generate less than 50,000 MWh a year, which can provide electricity for a city block;

A medium electricity plant generate 50,000~200,000 MWh a year, which can provide about 10,000 households;

A large electricity plant generate more than 200,000 MWh a year, which can provide electricity for the total Bellevue.

The largest plant capacity in 2015 in U.S. is about 3500000 MWh a year.

**Machine Learning with Random Forest**

**Goal:** According to the input weather data, filter out the impossible clean energy resources.

**Method:** Random Forest (Classification)

**Procedure:**

**1. Data mining**

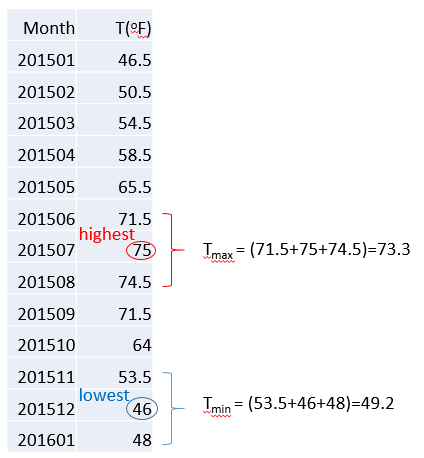
* Pre-calculate highest and lowest temperature weather data, termed as summer and winter temp.

Find the highest and lowest temperature and their accordingly months,

Situation one:

If it contains all the data for 12 months, use the temperature data of nearby 3 months to calculate the average temperature as the final Ts and Tw.

**Eg**.For 2015,CA:

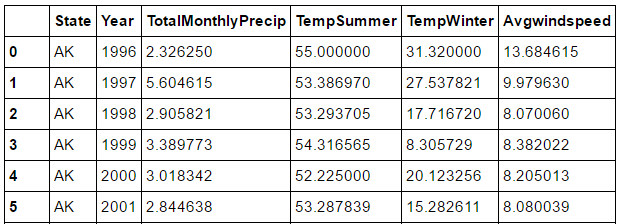


Situation two:

If it doesn’t contain all the data for 12 months, directly use the highest and lowest temperature as their Ts and Tw.

Final DataFrame to be produced prior feeding in for RF training:

Make the final weather dataframe containing Precipitation, Windspeed and Ts & Tw for each state in 1996--2016



**2. Random Forest**

Results: get all the possible states and weight percent that fits the user input weather data

**Eg.**

When Precip = 2;Ts = 67; Tw = 35; Windspeed = 6.5

Result: 'CT': 0.01, 'OR': 0.12, 'WA': 0.87

Equation to be applied will then be:

State1(source1 total electricity output) \* RF\_state1\_weight + State2(source1 total electricity output) \* RF\_state2\_weight + …

State1(source2 total electricity output) \* RF\_state1\_weight + State2(source2 total electricity output) \* RF\_state2\_weight + …

Output:

6 by 1 matrix containing all the sources with the sum weighted electricity total output in MW.

**3. Use P-value to filter out the impossible resources**

Based on the 6 by 1 matrix output from RF (sample mean), compared with the average electricity output of the states with its respective sources (population mean) using P-value at alpha = 0.05.

Output:

Filtered energy sources that are suitable for the location, with 95% confidence interval.

**Regression Analysis**

**Goal:** With all the suggested energy source types, prioritize clean energy and fill in the differences with conventional. Each suggested source will be analyzed with cost to give an estimate of the user how much will he/she saved over the course of 10 years. Using regression (logistic likely) to estimate the cost sector with 96-16 cost database constructed for this project.

**Method:** Regression (Logistic)

**Procedure:**

**1. Construct the cost dataframe**

Calculated average cost of certain resource by the voting weight of states

**Eg.** 4 votes for WA and 2 votes for CA,;The cost of NG in CA is 5.4, in WA is 5.09

so avgcost\_NG = 5.4\*33.3%+5.09\*66.7% = 5.19

|  |  |  |  |
| --- | --- | --- | --- |
| All the Possible Resources | Cost(cents/MWh) | CO2 Tax  (Cents/MetricTon) | Sales Price  (GovernPrice)  (cents/kWh)  (industrial) |
| Solar |  |  |  |
| Wind |  |  |  |
| Coal |  |  |  |
| NG |  |  |  |
| Petroleum |  |  |  |

\*cost sector takes in installation cost, maintenance cost, and operation cost into account and it is an average number.

\*CO2 tax values are based on a federal proposal in requiring all CO2 tax law in 2015.

2. Calculate the predicted amount saved of different resource

* Fit the cost model (logistic model, using F-statistics to filter out useless coefficients to get better fit) and get predicted cost and sales price for 2018-2028

Cost of buying electricity from Government = Sales Price (GovernPrice)\*Capacity

Cost of self-generated electricity = self-generation cost (material, etc.) \* Capacity

* Get the predicted Carbon Dioxide Tax for 2018-2028:

Carbon Dioxide Tax Estimate Source:

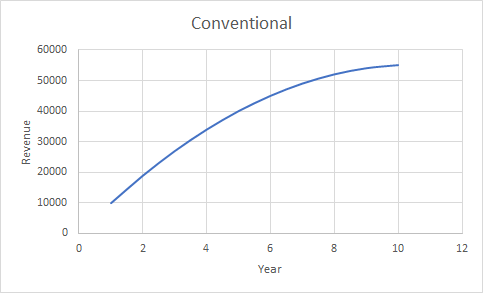
https://www.c2es.org/publications/options-considerations-federal-carbon-tax

* Calculate the amount saved in 2018 – 2028

Amount saved = Cost of Buying electricity from Government - Cost of self-generated electricity – Carbon Dioxide Tax \* emission amount

Output:

For Clean energies (Wind, Solar, and Hydro), gives an estimate amount saved over the course of 10 years, the expected output graph should be looking as below with data points supporting.



For conventional energies (Petro, NG, and Coal), gives an estimate amount saved over the course of 10 years with it being used to sub in the differences. The expected output graph should look like below. A secondary graph calculating the amount CO2 emitted will also be plot, this additional amount of CO2 will be added onto the total CO2 emission to give an reference on how much CO2 the user will be contributing to if choose conventional energy source.

