Analysis and Cost of Solar Power over Time

Capstone 1 Proposal

Summary

In this project, I will analyze a large dataset characterizing over 1 million solar power installations, targeting the cost over time of the roof mounted, residential installation to support the residential solar customer by predicting the future cost of a solar installation.

Structure

This project breaks down into the following steps:

1) Acquire the data
2) Develop a mathematical model of the retail cost of a solar install to facilitate identification of the variables of interest within the data.
3) Identify the data within that dataset that can be used and. massage into a format that can be used for the following steps.
4) Perform a data analysis to find a function that relates installation cost to time and use that function to predict the cost of a solar installation in the near term.
5) Validate the function by backtesting on current and prior years.

Data Acquisition

National Renewable Energy Laboratory has a dataset containing relevant information concerning more that 1.02 million solar installations (see https://openpv.nrel.gov/search). This dataset contains 82 variables. Most of these variables are not pertinent to this analysis. There is little documentation of many of the fields, though some are self-describing.

In addition, NREL also maintains a public dataset that is somewhat smaller and better documented. This data has also been cleaned to the extent that certain fields have been canonicalized.

Cost Model for Solar Installation

In market cycles we have come to expect that as the number of adopters grows, the cost of adoption drops as economies of scale come into play (producer cost of production drops) and competition forces producers to operate efficiently.

In technology cycles, these effects are often exaggerated by innovation and discovery. Moore's Law is a well-known example: for many years now processor power/cost has roughly doubled every 18 months. This is an example of an exponential decline in cost per unit of processing power. In solar energy, there is an analogous formulation, **Swanson's law,** i.e., the price of solar photovoltaic modules tends to drop 20 percent for every doubling of cumulative shipped volume.

For the customer, if the installation price is falling rapidly, waiting for the price to drop may be beneficial. On the other hand, if price is flat or increasing, buying now may be the best option. This study aims to educate that decision.

The cost of a solar installation has many components. The most costly single component has historically been the photovoltaic modules, but the other costs combine to perhaps 2-3 times the module.

At a minimum, the following items must be considered in an analysis

- Solar Panels
- Other hardware
  - Inverter
  - Racking
  - Cables, conduit, interconnection hardware, etc
- Installation labor
- Installer/Integrator cost and profit

I propose to isolate the variables above from the dataset and model them individually as functions of time. I will then combine them into a function that will model the total cost as a function of time.