

# Metis Data Science Boot Camp

## San Francisco, Winter 2020

### Final Project Proposal

Michael Boals

#### Introduction:

Demand planning is a critical component of any business that provides a product. Demand planning is the foundation for supply chain and logistics planning. However, demand planning is a complex problem that challenging to get right. Company growth and marketing plans combine with external factors such as seasons, cultural trends and even weather can impact product demand. This combined with the time series nature of the demand planning problem make it a good candidate for machine learning. In this project I will create a predictive model for regional energy demand in the US.

#### Data :

The Energy Information Association (<https://www.eia.gov/>) publishes periodic data for the energy demands across the planet. They have a detailed dataset containing the hourly energy demands for all the energy suppliers across the US. The dataset has 5 years' worth of hourly demand data and is 1.4 GB in size. I plan to combine this with historical regional weather data from the Dark Sky API. I would also like to include census data by region in the model. Each datafile has 4 years of demand data, I will use the first three years for training data and reserve the last year for testing.

The size of this dataset will most likely require the use of big data tools. I plan to use cloud computing services in combination with pySpark for my work. I also plan to convert the datafiles from their text based json format and store them in an SQL database.

#### Prediction Timeframe:

Future demand prediction becomes more challenging the farther in the future we try to predict. If I'm able to make accurate predictions in of future demand I will explore the relationship between model accuracy and the amount of time in the future we are predicting.

#### MVP:

Create a model that will predict the energy demand a week ahead based on the past years and previous weeks demand for a single generator.

#### Project Goal:

Predict demand 2-4 weeks in advance based on demand in past years, past weeks, national census and weather forecast.