**Predicting Peak Demand Periods for Electrical Power in Ontario**

**1. Opportunity**

I am employed at a glass making facility - O-I Brampton - in Ontario. I am working with an Energy Storage Supplier on a project that will save over C$1,000,000 per year by predicting and managing electricity demand. The system has 2 steps:-

1) Predict the 5 province-wide, peak electrical demand hours within a year.

2) Reduce the facility’s electrical demand during these peaks. This is done through large Tesla lithium-ion batteries.

The Supplier have developed their own proprietary prediction model.. However, it occurs to me that if I can develop a good prediction model, then this could be used by large electrical users to identify periods to manually reduce load, In fact, a few vendors already offer this as a SaaS. Therefore, this project focuses on the first part of the system - Peak Prediction.

**2. Size of opportunity**

C$1 million per year is a significant saving – How is this calculated? The province identifies the 5 peak total demand hours from the previous year. They compare them to O-I Brampton's demand during those peaks. The higher the proportion of total demand consumed by O-I, the higher the cost for electricity. So, if Brampton can predict the peak periods, and reduce demand, Brampton will save. The province benefits greatly from reduced peak load because they will ultimately need to build fewer power stations.

**3. Objective**

It is easy to predict all the peak electricity demand days in Ontario. Simply predict all 365 days of the year as peaks! Clearly, this is absurd - We have to reduce the facility demand to get benefits, and there is a cost to reducing demand. So, we need to refine the problem statement. **We are trying to predict all 5 peaks, while reducing the number of guesses it takes to catch those peaks.**

**4. Data Requirements for Modeling**

The dependent variable is total Ontario power demand. This could be analyzed by day, or hour.

Several variables influence power demand. Weather conditions drive demand through the use of air conditioning and heating. Additionally, the balance between commercial and household load is a major factor. This is driven by day of the week, hour of the day, and statutory holidays.

Based on my research, these are the major drivers of Ontario’s power demand.

**5. Available Data Sets**

1. Historical Electrical Demand

The Independent Electricity System Operator (IESO) manages Ontario’s power system. Their job is to ensure that supply matches demand. They publish several useful historical data sets. One such data set titled “Hourly Demand Report” details hourly energy demand by year. Their most reliable data goes from 2019 back to 2002. This gives about 365 days x 24 hours x 18 years = 157,680 data points covering about 6,500 days. Additionally, they publish zonal data, showing how demand is split between 10 zones. The zonal data comprises about 140,000 hourly data points, covering 5,840 days.<http://www.ieso.ca/Power-Data/Data-Directory>

2. Commercial/Household Load

day of the week can be calculated from dates. Also, several websites identify historical statutory holidays.<https://www.timeanddate.com/holidays/canada/2002>

3. Weather

Historical weather data is available on-line from the Canadian Ministry of Environment. This covers many Ontario weather stations going back several decades. However, this data does not appear to be stored in the most intuitive way.<http://climate.weather.gc.ca/historical_data/search_historic_data_e.html>

**6. Data Limitations**

1. Historical Actual vs Forecast Weather Data

The ideal data for this project is the data that would be used in a live prediction system. This means that instead of using historical weather data, I should be using historical weather forecasts. Unfortunately, I cannot find historical weather forecast data, so I will use actual weather data.

2. IESO Forecasting Data.

The IESO provides Peak Demand forecasting for 34 days ahead in their “Adequacy Report”. This forecast is not completely reliable, but would most likely form part of a live model. I have emailed IESO requesting this data, but they have not replied so far. If the historical forecast data is not available, I will not be able to use it in modeling as part of the prediction modeling.

**7. How I will solve this problem**

1. Identify the relevant data files to download

2. Download the data

3. Compile a full data set by merging the raw datasets together using dates and times as a key

4. Define how I will assess models so their performance can be compared

5. Perform Exploratory Data Analysis to identify pertinent features.

6. Build some regression models of demand using time series cross validation

7. Tweak the models and assess how they perform

8. Capture the results

**8. Deliverables**

1. Clean Data File, and defined data cleaning steps

2. Code in a github repository including model definitions and parameters

3. Report

4. Slide Deck

5. A proposal for how to commercialize the best model

Capstone Project 1 Proposal Draft v1 David MacGillivray 3 June 2019