****

**Description of data files**

The analysis should be based primarily on the given data.

* The file named **PRICE.csv**  contains hourly location based marginal prices (LMBP) and their components for electric power load for each hour of each day of the year, for each of 11 zones in and feeding into New York State (NYS). The data covers the time period from Nov. 1, 2011 to Nov. 30, 2018. There are 6 fields in this data file:

1. DATE\_TIME - date (DD/MM/YYYY) and time (hh:00) stamp.
2. ZONE - one of the geographical zones in and around NYS.
3. PTID - unique location identification code.
4. LBMP - location based marginal price in USD/MWh. This is the cost of the next mega-watt (MW) of electricity demanded by load (customers) in an hour (h) in a zone in and around NYS. LMBP is composed of the marginal cost of energy, the marginal cost of losses and the marginal cost of congestion.
5. MCL - marginal cost of losses in USD/MWh. This is the cost of transmission loss of energy (MWh) in a zone in an hour and one of three components of LMBP
6. MCC - marginal cost of congestion in USD/MWh. This is the cost of transmission congestion and one of three components of LMBP.

* The file named **WEATHER.csv** contains daily dry and wet bulb temperature observations and forecasts from and for weather stations in NYS. The time period covers Mar 1, 2011 to Dec. 16, 2018. There are 8 fields in this data file:

1. VINTAGE - takes on one of two nominal values: “Actual” or “Forecast”. When VINTAGE = “Actual” then recorded temperatures are those observed on a VINTAGE\_DATE. When VINTAGE = “Forecast” then recorded temperatures are those forecasted on the VINTAGE\_Date for a FORECAST\_DATE.
2. VINTAGE\_DATE - date of an observation or the publishing of a forecast of a temperature.
3. FORECAST\_DATE - date of the expected (forecasted) temperature.
4. STATION\_ID - three letter weather station identifier from NOAA.
5. MAX\_TEMP - maximum temperature in degrees Fahrenheit observed on the VINTAGE\_DATE or forecasted on the VINTAGE\_DATE for the FORECAST\_DATE.
6. MIN\_TEMP - minimum temperature in degrees Fahrenheit observed on the VINTAGE\_DATE or forecasted on the VINTAGE\_DATE for the FORECAST\_DATE.
7. MAX\_WET\_BULB - maximum wet bulb temperature in degrees Fahrenheit observed on the VINTAGE\_DATE or forecasted on the VINTAGE\_DATE for the FORECAST\_DATE.
8. MIN\_WET\_BULB - minimum wet bulb temperature in degrees Fahrenheit observed on the VINTAGE\_DATE or forecasted on the VINTAGE\_DATE for the FORECAST\_DATE.

* The file named **LOAD.csv** contains hourly electricity energy load data for each day of the year, for each of 11 zones in NYS. The data covers the time period of Nov. 1, 2011 to Dec 15, 2018 Loas is electric energy (mega-watt-hours or MWh) consumed in NYS that was transmitted through the grid. It does not include behind the meter electric generation not transmitted through the grid, such as produced by private solar panels. Load includes power imported from out of state. There are 5 fields in this data file.

1. DATE\_TIME - date (DD/MM/YYYY) and time (hh:00) stamp.
2. TIME\_ZONE - takes on one of two nominal values: EST (Eastern Standard TIme) or EDT (Eastern Daylight Time).
3. ZONE - one of the geographical zones in and around NYS.
4. PTID - unique location identification code.
5. LOAD - electric energy transmitted through the NYS transmission grid in mega-watt-hours (MWh).

* The file named **ADDITIONAL INFORMATION.xlsx** contains these worksheets:
  + Historical Demand Impacts. This includes estimated historical impact on grid generation from Energy Efficiency and Solar and Non-Solar Reduction in Grid Energy (Behind the Meter).
  + Electric Car Impact.
  + 2017 Generation Facilities.
  + Planned Future Generation.
  + Est. Production Costs.