1. The latitude and longitude information for some plants was missing from the data provided in the December 2018 EIA-860M Form. Some of the missing data was filled in by web searches, with most results coming from EnergyJustice.net.
2. Plant ID #99999 (Plant Name: State-Fuel Level Increment) have been excluded. There is no location information for these plants. See the following paper for the handling of a similar problem

<http://pubs.geothermal-library.org/lib/grc/1030635.pdf>

1. Net Generation Classes
   1. 5000, 50000, 100,000 > 1,000,000 > 10,000,000

List of Energy Types

|  |  |  |
| --- | --- | --- |
| Fuel Type | Code | Color Code |
| Natural Gas | NG | #F78B29 |
| Coal | COL | #99979A |
| Nuclear | NUC | #CF4A9B |
| Hydroelectric | HYD | #0081C5 |
| Oil | OIL | #EE1C25 |
| Wind | WND | #0FB14C |
| Solar | SOL | #D7C944 |
| Other | OTH | #FFEFD6 |

Steps

1. Fuel Types
   1. Define Fuel types to be modeled
   2. Assign Codes for each Fuel Type
   3. Develop table for conversion of fuel types for EIA-860M and EIA-923
2. EIA-860M
   1. Compile all Generators into one file

Summary of Actions

1. Gather raw data: The December 2018 EIA-860M report and the EIA-923 (or equivalent) annual report files from between 2001 and 2018 were gathered from the Energy Information Administration (EIA) website.
2. The EIA-860M and EIA-923 files were cleaned and prepared via the following actions
   1. EIA-860M
      1. There were seven separate worksheets in the raw EIA-860M Excel file. Only 4 of those worksheets were determined to be needed in this analysis. The following four worksheets were uploaded to SAS, each as a separate CSV file.
         1. Operating
         2. Retired
         3. Operating\_PR
         4. Retired\_PR
      2. Non-essential columns were removed. The following columns were maintained:
         1. Plant ID
         2. Plant Name
         3. Entity ID
         4. Entity Name
         5. Generator
         6. Nameplate Capacity (MW)
         7. Energy Source Code
         8. Operating Month
         9. Operating Year
         10. County
         11. State (originally labeled Plant State in the raw files)
         12. Latitude
         13. Longitude
      3. There were 11 plants that had missing or incomplete Latitude and Longitude data. These plants were removed. Further information about these plants were collected for possible future use. See below.
         1. A table of plants with missing latitude and longitude data was created. There are 11 plants with missing or incomplete data.
         2. The missing latitude and longitude data for the 11 plants were obtained manually. Base on the manual research all of these plants appear to be retired. A table of these 11 plants with the latitude and longitude data was uploaded into the SAS project. Currently it is not being used but may employed later to fill in missing data.
         3. The compiled EIA-860M table was extracted into two tables, one for plant location information, one for generator fuel, capacity, and age information.
   2. EIA-923 (or equivalent)
      1. From 2008 to the present the relevant report was EIA-923. From 2001 to 2007 the report was EIA-906.
      2. The relevant report for each year was extracted from the raw excel file into a csv file.
         1. EIA-923 (2008 – 2018)
            1. The relevant worksheet in the Excel file is the Generator and Fuels worksheet. If that worksheet is not part of the file then the Generator worksheet is the same file.
         2. EIA-906 (2001 – 2007)
            1. The files for EIA-906 contained an Excel file with only one worksheet. This worksheet is equivalent to EIA-923, Generator and Fuels
      3. Import each year’s csv file into a SAS project.
3. The EIA-86M Report for December 2018 is provided as an Excel spreadsheet with the following seven workbooks.
   1. Operating
   2. Planned
   3. Retired
   4. Canceled or Postponed
   5. Operating\_PR
   6. Retired\_PR
   7. Planned\_PR

Data from the Operating and Retired workbooks were combined into a single csv file. The file maintains the following columns

* Entity ID
* Entity Name
* Plant ID
* Plant Name
* Plant State
* Nameplate Capacity (MW)
* Energy Source Code
* Operating Year
* County
* Latitude
* Longitude

This work was performed in SAS Enterprise

Missing Data

See this website for some of the missing plants. They are small hydroelectric plants that may not qualify for >1MW and may have been retired.

<https://www.energy.ca.gov/almanac/renewables_data/hydro/>