

**Technical Report**

|  |
| --- |
| **No.1 Wind Turbines** |

Excel

* Renaming of columns - refer to Data Dictionary

Python

* Handling nulls and removing columns – refer to Data Dictionary; and Juypter Notebook for code
* Assumption – based on data distribution, we have applied
* mean value to t\_rsa (7618.5); and
* median values to p\_cap (158.0), t\_cap (2000), t\_hh (80.0), t\_rd (100.0) and t\_ttlh (130.1).

Tableau

* Join with EIA Operators (“Cleaned\_EIA.csv”) using with eia\_id -> eia\_id
* Join with “Wind\_Speed\_Data.xlsx” using t\_state -> state\_name

|  |
| --- |
| **No.2 EIA Operators** |

Python

* Handling nulls and removing columns – refer to Data Dictionary; and Juypter Notebook for code
* Assumption – The number of Operators was reduced (story 02) by including only Operators with > 10 data points.

Tableau

* Join with “Cleaned\_Windturbine.csv” using with eia\_id -> eia\_id

|  |
| --- |
| **No.3 Wind Speed** |

A third dataset was used to represent average wind speeds by US State.

Excel

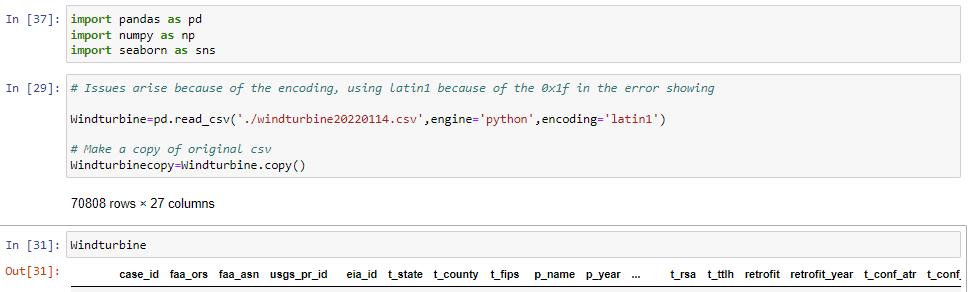
* Excel was used to insert wrangle data to include US State abbreviation and match with **Wind Turbines.**

Tableau

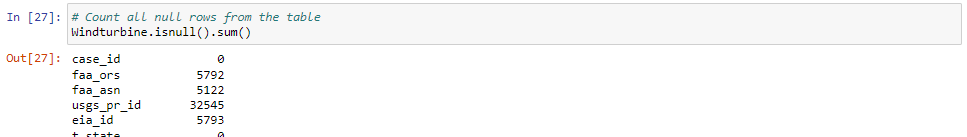
* Join with “Wind\_Speed\_Data.xlsx” using state\_name -> t\_state

|  |
| --- |
| **Sample Analysis of Wind Turbine and Operator datasets using Python** |

**Size of dataframe -** Windturbine

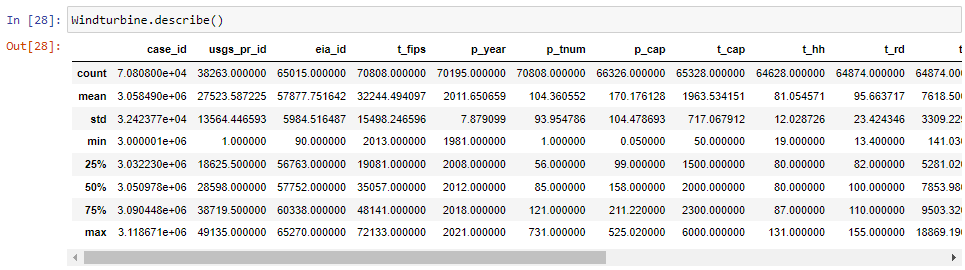
****

**Identifying nulls -** Windturbine

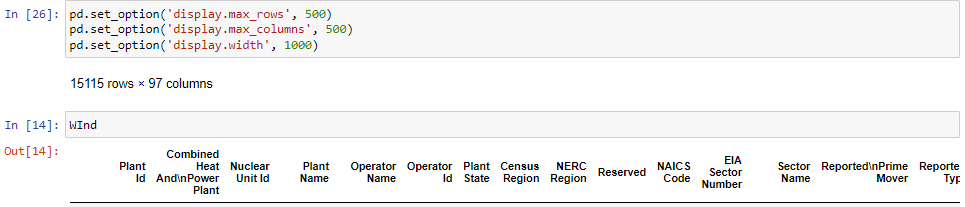
****

**Descriptive statistics -** Windturbine

|  |  |
| --- | --- |
|  |  |

****

**Size of dataframe -** CleanforGenerationandFuelData



**Identifying nulls -** CleanforGenerationandFuelData

