**Inputs & Info**

* Contains API # input which changes the well API along most of the tabs

**Data Summaries**

* Look at histograms and gradient maps for most of the inputs parameters in the modeling in the following tabs

**Clustering**

* There are two options for using this tab
  + Option 1: Create Clusters using the inputs on this tab
  + Option 2: Import the Clusters from Kriging TL tab using the toggle switch in the Kriging TL tab
* The Cluster Detail # input on this tab controls the Cluster numbers in the other tabs (i.e. Rev/Cost tab, Power Law tab)
* Allows us to create and analyze clusters by looking at clusters on map, seeing density charts for oil/gas/BOE output, average summaries for inputs parameters, and variable importance for predicting oil outputs

**Rev/Cost Analysis**

* The Clusters from the Clustering tab are fed into this tab (including Kriging TL clusters if that toggle is switched on)
* Cluster Detail # input on Clustering tab controls the Cluster outputs on this tab
* API # from Inputs & Info controls API in charts on this tab
* Allows us to look at the Revenue/Cost/Profit/NPV analysis for a particular well API or Cluster of wells

**Power Law Fits**

* The Clusters from the Clustering tab are fed into this tab (including Kriging TL clusters if that toggle is switched on)
* Cluster Detail # input on Clustering tab controls the Cluster outputs on this tab
* Allows us to create Power Laws and Hyperbolic Decline curves to model oil/gas/BOE output for both individual well APIs and Clusters of wells based on an input of the first few months of data.

**ML Production Pred**

* There are two options for using this tab
  + Option 1: Create Clusters using the inputs on this tab
  + Option 2: Import the Clusters from Kriging TL tab using the toggle switch in the Kriging TL tab
* If doing Option 1, this tab is meant to try out different numbers of clusters and see how building an ML model on one cluster can be extrapolated to another clusters and see how much accuracy is still maintained
  + The top 4 charts check the performance with a random test/train split on the entire data set
  + Next chart is for determining optimal number of clusters
  + Last 8 charts are used for checking the performance of the model on the training data (i.e. training cluster predicting on training cluster) and on the testing data (i.e. model built on training cluster and predicting on testing cluster)
* If using Option 2, Min TL Clusters, Max TL Clusters and Number of Clusters are not used as inputs as these are all set to the number of Clusters used in the Kriging TL tab.
  + Top 4 charts from option 1 are not affected by importing Clusters from Kriging TL tab. These charts are just a separate functionality.

**Cluster TL**

* This tab builds on the ML Production Pred tab and takes the Clusters built in the ML Production Pred tab (including if they are imported from the Kriging TL tab).
* The purpose of this tab is to see if adding a few wells from the new location to the training data mainly consisting of wells from a separate location can improve the model accuracy on the testing location
* Field A is the training set and multiple clusters from the ML Production tab can be assigned to Field A
* Field B is the testing set and multiple clusters from the ML Production tab can be assigned to Field B
* Hence the training set become Field A plus a user input proportion of Field B after a proportion of Field B has been held out as the Validation Set. This ensures that no wells in the validation set are used to build the model.

**Kriging TL**

* Take a small sample of wells (uploaded by user or use sample data downloaded from the app) and use kriging to determine the geological inputs of the surrounding area
* Next create clustering of the Eagle Ford data and find which cluster is most similar to the uploaded data
  + These clusters can be imported into other tabs for further analysis using the toggle switch
* Finally, build a RF model on the selected cluster and use the RF model to make prediction on the entire upload area with inputs determined by the kriged values

**Ensemble Learning**

* Determine which type of modeling is best when there is a lot of data overlap in the train/test sets compared to little to no overlap in train/test sets
* Determine if blended or combinations of types of models perform better than an individual model given different data overlap scenarios