**Help Document (Housing market prediction)**

**Packages/Libraries used:**

* **Pandas** - Easily handles missing data. It uses Series for one-dimensional data structure and Data Frame for multi-dimensional data structure. It provides an efficient way to slice the data.
* **Numpy** - NumPy arrays are faster and more compact than Python lists. An array consumes less memory and is convenient to use. NumPy uses much less memory to store data and it provides a mechanism of specifying the data types.
* **Matplotlib** - Is a cross-platform, data visualization and graphical plotting library for Python and its numerical extension NumPy. As such, it offers a viable open-source alternative to MATLAB.
* **Seaborn** - It is used for data visualization and exploratory data analysis. (Here in this model, we use seaborn to project the desired graphs).
* **sklearn.preprocessing** – It has several common utility functions and transformer classes to change raw feature vectors into a representation. (Used to implement Min-Max Scaler in training).
* **sklearn.metrics** -  It implements several loss, score, and utility functions to measure classification performance. (Used to import MSE and MAPE to calculate errors)
* **sklearn.model\_selection** - This method is used for setting a blueprint to analyze data and then using it to measure new data. (It is used to import train test split to train the dataset).
* **sklearn.ensemble** – It is used to combine the predictions of several base estimators built with a given learning algorithm in order to improve generalizability / robustness over a single estimator. (We use it import Random Forest Regressor to train the model).

**For the Implementation:**

The model has been split into different sections as follows for an easier implementation:

1. IMPORTING THE REQUIRED LIBRARIES AND DATASET AS .CSV file
2. CLEANING THE DATA
3. TRAINING THE DATA
4. DATAFRAMES FOR REPLACEMENT VALUES for the required set.
5. EDITING THE OUTLIERS
6. TRANINING THE MODEL to check for the error variations using MSE and MAPE
7. EDITING ALL OUTLIERS
8. FILLING IN NULL VALUES
9. ANAYLISING THE TRAINED DATA