**DATA ANALYSIS WITH FUEL ECONOMY DATASET**

**Dataset:**

## EPA Fuel Economy Testing:

## <https://www.epa.gov/compliance-and-fuel-economy-data/data-cars-used-testing-fuel-> economy

We use 2010 and 2012 datasets.

## DOE Fuel Economy Data:

<https://www.fueleconomy.gov/feg/download.shtml/>

Contents in the code file “**Datalytics.ipnyb**” :

1. Knowing the data attributes
2. Data Preprocessing
3. Data Visualization
4. Training and Testing Data
5. Developing the model
6. Prediction
7. Conclusions

**Data :** folder contains all the datasets before and after cleaning and updation.

**Code:** Python Jupyter Notebook

Libraries used:

* Numpy
* Seaborn
* Matplotlib
* Pandas
* Pyplot
* Collections
* sklearn.model : train\_test\_split
* sklearn.preprocessing : StandardScaler
* from sklearn.tree : DecisionTreeRegressor
* from sklearn.linear\_model : LinearRegression
* from sklearn.datasets : make\_regression
* from sklearn.ensemble : RandomForestRegressor
* from sklearn.neighbors : KNeighborsRegressor

Attributes in the datasets:

* Model Year
* Vehicle Manufacturer Name
* Veh Mfr Code
* Represented Test Veh Make
* Represented Test Veh Model
* Test Vehicle ID
* Test Veh Configuration
* Test Veh Displacement (L)
* Actual Tested Testgroup
* Vehicle Type
* Rated Horsepower
* # of Cylinders and Rotors
* Engine Code
* Tested Transmission Type Code
* Tested Transmission Type
* # of Gears
* Transmission Lockup?
* Drive System Code
* Drive System Description
* Transmission Overdrive Code
* Transmission Overdrive Desc
* Equivalent Test Weight (lbs.)
* Axle Ratio
* N/V Ratio
* Shift Indicator Light Use Cd
* Shift Indicator Light Use Desc
* Test Number
* Test Originator
* Analytically Derived FE?
* ADFE Test Number
* ADFE Total Road Load HP
* ADFE Equiv. Test Weight (lbs.)
* ADFE N/V Ratio
* Test Procedure Cd
* Test Procedure Description
* Test Fuel Type Cd
* Test Fuel Type Description
* Test Category
* THC (g/mi)
* CO (g/mi)
* CO2 (g/mi)
* NOx (g/mi)
* PM (g/mi)
* CH4 (g/mi)
* N2O (g/mi)
* RND\_ADJ\_FE
* FE\_UNIT
* FE Bag 1
* FE Bag 2
* FE Bag 3
* FE Bag 4
* DT-Inertia Work Ratio Rating
* DT-Absolute Speed Change Ratg
* DT-Energy Economy Rating
* Target Coef A (lbf)
* Target Coef B (lbf/mph)
* Target Coef C (lbf/mph\*\*2)
* Set Coef A (lbf)
* Set Coef B (lbf/mph)
* Set Coef C (lbf/mph\*\*2)
* Aftertreatment Device Cd
* Aftertreatment Device Desc
* Police - Emergency Vehicle?
* Averaging Group Id
* Averaging Weighting Factor
* Averaging Method Cd
* Averging Method Desc