1. **Dataset Link :**

" [Fuel Efficiency Data](#)
"

2. **Dataset Feature :**

- city_mpg
- class
- combination_mpg
- cylinders
- displacement
- drive
- fuel_type
- highway_mpg
- make
- model
- transmission
- year

3. **Dataset Outline :**

- Dataset Name : Car Data
- Dataset Description :
  This dataset contains various information about various vehicles . Such as
  - ❖ Engine information
  - ❖ Fuel consumption (MPG)
  - ❖ Transmission
  - ❖ Drive type

- ❖ Year of manufacture
- ❖ etc. The main purpose of this dataset is to analyze fuel efficiency of vehicles
- Number of Records : Total vehicle information: 550
- Feature Categories :

Categorical Features (Label/Type) :
- ❖ make
- ❖ model
- ❖ class
- ❖ drive
- ❖ fuel_type
- ❖ transmission

Numerical Features (Number) :
- ❖ city_mpg
- ❖ highway_mpg
- ❖ combination_mpg
- ❖ cylinders
- ❖ displacement
- ❖ year

- Target Variable : combination_mpg

4. **<u>Which advanced type model is applicable in these dataset for better results to published thesis papers? :</u>**

- CatBoost : Handles categorical data directly. No encoder required
- XGBoost : Provides high accuracy in numerical data , widely proven

- Combination : Comparing two models which make the paper stronger