Analysis of Ward's Automotive Yearbook.

Automotive yearbook shows the Model Import Car and Truck Specifications, not only this but also show the all specifications in details with all milage, horse powers etc. It provide the importing of car and trucks both with all specifications and facilities.

Dataset:

The dataset used for this project was provided by ma'am. The basic idea of analyzing the Automotive Yearbook was to get fair idea of Automotive sectors. What are the specifications, horsepower, engine, number of engine, pricing, etc.-etc. In this dataset it show the all type of automotive vehicle with all the required specifications. Toyota are the high in number in import for automotive compare to others.

With the help of this dataset it was easy to understand the method of using and plotting the data frame and graph.

These kind of analysis can be done using the data, by studying the factors such as :

- Year book of any Company, Franchise, Industries or many more
- Record of import or export .
- What type of automotive are import and export by industries.
- The specification of automotive and high demand of automotive in brand, horsepower, body style etc.-etc.

Tools & Libraries :

- Python • Jupyter Notebook • Pandas • Numpy • Seaborn • Matplotlib
- Plotly & Cufflink

Data Description:

The dataset contains the following Columns:

Make : Brand

Symbolling : Means import and cancel

Fuel-type : Gas and diesel

Aspiration : Std and Turbo

Number of doors: The number of doors are 2 and 4. There are some which is not mention in dataset

Body-style: 5 type of body styles (Sedan, Convertible, Hatchback, Wagon, Hardtop)

Drive wheels: 3 types of drive wheels (rwd, fwd,4wd)

Engine location: 2 types of engine location (front and rear)

Number of cylinders: there are 8 types of number of doors (4,6,5,3,12,2,8)

Not only this but also all this are mention in dataset which is useful and helpful to understand the dataset

Wheel base, length, width, height, curve wave, engine-type, stroke compression, horsepower, pricing.

Data Cleaning:

- I made the following changes and created the following variables:
- I made the new columns which represent the brand name .
- Made the sorting of price etc.
- Compression between fuel, aspiration, number of doors, body style, drive wheels, engine location, and more.
- Percentage of aspiration .
- Plotted the graph for better understanding and easy to understand.
- Comparing between fuel, aspiration, etc. etc.

EDA

This are the some of highlight which I have highlighted in my jupyter notebook of the data.

- Types of fuels automotive import?
  - Gas and Diesel .

- All Types of Brand?
  - There are 22 brands of automotive.

- Types of Body style ?
  - 5 types of body-style. And Hatchback are highest in number in terms of manufacture

- Highest importing automotive ?
  - Toyota brand are the highest is number.

- Engine location ?
  - In this 2 types of location is mention front and rear. Front is more in number by brands.

- Total Number of doors ?
  - 2 and 4 are the number of doors. But there are unknown number of doors which is not mention by brand. 4 number of doors are more in manufacture by brand.

- Total unknown normalized losses of automotive ?
  - Unknown normalized losses are losses which value is not mention by brand. Toyota is highest in number of unknown normalized losses.

- Type of aspiration of Brand ?
  - Aspiration are turbo and std. Std is more compare to turbo.

- Highest Total number of cylinders of brand?
  - 7 types of number of cylinders in particular automotive.
    (Two, three, Four, five, six, eight, twelve).In this four number of cylinders are high in number