

Analysis Car's Data.

Project Description:

- This Project Aims to analysis Car Data Using Big Data analysis tools.
- Analysis old Car Data and Create Conclusion to help new Upcoming Car's model in market.
- Create Bot To help Find Data for Special Purpose.

Technologies Used:

- Python
- Sql
- Java
- Apache Spark (Spark.sql).
- Anaconda (jupyter Notebook).
- Hadoop(winutil.ssh)

#List of Features

This Project analyses the old Car Data to get the

- Total Count of Sell Cars Brand Wise and find which is most seller also result show in chart?
- Which is most Popular model in first top of seller also result show in chart?
- Create bot for get input brand name from user and show all model in those Brand and show top seller model in chart Format.
- All Car brand's top sell model compared to each other and show in chart?
- mostly used Fuel Type and compare and show in chart
- find which Engine Type is most Popular and compared fuel type wise Engine Type?
- Which is most Popular transmission?
- Find Year by Cars count also find?
- Create bot to get input (lower price bond, high price bond, fuel type) from user and show all cars in those price bond?

#To-Do List:

- More queries to be added.

#Getting Started:

1. Set Platform.

Download Spark 2.7.1 or above, Anaconda, Hadoop-Winutil.ssh, data set

2. Install Platform Related Software. Set also Environment Variable.

3. Using Anaconda Prompt Install Following Packages.

pyspark, matplotlib, pandas, pylab

- 4.Create Spark Session in our case ("Car_data").
5. Upload CSV file in Data Frame using (.load).
- 6.Create Query Using Spark.sql(" ")
- 7.Convert Query Output into Bar Chart using matplotlib
- 8.Start executing Queries.

#Output

- My Car_Data Table Schema is Following:

col_name	data_type	comment
carID	int	null
brand	string	null
model	string	null
year	int	null
transmission	string	null
mileage	int	null
fuelType	string	null
tax	int	null
mpg	double	null
engineSize	double	null
price	int	null

root

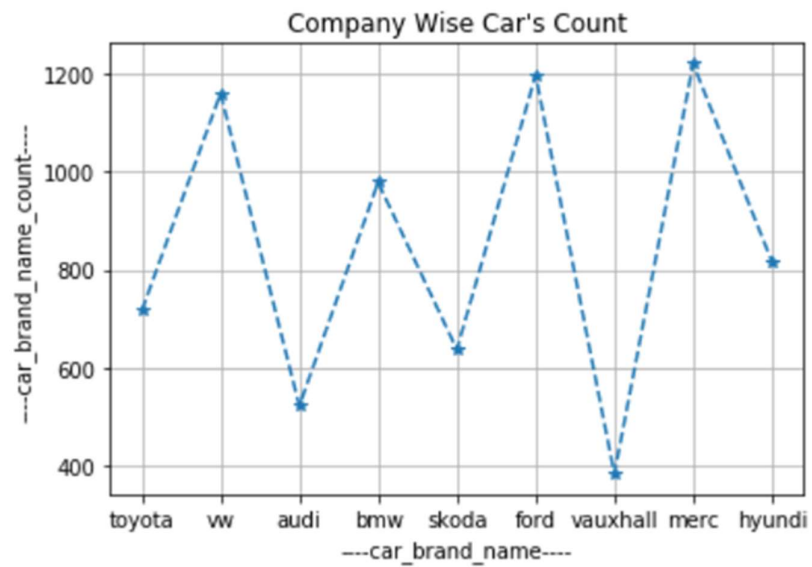
```

|-- carID: integer (nullable = true)
|-- brand: string (nullable = true)
|-- model: string (nullable = true)
|-- year: integer (nullable = true)
|-- transmission: string (nullable = true)
|-- mileage: integer (nullable = true)
|-- fuelType: string (nullable = true)
|-- tax: integer (nullable = true)
|-- mpg: double (nullable = true)
|-- engineSize: double (nullable = true)
|-- price: integer (nullable = true)

```

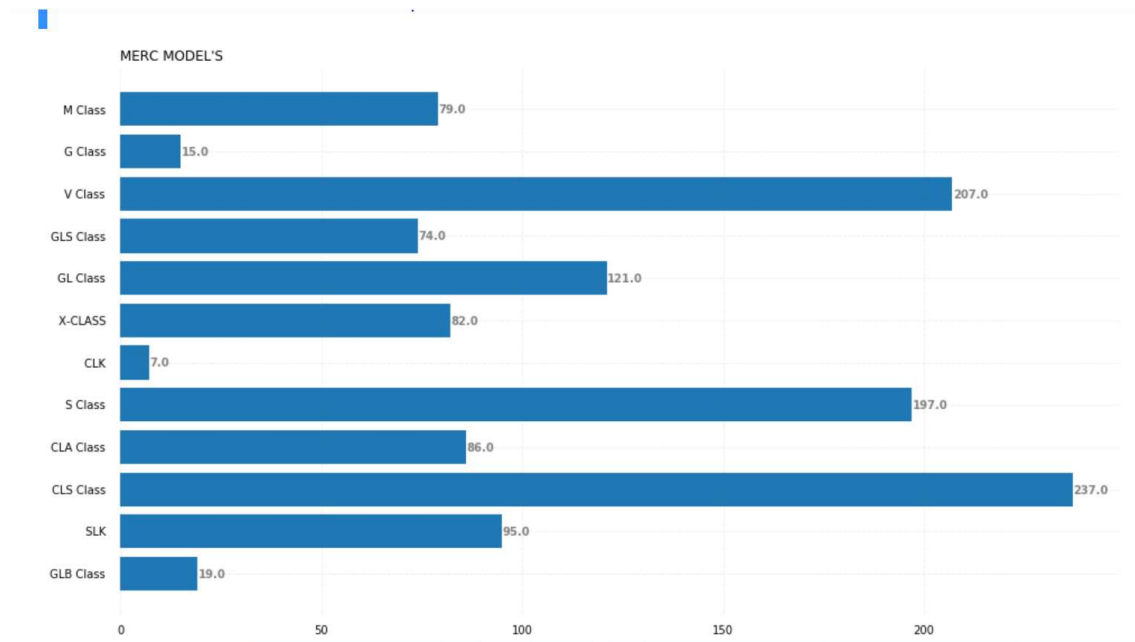
- Q1. Most Popular Car Manufacture Brand is:

brand	Count
toyota	717
vw	1159
audi	523
bmw	978
skoda	638
ford	1196
vauxhall	385
merc	1219
hyundi	817

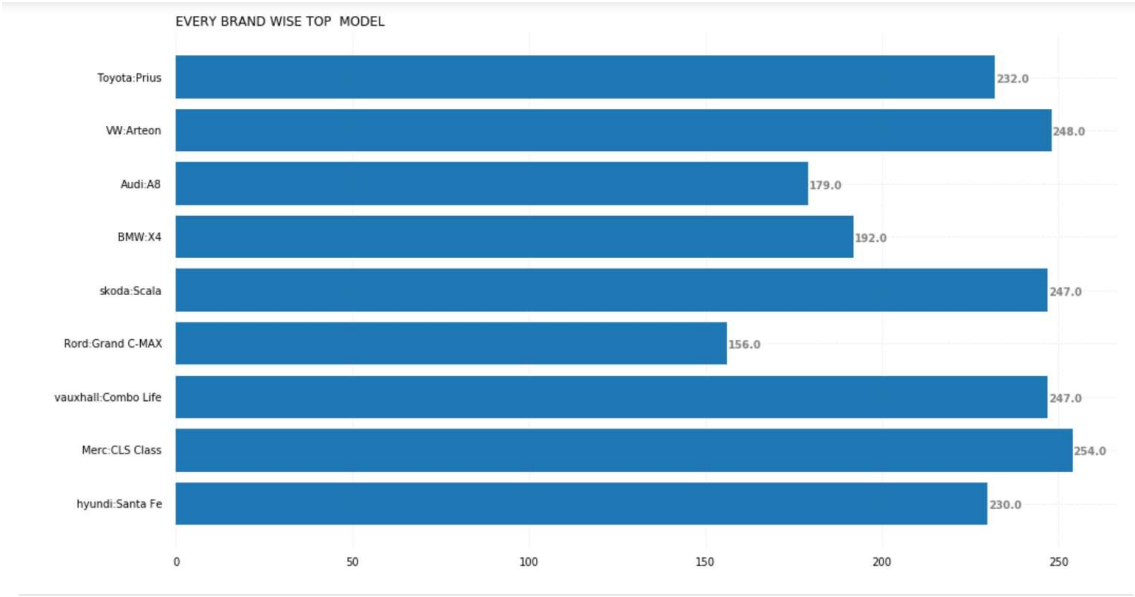


Q2. Which is most Popular model in merc ?

model	model_Count
M Class	79
G Class	15
V Class	207
GLS Class	74
GL Class	121
X-CLASS	82
CLK	7
S Class	197
CLA Class	86
CLS Class	237
SLK	95
GLB Class	19

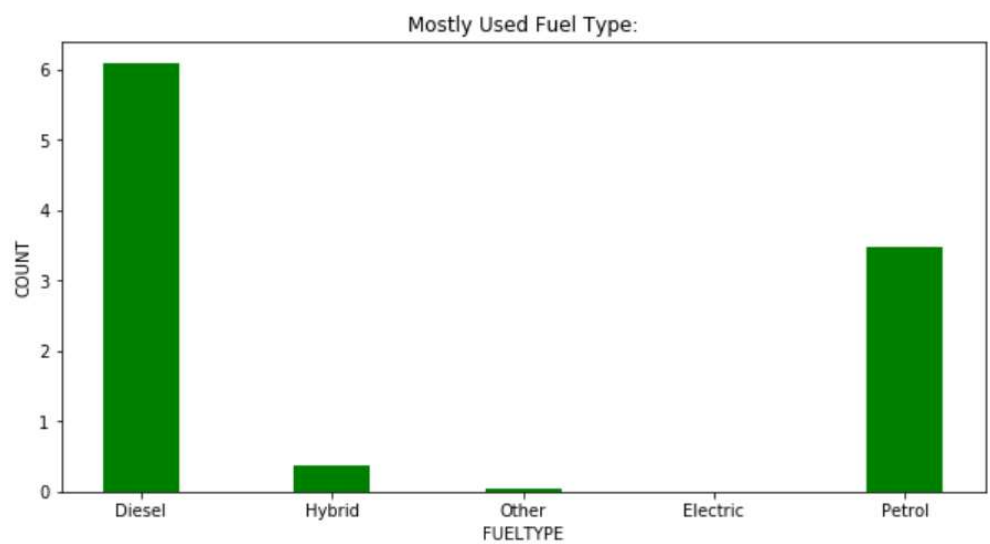


q.4 Shown Every Brand Top Sell Model Cars ?

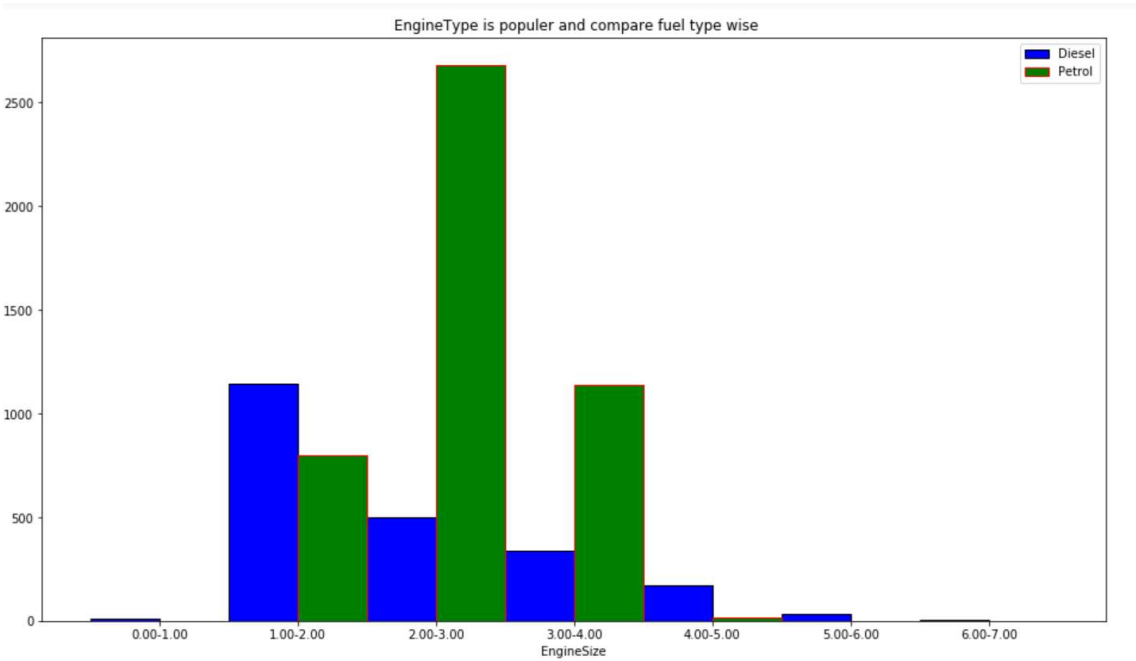


#Q5. Mostly used Fuel Type.

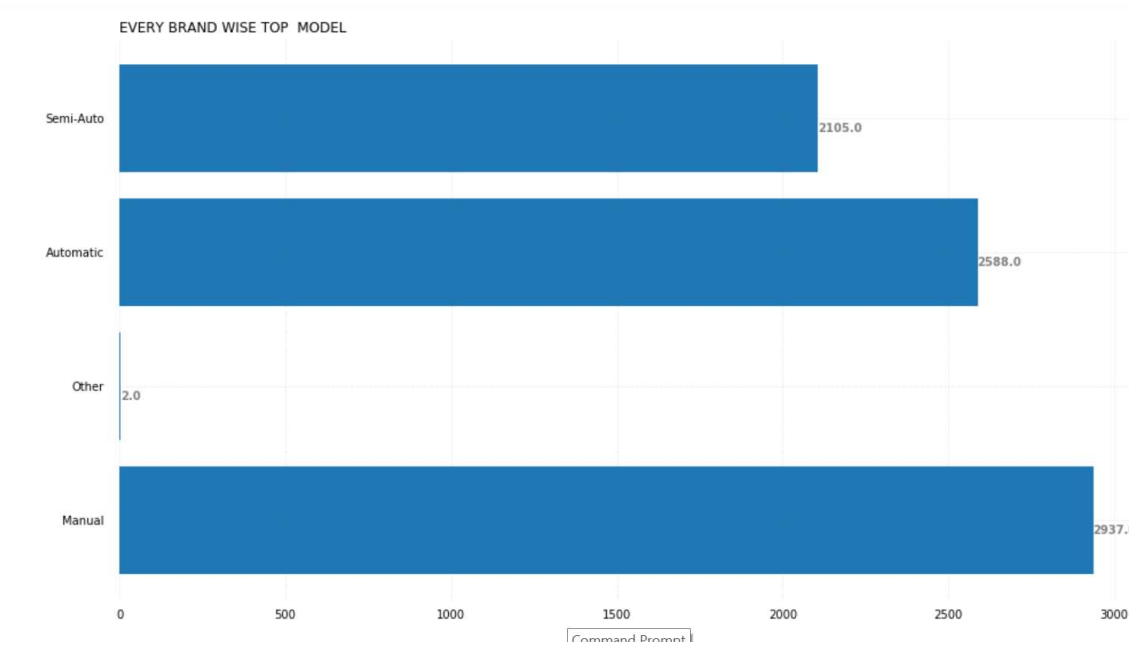
+-----+-----+



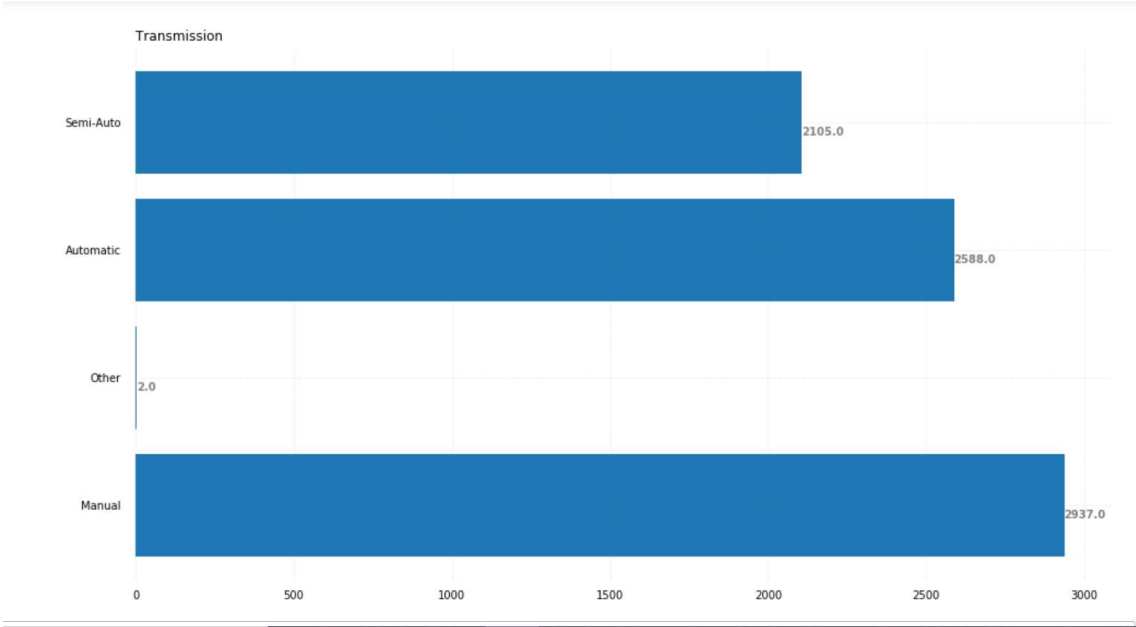
#Q6. find which Engine Type is popular and compare different Engine Type with fuel type wise?



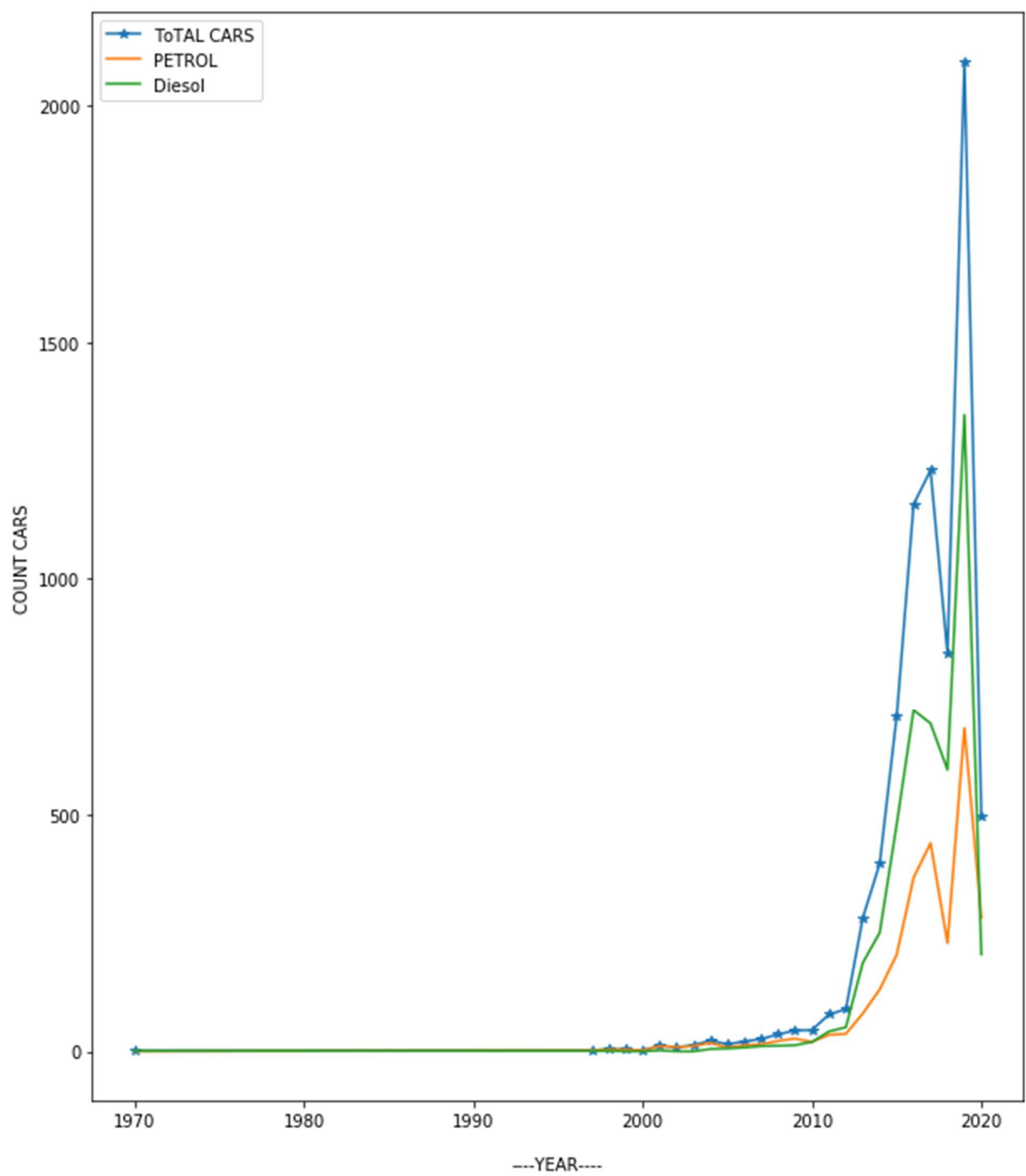
Q7. Which is most Popular transmission?



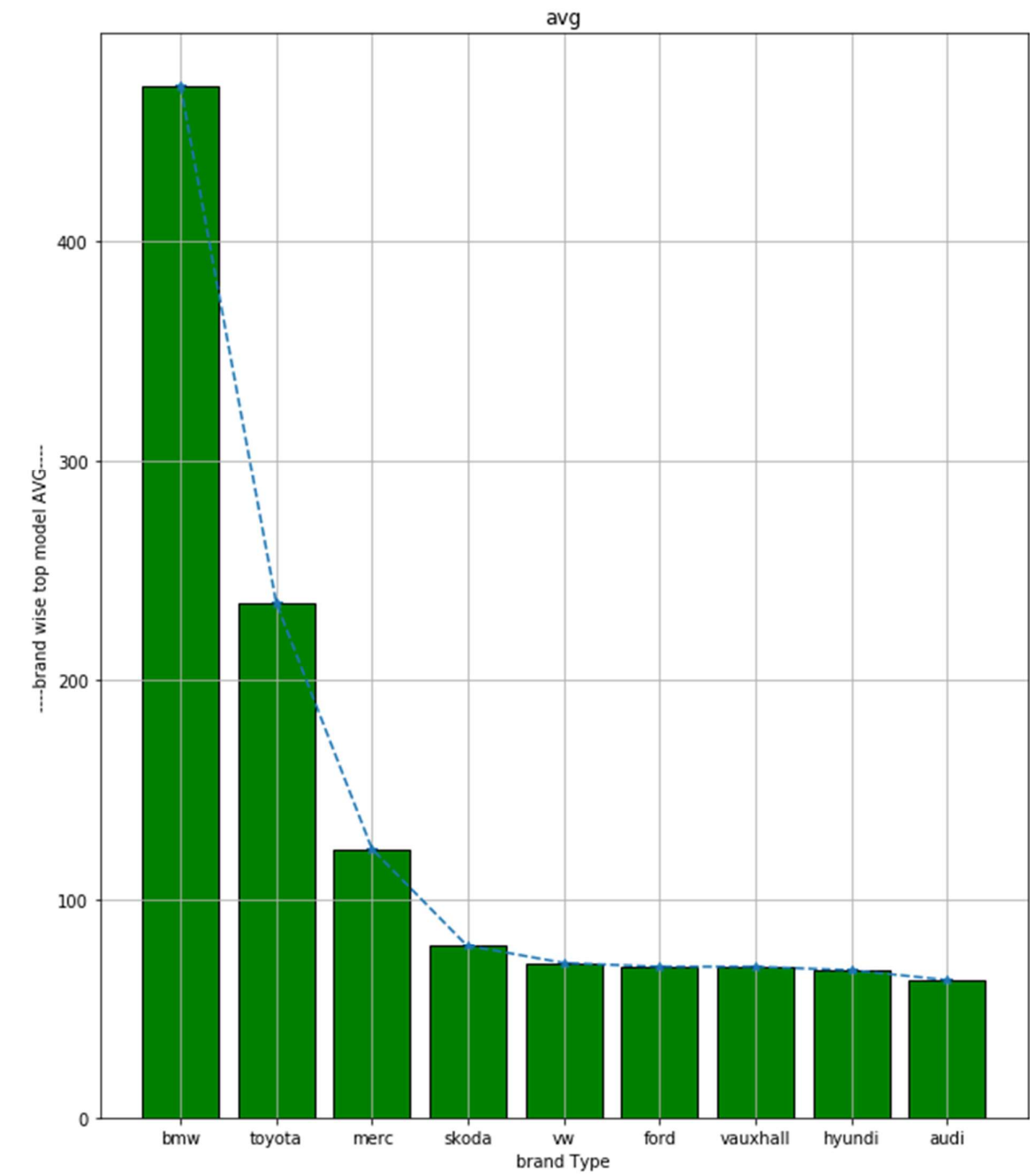
Q7.Which is most Popular transmission



#Q8. find how many car's Manufacture in year's and compare petrol and Diesel wise count ?

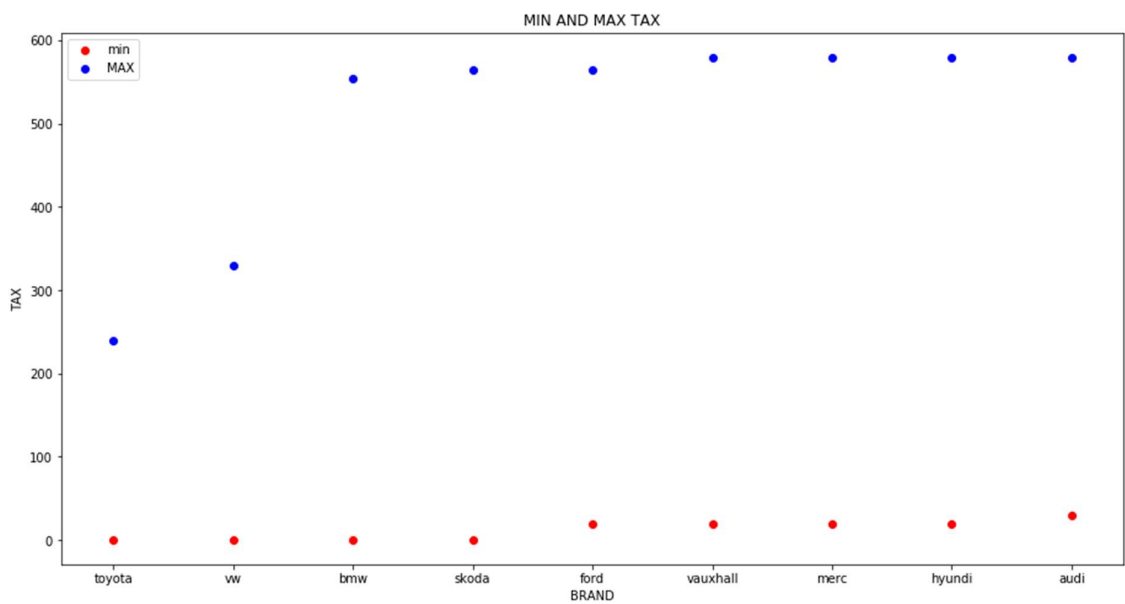


#Q9. Compare top Model AVG Brand wise?



Q10.Find MIN and max tax each model compares with brand ?

brand	max(model)	tax	brand	max(model)	TAX
vw	Tiguan Allspace	0	skoda	Yeti	240
bmw	i8	0	vw	Tiguan Allspace	330
skoda	Yeti	0	hyundi	Santa Fe	555
toyota	Verso	0	vauxhall	Zafira Tourer	565
hyundi	Santa Fe	20	toyota	Verso	565
merc	X-CLASS	20	merc	X-CLASS	580
vauxhall	Zafira Tourer	20	bmw	i8	580
ford	Tourneo Custom	20	ford	Tourneo Custom	580
audi	SQ7	30	audi	SQ7	580



Q11. Create Bot For Help to find cars?

Lowest Price is 450\$ Higt 159999\$:::
Enter lowest Price in \$: 5555

Enter higt Price Price in \$: 10000

We have available Transmissions::

Manual,
Automatic,
Semi-Automatic

Enter Transmission ?Manual

We have available Fueltype

Petrol,
Diesel,
Electric,
Other

Enter FuelTypeDiesel

In Your Price Bond Range Cars:

carID	brand	model	year	transmission	mileage	fuelType	tax	mpg	engineSize	price
17409	vauxhall	Antara	2012	Manual	63725	Diesel	240	42.8	2.2	5695
19591	hyundi	IX35	2013	Manual	91001	Diesel	150	48.7	1.7	5698
12432	skoda	Yeti	2011	Manual	93579	Diesel	200	46.3	2.0	5699
14709	hyundi	IX20	2012	Manual	32000	Diesel	30	64.2	1.6	5750
18539	skoda	Rapid	2016	Manual	74129	Diesel	0	78.5	1.4	5790
14955	skoda	Rapid	2015	Manual	78791	Diesel	20	70.6	1.6	5795
16539	vauxhall	Antara	2012	Manual	50478	Diesel	235	42.8	2.2	5800
16008	skoda	Rapid	2014	Manual	29000	Diesel	30	64.2	1.6	5890
18634	skoda	Rapid	2015	Manual	53549	Diesel	30	64.2	1.6	5950
13686	vauxhall	Zafira Tourer	2012	Manual	105362	Diesel	145	54.3	2.0	5950

- END