Car Data

Increase potential for your car sales....

Input Data

1. Input data gives **information** about the **car attributes.**
2. Some of the attributes are model, model type, car color, car mileage etc.
3. There are some **missing value** in car model columns.
4. After changing the data format there are many nan value but the number of **records** on the basis of model type are kept **intact**.

Challenges with Dataset

Unprocessed data

1. Structure of DataFrame
2. Null Values

Different Values

1. Value format is different

Multiple Columns

1. Different Column Name
2. Multiple column other than target data columns

Solution

Pre-processing

1. Structure of dataframe made similar to target data.
2. Splitting the information into multiple columns.

Normalization

1. Format of values is changed in accordance.
2. Letter cases are changed.
3. Language could also be translated.

Integration

1. Columns are renamed.
2. Extra columns are removed.
3. Number of records are unchanged.

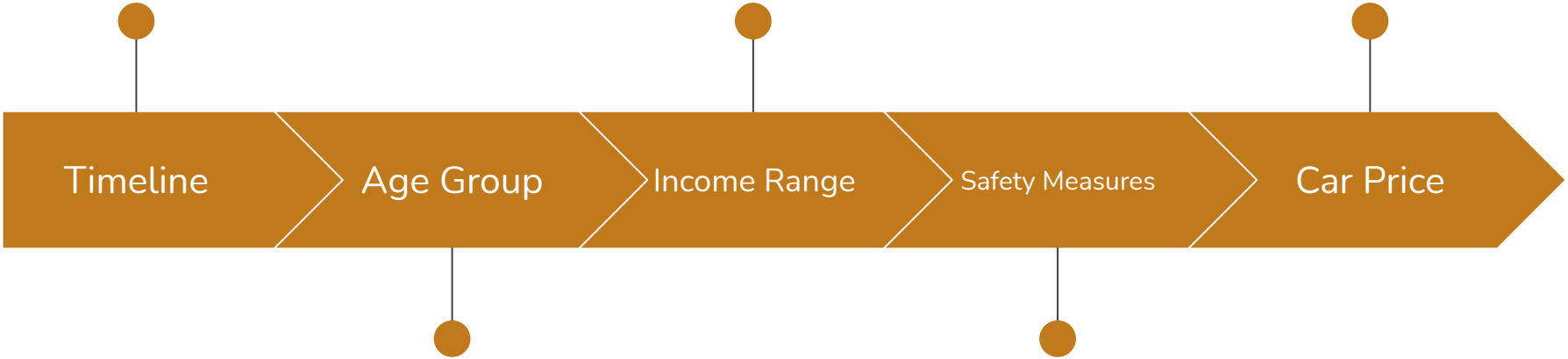
Potential Analysis

with more data

Car sales according to brand in previous year

Income range of consumers

Price of car according to model

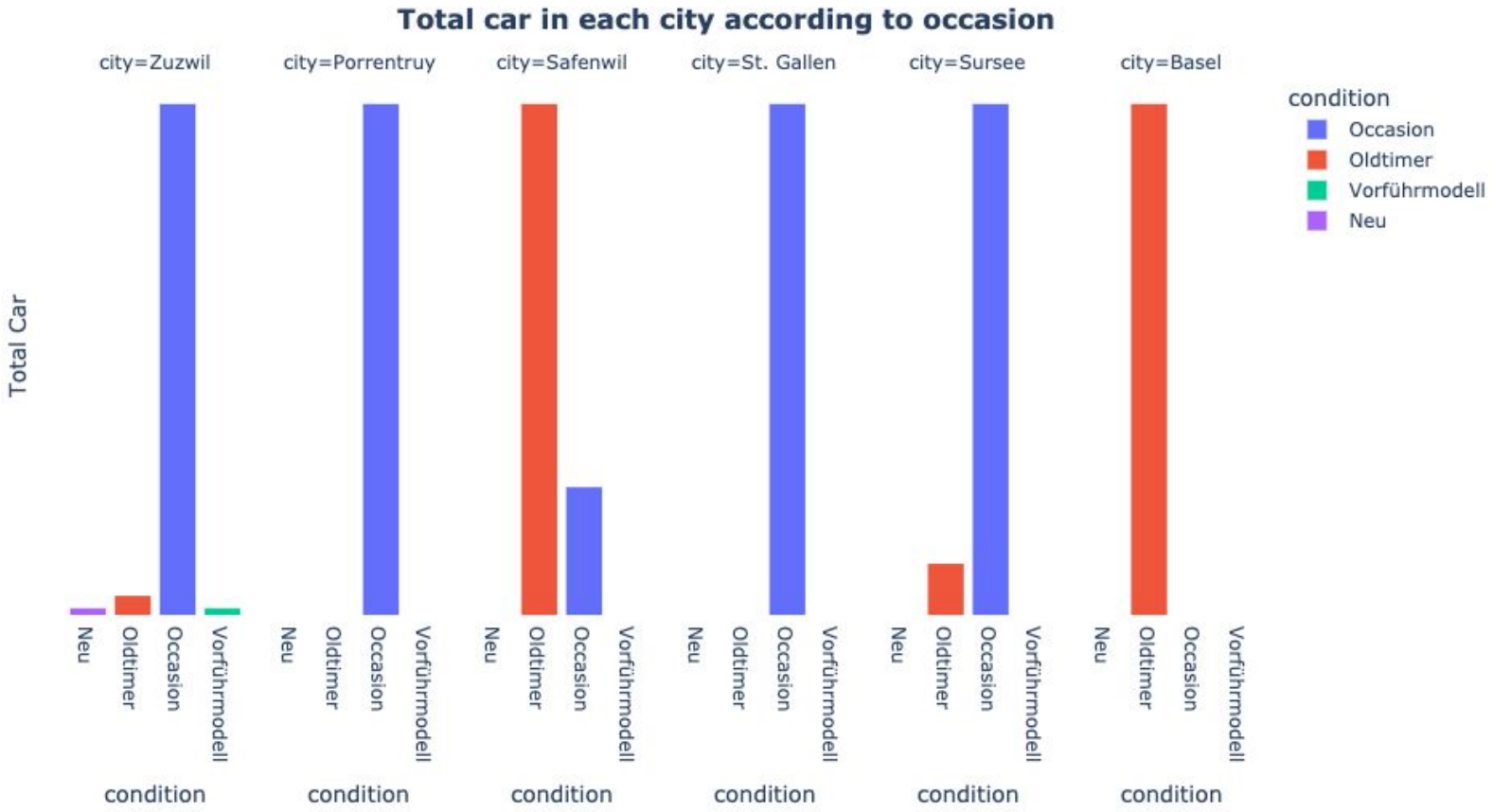


Age group of purchasers

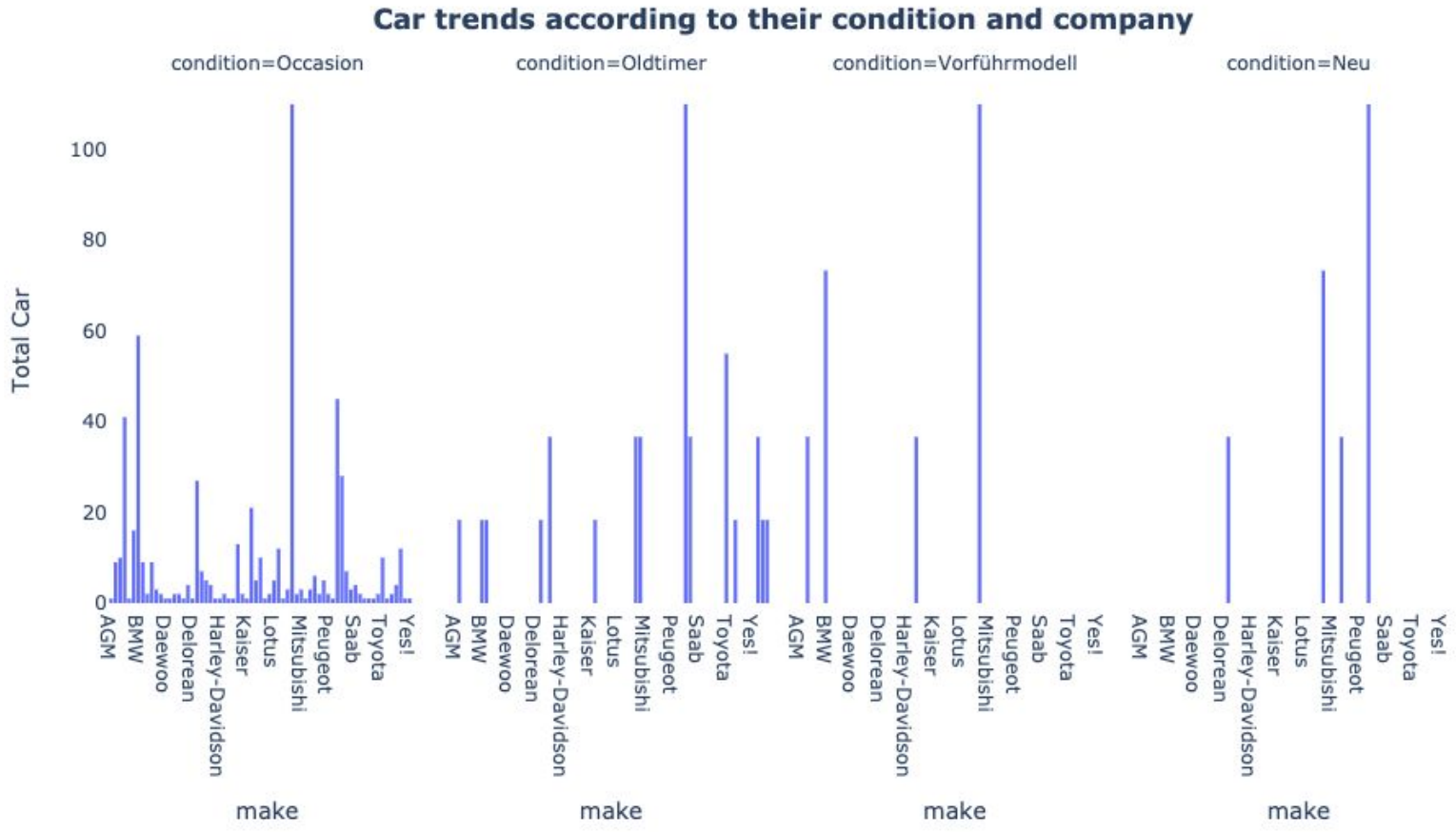
Car sold according to safety measure

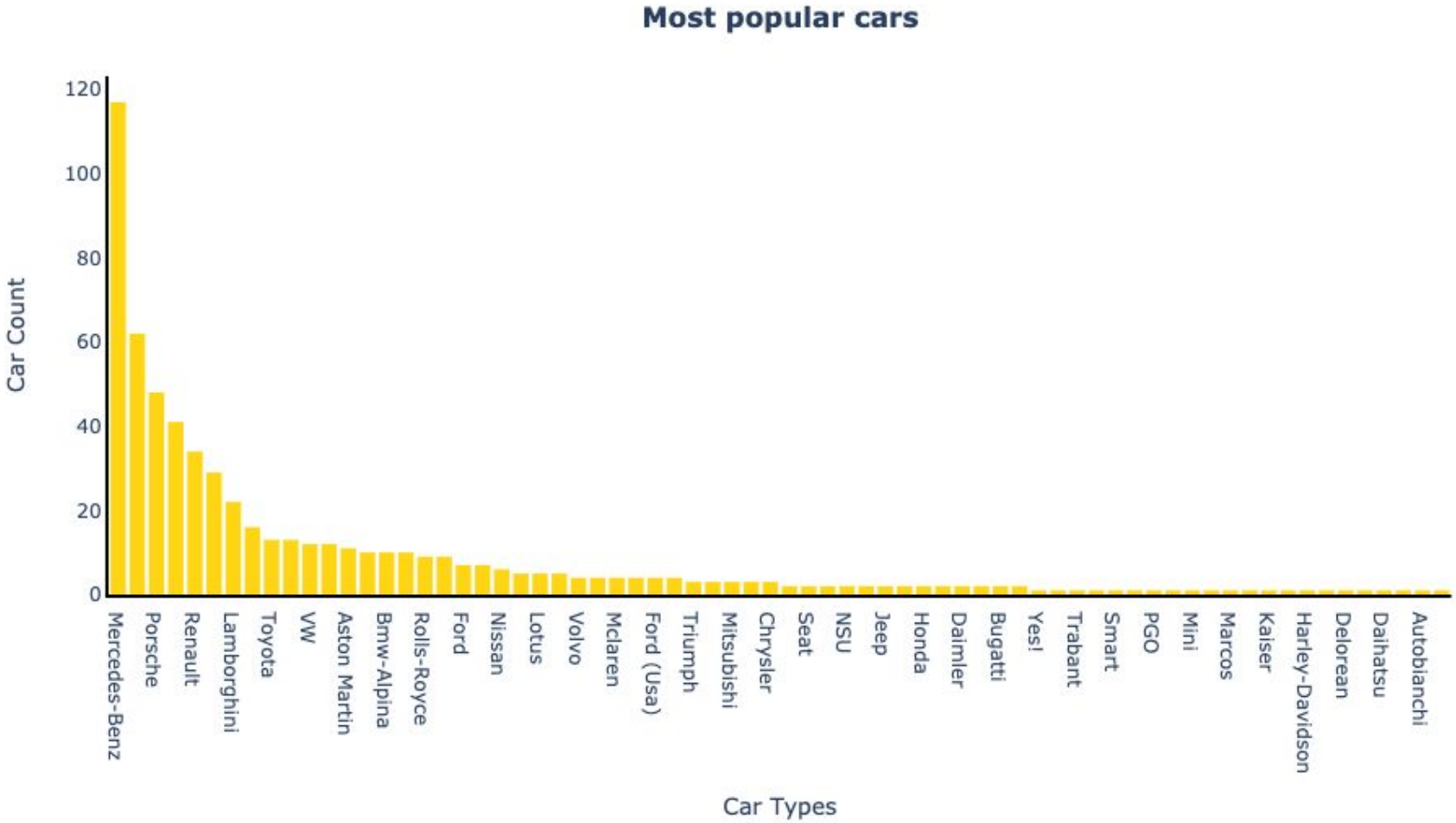
Key Facts

Although occasion is popular in Zuzwil but interestingly Basel favours Oldtimer!

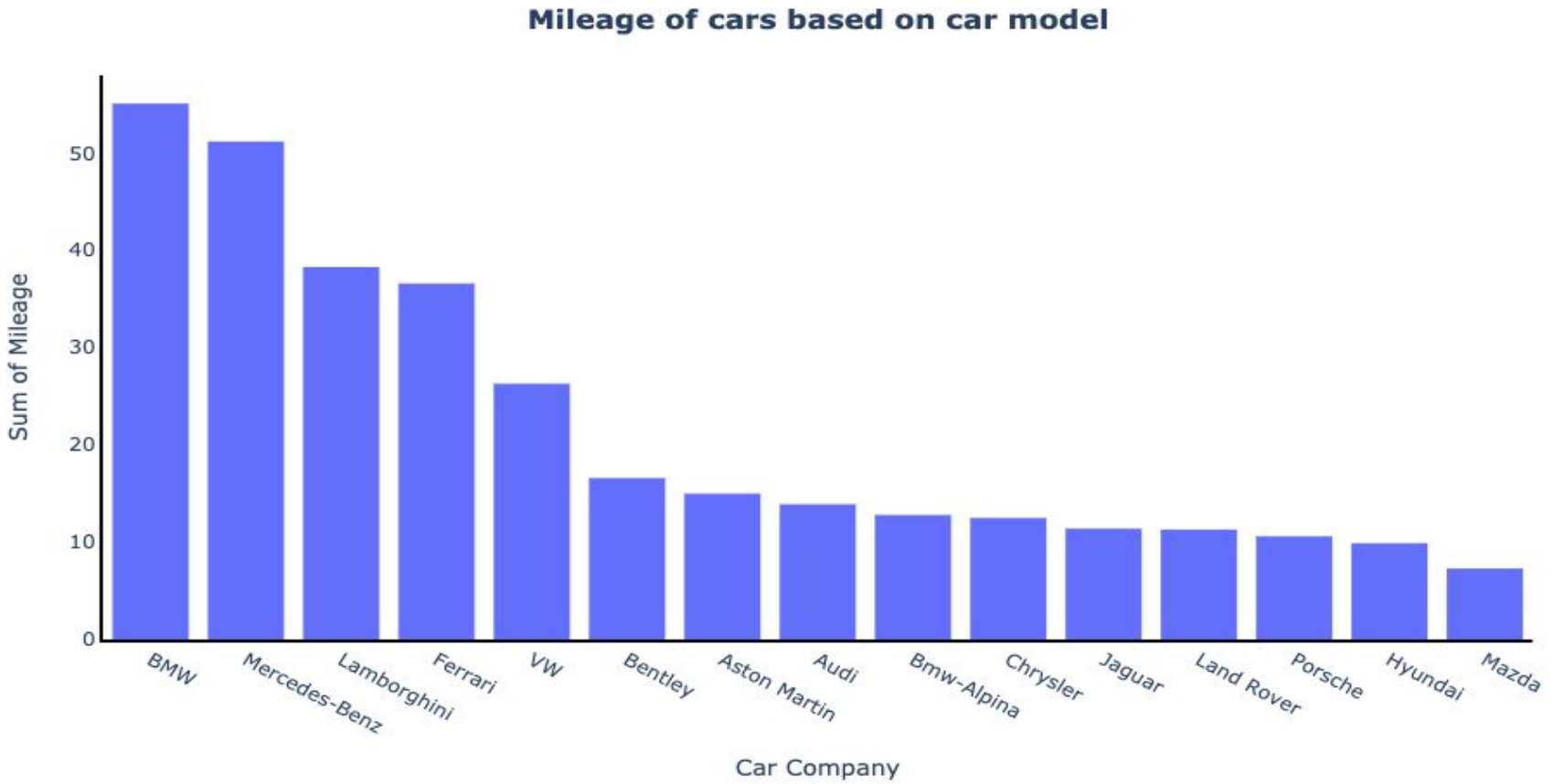


Neu and Vorführmodell are among the least favourite.

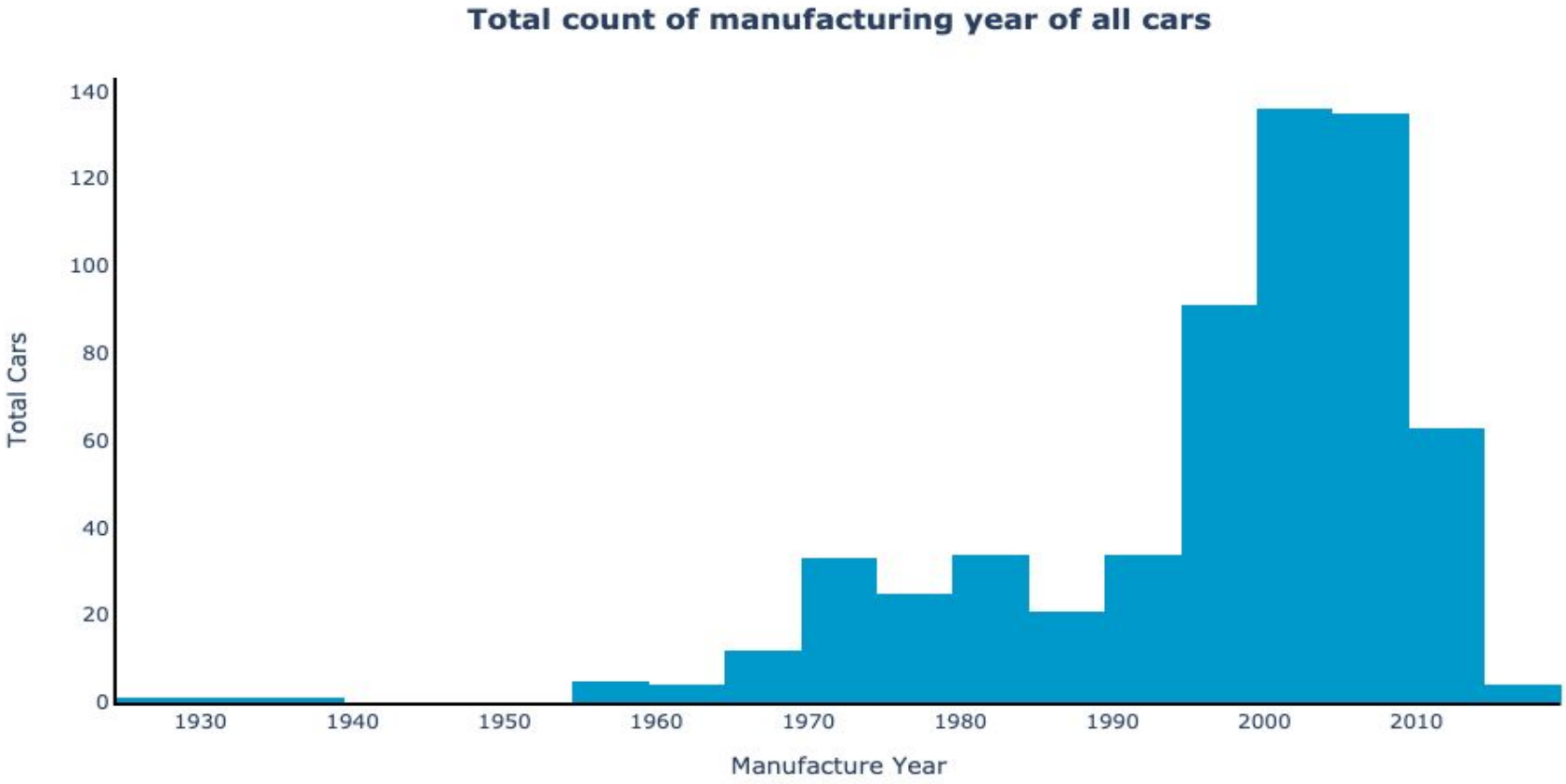



Most popular cars
Car Count
120
100
80
60
40
20
0
Mercedes-Benz
Porsche
Renault
Lamborghini
Toyota
VW
Aston Martin
Bmw-Alpina
Rolls-Royce
Ford
Nissan
Lotus
Volvo
Mclaren
Ford (Usa)
Triumph
Mitsubishi
Chrysler
Seat
NSU
Jeep
Honda
Daimler
Bugatti
Yes!
Trabant
Smart
PGO
Mini
Marcos
Kaiser
Harley-Davidson
Delorean
Daihatsu
Autobianchi
Car Types

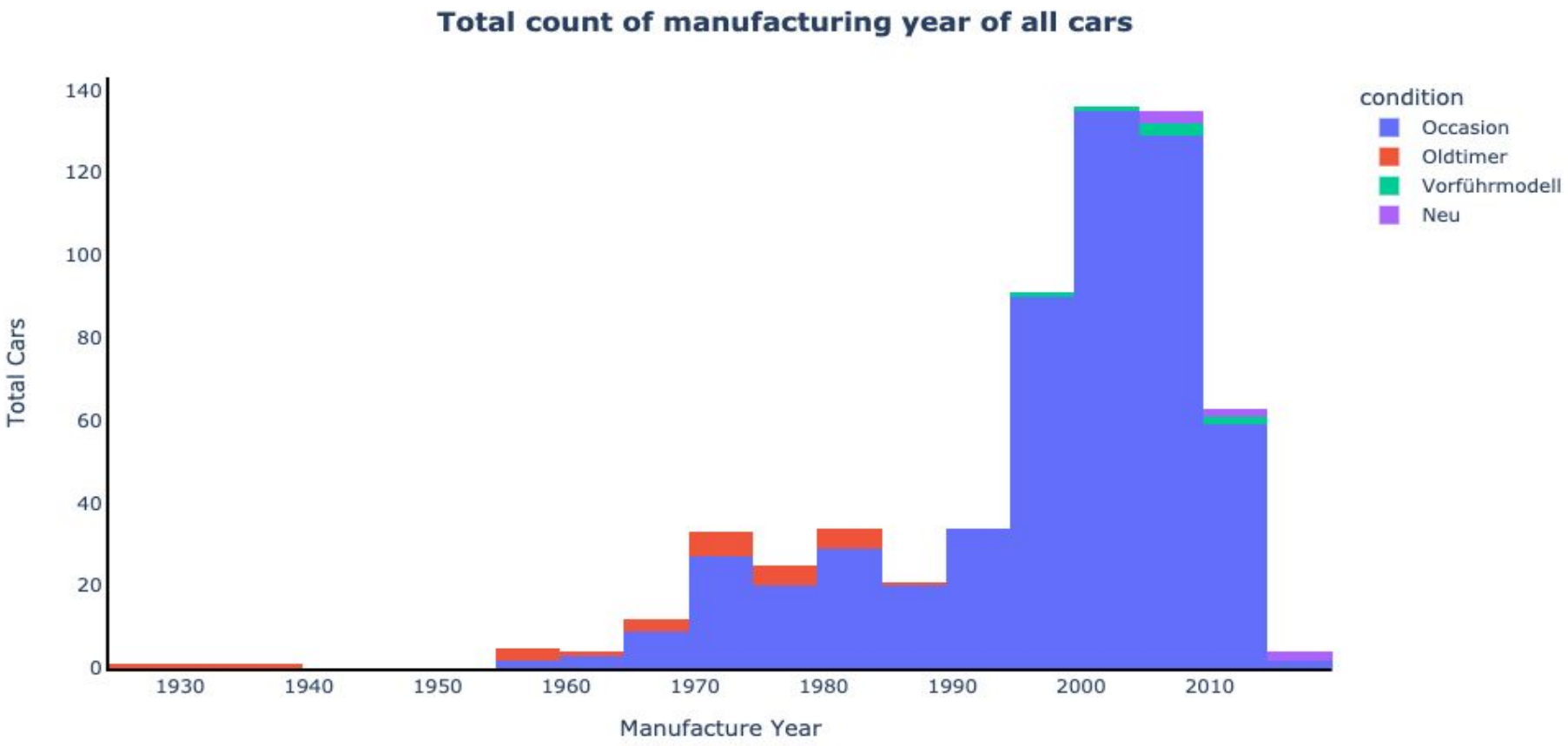
Does mileage matters? Although BMW models in all give more mileage but it is after Mercedes-Benz in popularity



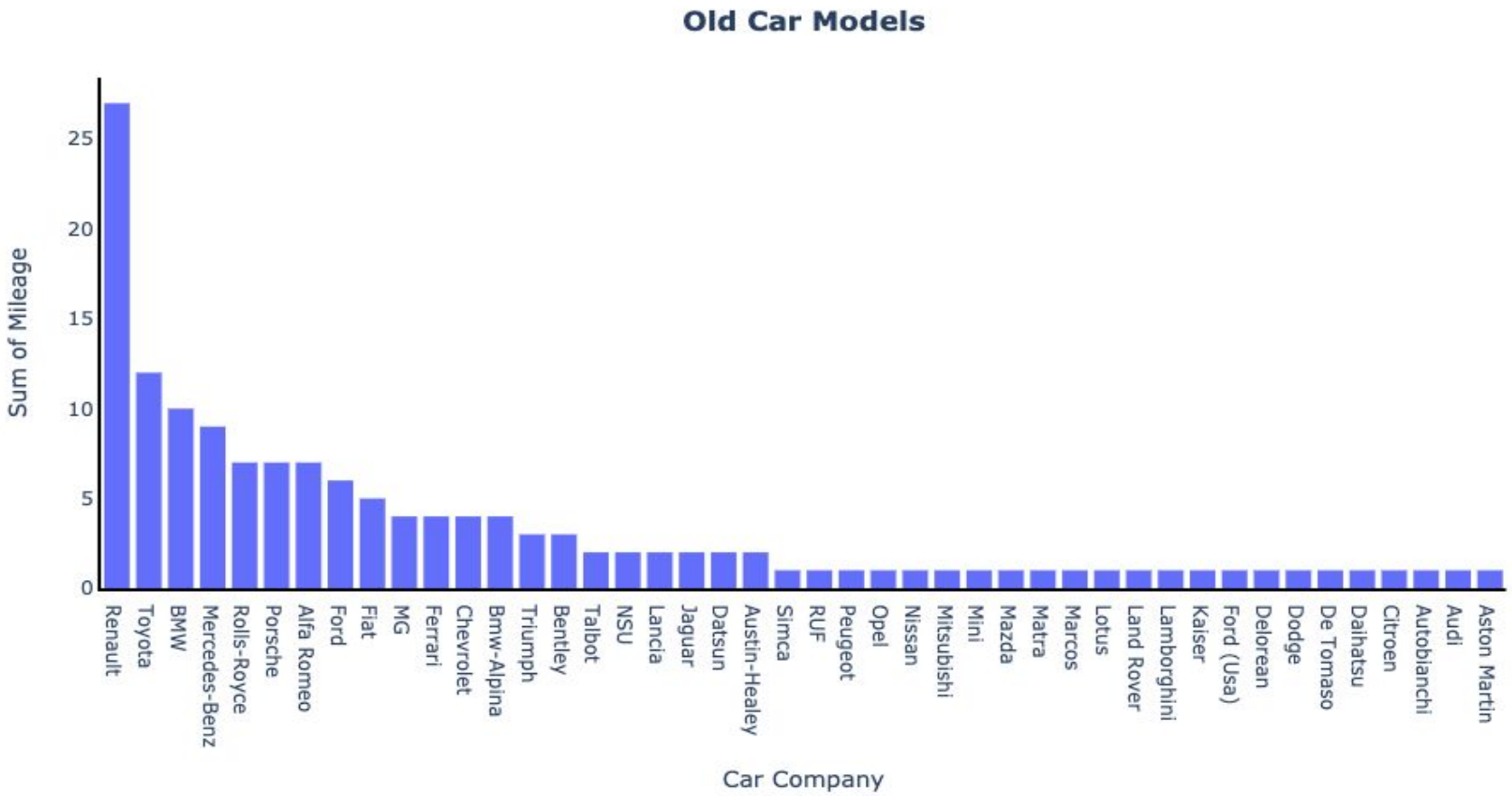
People still keep old models. Lets see their condition...



Oldtimers dominate till 1990.



Interestingly in cars manufactured before 1990 Renault is more popular



Impact

1. Data can easily be used in required format.
2. Area of business could be easily identified
3. E-commerce shop could deny the proposals of models which are not profitable.

Queries