**Car Price Prediction**

**Objective of the Project**

Here, In this Linear Regression project we are trying to predict a new car’s price with target variable as Car Selling Price based on some predictors like Transmission, Fuel Type, Car NAme, Owner, Year.

**Database Used**

No-Sql database used. Here we use Mongo-DB.

We used Spark as Apache Spark is a powerful and distributed data processing framework that can be used in conjunction with MongoDB to provide scalable and high-performance analytics capabilities.

It helps us to built-in connectors and libraries to seamlessly interact with MongoDB. These connectors allow you to read data from MongoDB collections directly into Spark, perform complex analytics and transformations, and write the results back to MongoDB.

I also perform all the EDA part of the dataset including univariate and bivariate analysis.

**Model Used**

With the help of Spark distributed data processing framework, I performed the Linear regression model and got my R2 as .45 which is not a good fit for the model. So, we need a further investigation regarding the process.