When it comes to developing a machine Learning model, we need to check the pre-requisites by exploring the data, cleaning the data, pre-processing the data, analysing the data and then finally building a ML model.
In this dataset which contains the details related to selling, payments, year of purchasing, price, vehicle types, engine type and many more.

1. At first, we need to conduct an Exploratory Data Analysis (EDA) and identify the crucial featured that will be utilized in the model.
* We can perform this task by loading the dataset and inspecting the data by importing libraries such as pandas, sklearn, matplotlib.
* Identifying the dependent and independent variables (used to predict target variables).
* Removing the duplicate values can help us give a more efficient model.
* Selecting the most relevant features that are related with the target variable. After analysing the above data, based on the knowledge I know, I can come to a conclusion of setting price as my target in the model.

2. To Justify the selection of the feature relevant to the target price.
We know that the price of the car model or any other vehicle is decided by
* The brand you choose like luxury cars like Benz or BMW or cars like Hyundai or Kia.
* Fuel type of the vehicle whether it is electric, petrol, diesel, the year of purchase.
* Transmission of a car whether it is manual driven or automatic.
* Engine Power
* Demand for the vehicle.
* Mileage of the vehicle and many more like customization of the cars if needed.
Based on these factors or variables we might be able to build a model which will be able to predict the price of the car when you specify the features you need and at what time you need it to be delivered. Therefore, I choose price as my target feature variable and the other features as my attributes.

3. The potential changes or limitations I would encounter during the featuring process may be:
* The missing values present in the dataset. As we can observe that in the given dataset for some instances there are no mentioned vehicle types, and in some the gearbox values. This could lead to either removal of a feature or imputing the missing values.
* Some features may appear to be relevant to the problem but may not add much value to the model like no.of pictures.
* When it comes to a large database which contains a large number of features, the complexity of the model also increases, making it difficult to identify the most important features. We need to use relevant techniques to identify the most relevant features.

4. Based on the Exploratory Data Analysis (EDA) and the features that are selected, proposing a good fit model that makes prediction of the approximate values of the car prices could be a regression model such as decision tree regression, linear regression. They can take account of the selected features such as mileage, vehicletype, fueltype, gearbox etc. to predict the price of the car.
* According to the research or data from google we can also be using some powerful algorithms like Random forests which can handle both numerical and categorical features and have a strong ability to capture non-linear regression between features and the target variable.
* It is also important to note that the accuracy of the model's prediction where the selected features may not fully capture all relevant features and need to be incorporated to further improves the model's performance.