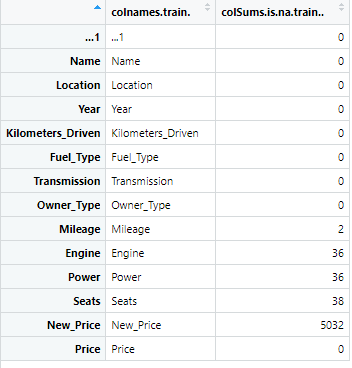
1. Below is a table of all missing value totals for each category in the table. Please ignore …1 as that is the row entry



This puts us in an odd situation when it comes to replacing values. The biggest issue comes with the New\_Price column. Roughly 8**6.06%** of entries in the database do not have a New\_Price recorded, which creates a high chance that we do not even have a chance to create an average to estimate a New\_Price for any task that uses the dataset. **New\_Price** needs to be dropped or supplemented with data from an outside source.

As for the two rows **missing mileage**, they would need to be dropped for any tasks that require a mileage estimate. The dataset provided no other rows with matching Name and Year, thus not giving us any data points to find any type of estimate of mileage for those cars. Again, an alternative would be supplemented by an outside data source.

As for **Engine**, **Power, and Seats** dropping them is mostly justified because they share the same rows, and losing such a small number of rows in a dataset seems negligible. Why not try to estimate when there are a few duplicate cars that share a year and name? Because thee attributes can still vary amongst cars of the same name, model, year, and make due to manufacturing and modifications.