Removing/Imputing Nulls:

For the Imputing missing values I've determined that if the vehicle displacement is available in another row for the given car that value should be used instead of the mean of vehicle displacement because displacement is the total volume of air and fuel that an engine's pistons move in one cycle.[From google] This isnt likely to change drastically for the same car only a few years apart so calculating the mean for this could lead to bad data.[This was copied out of the comments in my code where I originally answered these questions and made these comments]

Enrich Data:

Question I had while merging the data: Both dataset have multiple rows for a given name so how does it know which name in one dataset maps to the same name in the other dataset? Does it just map to the first one that appears in the dataset?[Also copied from my comments in my code]

Data Manipulation:

The US produces the majority of vehicles in this dataset while Europe produces the least.

The US vehicles have the lowest mpg average while the Japanese vehicles have the highest average.

The US vehicles have the highest average weight while the Japanese vehicles have the lowest.[all copied from my code comments]

Structuring the data:

While a lot of the features seem to have a similar scale there are a few outliers in the data and because of that I would choose standardization to fully capture the relative scale.[Copied from where I answered this question at the end of my jupyter notebook]