Melbourne Dataset proect

1. Using methods for MNAR and MAR identification, find which variables belong to either.
2. Use a host of imputation methods on the data, produce visuals and statistics to assess the performance of the different cleaning techniques
3. Assess the distribution of numerical variables, using two visualisation methods.
   1. Do any of the variables follow a normal distribution?
   2. Assess the level of skew in the variables
   3. Identify visually and statistically any anomalies in the data
4. Load the data column as a date variable and create three new variables:
   1. Month
   2. Season
   3. Year
5. Produce a heat map for the numerical data
   1. Which variables show interesting relationships
   2. Is there any suggestion of multicollinearity in the dataset
6. Turn the price variable into a cut and produce a countplot and crosstable with other grouping varibles: Council Area, Region Name, Method,type, Car (you may wish to covert it into a string)
   1. Assess key trends. Does there seem to profiles within the data?
7. Use methods for calculating effect size for numerical variables and assess the association effect on each pair (also include level of significance)
8. Use methods for calculating effect size for Categorical variables and assess the association effect on each pair (also include level of significance)
9. Use methods for calculating effect size for Categorical variables and assess the association effect on each pair (also include level of significance)
10. Produce your final analysis on the data considering all points made earlier.