1. Data Preprocessing:
   1. Handle any other data quality issues, such as outliers or inconsistencies.
   2. Convert categorical variables into numerical representations if necessary.
2. Feature Selection:
   1. Identify the relevant features that are likely to be informative for imputing the missing values.
   2. Remove any irrelevant or redundant features to reduce noise and improve efficiency.
3. Data Removal
   1. To simulate missing data and test the imputation techniques, you can randomly remove some values from your complete dataset.
4. Splitting the Data:
   1. Split the dataset into training and testing sets, ensuring that both sets contain missing values.
   2. The training set will be used to train the imputation model, while the testing set will be used to evaluate its performance.
5. Imputation Model Training:
   1. Choose an appropriate imputation technique based on the nature of your data and the missing value patterns.
   2. Train the imputation model using the training dataset.
6. Imputation and Evaluation:
   1. Apply the trained imputation model to the testing dataset to fill in the missing values.
   2. Evaluate the performance of the imputation model by comparing the imputed values with the actual values (if available) or by assessing the overall quality of the imputed dataset.
7. Iteration and Refinement:
   1. If necessary, iterate and refine the imputation process by trying different techniques, adjusting hyperparameters, or incorporating domain knowledge.