

1.

MAE_01_mean = 0.0068

Runtime_01_mean = 5.21

MAE_01_hd = 0.0042

Runtime_01_hd = 1433824.16

MAE_10_mean = 0.0691

Runtime_10_mean = 8.65

MAE_10_hd = 0.0457

Runtime_10_hd = 2835277.88

2. The ten percent error using the mean imputation has the biggest error. This is because the ten percent error has more missing values meaning that the algorithm has to do more imputations which will naturally be incorrect most of the time. The mean imputation is less accurate than the hot deck method because it uses the same value for every missing value instead of trying to find a value that would be close to the real value

3. The mean is faster because it requires less computations. The mean method only needs to read every row once but the hot deck method needs to read every row once for every row. The computational complexity of the mean method is $O(n)$ and the computation complexity of the hot deck method is $O(n^2)$

4. The hot deck method for the 10 percent error is the slowest because it has more errors than the 01 percent error and because of how the algorithm works that means it has to loop through the entire dataset for every error it finds.

5. The computational complexity of a median imputation algorithm would be greater than the mean imputation because it would require sorting the feature. Sorting takes $O(n \cdot \log n)$ time and mean takes $O(n)$. However, the median imputation algorithm would take less time than the hot deck method because the hot deck method runs in $O(n^2)$.