LAB SHEET – 1

DATA PREPROCESSING

1. Import the necessary libraries for this class and create a DataFrame. https://www.kaggle.com/datasets/nasa/meteorite-landings
2. Find the number of rows and columns in the DataFrame.
3. Descriptive statistics summary of the 'year' values in the 'met_df' DataFrame.
4. Slice the dataframe and create a new one such that
   a. The rows containing the year values less than 860 and greater than 2016.
   b. The rows having the 'reclong' values greater than or equal to -180 degrees and less than or equal to 180 degrees.
   c. The rows containing the 0 'reclat' and 0 'reclong' values from the previously filtered dataframe.
5. Check whether the last created DataFrame has missing values or not.
6. Retrieve all the rows containing the missing 'mass' values in the DataFrame.
7. Get descriptive statistics for the 'mass' column in the 'correct_lat_long_df' DataFrame.
8. Create an array containing the indices of the rows having the NaN values in the mass column.
9. Retrieve the missing 'mass' values from the DataFrame.
10. Replace the missing values in the 'mass' column in the DataFrame with median of mass.
11. Check whether all the missing mass values have been replaced by the median of the mass values or not.
12. Create a descriptive statistics summary of the 'mass' column in the above DataFrame containing the missing rows.
13. Create a descriptive statistics summary of the 'mass' column in the DataFrame.