

Assignment - Day 2 - Preprocessing

Instructions - Read the **NOTE** provided at the end of this document.

Q.1 Import numpy, pandas, matplotlib and other necessary libraries.

Q.2 Import the dataset from the following link:

https://drive.google.com/drive/folders/1oKxAXup3JvGEjSnfAXZ-iSf6H9Tz_VqD?usp=sharing

Q.3 Print the first and last 5 rows of the dataset.

Q.4 Print the description of the dataset. (describe()).

Q.5 Check if this dataset contains any null values and print the count of null values.

Q.6 Drop the unwanted columns from the dataset and print the first 10 rows of the data.

Q.7 Print the information of the dataset. (info()).

Q.8 Handle the outliers in the data frame column using InterQuartile Range (IQR) method.

Q.9 For all the numeric values columns, fill the outliers with the median values.

Q.10 Print the first 10 rows of the dataset and check and analyze it.

Q.11 Again, check if this dataset contains any null values and print the count of null values.

Q.12 Fill all the null values with the median values in the dataset and print the first 10 rows of the dataset.

Q.13 Again, check if this dataset contains any null values and print the count of null values.

Q.14 Check the information of the dataset once again. (info()).

[**NOTE:-** You can take Reference from resources, but the code should be completely yours, with output. Use Jupyter Notebook or google colab. Mention your 'Roll_Name' in the filename, send the files in ".ipynb (python notebook)" format. **Upload the assignment in your google drive(if jupyter notebook else if google colab just share the link) and copy and submit the shareable link in the website. Remember to use the same .ipynb file for each day's assignment, as all the assignments will be related.]**