

EXPLORATORY DATA ANALYSIS

After loading the data into the notebook and converting it into a dataframe. You need to perform data analysis to know your dataset.

In this particular case the data set had 24902 unique values with 36 features. Total number of observations as per the dataset being 141658 , huge number of missing values are also available hence need to mark all the missing values and replace them with NaN . Then replace NaN with the most appropriate value possible .

As per correlation matrix sys_mod_count is highly correlated to closing time and resolving time. priority and urgency are also highly correlated with each other. Priority and impact too have high correlation. Correlation between system_mod_count and reassignment_count is positive as well.