Assignment 10

1. Take a sample dataset containing multiple categorical variables and perform target encoding on it.
2. What is Base N encoding?

Ans:

Base-N encoder **encodes the categories into arrays of their base-N representation**. A base of 1 is equivalent to one-hot encoding (not really base-1, but useful), a base of 2 is equivalent to binary encoding. N=number of actual categories is equivalent to vanilla ordinal encoding.

1. Name suitable feature selection techniques. How you will do feature selection using Regularization?
2. What are different imputation techniques for handling categorical and numerical features?
3. What is the difference between univariate and multivariate imputation?