# I am working on ipums\_la\_99-small dataset. In my dataset there are 61 columns and 8864 rows. In my dataset some columns are numeric and some are objects. In my dataset there are missing values in some of the columns which are in the form of ‘?’. So, first of all I converted the ‘?’ into Nan. Then I fix the missing values. I have also, drop some of the columns which are not required. Then, I converted CONVERTING THE TWO CATEGORIES AS "0" AND "1

**FOR EXAMPLE:- FEMALE AS "1" AND MALE AS "0". Then, I pass the numerical columns in the list and then I normalize the numeric columns. Then I limit to categorical data using df.select\_dtype () and taking all the object data types as categorical columns without giving names. Then I check the original shape of categorical columns. Then, I view the categorical columns using df.columns (). Then, I do Onehotencoding. Then, I do Different Classification Methods. First, I do Logistic Regression and I found that the accuracy is 82.5% Then, finally, I do Random Forest and found the accuracy is 100%**