In a simple term,

Let’s say, we have a data set with features X is [ID, Surname, Age, Country] as follows

ID	Surname	Age	Country
1	Mitchell	42	France
2	Tamil	18	India
3	Scott	22	Germany
4	Henderson	44	France
5	Sharma	23	India

categorical column called “Country” and its values are - **[India, Germany, France]**

In ML regression models, predictions will do the good job if categorical values are converted into numerical (binary vectors) values. Encoding categorical data technique to apply for the above categorical set and the values a.k.a dummy variables will become

ID	India	Germany	France
1	0	0	1
2	1	0	0
3	0	1	0
4	0	0	1
5	1	0	0

This is called - One hot encoding technique.

Dummy Variable Trap :

With one hot encoding conversion, we have three columns in place. By including dummy variables in a regression model, we should consider to drop a column - “Dummy variable trap” N - 1 dummy variables (It is always a good practise to use minimal set of features to apply ML techniques and create regression model.) . In the above dummy variables table, we should consider to drop any of one columns. Let’s drop the column “Germany” and the final table looks like in below

1	Mitchell	42	0	1
2	Tamil	18	1	0
3	Scott	22	0	0
4	Henderson	44	0	1
5	Sharma	23	1	0

Thumb Rule :- Which dummy variable column do we need drop? The answer is - we can drop any of one dummy variables column. It can predict the dropped column's value based on other two columns. Let's take the record no 3 in the above table, both dummy variable values are '0'. So obviously another dummy variable column value is '1' and categorical value is 'Germany'.

Technically, the dummy variable trap is a scenario in which the independent variables are multi-collinear - two or more variables are highly correlated.

S