

I Prefer you to once open the Car dataset before reading this file you will get a better understanding on it.

What is Dependent Data and What is Independent Data?

Dependent Data:-

What do you Understand by dependent Data is which is dependent on another

Like in our Cars dataset if you are going do it from my repository

You are going to find this columns

`['brand', 'km_driven', 'fuel', 'owner', 'selling_price']`

These are our column Name in this just think of which is dependent one , You will find out like selling price is dependent on , which brand your Car belongs to , how much distance you car has Covered , fuel : what is average, owner: is it third , first or second

So In this Dataset selling price is dependent on our Independent Data . That's why our dependent data is `["Selling_price"]`

Independent Data:-

When you have find out the dependent colour other all are our Independent Data

In Our Car Dataset independent Data is `['brand', 'km_driven', 'fuel', 'owner']`

Categorical Columns are `['brand', 'fuel', 'owner']`

and Numerical Columns are `['selling_price', 'km_driven']`

Categorical To Numerical Data Techniques

we have 4 methods to Convert Categorical Data to Numerical Data

1)Get Dummies :-

Only Useful when in Categorical Columns we have 2-3 unique Values. if more unique items are their.we fall into the dummy trap and our Columns Number will get increased.

2)One Hot Encoder:-

One hot Encoder is the technique is also same as get dummies in this also you will get many rows but learn this technique also as it comes handy while doing other algorithm like SVM, Decision Tree and many more.

3)Ordinal Encoder:-

This one is more preferable when you have many unique values in a single column if you have like 5 unique values so the Ordinal Encoding will be like 1,2,3,4,5. This is one good Technique. Ordinal Encoder is for Independent Data and for 2D (n_samples,n_feature)

4)Label Encoder :-

This is the also as same as Ordinal Encoder but it returns array while Ordinal Encoder array in array , just Remember like It is for our Independent Column that's it. It is for 1D array (n_samples)

Nominal V/s Ordinal Data

Nominal Data:-

Nominal scale is a naming scale, where variables are simply “named” or labeled, with no specific order. Ordinal scale has all its variables in a specific order, beyond just naming them like sex(Female and Male),

Political preference(1- Independent ,2- Democrat ,3- Republican)

just you can remember like also they don't have many Unique Values That's it.

Ordinal Data:- also called the categorical variable scale, is defined as a scale used for labeling variables into distinct classifications and doesn't involve a quantitative value or order . Remember Like if more Unique Values are there use Ordinal Encoding for Independent Data

Label Encoder For dependent Data

It is Majorly used in surveys. Ex:- Survey of Phone think now how many companies , different types of model.

I think you all have got what is Nominal Data , Ordinal Data