

Categorical data is non-numerical data.

It sometimes have information about hierarchy.

This data can be stored in categories on the basis of relevance.

To pass this data to model, encoding techniques like one hot or ordinal encoding (order is imp.) , nominal encoding,

In case of multi-linear regression, each category is treated as separate line (with diff. co-eff. &  $x$ )

ex- If we are using one hot &  $x_1, x_2, x_3$  are features.

$x_3$  is to be encoded using one hot encoding.

$\swarrow \searrow$   
 $x_4 \quad x_5$

$$y = w_1 x_1 + w_2 x_2 + w_4 x_4 + w_5 x_5$$