What is Categorical Data?

Categorical data are variables that contain label values rather than numeric values.The number of possible values is often limited to a fixed set. Categorical variables are often called nominal.

What is the Problem with Categorical Data?

- Many machine learning algorithms cannot operate on label data directly. They require all input variables and output variables to be numeric.
- In general, this is mostly a constraint of the efficient implementation of machine learning algorithms rather than hard limitations on the algorithms themselves.
- This means that categorical data must be converted to a numerical form. If the categorical variable is an output variable, you may also want to convert predictions by the model back into a categorical form in order to present them or use them in some application.

How to Convert Categorical Data to Numerical Data?

This involves two steps

- Integer Encoding - As a first step, each unique category value is assigned an integer value. The integer values have a natural ordered relationship between each other and machine learning algorithms may be able to understand and harness this relationship.
- One-Hot Encoding

How does the term "one-hot encoding" relate to the preprocessing of categorical variables in machine learning?

- For categorical variables where no such ordinal relationship exists, the integer encoding is not enough.
- In fact, using this encoding and allowing the model to assume a natural ordering between categories may result in poor performance or unexpected results (predictions halfway between categories).
- In this case, a one-hot encoding can be applied to the integer representation. This is where the integer encoded variable is removed and a new binary variable is added for each unique integer value.