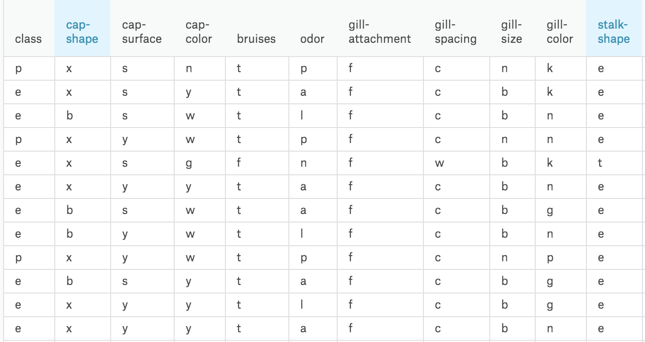
**Data Understanding**

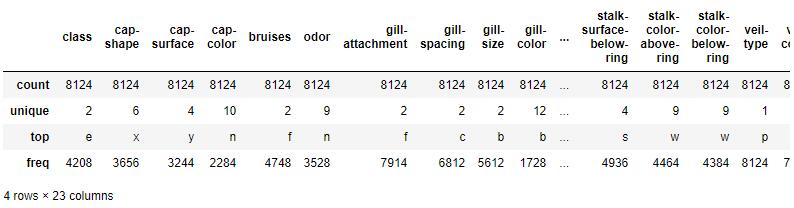
Raw data from Kaggle

Dataset with 23 species of gilled mushrooms

Over 8000 variables and 23 attributes

Each attribute influences on the outcome of a single type of mushroom is edible or poisonous.





**Data Preparation**

First, check data by null check, data should have no null values, otherwise, there will be problems in building the model.

Second, since the dataset has string values, dataset needed to convert all the unique values into integers then label encoding is performed on the data.

Third, use data balance to make sure data prediction will be unbiased towards the more frequent class.

Last, split the data into train data and test data according to the ratio of 8 to 2 to build and test the model.

