**Handling missing values in Categorical Variables**

1. Delete rows with missing value

* Need to see well if doing so will not make use lose important information.
* Can do it with very large datasets – millions of rows

1. Replace with the most frequent values

* Mode replacement

1. Apply Classifier algorithm to predict missing value

* Use other features are independent features even the output feature – Train dataset
* Row with missing data will be the test dataset
* Give only the test data records and get predictions
* Can apply KNN

1. Apply unsupervised ML

* Clustering techniques
* Take only the features (not the output) with all data – say we have 2 features 🡪 take these 2 columns and group them in clusters
* See where the missing data falls to know the missing data – in which cluster does it fall?