Since overfitting is bad, it is best to find a way to limit features. The process of reducing features is called dimensionality reduction. There are two options for coping with too many features: elimination and extraction.

**Feature Elimination**

Your first idea is to remove a good amount of features so the model won't be run using every column. This is called **feature elimination**.

Feature elimination means what you think: You remove, or eliminate, a feature from the dataset. In our school supply example, you remove features that aren't relevant to what we're looking for, such as name, address, and zip code. This simple method increases and maintains interpretability.

The downside is, once you remove that feature, you can no longer glean information from it. If we want to know the likelihood of people buying school supplies, but we removed the zip code feature, then we'd miss a detail that could help us understand when certain residents tend to purchase school supplies.

**Feature Extraction**

Feature extraction combines all features into a new set that is ordered by how well they predict our original variable.

In other words, feature extraction reduces the number of dimensions by transforming a large set of variables into a smaller one. This smaller set of variables contains most of the important information from the original large set.

**NOTE**

Sometimes, you need to use both feature elimination and extraction. For instance, the customer name feature doesn't inform us about whether or not customers will purchase school supplies. So, we would eliminate that feature during the preprocessing stage, then apply extraction on the remaining features.