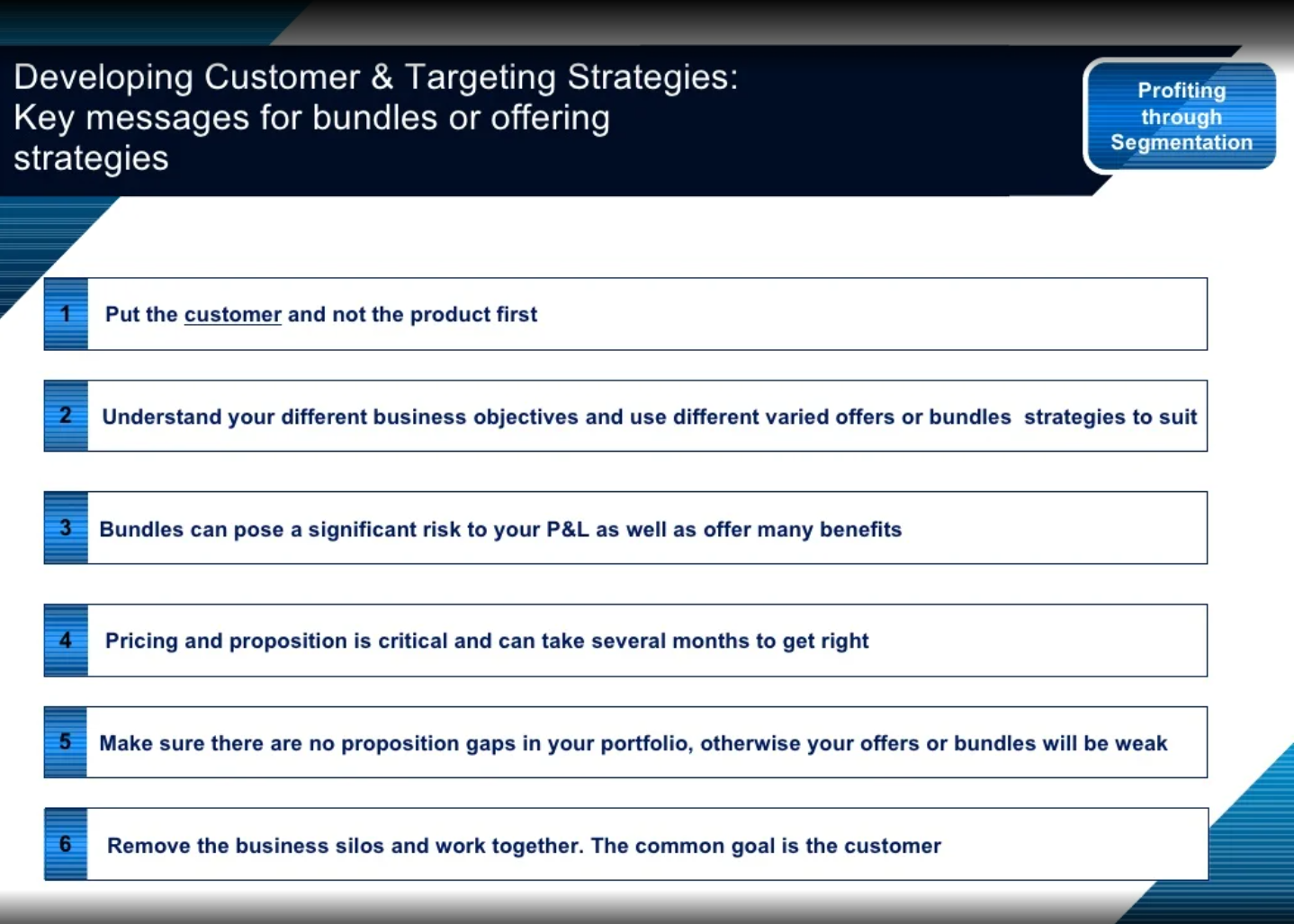
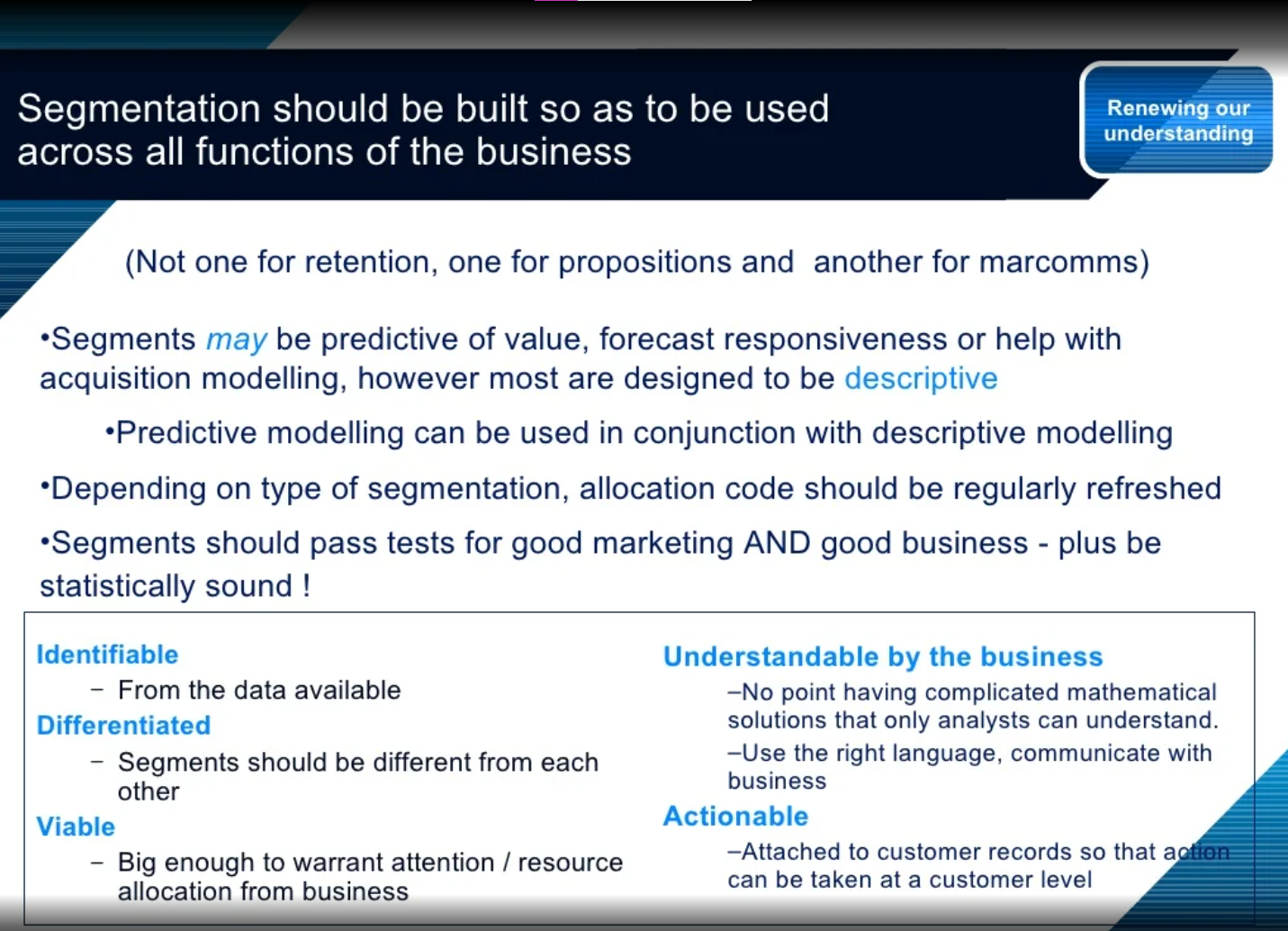
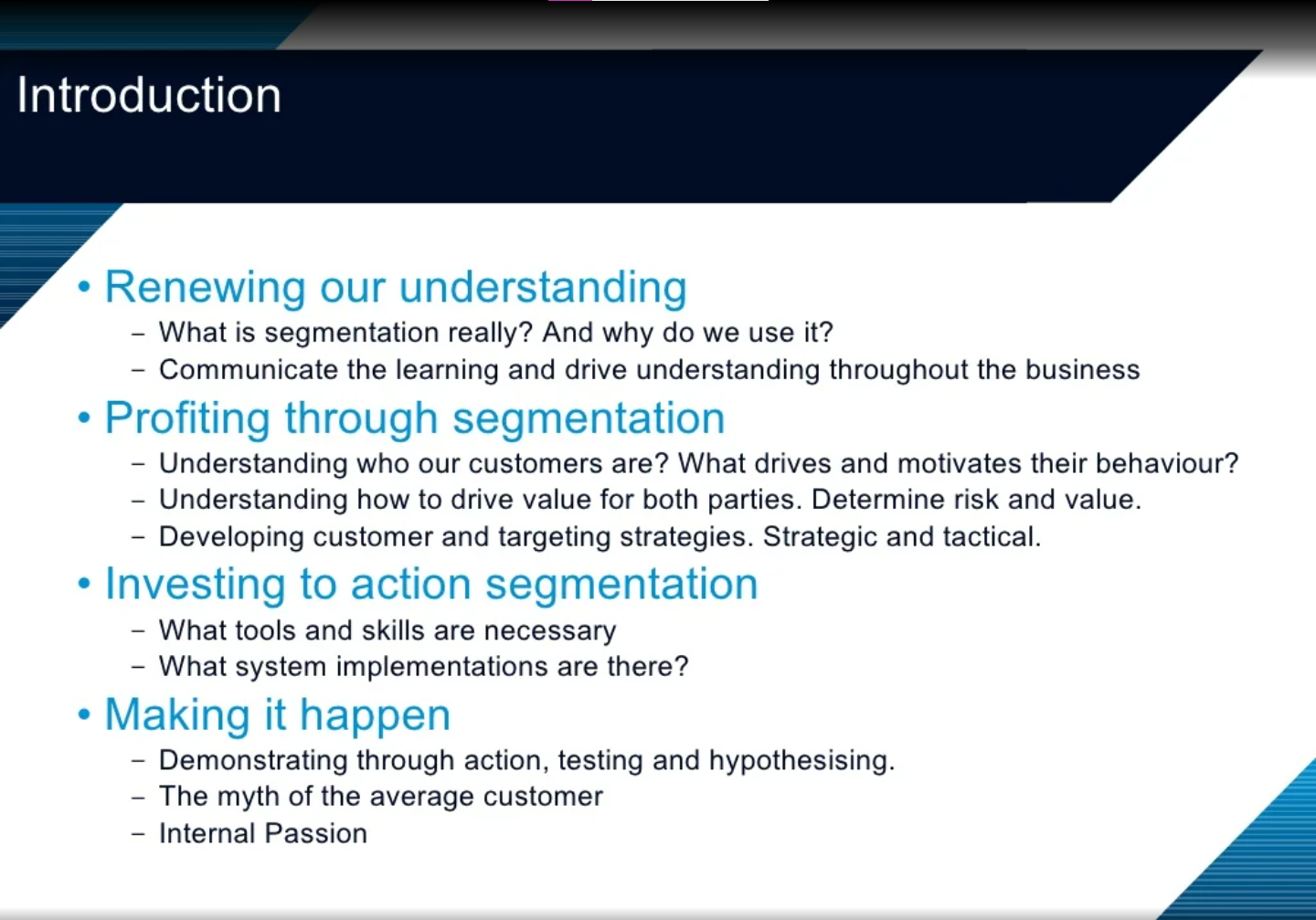
K means clustering is one of the most popular clustering algorithms and usually the first thing practitioners apply when solving clustering tasks to get an idea of the structure of the dataset.

The goal of K means is to group data points into distinct non-overlapping subgroups.

One of the major applications of K means clustering is segmentation of customers to get a better understanding of them which in turn could be used to increase the revenue of the company.







Why silhouette method is better than elbow method?

The major difference between elbow and silhouette scores is that **elbow only calculates the euclidean distance whereas silhouette takes into account variables such as variance, skewness, high-low differences, etc.** Elbow and Silhouette methods are used to find the optimal number of clusters. Ambiguity arises for the elbow method to pick the value of k. **Silhouette analysis can be used to study the separation distance between the resulting clusters** and can be considered a better method compared to the Elbow method.

When would you use a silhouette?

Silhouette analysis can be used **to study the separation distance between the resulting clusters**. The silhouette plot displays a measure of how close each point in one cluster is to points in the neighboring clusters and thus provides a way to assess parameters like number of clusters visually.

Which silhouette score is best?

The value of the silhouette coefﬁcient is between [-1, 1]. **A score of 1 denotes the best** meaning that the data point i is very compact within the cluster to which it belongs and far away from the other clusters. The worst value is -1