**7-step marketing analytics process**.

**Step 1: Business Problem Understanding**

Online fragrance and cosmetic commerce is expanding with rapid speed over $16.2 billion annually, as buyers increasingly shift from in-store purchase to online transaction. Growth has also introduced fiercer competition as weblogs and influencers push online sales through product-centric web postings. The Chief Marketing Officer of a leading beauty care company needs to know more about online purchasing behavior to continue refining marketing further. Historically, the company centered on intermittent "big spenders," but today this is contested. The problem of the business at its source is to be able to effectively segment customers using behavioral patterns for more targeted and effective marketing strategies.

### **Step 2: Data Understanding and Collection**

To satisfy the business requirement, customer behavior information was gathered from web transactions. The data set, referred to as ClusterData from the MKTG525 library, includes fields like CustomerID, Age, Annual Income, Spending Score, and Product\_Description. These are basic demographic and buying behaviors that are essential to the construction of customer segments. An initial analysis was conducted to examine the range, distribution, and interrelation of the variables, creating a solid ground for effective modeling.

**Step 3: Data Preparation and Feature Selection**

Prior to performing cluster analysis, the dataset was cleaned and prepared thoroughly. Outliers in the Annual\_Income variable were identified and deleted with the Interquartile Range (IQR) technique to avoid distortion of the findings. The resultant dataset, which was titled **no\_outliers**, was subsequently standardized using the proc stdize method with the range technique to guarantee that all variables (Age, Annual\_Income, and Spending\_Score) had an equal weightage to the clustering process. These three variables were specifically selected as K-means cluster attributes, as they yield explicit information about customer behavior and purchasing power.

**Step 4: Modeling Development**

K-means clustering was also utilized to experiment with various segmentation scenarios. Models were developed with K from K=3 to K=7, and standardized data were utilized. The output of each model was evaluated based on the anticipated total R-Squared values since these indicate the goodness of fit of the data clustered together. After comparing and contrasting, it was discovered that K=7 was the optimal number of clusters, which achieved the highest value of around 0.75753. This result reflects high homogeneity within clusters and high differentiation among them.

**Step 5: Model Evaluation and Interpretation**

The last model with K=7 clusters was explored for interpretation purposes. Each cluster was established through summary statistics like mean, standard deviation, minimum, maximum, and median values of Age, Annual Income, and Spending Score. The statistics unveiled unique customer segments, e.g., high-income low-spending, young adults with high spending, and average-income customers with medium spending. The data that are obtained from this segmentation allow the company to understand more about the variations of customers and determine certain practices that may be able to shape marketing methods.

### **Step 6: Model and Results Communication**

The results of the clustering analysis were summarized in a table that specifies each of the seven segments' attributes. The results will be shared with stakeholders such as the marketing and product teams to inform the creation of focused marketing campaigns. For instance, a segment that has been identified as comprising young consumers with high spend scores can be targeted with high-end product lines endorsed by influencers, whereas more conservative segments can be targeted more effectively using discount promotions.

### **Step 7: Model Deployment**

The customer segments can be fed to the CRM or campaign management system of the organization for ready marketing deployment. Each individual in the data set has been tagged with a cluster label, so personalized email campaigns, promotions, and recommendations can be done. This type of segmentation can also be utilized for testing new marketing campaigns within particular clusters, which results in increased engagement, increased return on investment, and better marketing resource allocation.