**MARKET SEGMENTATION:**

* A Market planning is a logical sequence and series of activities helps to organization to reach the customer locally as well as globally , The plan two components Strategic Marketing Planning , Tactical Marketing Planning.
* Market segmentation is a decision marketing tool, its offering an opportunity to think and rethink, leads critical new insights
* Implementation Market segmentation requires a substantial investment by the organization. The Evaluation of the success of the segmentation strategy and the continuous monitoring of market dynamics.
* Market segmentation also leads to tangible benefits, including a better understanding of differences between consumers, which improves the match of organisational strengths and consumer needs.
* The strategic plan outlines the long-term direction of an organisation, but does not provide much detail on short-term marketing action required to move in this long-term direction.
* Tactical marketing is responsible for the equipment: the quality of the walking shoes, food, water, a raincoat.

**MARKET SEGMANTATION ANALYSIS:**

* The process of grouping consumers into naturally existing or artificially created segments of consumers who share similar product preferences or characteristics.
* Market Segmentation having three approaches the creation of segments from existing consumer classifications, and the emergence of segments from qualitative research.
* Commonsense and data-driven segmentation are two extremes, the two pure forms of segmentation approaches based on the nature of the segmentation criterion.
* The basic structure is the same for both commonsense and data-driven market segmentation: an organisation needs to weigh up the advantages and disadvantages of pursuing a segmentation strategy
* The ten steps of market segmentation analysis are the same for commonsense and data-driven segmentation, different tasks need to be completed for each one of those approaches.

**DECIDING (NOT) TO SEGMENTATION:**

* Market segmentation has developed to be a key marketing strategy applied in many organisations.
* The commitment to market segmentation goes hand in hand with the willingness and ability of the organisation to make substantial changes and investments.
* The first group of barriers relates to senior management. Lack of leadership, pro-active championing, commitment and involvement in the market segmentation process by senior leadership undermines the success of market segmentation.
* Process-related barriers include not having clarified the objectives of the market segmentation exercise, lack of planning or bad planning, a lack of structured processes to guide the team through all steps of the market segmentation process.
* The purpose and dedication is required, tempered by patience and a willingness to appreciate the inevitable problems which will be encountered in implementing the conclusions.

**EXTRACTING SEGMENTS:**

**Grouping Consumers:**

* Consumer datasets are typically not well structured. Consumers come in all shapes and forms The combination of exploratory methods and unstructured consumer data means that results from any method used to extract market segments from such data will strongly depend on the assumptions made on the structure of the segments implied by the method.
* The result of a market segmentation analysis, therefore is determined as much by the underlying data as it is by the extraction algorithm chosen.

**Distance-Based Methods:**

* Distance-based methods use a particular notion of similarity or distance between observations, and try to find groups of similar observations. So-called *model-based methods* are described second
* **Distance Measures:**
* A distance measure has to comply with a few criteria. One criterion is symmetrythat is ***d(*x*,* y*)* = *d(*y*,* x*).***
* A second criterion is that the distance of a vector to itself and only to itself is ***d(*x*,* y*)* = 0 ⇔ x = y*.***
* Euclidean distance is the most common distance measure used in market segmentation analysis. Euclidean distance corresponds to the direct “straight-line” distance between two points in two-dimensional space,
* **Hierarchical Methods :** Hierarchical clustering methods are the most intuitive way of grouping data.
* *Divisive* hierarchical clustering methods splits it into two market segments each of the segments is again split into two segments. This process continues until each consumer has their own market segment.
* *Agglomerative* hierarchical clustering approaches the task from the other end. The starting point is each consumer representing their own market segment (*n* singleton
* clusters). Step-by-step, the two market segments closest to one another are merged until the complete data set forms one large market segments.

**Partitioning Methods:**

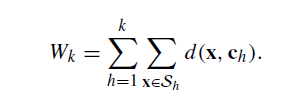
* A partitioning clustering algorithm aiming to extract five market segments, in contrast, would only have to calculate between 5 and 5000 distances at each step of the iterative or stepwise process (the exact number depends on the algorithm used).
* ***k*-Means and *k*-Centroid Clustering**: K-means is an iterative, centroid-based clustering algorithm that partitions a dataset into similar groups based on the distance between their centroids. The centroid, or cluster center, is either the mean or median of all the points within the cluster depending on the characteristics of the data.
* **Improved K-Means:** The simplest improvement is to initialise *k*-means using “smart” starting values, rather than randomly drawing *k* consumers from the data set and using them as starting points.
* **Hard Competitive Learning:**Hard competitive learning, also known as learning vector quantisation hard competitive learning also minimises the sum of distances from each consumer contained in the data set to their closest representative (centroid), the process by which this is achieved is slightly different.
* Hard competitive learning randomly picks one consumer and moves this consumer’s closest segment representative a small step into the direction of the randomly chosen consumer.
* **Neural Gas and Topology Representing Networks:** designed to adaptively organize nodes in a continuous data space, aiming to minimize the quantization error. Unlike K-Means, it does not restrict data points to fixed clusters but distributes neurons across the input space more flexibly.
* **Topology Representing Networks (TRN)** are a type of neural network that combine clustering with topological mapping. These networks aim to create a low-dimensional representation of input data while preserving its topological structure.
* **Neural Networks:** The network has three layers. The input layer takes the data as input. The output layer gives the response of the network. In the case of clustering this is the same as the input. In-between the input and output layer is the so-called hidden layer. It is named hidden because it has no connections to the outside of the network. The input layer has one so-called node for every segmentation variable.
* **Two-Step Clustering:** The two-step cluster analysis uses both distance measures and a model-based clustering approach. It can handle datasets with a mixture of continuous and categorical variables by applying either log-likelihood or Euclidean distance measures, making it highly versatile for different data types and research purposes.
* **Bagged Clustering**: enhances the reliability and stability of customer segmentation by aggregating the results of multiple clustering models. It is particularly effective in market segmentation, where diverse customer behaviour and noisy data can lead to inconsistent cluster assignments with single clustering methods.

**Model-Based Methods**

* model-based segment extraction methods do not use similarities or distances to assess which consumers should be assigned to the same market segment.
* **Finite Mixtures of Distributions:** The simplest case of model-based clustering has no independent variables *x*, and simply fits distribution to *y*. To compare this with distance-based methods, finite mixtures of distributions basically use the same segmentation variables.
* **Normal Distributions:** The Normal Distribution, also known as the Gaussian distribution, is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean. It is one of the most widely used probability distributions in statistics and is crucial for many statistical analyses.
* **Binary Distributions**: A Binary Distribution refers to a type of probability distribution where the outcome of a random variable can take only two possible values, typically labelled as 000 and 111, "success" and "failure," or "yes" and "no." It is the foundation for modelling binary data and is often associated with the Bernoulli Distribution.
* **Finite Mixtures of Regressions:** Finite mixtures of distributions are similar to distance-based clustering methods and – in many cases – result in similar solutions. Compared to hierarchical or partitioning clustering methods, mixture models sometimes produce more useful, and sometimes less useful solutions.
* **Biclustering Algorithms:** Biclustering algorithms are advanced clustering techniques designed to simultaneously cluster rows and columns of a matrix (or dataset). Unlike traditional clustering, which groups only rows (data points) or columns (features), biclustering identifies subsets of rows that are related to subsets of columns, revealing local patterns in the data.

**Data Structure Analysis:**

* **Cluster indices** provide insight into particular aspects of the market segmentation solution. Which kind of insight, depends on the nature of the cluster index used. Generally, two groups of cluster indices are distinguished: internal cluster indices and external cluster indices.
* **Internal Cluster Indices:** A very simple internal cluster index measuring compactness of clusters results from calculating the sum of distances between each segment member and their segment
* representative. Then the sum of within-cluster distances*Wk* for a segmentation solution with *k* segments is calculated using the following formula where we denote the set of observations assigned to segment number *h* by S*h* and their segment representative by c*h*:



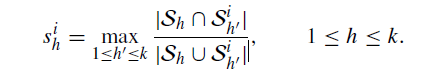
* **External Cluster Indices:** metrics used to evaluate the quality of a clustering solution by comparing it to a pre-existing, ground-truth clustering or labelled dataset. These indices quantify how well the clusters produced by a clustering algorithm align with the true class labels or predefined clusters.
* **Gorge Plots:** type of visualization used in clustering and dimensionality reduction to illustrate the performance or effectiveness of a clustering algorithm. They typically display the gap or difference between within-cluster variance and between-cluster variance, helping to evaluate the compactness and separation of clusters in the data.

**Global Stability Analysis:**

* evaluate the overall stability of a system across its entire phase space, rather than focusing on stability at individual equilibrium points or local regions. It determines whether a system, once perturbed, will return to a steady state or exhibit stable behaviour over time, regardless of the initial conditions.
* Resampling methods offer insight into the stability of a market segmentation solution across repeated calculations. To assess the global stability of any given segmentation solution, several new data sets are generated using resampling methods, and a number of segmentation solutions are extracted.

**Segment Level Stability Analysis:**

* *segment level* stability of market segments contained in those solutions to protect against discarding solutions containing interesting individual segments from being prematurely discarded. After all, most organisations only need one single target segment.
* **Segment Level Stability Within Solutions (SLS*W*):** metric used in clustering and market segmentation to evaluate the internal consistency or stability of segments within a single clustering solution. It assesses whether the same data points consistently belong to the same segments under slight variations in the data or clustering process.
* Compute a partition of the data (a market segmentation solution) extracting *k* segments S1*, . . . ,*S*k* using the algorithm of choice (for example, a partitioning clustering algorithm or a finite mixture model).
* Draw *b* bootstrap samples from the sample of consumers including as many cases as there are consumers in the original data set (*b* = 100 bootstrap samples works well).
* Cluster all *b* bootstrap samples into *k* segments. Based on these segmentation solutions, assign the observations in the original data set to segments S*i* 1*, . . . ,*S*i K* for *i* = 1*, . . . , b*.
* For each bootstrap segment S*i* 1*, . . . ,*S*i k*, compute the maximum agreement with the original segments S1*, . . . ,*S*k* as measured by the Jaccard index:



* The Jaccard index is the ratio between the number of observations contained in both segments, and the number of observations contained in at least one of the two segments.
* Create and inspect boxplots of the *si h* values across bootstrap samples to assess the segment level stability within solutions (SLS*W*). Segments with higher segment level stability within solutions (SLS*W*) are more attractive.
* **Segment Level Stability Across Solutions (SLSA)** is a metric used to assess the consistency of segments across different clustering solutions. It evaluates how well the segments identified by different clustering algorithms or configurations align with each other, ensuring reliability and robustness of the segmentation process.