**SUMMARY**

NAME: PAVAN KUMAR CHEERLAVANCHA

**STEP 1: Deciding (NOT) to Segment:**

**3.1 Implications of Committing to Market Segmentation:**

Market segmentation is a big deal in marketing, but it's not always the best choice for every organization. Before going into it, you want to understand what you're getting into. It is a long-term thing, and it's not cheap. There are costs involved in researching, surveying, designing, and promoting to different segments of the market if we need to create new products, modify existing ones, change prices etc. We will know how organization works internally. Instead of organizing around products, should organize around market segments. Since committing to market segmentation is a big deal, top executives need to be the ones making the call. And once you decide to go for it, everyone in the organization needs to on board and keep working towards those goals.

**3.2 Implementation Barriers:**

If top-level leaders not actively supporting market segmentation, it's hard for the marketing team to make it work. Lack of resources or funding from senior management can also stop successful implementation.

Problems like resistance to change, poor communication, and short-term thinking within the company can block effective market segmentation. Having qualified marketing experts in the organization is important especially in larger markets. Lack of skilled data managers and analysts can also be difficult.

Managers may not accept using techniques they don't understand, so it's important to present segmentation analysis in an easy and understand way. Many barriers can be identified early on and addressed. Implementing market segmentation requires dedication, patience, and understanding.

**STEP 2: Specifying the Ideal Target Segment:**

**4.1 Segment Evaluation Criteria:**

In this step of market segmentation analysis depend on input from users. It's important to involve users throughout the process, not just at the beginning or end. Once a commitment to segmentation is made, the organization plays a big role in Step 2 by establishing criteria for evaluating segments. These criteria are divided into "knock-out" features that are essential and must be met, and "attractiveness" features that help compare remaining segment. They must decide which attractiveness criteria to use and how important each one is to the organization. While knock-out criteria automatically eliminate some segments, attractiveness criteria are chosen by the team and then used to decide how to apply each segment.

**4.2 Knock-Out Criteria:**

Knock-out criteria are used to decide if market segments from segmentation analysis qualify for further assessment using attractiveness criteria. The first set of knock-out criteria includes factors like similarity, uniqueness, size, alignment with the organization's strengths, identifiability, and accessibility. These criteria must be met for a segment to be considered. It's important for senior management, the segmentation team, and the advisory committee to understand these criteria. While most criteria are straightforward, some, like segment size, may need specific details clarified.

**4.3 Attractiveness Criteria:**

Unlike knock-out criteria which are yes or no, attractiveness criteria are rated on a scale. Each market segment is evaluated based on how well it meets each criterion. The overall attractiveness, considering all criteria, determines if a segment is chosen as a target in Step 8 of segmentation analysis.

**4.4 Implementing a Structured Process:**

The most common way to evaluate market segments for targeting is using a segment evaluation plot, which shows segment attractiveness and organizational competitiveness. These values are decided by the segmentation team because there's no universal set of criteria for all organizations. The factors for both attractiveness and competitiveness need to be discussed. It's recommended to use not more than six factors, determined by a team and possibly modified by an advisory committee. Including representatives from different parts of the organization is important because each unit has a different viewpoint and will be affected by the segmentation strategy. While the segment evaluation plot can't be completed in this step since there are no segments yet, selecting attractiveness criteria early ensures all relevant information is collected and makes selecting a target segment easier.

**STEP 3: Collecting Data:**

**5.1 Segmentation Variables:**

Commonsense segmentation means we use basic ideas to split the market. For example, if we're talking about vacations, we might just divide people by gender, like separating men and women. Then, we look at other details about these groups, like their age or what they like about vacations.

In data-driven segmentation instead of using one thing like gender, we use different data points to create groups we look at gender, age, how many vacations people take, and what they like or don't like about vacations. This helps us find more specific groups that might be interested in what we are selling. Both ways are useful, but data-driven segmentation gives us more detailed groups to work.

Whether we're using basic ideas or complex data analysis, having good quality data is important. It helps us put people in the right groups and describe those groups accurately. This is important because it helps us modify products, decide on prices, choose the best ways to sell.

To gather this data, we might use surveys. However, surveys may not always give us the most accurate picture, especially when people want to appear socially desirable. So, it's important to consider various sources and pick the one that best reflects real consumer behaviour.

**5.2 Segmentation Criteria:**

Before identifying segments or collecting data, organizations must decide which segmentation criterion to use. This refers to the type of information used for segmentation, like geographic, demographic, psychographic, or behavioural factors.

**5.2.1 Geographic Segmentation:**

Geographic information is the first way to divide up markets, where people's location determines their segment. If a country wants to attract tourists from nearby countries, it needs to consider language differences. Companies like Amazon and IKEA adjust their offerings based on where customers live, showing how geographic segmentation works in practice

The advantage of this approach is that it's easy to target communication to specific areas, like using local newspapers or TV stations. This is just because people live in the same area doesn't mean they have similar preferences.

**5.2.2 Socio Demographic Segmentation:**

Socio-demographic segmentation divides people based on factors like age, gender, income, and education, which can be very useful in certain industries. For example, luxury goods are often associated with high income, while cosmetics are often targeted by gender. However, these criteria may not always explain why people prefer certain products. While they can help identify segments easily.

**5.2.3 Psychographic Segmentation:**

When people are grouped based on their beliefs, interests, preferences, aspirations, or reasons for buying a product, it's called psychographic segmentation. This approach aims to understand the mindset of consumers. Benefit segmentation, focusing on what benefits people seek, and lifestyle segmentation, based on their activities and opinions, are common types of psychographic segmentation. Unlike simpler criteria like geography or demographics, psychographic factors are more complex and often require multiple variables to capture accurately.

**5.2.4 Behavioural Segmentation:**

Another way to segment consumers is by looking at their behaviour or actions. This includes things like how they buy a product, how much they spend each time, or how they search for information. Studies have shown that using behaviour as a segmentation criterion can be more effective than using geographic factors. The advantage of this approach is that it focuses on real actions rather than what people say they do, leading to more accurate segmentation.

**5.3 Data from Survey Studies:**

Survey data is cheap and easy to collect, making it a possible approach for any organisation.

**5.3.1 Choice of variables:**

Choosing the right variables based on common sense or data analysis, is important for correct survey. When we're using data, we need to include every factor that related to the groups we're trying to understand. Adding unnecessary stuff can make surveys long people won’t get correct answers. We need to be careful when collecting data and make sure we're not gathering useless information. Including too many unnecessary factors can mess up the whole process. They're called "noisy" or "masking" variables.

To make a good survey, it's helpful to do some exploratory research first, to get what's important. Then we can ask the right questions to get the best answers. Mixing both types of research makes sure we're not missing anything important.

**5.3.2 Response Options:**

When people can only choose between two options, like yes or no, it gives us what's called binary data. We represent these choices with 0s and 1s in a dataset. It's straightforward because there's a clear difference between the two choices, which makes it easy for us to analyse later.

When people can choose from a bunch of options without any specific order, like picking their occupation from a list, it's called nominal data. We can turn this into binary data by assigning a 0 or 1 to each option. If people can give a number as an answer, like their age or how many nights they stayed at a hotel, it's metric data. Most surveys ask people to rate something on a scale, like from strongly agree to strongly disagree. This gives us what's called ordinal data. We should give people either binary or metric options because it makes the analysis easier later.

**5.3.3 Response Style:**

When people answer survey, they might not always give honest or accurate answers. One common bias is called a response bias, where people tend to answer questions based on something other than what the question is actually asking. This can happen consistently over time, and it's called a response style

There are different types of response styles, like always picking extreme answers (like strongly agree or strongly disagree), always picking the middle option, or just agreeing with everything. So, when collecting data for market segmentation, it's really important to try to minimize these biases, we need to do extra analysis to make sure we're not being misled. Sometimes, we might even need to remove people from the data who seem to be answering based on a bias.

**5.3.4 Sample Size:**

When we do statistical analyses, we usually have recommendations for how big our sample should be. But when it comes to market segmentation, it's not as clear.

Imagine you're trying to group customers based on just two things, like price and quality. If you don't have enough customers in your sample, it's really hard to figure out how many different groups there are. But if you have plenty of customers, it's much easier to see the different groups.

But it's not just about having a big sample size. The quality of the data matters too. If there are problems with the way the questions are asked or if people aren't being honest, it can mess up the analysis.

Overall, to get accurate results from market segmentation, you need to make sure you have enough customers in your sample, each thing you're using to group them is well represented, and the data is good quality.

**5.4 Data from Internal Sources:**

Another way to gather data for market segmentation is through experiments. These experiments can happen either in real-life settings or in a controlled laboratory environment. For example, researchers might test how people react to different advertisements. The responses to these ads could then be used to group people into segments. For instance, people are presented with different product options that vary in specific features. They then choose which product they prefer based on these features.

**5.5 Data from Experimental Studies:**

Experimental data, such as from tests or studies, can be used to understand market segments better. These tests can be done in real-life situations or in controlled environments like labs. This reaction could be used to group people into segments based on their preferences. This knowledge can then help in dividing the market into segments.

**STEP 7: DESCRIBING SEGMENTS:**

**9.1 Developing a Complete Picture of Market Segments:**

Segment profiling is all about understanding how different groups of customers are different from one another. Describing Segments is similar to profiling, but instead of using the same variables, additional information about the customers in each segment is considered.

Profiling would involve looking at the differences between segments based on their travel motives. Describing the segments would involve considering things like their age, gender, past travel behaviour, and other factors that can help paint a fuller picture of each group.

We can analyse the differences between segments using either simple descriptions or more complex statistical methods. Visualizations can make this information easier to understand compared to tables and numbers.

**9.2 Using Visualisations to Describe Market Segment:**

There are many types of charts that can show the differences between descriptor variables. We'll focus on two basic methods that work well for different types of variables are for nominal and ordinal descriptor variables (such as gender, level of education, country of origin), or metric descriptor variables (such as age, number of nights at the tourist destinations, money spent on accommodation).

Using graphs to describe market segments has two big advantages. First, it makes it easier for both the person analyzing the data and the people using it to understand the results. Second, it includes information about whether the differences shown are actually important, which helps prevent reading too much into small differences.

**9.2.1 Nominal and Ordinal Descriptor Variables:**

When we want to describe the differences between different groups of customers based on one characteristic, like gender or education level, we start by comparing how many people belong to each group within each segment.

To do this, we put all the data into a table where one side shows the segments and the other shows the characteristic we're interested in. When we look table and see if there are any big differences between segments.

Another way to show this information is with a mosaic plot. This kind of plot not only shows the differences between segments but also gives an idea of how big each segment is overall.

Mosaic plots can also help us see if there's a relationship between more than two characteristics, like age, gender, and education level. They can even show if the differences we see are likely to happen just by chance or if they're significant. We do this by using colors to highlight where the observed differences are bigger or smaller than we'd expect if the characteristics were independent.

**9.2.2 Metric Descriptor Variables:**

Here "conditional" means that the plots are split into sections, with each section showing the results for a specific part of the data, like different groups of customers. These plots are great for seeing how different groups compare using numbers like age.

To make the plots easier to understand, we can label each section with the name of the group instead of just a number. We do this by creating a new label that combines the word "Segment" with the segment numbers. Then, we make a histogram for each group showing the age distribution.

By using a parallel box-and-whisker plot, we can see the spread of ages for each group separately. This helps us compare the age distributions between different segments.

We can also use a modified version of a plot called the segment level stability across solutions (SLSA) plot. This plot helps us track how a certain number changes over different segmentation solutions. In this modified version, we use different colors to show additional information about the variable being measured.

**9.3 Testing for Segment Differences in Descriptor Variables:**

Simple statistical tests can help us check if there are differences between groups of customers in certain characteristics. One straightforward method is to run separate tests for each characteristic we're interested in. When we divide customers into groups, like market segments, we can treat these groups like any other category, such as male or female.

We can use a test called the χ2-test to see if there's a connection between being in a certain segment and another category, like gender or education level. If the p-value from this test is small, usually less than 0.05, it suggests there are significant differences between segments.

We look at numerical characteristics, like age or how much money someone spends, we might use parallel boxplots to compare the segments. Another popular method is Analysis of Variance (ANOVA), which helps us see if there are significant differences in the average values across segments.

When we are comparing the average level of moral obligation to protect the environment among different segments, ANOVA can tell us if there are significant differences. If the p-value from ANOVA is small, it means at least two segments have different average levels of moral obligation.

To make sure we're not making too many mistakes when interpreting these tests, we need to adjust the p-values for multiple comparisons. One common method is Bonferroni correction, but there are other less conservative methods available, like the one proposed by Holm. These adjustments help us maintain a reasonable level of confidence in our results.

**9.4 Predicting Segments from Descriptor Variable:**

Understanding market segments better, we can use a method where we predict which group a customer belongs to based on certain details about them. We do this by using a model called regression. This model compares the group a customer is in (which we call the dependent variable) to different details about them (which we call independent variables). By using techniques from statistics and machine learning, we can figure out how accurately we can predict which group a customer belongs to based on these details. We can also learn which details are most important for figuring out the customer's group. Regression analysis is a key part of this process.

Regression analysis works by assuming that we can predict one thing (let's call it "y") based on other things (let's call them "x1," "x2," and so on). In the simplest form, known as linear regression, it's like drawing a straight line through some points on a graph. But sometimes, the thing we're trying to predict doesn't follow a normal pattern, like when it's categorical (for example, just yes or no). We might use a different distribution called the Bernoulli distribution. And because these distributions might not work with a simple straight line, we use something called a link function to help us make predictions. We can even use this for classification, like sorting customers into different groups based on their characteristics.

**9.4.1 Binary Logistic Regression:**

Binary logistic regression is a statistical method used to model the relationship between one or more independent variables (predictors) and a binary outcome variable (response), where the outcome variable has only two possible outcomes, typically coded as 0 and 1.

The goal of logistic regression is to model the probability that an observation falls into one of the two categories based on the values of the predictor variables. This is done using the logistic function (also called the sigmoid function), which maps any real-valued number to the range [0, 1].

The coefficients (also called weights or parameters) of the logistic regression model are estimated using a process called maximum likelihood estimation. This process finds the values of the coefficients that maximize the likelihood of observing the given data under the assumed logistic regression model

Once the coefficients are estimated, we can use the logistic function to predict the probability of an observation belonging to one of the two categories. If the predicted probability is greater than a chosen threshold (commonly 0.5), the observation is classified as belonging to the category with a value of 1; otherwise, it's classified as belonging to the category with a value of 0.

Logistic regression models are evaluated using various metrics such as accuracy, precision, recall, and the receiver operating characteristic (ROC) curve.

**9.4.2 Multinomial Logistic Regression:**

Multinomial logistic regression is an extension of binary logistic regression that allows for modelling outcomes with more than two categories. It's used when the dependent variable (response variable) has three or more unordered categories.

Multinomial logistic regression uses the concept of logits, which are the natural logarithm of the odds. Instead of modelling the probability directly, it models the log-odds of each category relative to a reference category.

The coefficients of the multinomial logistic regression model are estimated using maximum likelihood estimation, similar to binary logistic regression. The goal is to find the values of the coefficients that maximize the likelihood of observing the given data under the assumed model.

Once the coefficients are estimated, we can use them to predict the probabilities of each category for a given observation. The category with the highest predicted probability is then assigned as the predicted category for that observation.

Multinomial logistic regression models are evaluated using appropriate metrics depending on the specific problem, such as accuracy, confusion matrix, and classification report.

**9.4.3 Tree-Based Methods:**

Classification and regression trees (CARTs) are a way to predict outcomes like "yes" or "no" or categories, using information about other things. They're part of machine learning, where computers learn from data. CARTs have some benefits: they help pick out important info, they're easy to understand because you can see them visually, and they can handle lots of info. But sometimes, they can be a bit unpredictable. A small change in the data can make a completely different tree.

First, we have to split the data into groups based on one thing at a time. When we want each group to be as similar as possible when it comes to the outcome. Ideally, everyone in a group has the same outcome. This step-by-step splitting is called "recursive partitioning".

The starting point is the "root" node, which includes everyone. The ends of the branches are "terminal" nodes, where the tree stops splitting. To predict what group someone belongs to, you follow the branches down the tree based on their info until you reach a terminal node.

Different methods for making the tree can vary in a few ways:

* We can split the data into two groups at each step, or more.
* We can use different rules for picking which info to split on next.
* We can use different rules for deciding where to split.
* We can have different rules for when to stop splitting.
* We can handle the final predictions at the end of the branches.