INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NPTEL

NPTEL ONLINE CERTIFICATION COURSE

Marketing Research

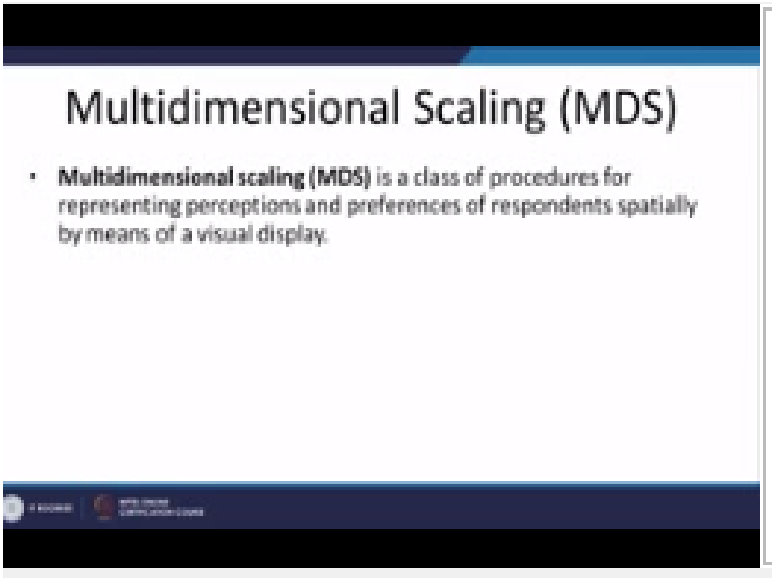
Lec – 39
Multi Dimensional Scaling

Dr. Jogendra Kumar Nayak
Department of Management Studies
Indian Institute of Technology Roorkee

Welcome to everyone to the session of marketing research and analysis. Today we will be discussing about a very important topic which is called multidimensional scaling. What is multidimensional scaling and why it is so useful mostly in the field of marketing is because it helps you in several ways. For example, to identify the image of a product, the position of a product in the minds of the consumer.

For example, to identify gaps in the market so that a company can may be fit in a new product or new brand in the space where there is a gap. So there are several uses which we will see later on, but multidimensional scaling is also referred as perceptual mapping. So it is something like the perception of individuals being mapped in the space so that we can see what do they feel about the various brands with the marketer wants to know about. So let us see how this is defined.

(Refer Slide Time: 01:32)



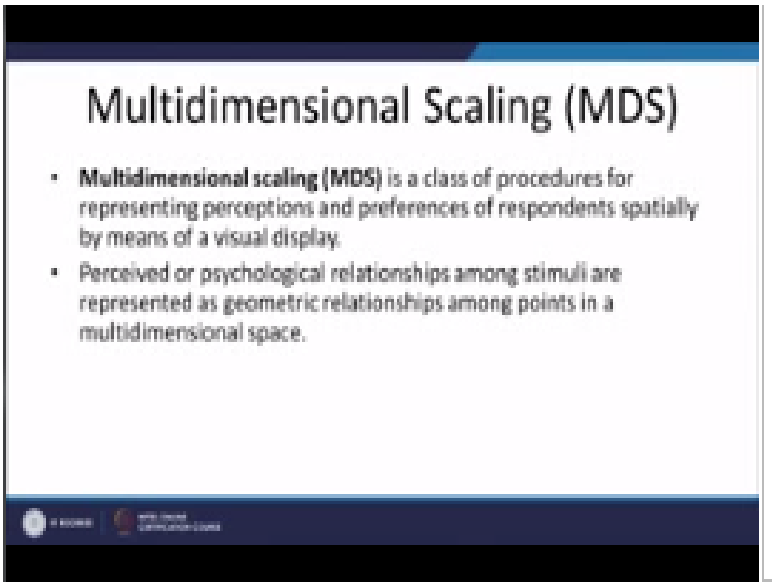
Multidimensional scaling is defined as a class of procedures for representing perceptions and preferences of respondents spatially by means of a visual display. So it represents the perceptions of people and the preferences of people. Now many times it has, we have seen that people might prefer or said something that they like a particular product or surveys, but at the end buy something else, so why is this happening? What is the difference? You like something but you don't purchase maybe that product.

So in order to understand these perception and preferences of people multidimensional scaling comes as a very, very handy tool. So let me give an example. For example let us say a company, automobile company for example, let us say on fuel and let us say on design. You could take any dimension. So maximum generally we talk about three dimensions because that's the possibility more than that it become highly complex, so on basis of this suppose you place different brands of cars.

Let us say Alto, this is just hypothetical I am placing let us say Mercedes Benz, for example this is let us say some other car, Honda City, so when the company is trying to well it takes the opinion of perception of people and places the products in the map in the 2D space, then there is a possibility to understand whether there is some kind of a gap, maybe the company finds, obviously this is the positive, negative, positive, negative.

So maybe this is the area were nothing is there. So the company can maybe target this place, or similarly the company can also find out some other space, other points in the space, so that it can understand whether there is any gap so that a new product could be placed out there.
So how do you do it is the question.

(Refer Slide Time: 03:42)

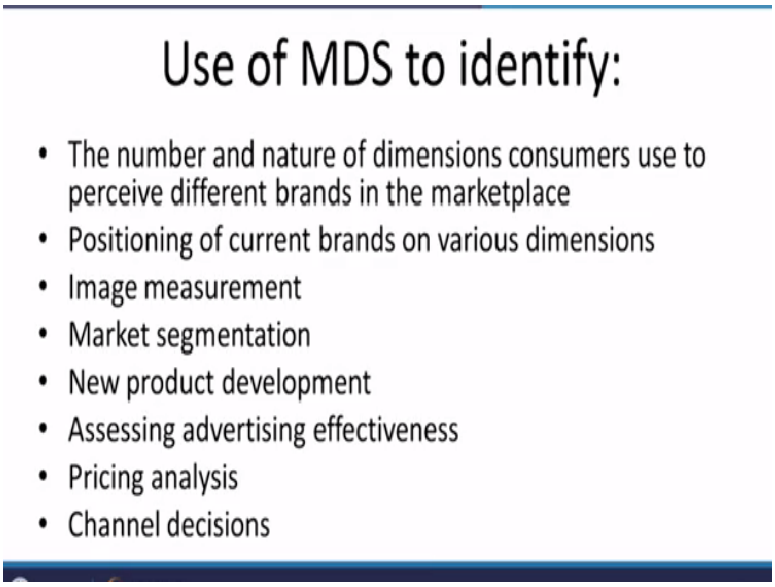


So let us see percept or psychological relationships among stimuli are represented as a geometric relationship among points in a multidimensional space. These geometric representations are often called spatial maps that is what I was trying to say. These are called spatial maps. The axis of the spatial map are assumed to denote the psychological bases or underlying dimensions, respondents use to form perception and preferences for stimuli. Another case I told you, design and fuel.

Suppose I would ask a large number of respondents and ask them which are the cars that they feel are very similar in nature. I am not asking them on what dimensions, but I am asking them, only requesting them one thing, they should not be comparing, you know finding similarities or dissimilarities on basis of too many different ways. So somewhere I am trying to restrict them on 2 or 3 dimensions only or factors only but not more than that.

So but I am not saying which dimensions are there, they are free do that. So when I am doing it automatically I will get a large number of similarities or dissimilarities, automatically what happens at the end of the day the respondent from the different similarities, we can understand on what basis, what parameter did these consumers these respondents make the similarity or dissimilarity choices. So these attributes are basically what is called as dimensions.

(Refer Slide Time: 05:21)



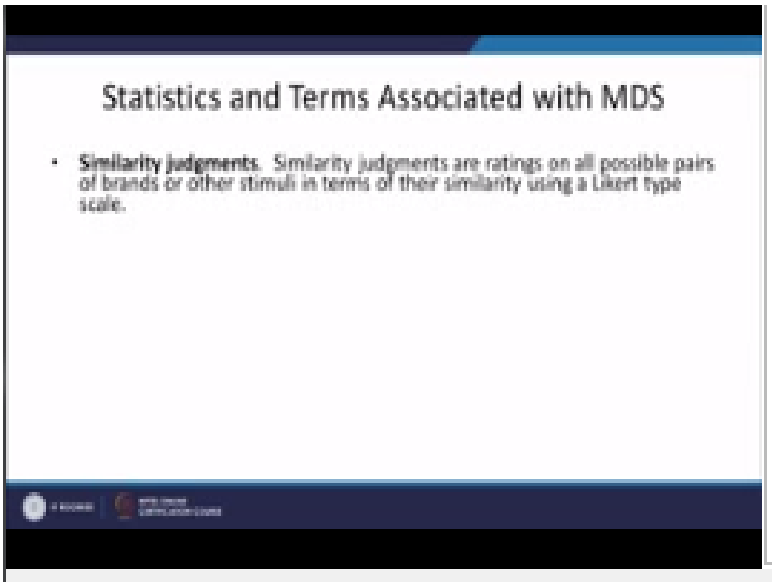
So the use of MDS is to identify as it is said you can see, the number and nature of dimensions consumers use to perceive different brands in the market place. So what are the dimensions, could be for example, in the car case, design, price, or it could be safety, could be enough technology, so various dimensions that they could be using.

Similarly next is what is the positioning of the current brands, so let us say the current brand is Honda City we are talking about. So what is the current position of the brand? So if the company can know, what is the current position of the brand then automatically if it is good they can be rest happy, but if there is something else, or they can define there is some mistake or there is some deficient, in that case those can improved.

Similarly MDS is also used to understand the image of various products in the market, to segment the market obviously when you identify the different target roots it helps to segment the market. New product development as I said, you identify the gaps and on basis of this gaps you can pitch in the product in the market.

To access the advertising effectiveness suppose a company has placed an ad and he wants to know how effective the ad is. The impact of pricing, analysis of pricing and the channel decisions. So these are all the different kinds of utilities, some of the utilities that MDS helps us as to the marketers basically. So what are the terms associated with MDS?

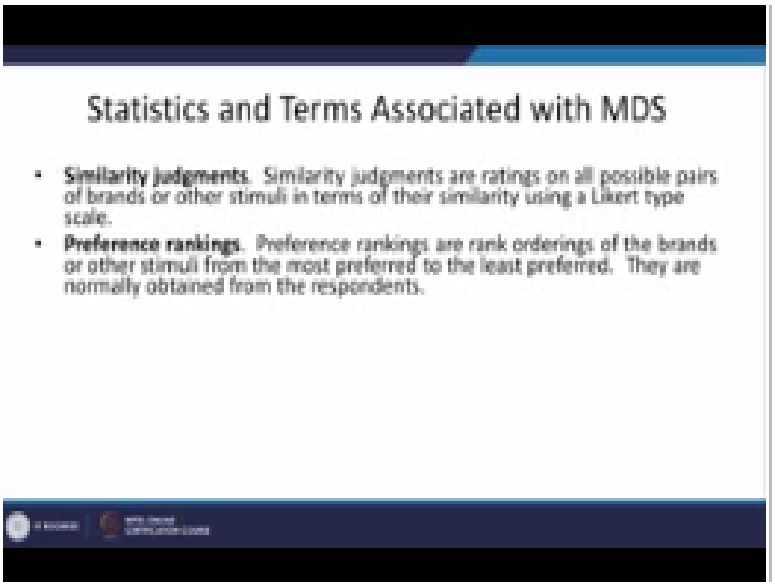
(Refer Slide Time: 06:58)



First it is called a similarity matrix, similarity or dissimilarity. So basically what we use is similarity judgements are ratings on all possible pairs of brands or other stimuli in terms of the similarity using a Likert scale. Let us say it looks like a co relation chart, A, B, C and D, so here what I am saying A, B, C and D so we are making similarities, we are trying to compare between the different various brands on basis of similarity.

If they are very similar let us say we will say it is one, thus if they are highly dissimilar then the rating could be 6, 7, 5, 4, whatever as per the scalier used. So preference rankings.

(Refer Slide Time: 07:50)



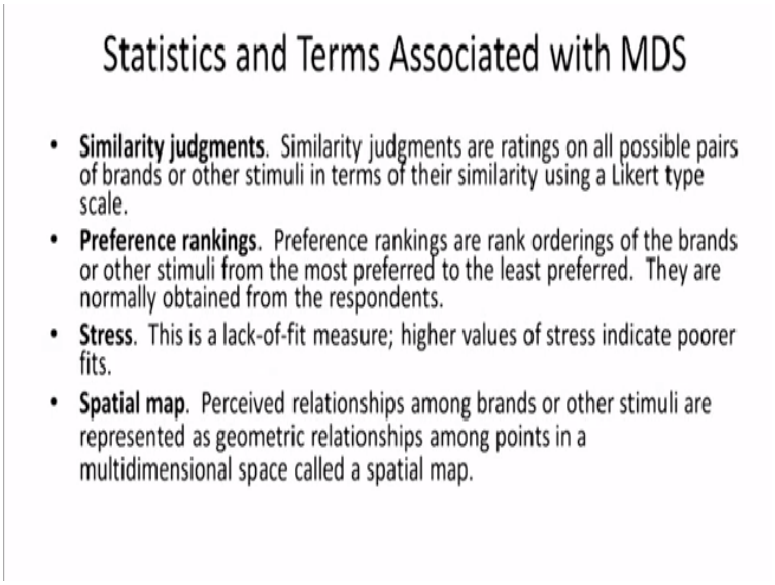
Similarity and then preference. Preference rankings are rank orderings, now in a scale you could say well, the similarity between two brands are like 5 or 1 or 7 but preference means in a set of products, let us say there are 15 products, or 18 products. Which one is the number one rank? What is the aspect of preference of the consumer and why it is important is that sometimes consumers on basis of similarity they provide similar rankings to two different products, but when it comes to preference it entirely changes.

So two products might look, for example Colgate and Pepsodent might look very similar to on basis of similarity, but on preference one of them might be wining easily over the other or may be Dabur or something like you know some ayurvedic stuff could be preferred over the Colgate or Pepsodent, could be possible.

So the question is similarity and preference are different things. Stress, this is the lack of fitness, higher value of stress indicate poor fits, now what is the stress, let me explain through numerical term, stress is nothing but it is calculated as stress is equal to distance, original distance minus d perceived divided by d average or d mean.

So it is a ratio of the original value, original distance. Now what is this distance we are talking about, the distance in the map or the spatial map. This distance is nothing but the equity distance. So we know how to calculate the equity distance that x2-x1 +y2-y1 square root over. So when we are talking about this value if d0-dp or d0 and dp, d original and d perceived are very close then what happens is that the stress value become low. And lower the stress value that means the original and the perceived are similar. Spatial map.

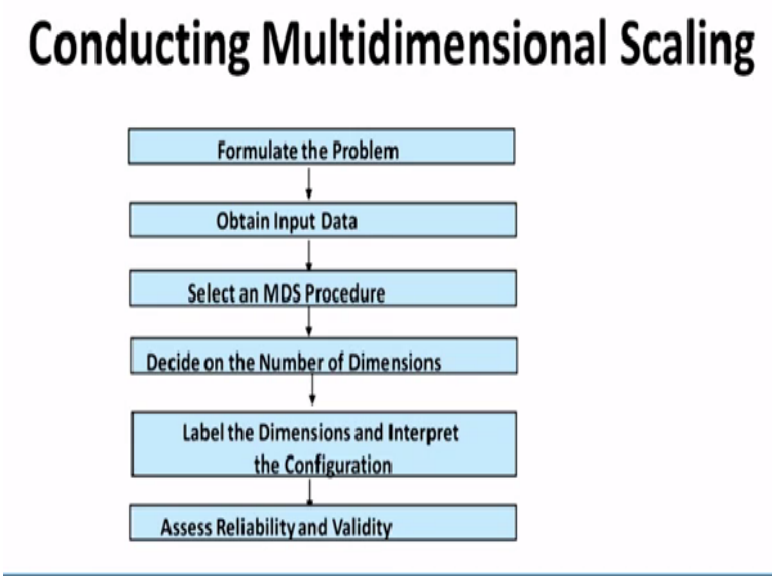
(Refer Slide Time: 10:02)



The perceived relationship among brands or other stimuli are represented as geometric relationships among points in a multidimensional space called a spatial map. And coordinates indicate the positioning of a brand are a stimulus. So with the coordinates tells us why Honda City is here as per its coordinates. Mercedes is here because of its coordinates. How do you conduct the multidimensional scaling?

First, to start, begin with is to formulate the problem. So what is your problem? Let us say your problem is to identify the different brands of snacks available in the market and check whether there is a scope for introducing a new snack.

(Refer Slide Time: 10:49)

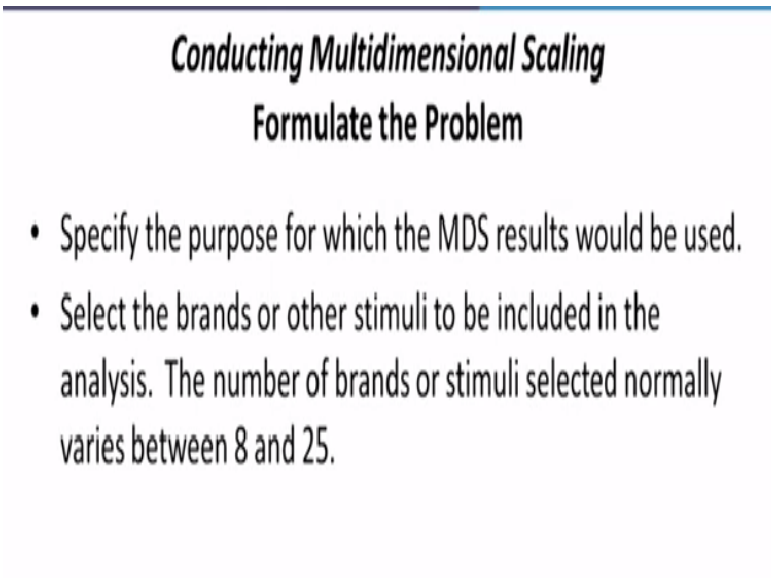


Obtain the input data, now input data is very, very important especially in this case. The input data you have to bring in is like asking people the similarities of among the different brands available. Select an MDS procedure, now what kind of approaches you are going to use we are going to see that decide the number of dimensions, how many dimensions you want to use 2, 3 should not be more than 3, as I said it becomes complicated.

Then label the dimensions and interpret the configuration you have to find out what dimensions people had used or respondents had used in order to create this similarity matrix. Finally to assess the reliability and validity. This is done as usual, if you remember we have been talking about having a set of samples, data set dividing into two parts, one is a treatment and other is the control and the other is the hold out group.

So you have the data, test the data on one set and then check on the whole out group sample test and check whether they are similar or dissimilar.

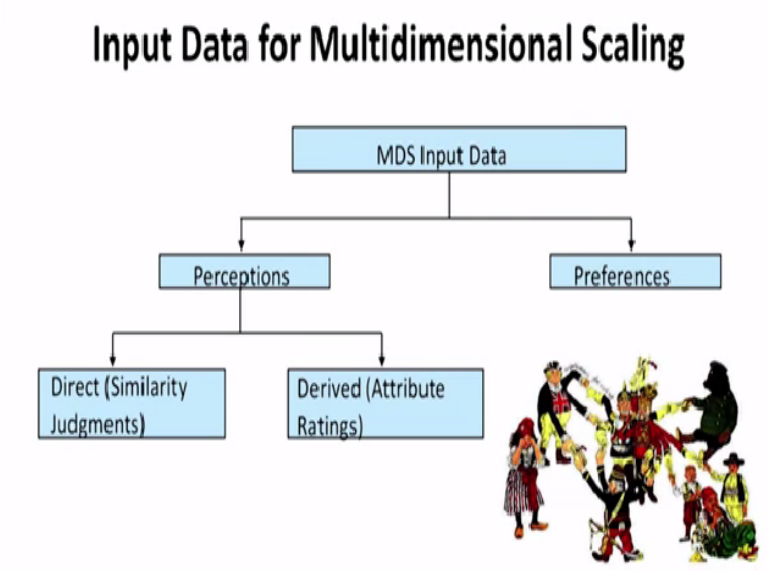
(Refer Slide Time: 12:01)



So specify the purpose of which the MDS as should be used. Select the brands or the stimuli, the brands or the stimuli to be included in the analysis. As it says if you can see the number of brands are stimulates selected normally varies between 8 and 25. So anything above 25 it is very, very difficult, because you understand, when the person has to give a preference he is comparing all the 25 brands, then he selects for the first one and then he selects among all 24 second one, third goes on it becomes tougher, it is very a tough challenge.

So too many is not very good. The choice of the number and specific brands are stimulate to be included should be based on the statement of the marketing research problem, and the judgement of the researcher. So what is the problem you are trying to study, so on basis of this problem you will decide what kind of a number we will have. So the input data now for the MDS scaling.

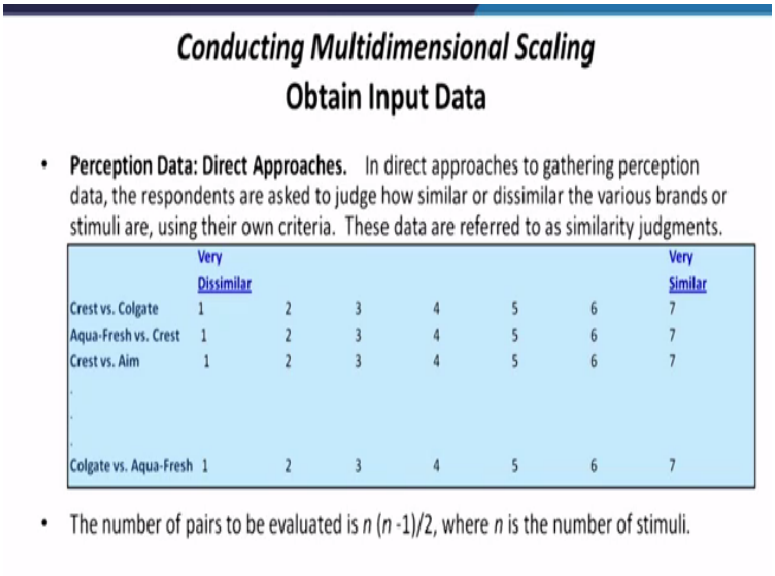
(Refer Slide Time: 13:01)



The input data is basically as you can see is based on perceptions or preferences. Now perceptions, as I said directly you can have a perception score by just asking people what is the similarity between the several branch, let us say Colgate, Hindustan, Pepsodent, let us say for example Dabur, Lal Dant Manjan, different brands of toothpaste I am talking about, Meswak. And then I am asking what and how they are similar, this is a direct, in a scale of may be 7 like scale you can say or something or the other way is derived attribute ratings.

Now attribute ratings by deriving. Now what I am doing is I am giving several attributes, now I will show you, there are several attributes could be let us say how it is cleaning the teeth, how it is, what kind of fragrance it is just having, how freshness it gives. So all these things are the different attributes, on base of this attribute the score the people give, I can derive a score. The next is the preferences, so preferences is the rank orders which we have already discussed.

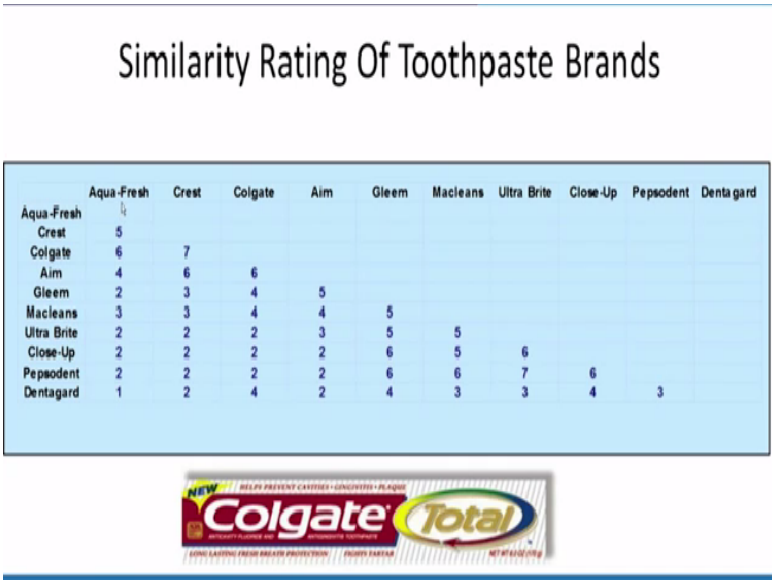
(Refer Slide Time: 14:12)



Now let us say, let us look at this perception data, now this is the direct approach, now as you can see, now Crest versus Colgate, Aqua Fresh versus Crest, Crest versus Aim. So very dissimilar is 1, very similar is suppose 7. Now it could be, you can use it up to your wish, but the way you like to. So we are saying highly similar is seven in this case and highly dissimilar is one. Sometimes you can also say highly similar is one and highly dissimilar is 7, it does not make much of a difference, and does not matter, but your understanding has to be clean.

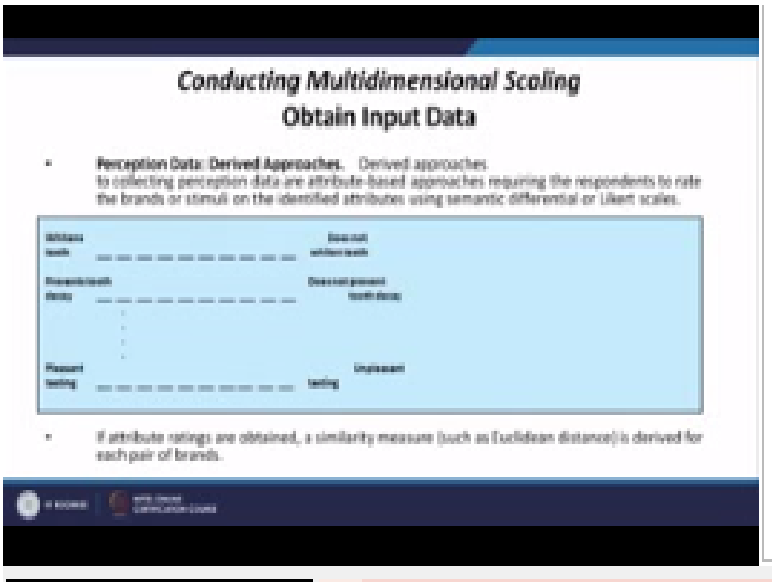
So the number of paste to be evaluated is obviously n into n-1/2, suppose there are 8 pairs so 8 x 7 n x n-1 / 2 = 28. So 28 combinations, so now just imagine if you have 25 as we said 8 to 25 so 25 x 24 /2 = 300 combinations. So that is a large number. So 300 being generally called in researchers as a magical number. So if you are doing data analysis or data collection also, if you achieve a number of 300 sample size is decent, study is decent.

(Refer Slide Time: 15:38)



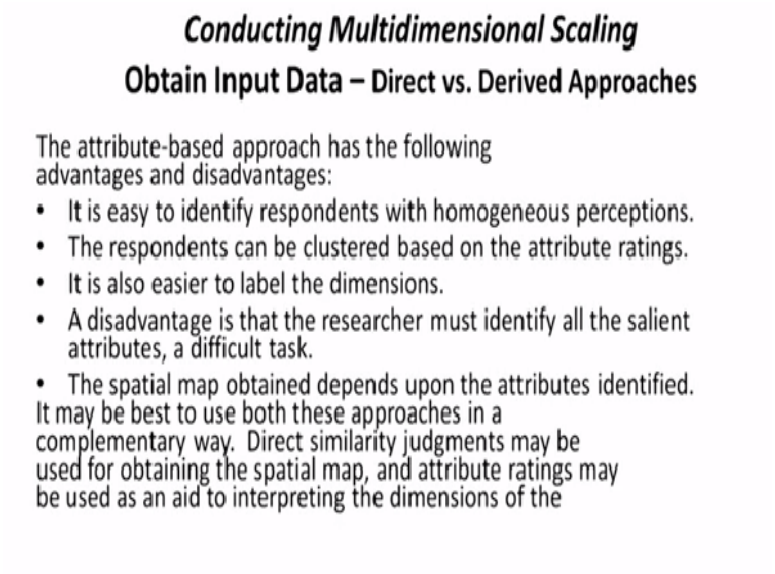
Similarly the rating, you can see the ratings have been marked here and it looks like a correlation chart. Now if you look at Aqua fresh to Aqua fresh it will be nothing, 0 because it is one, obviously one. Crest to crest, one. So it is be used in a lower triangular matrix or a upper triangular matrix. Now this is a lower triangular matrix, correlation matrix.

(Refer Slide Time: 16:00)



Now perception data, this is what I was talking about, the different perception attributes and once this data, if attributes are obtained the similarity measure is derived from each pair of brands. Now once the respondent on the basis of this attribute, once he collects the data then finally he can use this data as also for the input data. Now obtain input data direct versus derived approach. Now what is the difference?

(Refer Slide Time: 16:29)



If you have a direct it is easy. There is no doubt about it; the research does not have to identify a set of silent attributes. The direct approach has the following advantages, but the disadvantages are that the criteria are influenced with the brands or stimuli being evaluated. So this is disadvantage, the disadvantage is, that the criteria of selection is being influenced by the brands. So this is a disadvantage with the direct approach.

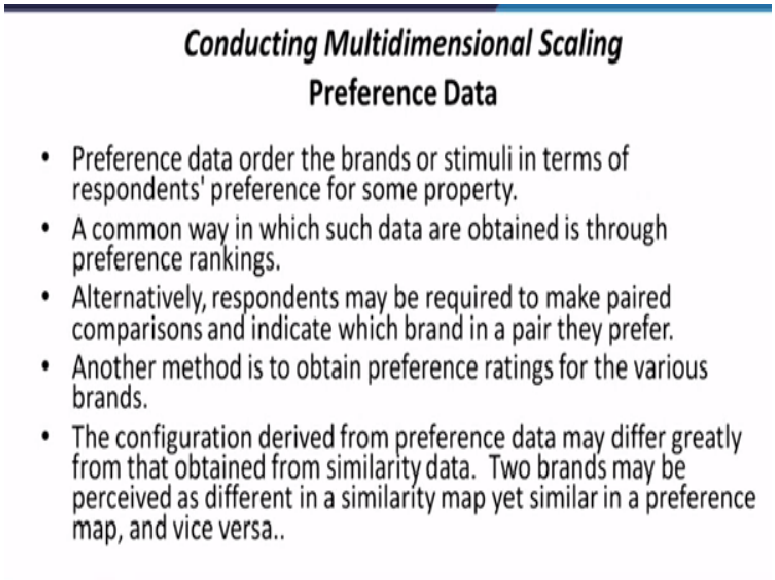
Further more it maybe difficult to label the dimensions of this spatial map. Because you cannot say on which dimension actually the person has evaluated so that is a clumsy thing. On the basis of the attribute based approach which was the second one we saw it is easy to identify respondents with homogeneous perceptions, obviously, people who have got a similar thought process can be easily grouped or identified.

The respondents can be clustered based on their attribute rating; obviously I just said that if similar scoring is being given then they can be clustered together. It is easier to label the dimensions because more or less the dimensions remain the same. The disadvantage is that the researcher must identify all the silent attributes which is a difficult task. Suppose you miss out some of the important attributes let us say as per teeth cleaning.

One of them could be like the amount of quantity let us say in the toothpaste. Now maybe that is an important attribute for people to select a brand. But now suppose your mistake then automatically the explanation power will go down, will drastically fall down. The spatial map obtained depends upon the attributes identified. It may be best to use both these approaches in a complimentary way, direct, similar judgements may be used for obtaining the spatial map and attribute ratings may be used in 8 to interpret the dimensions.

So what we are saying, let us club the both methods or use both the methods. The first method, direct method can be used for obtaining the spatial, that means this map. So to this map and the attribute rating can be used to identify the dimensions.

(Refer Slide Time: 18:54)



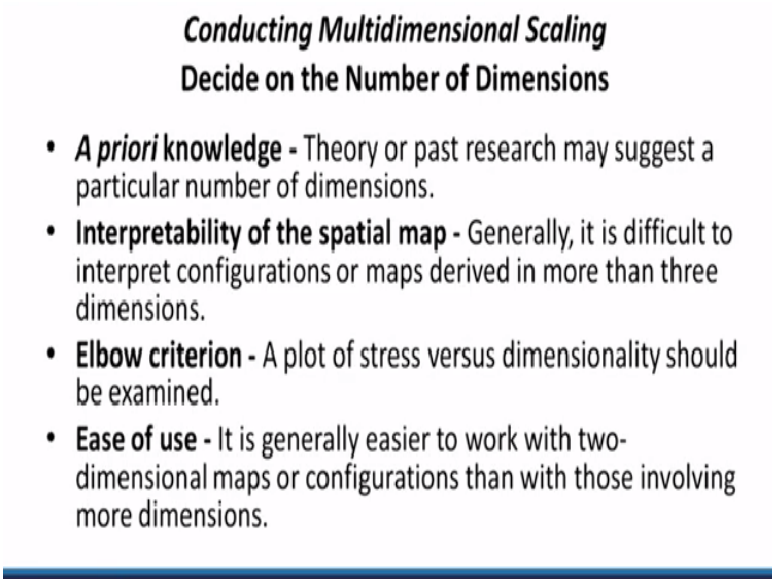
Preference data orders the brand as I said in terms of respondents preference for some property. So as I said the person could be respondent could be giving you a similarity of two different brands, but while you are waiting for a preference he is only selecting one of them and the second one which is highly similar he is keeping at a very low preference for some reason, we don't know.

A common being such data is through preference rankings. Alternatively respondents may be required to make paired comparisons and indicate which brand they prefer. Now for example, as I said Colgate versus Pepsodent, Colgate versus Crest, let us say Crest versus Dabur Lal Dant Manjan, so you have to make those combinations space. If you remember when we were talking about scale also we were making paired comparisons.

Paired comparisons, non paired comparisons, so we were comparing the two different brands. Another method is to obtain propagating for the various brands. The configuration derived from the preference data may differ greatly from that obtained from the similarity data. That is I have been repeating.

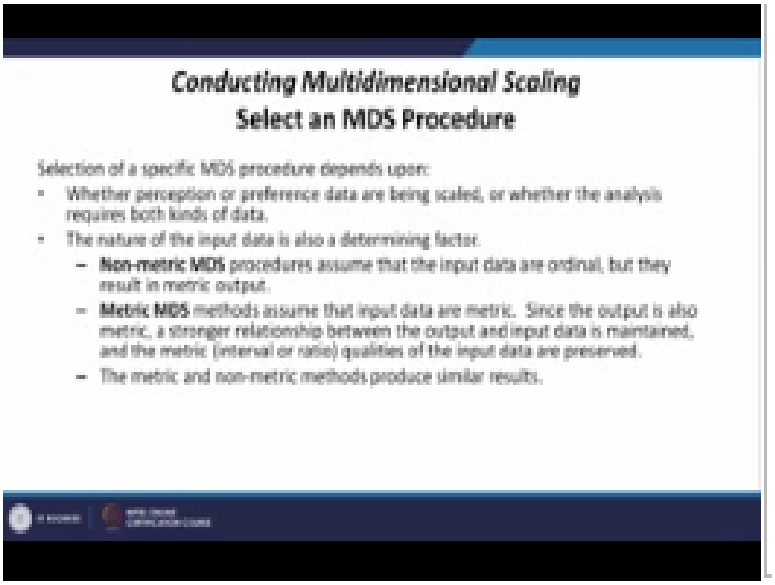
Two brands may be perceived as different in a similarity map x similar in a preference map. So people on basis of preference they like the same, two different products similarly, but in case of similarity matrix they are highly differently placed together, so that is the complexity.

(Refer Slide Time: 20:33)



Now how to conduct, selection of specific MDS depends upon whether perception or preference data are being scaled or whether the analysis requires both kinds of data. So we have seen sometimes perception based, sometimes preference based, similarity versus preference. So which one you are using, similarity we are using to find out spatial map and preference ratings are being used to find the dimensions.

(Refer Slide Time: 21:02)



The nature of the input data is a very important thing, although whether your data is metric or non-metric the results are not going to change much, but still if it is non-metric multidimensional scaling which is ordinal ranking basically it has seen that the input data are ordinal but the result is again metric output.

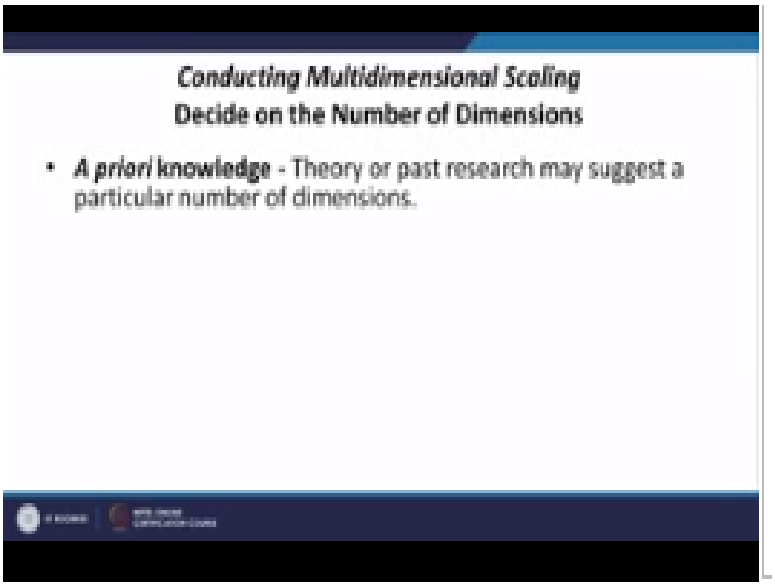
In the metric multidimensional scaling method the input data is metric and this is preferable. Why, because the output is going to be metric in nature, so since the output is also metric a stronger relationship between the output and the input data is maintained and the metric qualities of the input data preserved.

So that is the benefit. But as you can say at the end the metric and the non metric methods produce similar results. Another factor influencing the selection of the procedure is whether the MDS analysis will be conducted at the individual or the aggregate level. Now that is very important, now what does it mean? See you can conduct multidimensional scaling; the beauty of this method is that it can done for a individual case, even single respondent, but you can also do it for aggregate level.

That means you take the opinion of 300 respondents and take an average of it and use that as a score. So what are you using as a study that influences the process? So at the aggregate level it is assumed that all the respondents use the same dimensions? Now, individual level versus aggregate level. In the individual level you know what that the individual has, what kind of dimensions he has used to make the comparisons, but when you are using aggregate score of 300 or 200 or 100 people then there is a problem.

You don’t know which dimensions have they actually utilized to make those comparisons, so in such conditions it becomes slightly difficult to manage the aggregate part. But aggregate is obviously more preferable because it is giving a wide opinion in a result of too many people at one go, so that is more preferable.

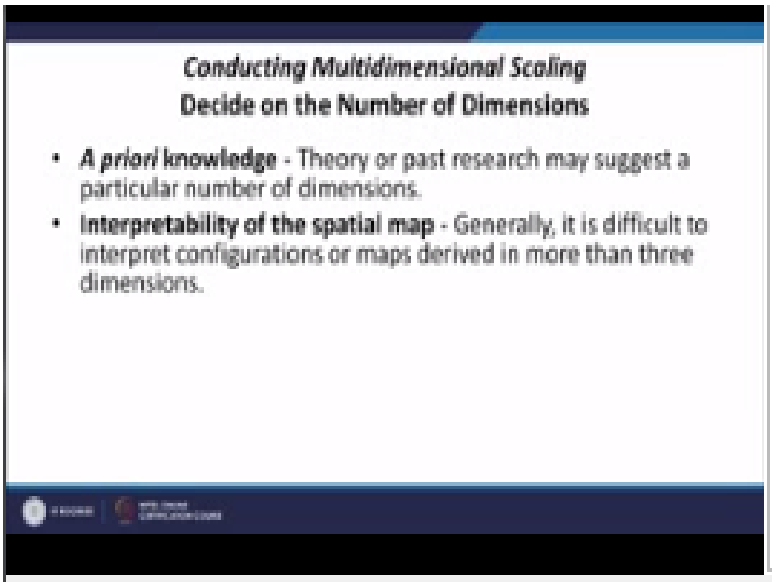
(Refer Slide Time: 23:18)



When you conduct the multidimensional scaling you have to decide on the number of dimensions as I said, but the number of dimensions should be limited to 2 or 3, not much. A priori knowledge, that means theory or past research may suggest you a particular number of dimensions.

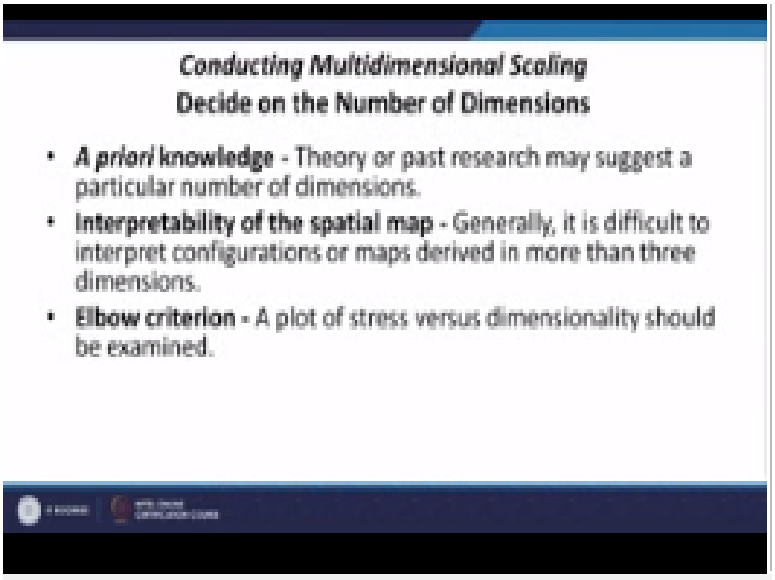
So if you say on what basis are the cars being compared, so you have some knowledge? Cars can be compared on the basis of the brand, from where they come, the price, the efficiency. So let us take 2 or 3 important ones.

(Refer Slide Time: 23:51)



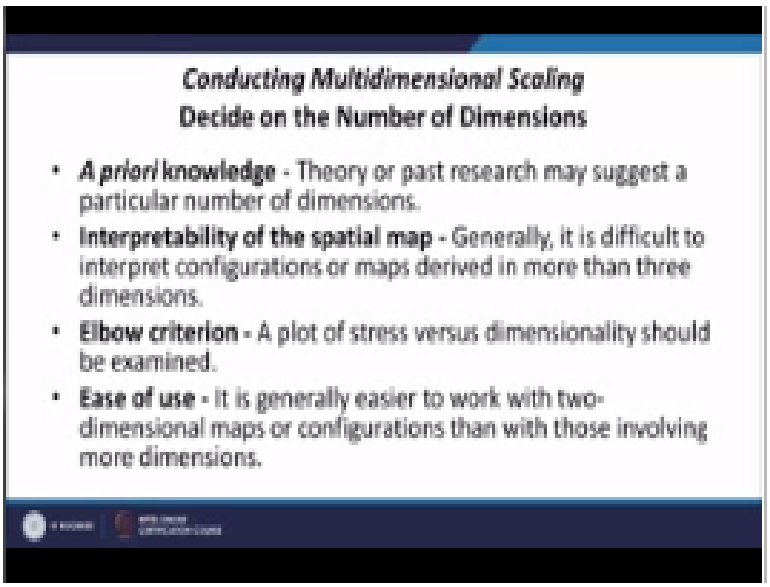
Generally interpretability of the spatial map, now to interpret the spatial map that is the perceptual map. Generally it is difficult to interpret maps derived in more than 3 dimensions that is what I was talking about. If you have more than 3 dimensions x, y, and z, then it is difficult to interpret.

(Refer Slide Time: 24:11)



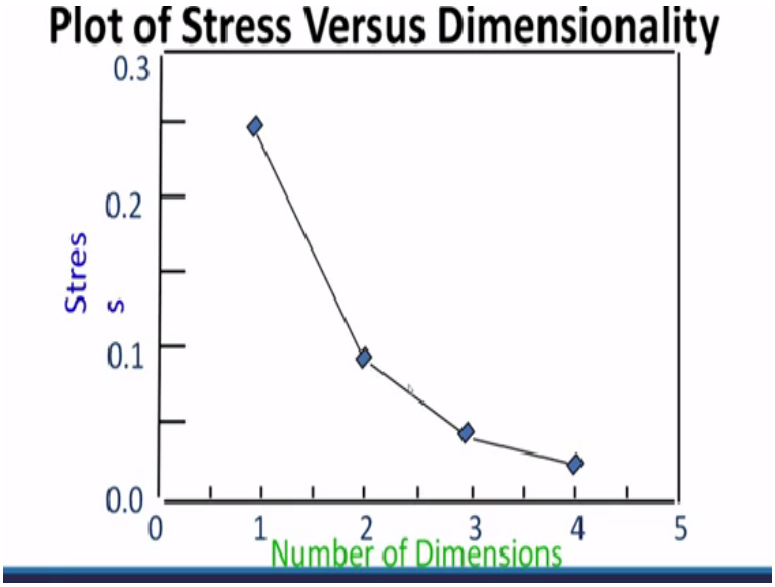
Elbow criteria, now this is if you remember factor analysis I explained, there is something called a scree plot, so a scree plot is similar to a scree plot where you find out the dimensions, number of dimensions that should be examined. I will show you in the next slide.

(Refer Slide Time: 24:26)



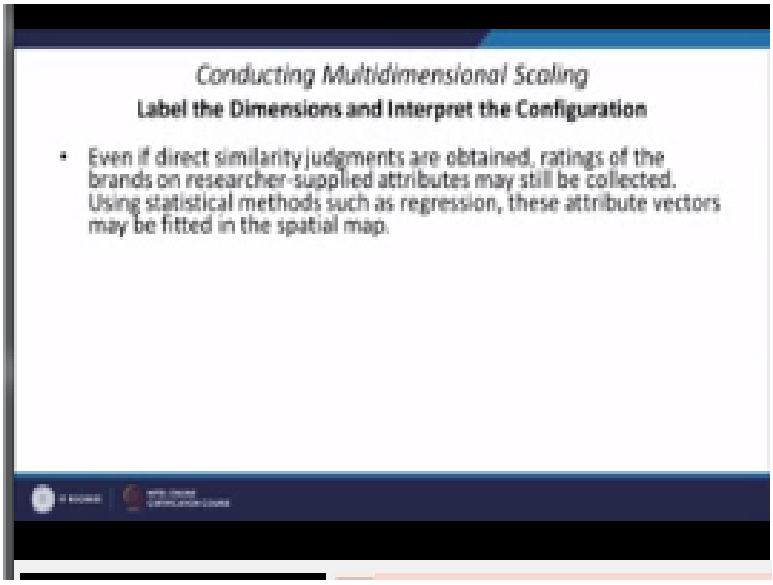
Ease of use, it is generally easier to work with 2 dimension maps then those involving more than 2 dimensions. So as you increase the dimensions the complexity increases, now this is what I was talking about.

(Refer Slide Time: 24:41)



A plot of stress, so if you remember, now if you look at the changes, now if you look at the change here. Now for example the change, the sharpness in the change, we will decide how many dimensions should be there. So in this case there could be 2 or maximum 3 dimensions. Now stress value, the higher the stress value the poor the fit that I explained also through this formula.

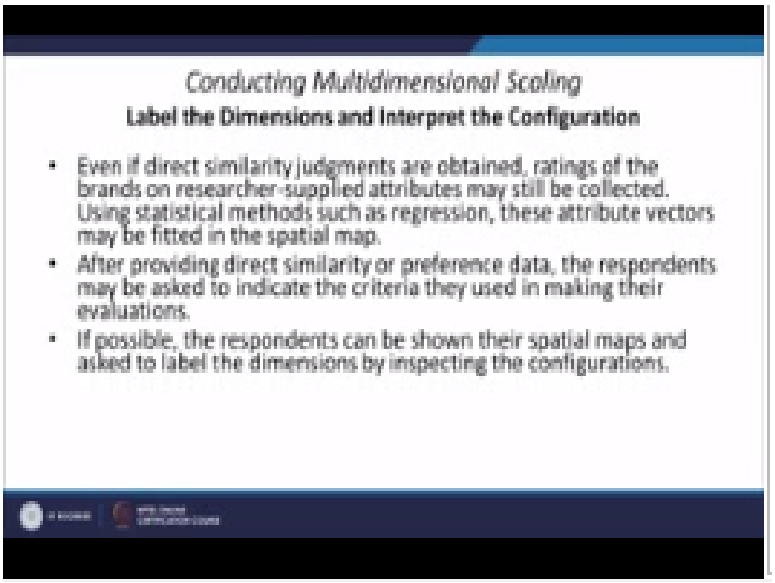
(Refer Slide Time: 25:11)



Now Finally coming to the labelling the dimensions and interpreting the configuration. Even if direct similarity judgements as obtained ratings of the brands on research supplied attribute must still be collected. Now what is it saying, even if direct similarity judgements are obtained from the people, respondents, ratings of the brands on researcher supplied attributes will still be collected using statistical method such as regression, so this is by using the beta ways for example in the regression, you use the beta ways.

Similarly you can calculate the beta ways here and those beta ways you can find out through this strategical methods what are the attributes which could be fitted on to this same spatial map which people are giving importance to.

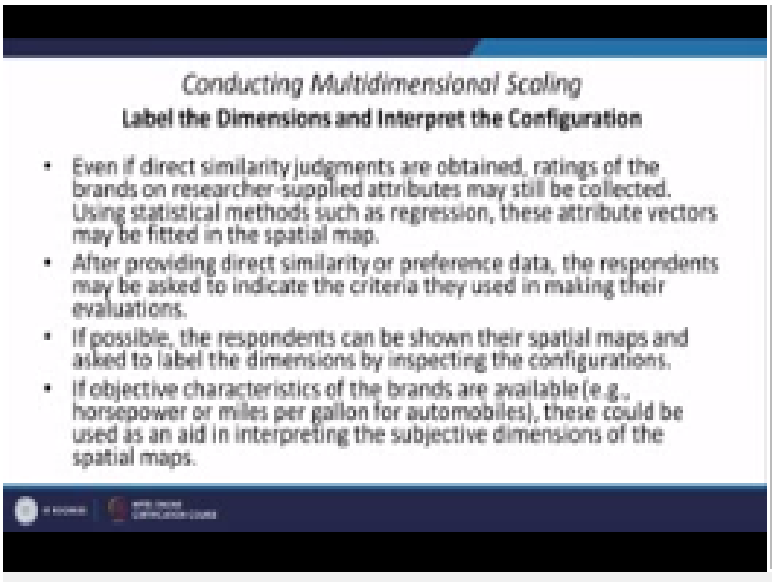
(Refer Slide Time: 26:05)



After providing direct similarity or preference data the respondents may be asked to indicate the criteria they used in making the evaluation. So once you have, after providing the similarity or data people have given to you then you can even ask the people on what basis did they make the evaluations, if possible respondents can be shown the spatial maps and asked to label the dimensions.

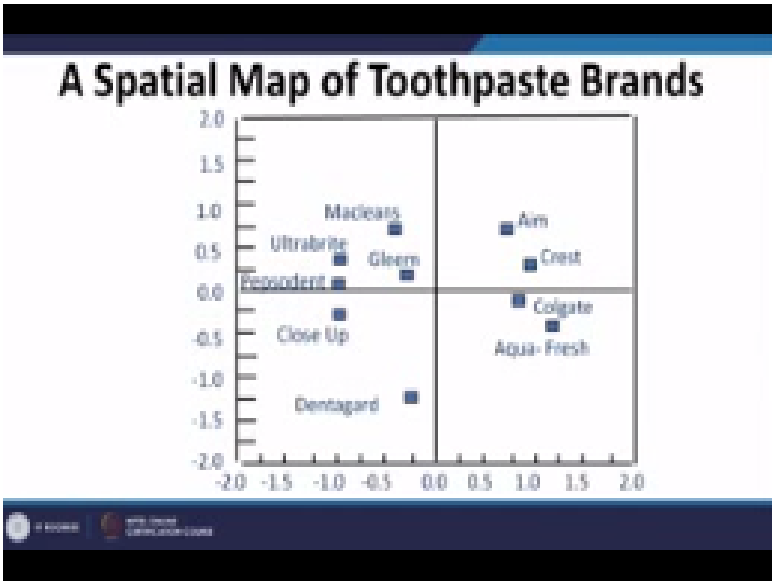
So instead of you doing it, you can even take the opinion of people and ask them please provide the dimensions that you have used to make this analysis.

(Refer Slide Time: 26:38)



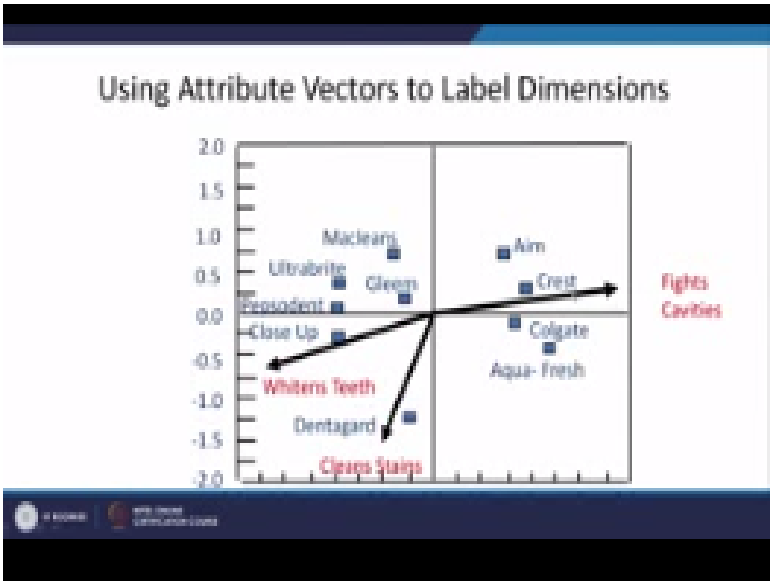
If objective characteristic of the brands are available these could be used as an aid in interpreting the subjective dimensions. Now this is very important, sometimes through our past research also we have some data available to us. We could use these data also to interpret the dimensions of the spatial maps.

(Refer Slide Time: 26:57)



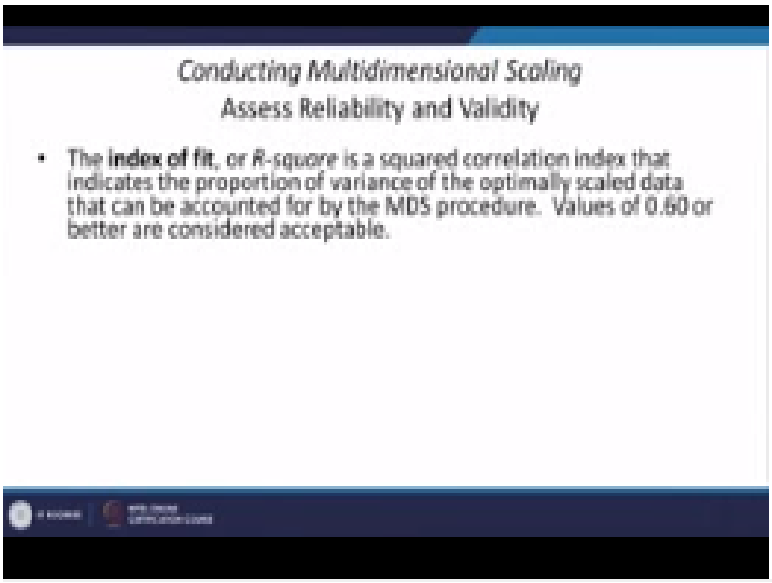
Now let us look at this case, now the spatial map of toothpaste brands. Now the difference, now this is a spatial map and if you look at the spatial map, if you look at this, the different, these are scores and if you look at Colgate, Agua Fresh, Crest, Aim, so it looks like this brand for example, Ultra Bright, Pepsodent is very different from Colgate. But as I said, although it is on similarity it is so different but on basis of preference it might not be, people might be purchasing it differently.

(Refer Slide Time: 27:40)



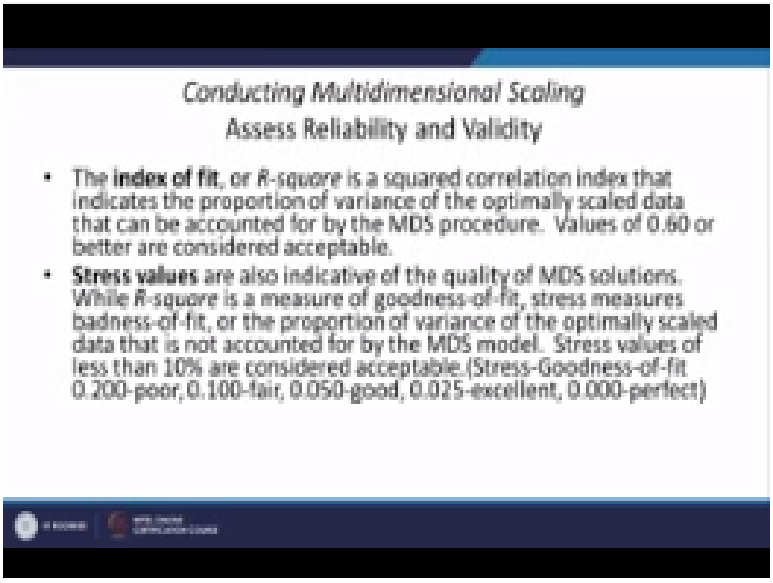
Now using the attribute vectors, for example we have taken three, they have calculated on bases of regression, 3 attributes, fight cavities, whitens teeth, clean stains. On basis of clean stains if you can see Denta Guard is the closest. In basis of whitens teeth it is Close Up. In basis of fight cavities it is Crest and Colgate. So these attributes can be used, this is the attributes which as I said if you remember on basis of the attributes also you can find out the scores, the similarity scores. So on basis of the attributes you can say how this brands are distributed in the spatial map.

(Refer Slide Time: 28:23)



Finally it has come to assess the reliability and validity. R-square, the index of R-square is one of the used values here. Is X squared correlation index, that we have also done in the regression chapter also when we did decision of regression. So that indicates the proportion of various optimally scale data can be accounted by the MDS. In fact the R-square tells you how much of explanation is happening. So value of point 6 are better consider exceptions.

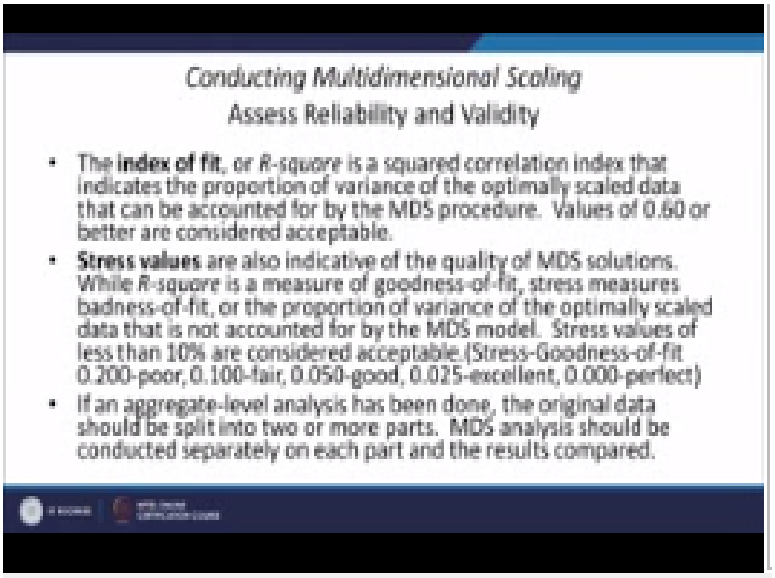
(Refer Slide Time: 28:51)



Stress values is very important because the stress values, higher the stress values poor the fit. As you can see stress values measures badness of fit or rather goodness of fit you can, stress measures the badness of the fit, because R-square measures the goodness of fit. So higher the stress, poor the fit. So stress values are less than 10% are considered acceptable. So goodness of fit if you see 0.2, if your stress is around 0.2, it is poor, the fit is poor.

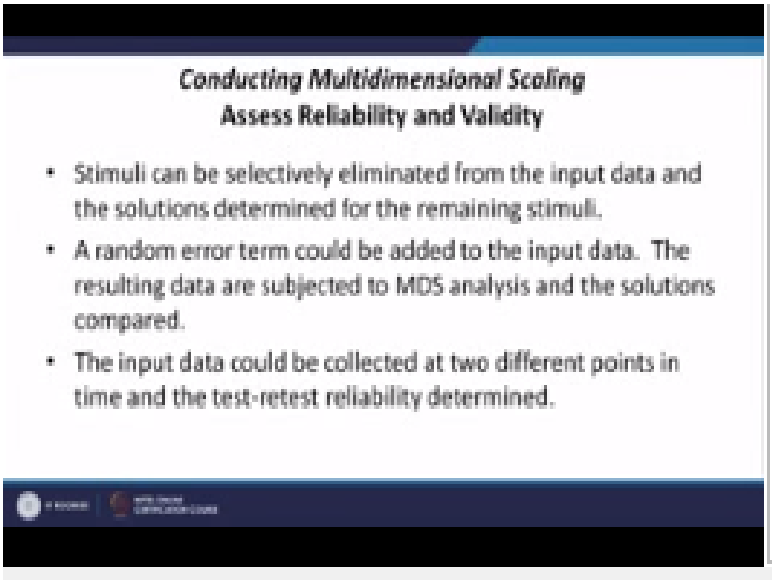
If it is 1 fair, if it is 0.050 good, 0.025 excellent, 0.000 is perfect. It will be 0.000 only when the original data and the perceived data are equal or same.

(Refer Slide Time: 29:42)



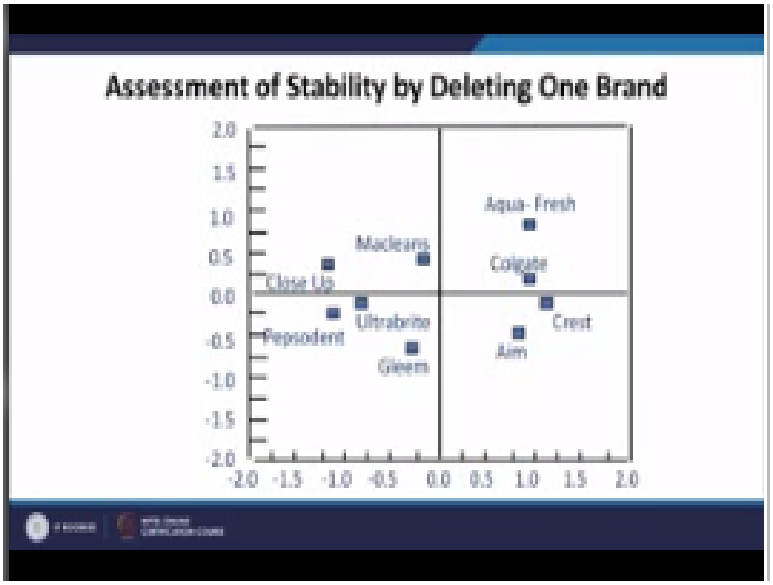
If we will aggregate the level analysis has been done, the original data should be split into 2 or more parts. Now suppose you have done the analysis on the aggregate data you don't know what dimensions people have used. So in order to ensure the validity split the data into 2 or more parts and make the analysis on separately on the both parts and compare the results.

(Refer Slide Time: 30:12)



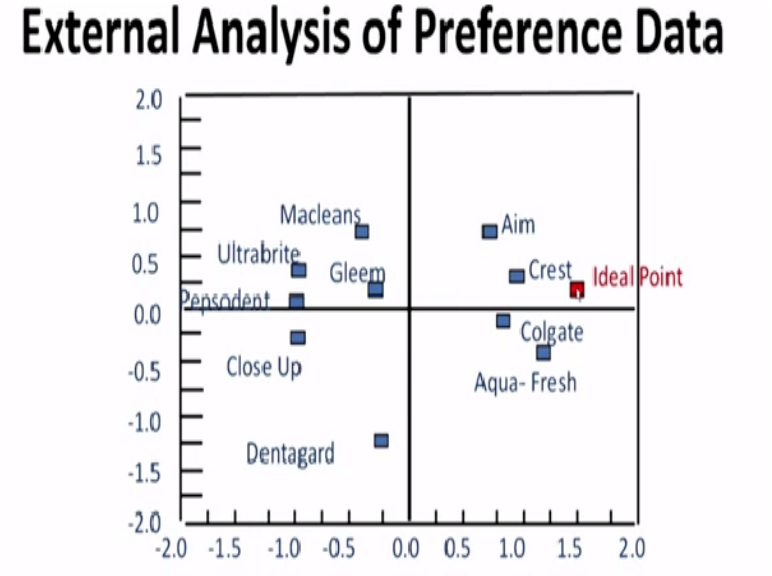
So these are some of the things which I am not going to discuss much.

(Refer Slide Time: 30:16)



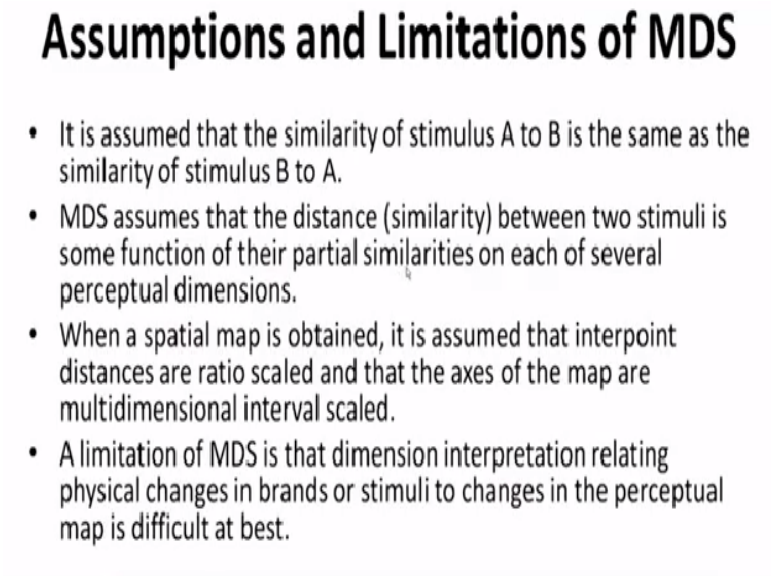
Finally one more thing, if you sometimes feel that one brand particularly, let us say you have used 10 brands and one brand is because of one brand the assessment is not being very poor, there is a instability. So it is better to identify the brand which is very different from the rest and to delete that and then check whether there is a better stress value or stability. So from earlier studies also you can see this is the ideal point.

(Refer Slide Time: 30:45)



Ideal point is something which is done, which is used through early past researches, from the past data that you have collected on similar studies, you can say what should be the ideal point for a good pressure toothpaste and on basis of this ideal point you can make also comparisons, no limitations.

(Refer Slide Time: 31:12)

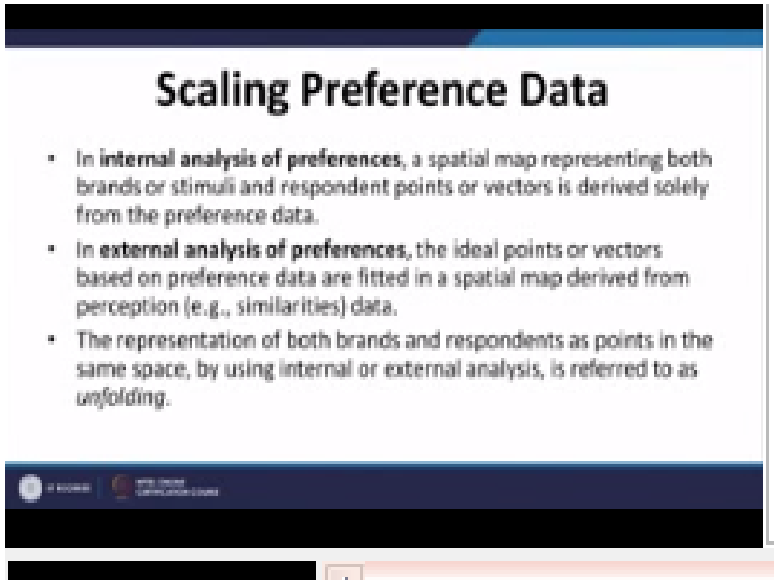


These limitations are okay, so I am not getting the limitations much because there is not much of limitations in this, so let us go to a few. A limitation of MDS is the dimension interpretation relating to physical changes in brands or stimulate to changes in the perceptual map is difficult at best. This is the limitation. So this limitation that the interpretation relating to physical changes is difficult to understand or to interpret.

Similarly MDS assumes that the distance between two stimuli is some function of that partial similarities, now this is the limitation, it assumes that the two stimuli, the distance, the distance between the 2 stimuli is some function of the partial similarities, but it might not be, it might not be, it could be something else. It could be some other reasons also.

So it is, one more is that the similarity of stimulus A to B is the same as the similarity of stimulus B to A, now this is crucial this is to understand. If A to B, the similarity between A to B is same as the similarity between B to A is actually, may not be correct all the time.

(Refer Slide Time: 32:37)



So well this is the last I think. So I am going to close it. So scaling preference data, there are 2 types of internal analysis of preferences and external analysis. Internal analysis is the one the spatial map that has been built on basis of the data that you have collected from the respondents and the external analysis is the ideal point of vectors based on preference data fit in a spatial map derived from perception from past data, some past record.

This representation of both of the brands and respondents as points in the same space by using internal and external analysis is referred to as unfolding which is I think, let us not get into that. But yes I can tell you, the external analysis of preference is the data which has been used, collected from the past and that is being used in the further analysis.

External analysis is preferred in most of the situations because already some research has been done and people have found out and those can be used to make a study. Well this is all we have for multidimensional scaling. I hope you have understood the meaning of multidimensional scaling, how you should have the input data and once you have the input data it is, actually.

Let us not get too much deep into it, just understand from the brief it is used to helpful to understand the similarity and through the similarity you can identify new opportunities, new products, the positioning of the product in the minds of the consumer and all these things. So when you are talking about MDS, multidimensional scaling you need to find out 2, 3 things, one, calculate R-square to check the goodness of and the stress which is also goodness of fit, but indirectly is the badness of fit, and see how many dimensions are available in the study.

So on basis of these dimensions where these products are placed in the spatial map and accordingly as a researcher or a marketer what action would you take in order to change the image, or change the opinion of the people, or bring in changes in your product to improve the products expectation in the market and all these things. So I hope you have understood multidimensional scaling and how to do it and what is the importance. Thank you very much.

For Further Details Contact

Coordinator, Educational Technology Cell
Indian Institute of Technology Roorkee
Roorkee-247 667
E-Mail: etcell.iitrke@gmail.com, etcell@iitr.ernet.in
Website: www.iitr.ac.in/centers/ETC., www.nptel.ac.in

Production Team
Sarath Koovery
Mohan Raj. S
Jithin. K
Pankaj saini

Graphics
Binoy. V. P

Camera
Arun. S

Online Editing
Arun. S

Video Editing
Arun. S

NPTEL Coordinator
Prof. B. K. Gandhi

An Educational Technology Cell
IIT Roorkee Production