



IIT Madras

ONLINE DEGREE

Computational Thinking
Professor. Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute

Professor. G Venkatesh
Indian Institute of Technology, Madras

Relations among customers based on their spending patterns (Part 2)

(Refer Slide Time: 0:26)

The image shows a wooden table with several papers and a laptop screen. The papers contain handwritten data tables. The laptop screen shows a 'Sun General' application window with a table of items and their counts.

Handwritten Table 1 (Top Left):

Item	Qty	Price	Total
Apples	10	8	80
Bananas	5	4	20
Oranges	3	5	15
Apples	10	8	80
Bananas	5	4	20
Oranges	3	5	15

Handwritten Table 2 (Bottom Left):

Item	Qty	Price	Total
Apples	10	8	80
Bananas	5	4	20
Oranges	3	5	15
Apples	10	8	80
Bananas	5	4	20
Oranges	3	5	15

Laptop Screen Table (Sun General):

Item	Category	Qty	Price	Total
Apples	Stationery	2	5	10
Bananas	Stationery	4	20	80
Oranges	Stationery	1	15	15
Graph Book	Stationery	1	20	20

Professor Madhavan Mukund: So we had done this grouping of these customers according to how much they have bought by calculating total items for each category, number of items. And, then we had segregated them based on those who had only 1 item food.

Professor G Venkatesh: 1 item, yeah.

Professor Madhavan Mukund: when those who had few items, we had binned them as few items and many items.

Professor G Venkatesh: less than 10.

Professor Madhavan Mukund: less than 10, more than 10.

Professor G Venkatesh: more than 10.

Professor Madhavan Mukund: And, then we had kind of suggested that within these groups, one could look for a distance between them by looking column wise at the separation between the

numbers. And, finding how close the customers are using them. So, that gives us the numerical way to justify that 2 customers are close to each other, right.

Professor G Venkatesh: So, should we go through that process, I mean, see, we, we do not need to do it for this I guess, because numbers are small.

Professor Madhavan Mukund: So, we can do it for...

Professor G Venkatesh: So, we can do it for these.

Professor Madhavan Mukund: These are where some interesting patterns might be there. So, what we need to do is maybe think of them as, you know, 6, so we can.

Professor G Venkatesh: 1, 2, 3, 4, 5, 6.

Professor Madhavan Mukund: Let us give them some...

Professor G Venkatesh: Index them, yeah...

Professor Madhavan Mukund: We will say, this is customer 1, 2, 3, 4. So, these are just an index that we are assigning...

Professor G Venkatesh: So, we are saying that these are 6 customers, and we want to compare every customer ...

Professor Madhavan Mukund: With every other.

Professor G Venkatesh: every, so, there are 6 customers you want to compare everybody, it means 6, each of them with others means, 6 into 5. 30.

Professor Madhavan Mukund: 30.

Professor G Venkatesh: But, again, A compared with B, and B compared with A are the same.

Professor Madhavan Mukund: Yeah, so that will be...

Professor G Venkatesh: 15 comparison should come, right? So, 15 pairs we want to take. We have to be careful about these pairs, right? We should not make

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: We should make sure that all the pairs we have counted.

Professor Madhavan Mukund: We do not miss out any...

Professor G Venkatesh: That is the first thing we have to. Now, for every pair, we have to compute, we want to compute a...

Professor Madhavan Mukund: Distance.

Professor G Venkatesh: Distance.

Professor Madhavan Mukund: So,...

Professor G Venkatesh: So, every pair we compute a distance, how we want to use it? Pairs which have the smallest distance.

Professor Madhavan Mukund: Will be the nearest customer...

Professor G Venkatesh: Nearest customers.

Professor Madhavan Mukund: That is what we want to....

Professor G Venkatesh: That is what we want to see.

Professor Madhavan Mukund: Patterns are similar across all the columns.

Professor G Venkatesh: Okay.

Professor Madhavan Mukund: So, distance, let us just take a simple distance, which is just we will take a difference.

Professor G Venkatesh: Difference. So, column wise?

Professor Madhavan Mukund: Column wise.

Professor G Venkatesh: So, 3 minus 3 is 0, like that? 13 minus 8 is 5.

Professor Madhavan Mukund: Right. So, we will always take the positive numbers. We will not take minus 5 plus 5, because, we...

Professor G Venkatesh: Subtract and take the absolute value, right?

Professor Madhavan Mukund: So, we do not want to cancel out that

Professor G Venkatesh: No.

Professor Madhavan Mukund: 1 person has more food and less utilities and...

Professor G Venkatesh: So, the distance column wise we are taking, taking the absolute values of those.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: Then,...

Professor Madhavan Mukund: Adding them.

Professor G Venkatesh: Adding the distance column wise.

Professor Madhavan Mukund: And, that will be the total distance.

Professor G Venkatesh: That will be total distance.

Professor Madhavan Mukund: Yeah, and we want to know

Professor G Venkatesh: There is a name for this. I mean, that is what we call a rectilinear distance.

Professor Madhavan Mukund: Yeah, I think or Manhattan distance.

Professor G Venkatesh: Manhattan distance.

Professor Madhavan Mukund: Yeah, I think they call it.

Professor G Venkatesh: Manhattan distance because...

Professor Madhavan Mukund: Because, it is like how much, if you have a grid of roads, how many roads you have to take to go from here to there.

Professor G Venkatesh: There, X, you go in the X direction and then you go in the Y direction like that.

Professor Madhavan Mukund: So, in each...

Professor G Venkatesh: But, here only thing is that we have 1, 2, 3, 4, 5.

Professor Madhavan Mukund: Yeah, just like, if you are living...

Professor G Venkatesh: 5 dimensions...

Professor Madhavan Mukund: 5 dimensional universe, we have to do it.

Professor G Venkatesh: Alright.

Professor Madhavan Mukund: And, yeah, now we want to, for every every pair as you said, we want to asso...we want to note down this distance. I mean...

Professor G Venkatesh: Note down.

Professor Madhavan Mukund: It is plain to say that, you know, it has been of 'srivatsan distance' and it is 'adwaith akshay' distance. So, it will be useful to note them against these indices. So, we want to (save) with the distance between 2 and 3 or distance in 3 and 6. So, that way...

Professor G Venkatesh: So, that is the good way to keep track. There should be different way of keeping track, no?

Professor Madhavan Mukund: so, so actually one way we could do is to just write down all these pairs.

Professor G Venkatesh: All pairs.

Professor Madhavan Mukund: And, then for every pair we write down the distance next to the pair.

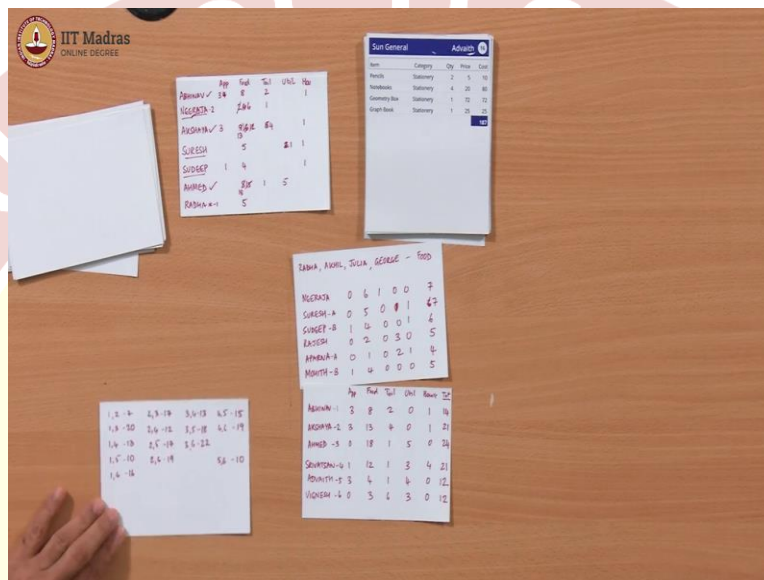
Professor G Venkatesh: Okay.

Professor Madhavan Mukund: So, that could be...

Professor G Venkatesh: Let us try that first, I mean let us do it.

Professor Madhavan Mukund: So, let us start with 1.

(Refer Slide Time: 3:54)



Professor G Venkatesh: 1. So, if you start with 1, what are the possible ways of pairing 1? 1 to 2,...

Professor Madhavan Mukund: So, one can.

Professor G Venkatesh: 1 with 2, 1 to 3

Professor Madhavan Mukund: 1 to 3

Professor G Venkatesh: 1 to 4, 1 to 5, 1 to 6. That is one pair, one set of pairs which you covered with 1.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: And, you will start with 2.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: Now, 2 with 1 is already covered, covered.

Professor Madhavan Mukund: Yeah, so, we do not need to look at 2 and 1

Professor G Venkatesh: 2, 2, so, 2 with 3.

Professor Madhavan Mukund: 2 to 2 is not obviously meaningful because that will...

Professor G Venkatesh: No, it will not help here, 2 to 3, 2 to 4, 2 to 5, 2 to 6. Then you will start with 3.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: 3 to 4, 3 to 5, 3 to 6; a way of systematically enumerating the pairs.

Professor Madhavan Mukund: And,...

Professor G Venkatesh: And, then 4 with 5, 4 with 6; then, finally 5 with 6.

Professor Madhavan Mukund: Finally, I will just write it here. 5 with 6, right. So, we should check so, 1, 2, 3, 4, 5, 9, 12, 15.

Professor G Venkatesh: 15,

Professor Madhavan Mukund: Exactly.

Professor G Venkatesh: Wonderful, wonderful.

Professor Madhavan Mukund: So, now what we can do is, now for every such pair, now we look at 1-2, we want to find the distance 1-2. So, we look at row 1, row 2 and do this distance.

Professor G Venkatesh: So, do it.

Professor Madhavan Mukund: And, then write it down. Yeah. So, we can do that. 1 and 2

Professor G Venkatesh: 3 minus 3, 0. 13 minus 8 is 5.

Professor Madhavan Mukund: 5.

Professor G Venkatesh: 4 minus 2 is 2.

Professor Madhavan Mukund: 7.

Professor G Venkatesh: So, 0 plus 5, 5 plus 2, 7. 0. 0. So, total distance is 7.

Professor Madhavan Mukund: So, this distance is 7. So, 1 and 3..

Professor G Venkatesh: So, It is correct? This is 0, this is 0,

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: This is 2, this is 5.

Professor Madhavan Mukund: That is 0.

Professor G Venkatesh: This is 0. So, 5 plus 2, 7. Okay. Now, we want to do 1 and 3.

Professor Madhavan Mukund: 1 and 3.

Professor G Venkatesh: 1 and 3. Okay, so, 3 minus 3, 0.

Professor Madhavan Mukund: 3.

Professor G Venkatesh: 13, 10.

Professor Madhavan Mukund: 13.

Professor G Venkatesh: 13, 1.

Professor Madhavan Mukund: 14.

Professor G Venkatesh: 5.

Professor Madhavan Mukund: 19.

Professor G Venkatesh: 1.

Professor Madhavan Mukund: 20

Professor G Venkatesh: 20.

Professor Madhavan Mukund: So, this distance is 20.

Professor G Venkatesh: Okay, then we do 1 and 4.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: Okay, 2,

Professor Madhavan Mukund: 2.

Professor G Venkatesh: 4.

Professor Madhavan Mukund: 6.

Professor G Venkatesh: 1.

Professor Madhavan Mukund: 6 plus 1, 7.

Professor G Venkatesh: 3.

Professor Madhavan Mukund: 10.

Professor G Venkatesh: 13.

Professor Madhavan Mukund: 13.

Professor G Venkatesh: 13 okay. See, we cannot, cannot subtract the totals.

Professor Madhavan Mukund: Yeah, yeah.

Professor G Venkatesh: Because of that plus minus, right.

Professor Madhavan Mukund: Yes, exactly. So, we could...

Professor G Venkatesh: So, here the distance, actually, 21 minus 14 is 7.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: We cannot do it, okay.

Professor Madhavan Mukund: Because here for instance, the apparel is 2 more than this.

Professor G Venkatesh: So, this is giving us some better.

Professor Madhavan Mukund: So, this is 4 more than that.

Professor G Venkatesh: Yeah.

Professor Madhavan Mukund: So, 3 more than that.

Professor G Venkatesh: We cannot cancel it. Now, you are comparing 1 with 5.

Professor Madhavan Mukund: 0.

Professor G Venkatesh: 0, 4.

Professor Madhavan Mukund: 4.

Professor G Venkatesh: 1.

Professor Madhavan Mukund: 5.

Professor G Venkatesh: 4.

Professor Madhavan Mukund: 9.

Professor G Venkatesh: 1.

Professor Madhavan Mukund: 10.

Professor G Venkatesh: 1 with 6, 3, 5.

Professor Madhavan Mukund: 8.

Professor G Venkatesh: 4.

Professor Madhavan Mukund: 12.

Professor G Venkatesh: 3.

Professor Madhavan Mukund: 15.

Professor G Venkatesh: 1.

Professor Madhavan Mukund: 16. So, now we are done with all comparisons involving 1.

Professor G Venkatesh: 1, okay. So, let us start with 2 now.



Professor Madhavan Mukund: Yeah, 2 and 3.

Professor G Venkatesh: 2 and 3. So, 3 and 0, 3.

Professor Madhavan Mukund: 3.

Professor G Venkatesh: 5.

Professor Madhavan Mukund: 8.

Professor G Venkatesh: 3.

Professor Madhavan Mukund: 11.

Professor G Venkatesh: 5.

Professor Madhavan Mukund: 16.

Professor G Venkatesh: 1.

Professor Madhavan Mukund: 17.

Professor G Venkatesh: Okay, now we do 2 and 4.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: Okay, 2.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 1.

Professor Madhavan Mukund: 3.

Professor G Venkatesh: 3.

Professor Madhavan Mukund: 6.

Professor G Venkatesh: 3.

Professor Madhavan Mukund: 9. 12



Professor G Venkatesh: Okay, now adding 2 and 5.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: 0.

Professor Madhavan Mukund: 0.

Professor G Venkatesh: 9.

Professor Madhavan Mukund: 9.

Professor G Venkatesh: 3.

Professor Madhavan Mukund: 12.

Professor G Venkatesh: 4.

Professor Madhavan Mukund: 16.

Professor G Venkatesh: 1.

Professor Madhavan Mukund: 17.

Professor G Venkatesh: Okay, now we are doing 2 and

Professor Madhavan Mukund: 6.

Professor G Venkatesh: 6. 3.

Professor Madhavan Mukund: Yeah.

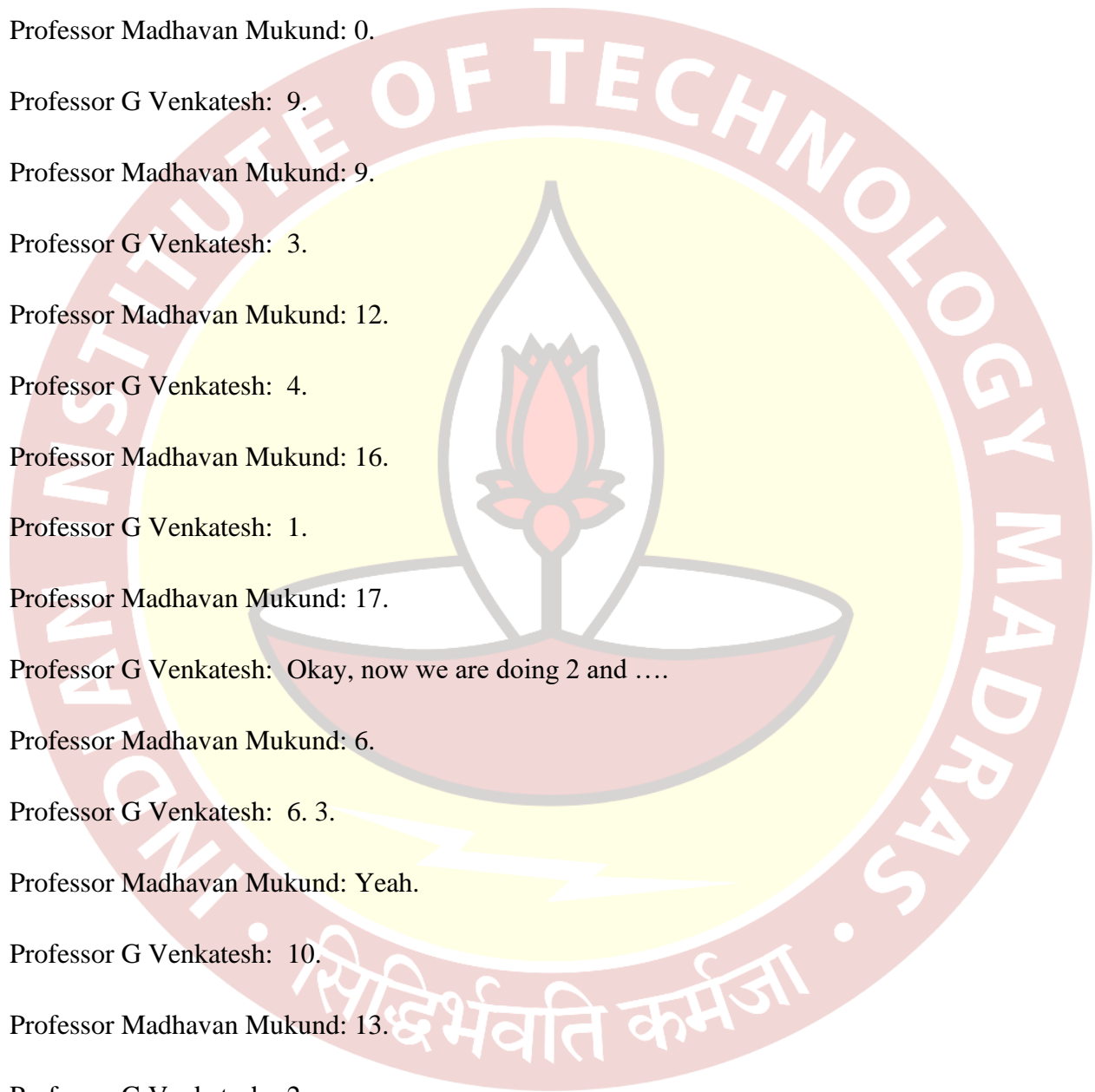
Professor G Venkatesh: 10.

Professor Madhavan Mukund: 13.

Professor G Venkatesh: 2.

Professor Madhavan Mukund: 15.

Professor G Venkatesh: 3.



Professor Madhavan Mukund: 18.

Professor G Venkatesh: 1.

Professor Madhavan Mukund: 19.

Professor G Venkatesh: So, we have done the 2.

Professor Madhavan Mukund: Yeah. Now 3.

Professor G Venkatesh: Go to 3.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: 1.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 6.

Professor Madhavan Mukund: 7.

Professor G Venkatesh: 0.

Professor Madhavan Mukund: 7.

Professor G Venkatesh: 2.

Professor Madhavan Mukund: 9.

Professor G Venkatesh: 4.

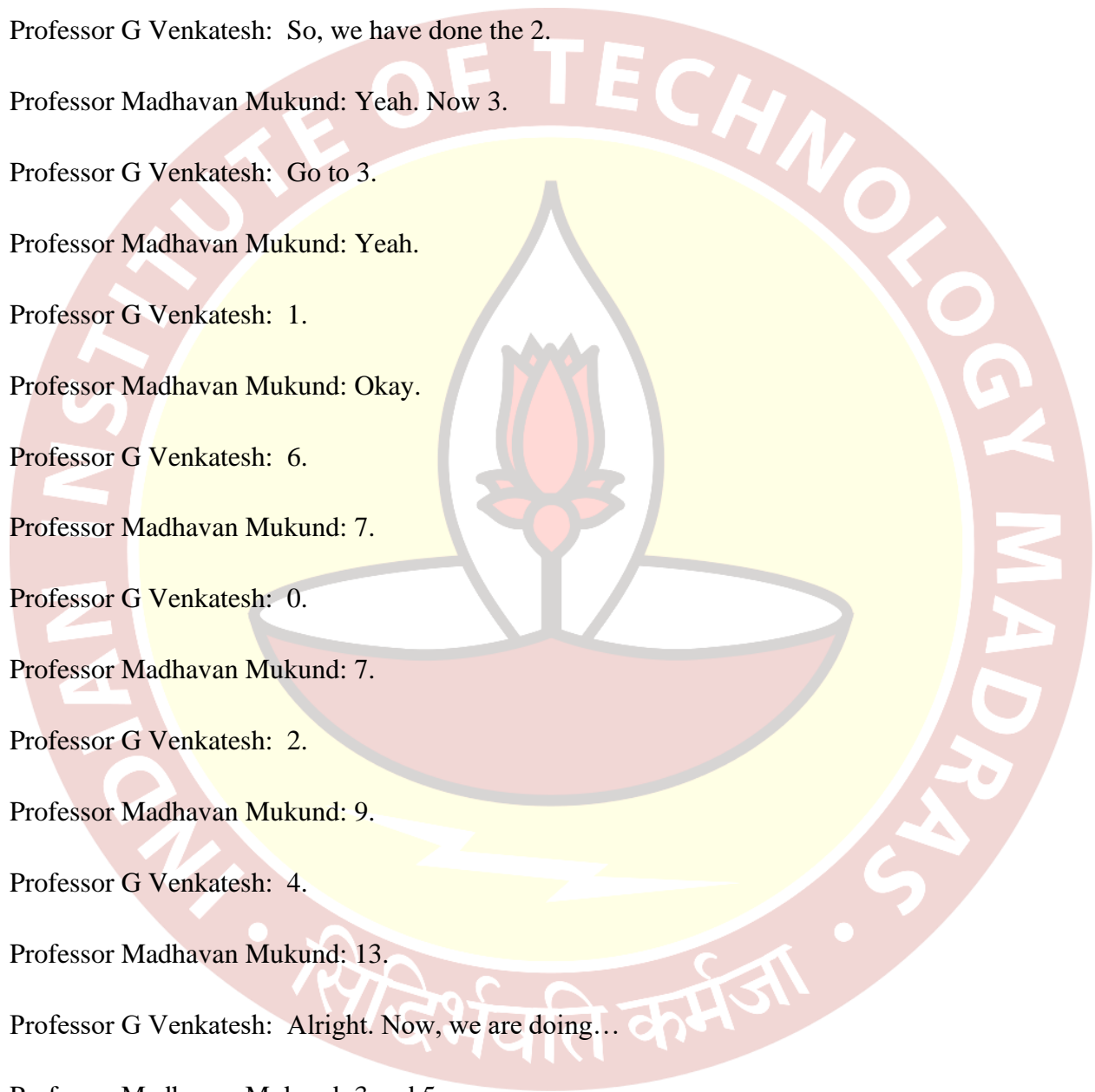
Professor Madhavan Mukund: 13.

Professor G Venkatesh: Alright. Now, we are doing...

Professor Madhavan Mukund: 3 and 5.

Professor G Venkatesh: 3 and 5. 3.

Professor Madhavan Mukund: 3.



Professor G Venkatesh: 14.

Professor Madhavan Mukund: 17.

Professor G Venkatesh: 0.

Professor Madhavan Mukund: 17.

Professor G Venkatesh: 1.

Professor Madhavan Mukund: 18.

Professor G Venkatesh: 0.

Professor Madhavan Mukund: 18, okay.

Professor G Venkatesh: Okay. Now, we are doing 3 and 6.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: 0.

Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 15.

Professor Madhavan Mukund: 15.

Professor G Venkatesh: 5.

Professor Madhavan Mukund: 20.

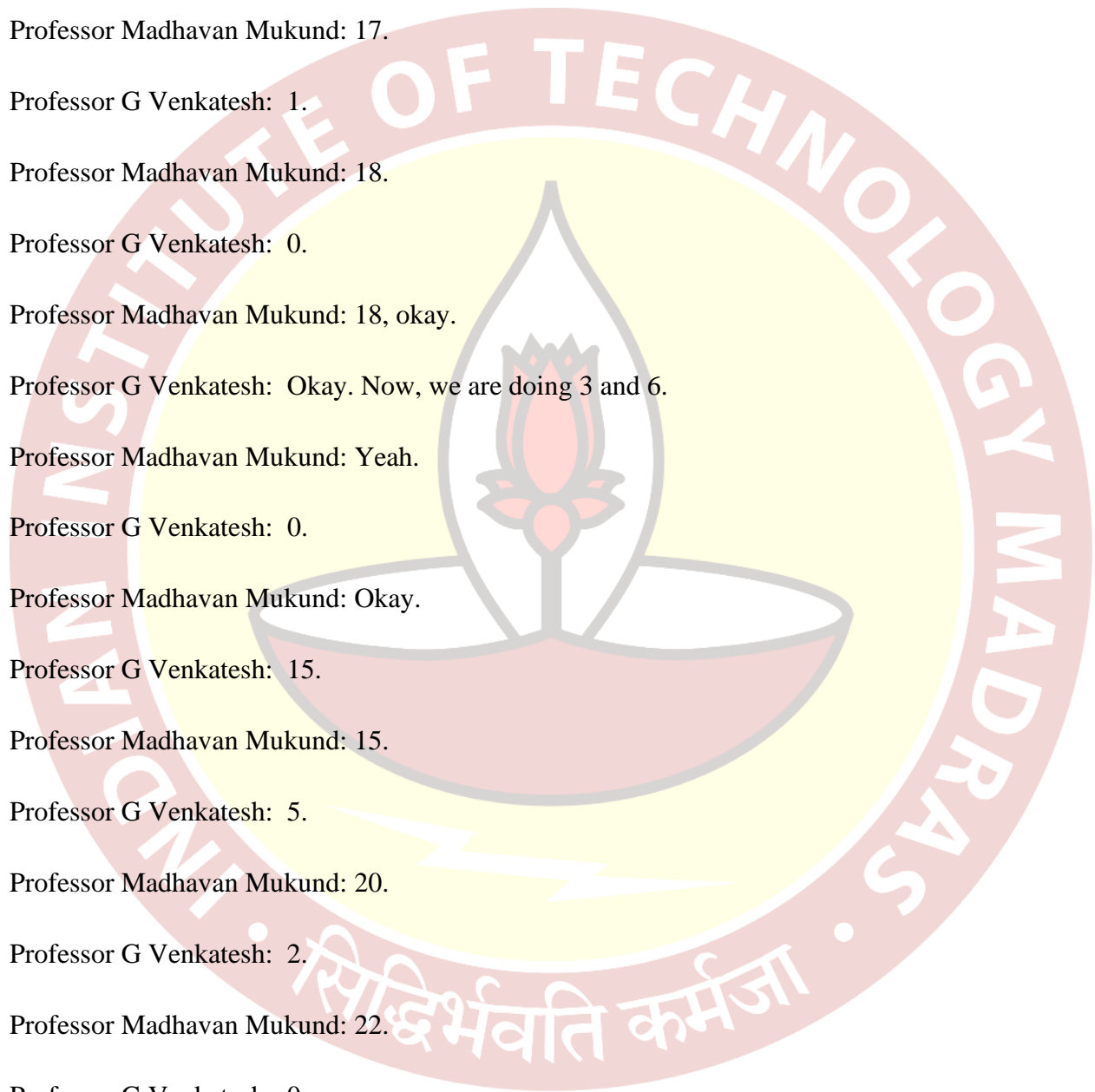
Professor G Venkatesh: 2.

Professor Madhavan Mukund: 22.

Professor G Venkatesh: 0.

Professor Madhavan Mukund: 4 and 5.

Professor G Venkatesh: 4 and 5, 2.



Professor Madhavan Mukund: Okay.

Professor G Venkatesh: 8.

Professor Madhavan Mukund: 10.

Professor G Venkatesh: 0.

Professor Madhavan Mukund: 0.

Professor G Venkatesh: So, it is still 10.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: 1.

Professor Madhavan Mukund: 11.

Professor G Venkatesh: 4.

Professor Madhavan Mukund: 15 okay.

Professor G Venkatesh: Now, we are doing 4 and 6.

Professor Madhavan Mukund: 6.

Professor G Venkatesh: 1.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: 9.

Professor Madhavan Mukund: 10.

Professor G Venkatesh: 5.

Professor Madhavan Mukund: 15.

Professor G Venkatesh: 0.

Professor Madhavan Mukund: 15.



Professor G Venkatesh: 4.

Professor Madhavan Mukund: 19, okay.

Professor G Venkatesh: Now, finally we have come to 5 and 6.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: 3, 1

Professor Madhavan Mukund: 4.

Professor G Venkatesh: 5.

Professor Madhavan Mukund: 9.

Professor G Venkatesh: 1

Professor Madhavan Mukund: 10.

Professor G Venkatesh: 0. Alright, so what do we see?

Professor Madhavan Mukund: So, the smaller the distance, the nearer they are. So, in this whole thing there is only 1 distance less than 10 and which is 1 and 2.

Professor G Venkatesh: Okay.

Professor Madhavan Mukund: So, 1 and 2 is closest then the next close smallest thing is 10, I think. 1 and 5 and 5 and 6 but of course, the thing is that 2 to 5 is very large, 17. 2 to 6 is also very large, 19. So, what I am saying is, if we group, say that 1-5 is close and 5-6 is close, 1-6 is different. So, we have to be careful about. So, 1-2 is close, for example. But if we try to group 1, 2 and 5 because 1-5 is close. Then 2-5 turns out to be 17.

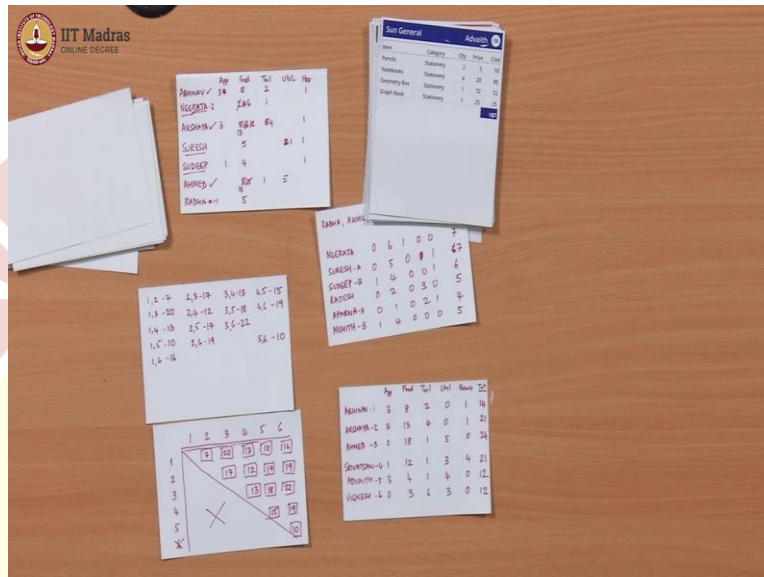
Professor G Venkatesh: Hmm.

Professor Madhavan Mukund: So, grouping into 3 is, then you have to somehow decide how to, may be you could take the total distance across all the pairs in that group.

Professor G Venkatesh: Hmm, Hmm, Hmm

Professor Madhavan Mukund: But, basically we want to do this. But I see something interesting here. Maybe before we identify. So, actually I think there is nicer way to write down this table.

(Refer Slide Time: 9:42)



Professor G Venkatesh: Hmm.

Professor Madhavan Mukund: Alright. So, it is like we have to pair up these 6 guys.

Professor G Venkatesh: Hmm. I see, okay. You are keeping 1, 2 in the rows and 1,2 in the column.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: Hmm.

Professor Madhavan Mukund: These 6 guys.

Professor G Venkatesh: Okay.

Professor Madhavan Mukund: Of course, we do not need to pair up anybody with themselves. So, we want to put something here, here, here, here and here. So, these are the 5 pairs.

Professor G Venkatesh: These is, this is what you have written here.

Professor Madhavan Mukund: In the first row. First thing then....

Professor G Venkatesh: So, those numbers can be filled in that I presume?

Professor Madhavan Mukund: Yeah, and then similarly we have...

Professor G Venkatesh: Oh! I see. So, you only get entries over here.

Professor Madhavan Mukund: Yeah, so, only entries above the diagonal of this matrix will come; okay and there will be exactly our 15 entries. And, then we do not need to put anything here because they will just fill the same.

Professor G Venkatesh: Yeah.

Professor Madhavan Mukund: And, now we can put the same thing here. So, we can say, 7, 18.

Professor G Venkatesh: So, it gives us some spatial way of seeing it?

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: May be something will come if you look at this, but I do not know. We will see.

Professor Madhavan Mukund: So, this is the same data, just arranged in a more, possibly logical way.

Professor G Venkatesh: Okay.

Professor Madhavan Mukund: So, if you want to pair, we will say 4-5, you go to the 4 row and 5 column and then you say, 4-5 is 18. Like that.

Professor G Venkatesh: Like that. Okay.

Professor Madhavan Mukund: So, Okay. So, 4-5,.

Professor G Venkatesh: This is what, this is, 5. This is 5. 1, 2, 3, 4, 5.

Professor Madhavan Mukund: This whole thing should be shifted down because 1 is not compared to itself. So, there is an empty row here. So, this is 1-2. 1, 2 and 2.

Professor G Venkatesh: Okay, Okay.

Professor Madhavan Mukund: So, there is no 6 6. So, there is an empty row here. So, 6 is not binned. So, we do not actually compare 6 with anything.

Professor G Venkatesh: Yeah, okay. So, this is 5.

Professor Madhavan Mukund: So, now we are saying that this is clearly close.

Professor G Venkatesh: So, 1-2 is a close fit and 5-6.

Professor Madhavan Mukund: 5-6 is a close fit.

Professor G Venkatesh: Somewhere here also. So, here is also 1. 1-5.

Professor Madhavan Mukund: And, 1-5 is a close fit.

Professor G Venkatesh: And, 5-6? 5 6, therefore.

Professor Madhavan Mukund: So, 1, 5, 6 may be a good

Professor G Venkatesh: So, look at 1, 5, 6 may be a decent group. Let us look at that. 1, 5, 6.

Professor Madhavan Mukund: But 1-6 is not so good. See, 1-6 is 16. So, why is 1 and 5 a good group? They have similar, in these 3 columns they are similar, they are different in that. So, it is a question of quantities. So, but I guess what we are really looking at and this is not so much grouping part of it, which is a separation question. But how to keep...

Professor G Venkatesh: Maybe if we have done better distance measurement.

Professor Madhavan Mukund: Correct.

Professor G Venkatesh: We did some very grossed up a lot of things. Now that we know that we have to pair things.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: For these people especially, we can ignore this gross numbers, go back to their original data. Take the original full data here and then find a distance directly from these cards. May be you get a better distance measure. And, we will learn something from that.

Professor Madhavan Mukund: Yeah and there are many assumptions we made, for instance, we looked at the number of line items. But we did not look at how many of each they are buying.

Professor G Venkatesh: Yeah.

Professor Madhavan Mukund: So...

Professor G Venkatesh: We could look at the price. Total cost.

Professor Madhavan Mukund: There could be somebody who buys large quantities of 1 particular item as opposed to somebody who buys only 1 quantity of that item. And, we are grouping them in the same bucket. So, that way how we collect this information that we start with is also important, I mean.

Professor G Venkatesh: So, they, I mean, I am, I am little curious about this, this method that they are using this, business of taking pairs. So, we said that there is this nested iteration. So, we are saying that there is iteration, outer iteration and there is another iteration which is inside this outer iteration. So, there are, this, we have nested the iterations.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: And, we, when we did it for those cards, we found that it was bit messy to do it. Because, you start with 1 card, you move to different pile.

Professor Madhavan Mukund: Yes

Professor G Venkatesh: But now for each of these cards, you have to measure again with the original piles. So, you have to go through the first pile once again.

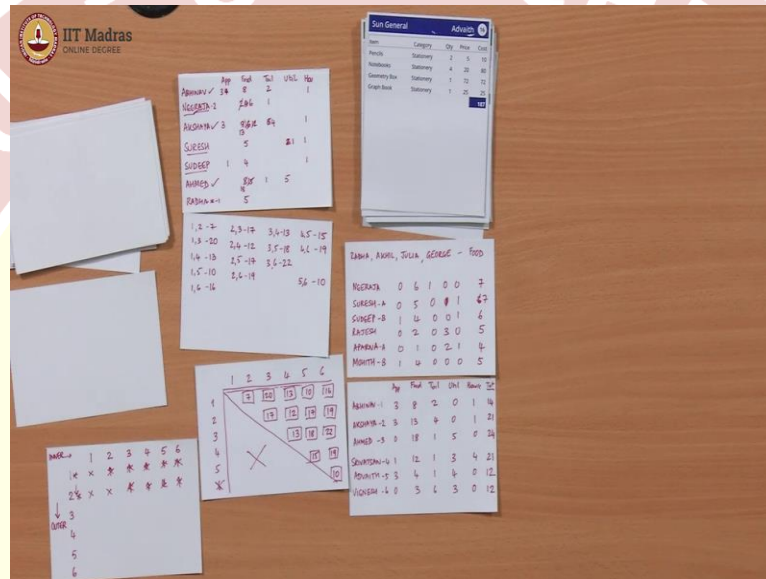
Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: So, you have to keep 2 cards aside. One which you have seen, not seen and then among those that you have not seen, again, those that you have considered, not considered.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: Seen, not seen is the outer iteration. Considered, not considered is the inner iteration. So, that business of keeping track of the where you are in the outer iteration and where you are on the inner iteration was messy actually. So, now we have a much more systematic way of keeping track of these inner and outer iterations pairs. Because you have just made all the pairs.

(Refer Slide Time: 14:25)



Professor Madhavan Mukund: Correct. So, let us say that this is the outer. So for every card, every customer, we are checking every other customer.

Professor G Venkatesh: Yeah.

Professor Madhavan Mukund: So this is the inner.

Professor G Venkatesh: Yeah.

Professor Madhavan Mukund: So, what we do is we first fix that we are here.

Professor G Venkatesh: Okay.

Professor Madhavan Mukund: We will mark that we are looking at this. And, then we start moving down this list. So, we know that we are not going to look at...

Professor G Venkatesh: Because it is already the same card.

Professor Madhavan Mukund: So, then we look at this item. So, this is a kind of, so we are fixing on customer 1 and second customer is 2. Then after we are done with that, then we move to customer 3.

Professor G Venkatesh: 3.

Professor Madhavan Mukund: Then after we are done with that, we move to customer 4. After, we done with that we move to 5 and so on. And, then finally when we reach the end of the list, right, then there are no more customers to look at. So, now we move to the next row.

Professor G Venkatesh: No more customers to look at in comparison for 1.

Professor Madhavan Mukund: For 1.

Professor G Venkatesh: The outer loop.

Professor Madhavan Mukund: The outer loop.

Professor G Venkatesh: They you start again with 2.

Professor Madhavan Mukund: You start with 2.

Professor G Venkatesh: Now, you go to 3.

Professor Madhavan Mukund: And, now we know that we should not look at 1

Professor G Venkatesh: We have already seen it.

Professor Madhavan Mukund: We have already seen it. We should not look at 2 because it is

Professor G Venkatesh: 2 and 2 is not.

Professor Madhavan Mukund: So, we will start with...

Professor G Venkatesh: 3.

Professor Madhavan Mukund: 3, then 4 and 5.

Professor G Venkatesh: So, it looks like we have 2 arrows, right. 1 arrow is keeping track of the first.

Professor Madhavan Mukund: Correct. So, and, then we have moving kind of, we keep track of 1 row and then we move across the columns. And then when we reached the end, we move...

Professor G Venkatesh: Move the first row again.

Professor Madhavan Mukund: Move this back.

Professor G Venkatesh: Move this back.

Professor Madhavan Mukund: Move again. Move this back.

Professor G Venkatesh: So, that is the systematic way of keeping track of the same, but same data set.

Professor Madhavan Mukund: Yes.

Professor G Venkatesh: Same data set but we are moving through this data set in 2 levels of iteration.

Professor Madhavan Mukund: Correct.

Professor G Venkatesh: Nested.

Professor Madhavan Mukund: So, both the rows and columns corresponds in this case to the same data set.

Professor G Venkatesh: Same data set.

Professor Madhavan Mukund: To the same list of names.

Professor G Venkatesh: Same list of names.

Professor Madhavan Mukund: But we are just processing them in a systematic.

Professor G Venkatesh: Processing them. We are trying to find out pairs.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: And, there is a systematic way of doing these pairs by keeping track of this.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: Matrix, row column thing.

Professor Madhavan Mukund: And, when it is 2 nested iterations, we can of course visualize it easily. If it becomes more like, when since we wanted all triples

Professor G Venkatesh: Triples .

Professor Madhavan Mukund: Then, ...

Professor G Venkatesh: Then we need 3.

Professor Madhavan Mukund: Then we need to draw.

Professor G Venkatesh: Then we need to draw cube.

Professor Madhavan Mukund: Cube and then it will be difficult to draw on a piece of paper and keep track. But, mentally we can imagine.

Professor G Venkatesh: Yeah.

Professor Madhavan Mukund: That we are keeping track of 3 dimensions.

Professor G Venkatesh: 3 arrows we will have, right. First arrow we will move. So, we do not have to...

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: draw it as rows and columns.

Professor Madhavan Mukund: Yeah. So, 1 arrow will move this way and 1 arrow will move that way and 1 arrow will move up.

Professor G Venkatesh: Up.

Professor Madhavan Mukund: So, up to 3 we can....

Professor G Venkatesh: When, when 1 is done then you will move this and

Professor Madhavan Mukund: Yeah, so, then you will move to the next one here. Then you will go all the way down, for each of these you will go the way up, and then so on. So, you keep going up.

Professor G Venkatesh: Okay.

Professor Madhavan Mukund: And, then come back.

Professor G Venkatesh: So far, luckily we have not had to do 3 iterations,...

Professor Madhavan Mukund: But, 2 iterations is..

Professor G Venkatesh: 2 iterations it seems to be...okay, nice. I think this is nice.

Professor Madhavan Mukund: And, then once you can do these 2 iterations, what you do at each step of iteration can vary, even within the problem as we said we could define the distance differently, or we could calculate some other measure. So, 1, and then the important thing is that once you have it indexed in this way, then you can note it down neatly as a as table. So, you do not lose track of... So, in this measure, for example, we had this kind of awkward thing where we have to keep track of the pair and the..so we do not have to worry about that anymore. Because the pair is...which pair it is, is determined by its position.

Professor G Venkatesh: Position in the table.

Professor Madhavan Mukund: In the table. So, this is kind of implicit way of keeping track of that the pair associated with each of these numbers.

Professor G Venkatesh: So, we had 2, I mean, here, the data itself we organized as tables here but the difference is in this table and this table. In this table, row was names.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: Column was categories

Professor Madhavan Mukund: Yeah, so the rows and the columns are of different types.

Professor G Venkatesh: Different types, whereas in this table, they are both names.

Professor Madhavan Mukund: Correct.

Professor G Venkatesh: Because we are comparing names with names.

Professor Madhavan Mukund: Yeah, yeah.

Professor G Venkatesh: So, you can make all kinds of tables it looks like. Tables are beautiful way of...

Professor Madhavan Mukund: Yeah, tables are certainly a very...

Professor G Venkatesh: Beautiful way of organizing.

Professor Madhavan Mukund: Useful way to organize.

Professor G Venkatesh: Organize information.

Professor Madhavan Mukund: Because, then by looking at the row and column, for each row and each column, basically that is what we are doing here. Here we are saying for customer 5 and customer 4, what is the distance? Here, we are saying for customer adwaith and item toiletries what is the number?

Professor G Venkatesh: Yeah.

Professor Madhavan Mukund: So, basically by looking at that position in the table, by looking at row number, row identity and column identity, we know what that quantity is talking about. So, that is a very useful way of arranging data. So, tables, tables are, I think are quite, that is why I guess you see them all over the place. Because, everybody can easily read them off and also update when you want to update.

Professor G Venkatesh: But to make the thing is that this table I understand because we knew that we had 6 items, so, we could make this table easily. This one, we do not know how many categories are there.

Professor Madhavan Mukund: Correct.

Professor G Venkatesh: First, you could first pass, make a pass; find out the number of categories. Find out the number of names.

Professor Madhavan Mukund: Yeah.

Professor G Venkatesh: Then once we know the names.

Professor Madhavan Mukund: But we do that even when we are making tables because we have to recognize whether this customer is new one or it is the second bill for an existing customer.

Professor G Venkatesh: correct.

Professor Madhavan Mukund: And, we also made some decisions like, we said that we would put, you know, batteries which were utilities and stationery and electronics all into one.

Professor G Venkatesh: 1, 1, 1 category. 1 thing. 1 row, 1 column.

Professor Madhavan Mukund: So, it is some kind of grouping.

Professor G Venkatesh: Grouping.

Professor Madhavan Mukund: to decide the column names. But, yeah, as you said, if you can do 1.

Professor G Venkatesh: 1 pass to do that.

Professor Madhavan Mukund: Just to check what are the different column names that could come and how you would want to collapse them then after that keeping track of them becomes easy.

Professor G Venkatesh: Alright.

