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 1)

Professor Madhavan Mukund: So, the shopping bill data last time we were actually trying to look at Customer Behavior at work whether they go to the same shop or not. This loyal customer problem but maybe from the shops point of view, they might want to understand what type of customers are there I mean, are the customers who buy similar things or the customers buy different things in terms of targeting them for offers or for attracting them in some new products

Professor G Venkatesh: So to customers, so, customer customers tend to be similar in terms of their purchases, like for example, parents with children are going to school, will buy, will tend to buy stationery books, pens, school materials, they will all buy. So, they will all vary. So, the shop can determine Oh, this looks like

Professor Madhavan Mukund: And for example,

Professor G Venkatesh: children going to school…

Professor Madhavan Mukund: Large families with many people staying probably their food food.

Professor G Venkatesh: Food bill might be very large.

Professor Madhavan Mukund: Depending on that people can decide whether this is a small unit or a large unit.

Professor G Venkatesh: Large Unit. So, I mean the way the shop can use this information is that if they find that one family which is buying stationary school going children family, they come and buy a certain kind of stationary item. Then when they see another family which is likely to buy which is buying stationery then they could try and promote this item. So it is a promote to me,…

Professor Madhavan Mukund: I think they would definitely benefit also if they want to make an offer in many ways.

Professor G Venkatesh: it needs a bundles in terms of offers combining things.

Professor Madhavan Mukund: Visible or in terms of combining it or offering some special price or some, some season.

Professor G Venkatesh: So, this is I guess is what most companies do their marketing department does this right they try and find out they try to segment the customers they call Segmentation right? They try to segment the customers into categories and say these are customers who will pay a lot of money or they are likely to buy this kind of product or service. So, this method of segmentation, psychographic segmentation method.

Professor Madhavan Mukund: And I think they advertise to different groups in different ways.

Professor G Venkatesh: different ways on different channels. So, some might be on children's channel.

Professor Madhavan Mukund: yeah or they might be social media for a certain….

Professor G Venkatesh: or through social media for those people are socially correct.

Professor Madhavan Mukund: Magazines or print.

Professor G Venkatesh: Print might be a better thing and so on yeah, yeah.

Professor Madhavan Mukund: I think overall this is a very useful thing. So even in packaging, I think they make the packaging attractive person they might make it bright colors…

Professor G Venkatesh: Or put a Mickey Mouse or. So, that it will attract children. You can do that you can do that. So, so when so, what we want to do is find out if 2 customers are similar, similar?

Professor Madhavan Mukund: So, we want to group these customer so at least we need to find one or 2 out to customers are similar.

Professor G Venkatesh: So, similar meaning that they have similar purchasing patterns they buy similar things

Professor Madhavan Mukund: Buy similar things exactly.

Professor G Venkatesh: But what does it mean?

Professor Madhavan Mukund: and what level because I think when this bill we have…

Professor G Venkatesh: Only evidence is we have bills.

Professor Madhavan Mukund: So, we have the exact item which is bought like shoes or socks. Then we have the category in this particular bill it is all footwear and apparel but there also be we have seen Food and Cosmetics and all that. So, there is a category there is a quantity and there is a price. The cost is just the price times the quantity so I think if we narrow it down to items it might be very because somebody might be buying one type of like shirt another person

Professor G Venkatesh: I think apparel is…

Professor Madhavan Mukund: So, I think we should look at the level of categories.

Professor G Venkatesh: even in the category maybe just look at Apparel

Professor Madhavan Mukund: So, the higher group category. So, what type I mean I think we have Packaged food and other food and so we can just get as a food, food.

Professor G Venkatesh: So, we can just say apparel. So, let us look at the cards maybe we can find the categories.

Professor Madhavan Mukund: But if we just look at which categories that might be too I mean everybody might be buying some apparel some food some stationery.

Professor G Venkatesh: See how many items?

Professor Madhavan Mukund: I want to probably understand how much?

Professor G Venkatesh: either we look at the total amount or look at quantity we can start with quantity?

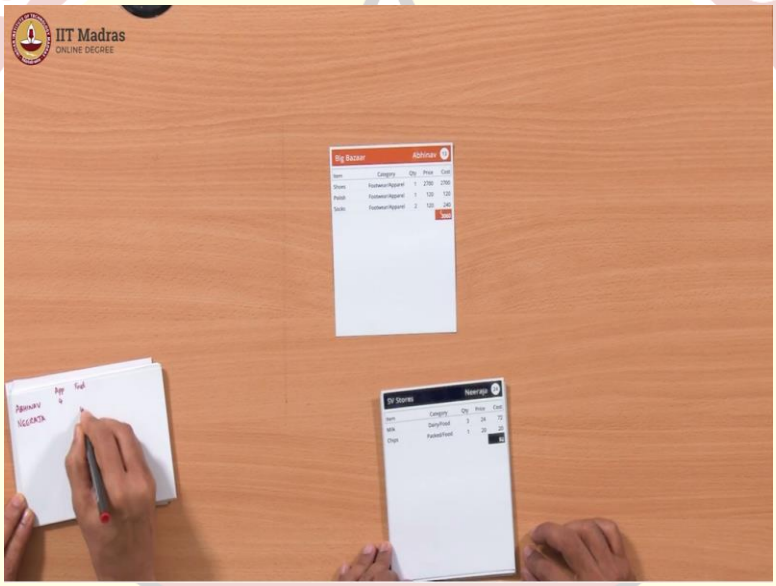
Professor Madhavan Mukund: Quantities probably better because total might be skewed by some high price.

Professor G Venkatesh: one item,

Professor Madhavan Mukund: So, let us do that. So, let us…

Professor G Venkatesh: So, we want to count the number of items of each category that the person has bought.

(Refer Slide Time: 04:06)



Professor Madhavan Mukund: Across all the bills?

Professor G Venkatesh: Across all the bills.

Professor Madhavan Mukund: So we do not care…

Professor G Venkatesh: Person by person,

Professor Madhavan Mukund: which shop they go to for each person.

Professor G Venkatesh: They do not care…

Professor Madhavan Mukund: We want to know their overall shopping behavior.

Professor G Venkatesh: Hopefully

Professor Madhavan Mukund: Okay, so we will write this down for each student so the first category, first so customers

Professor G Venkatesh: First customer is Abhinav, apparel is one category

Professor Madhavan Mukund: So I will just write,

Professor G Venkatesh: A you can write.

Professor Madhavan Mukund: And in this bill there are 1 4, 4 items currently for.

Professor G Venkatesh: So, we are as we go along with finding the categories, let us do that.

Professor Madhavan Mukund: So, Abhinav again

Professor G Venkatesh: Here we have got Abhinav again we have got food. In fact this looks like it is clustered by customer already.

Professor Madhavan Mukund: I think last time we had been them and we let us shuffle them so that we do not know this. And…

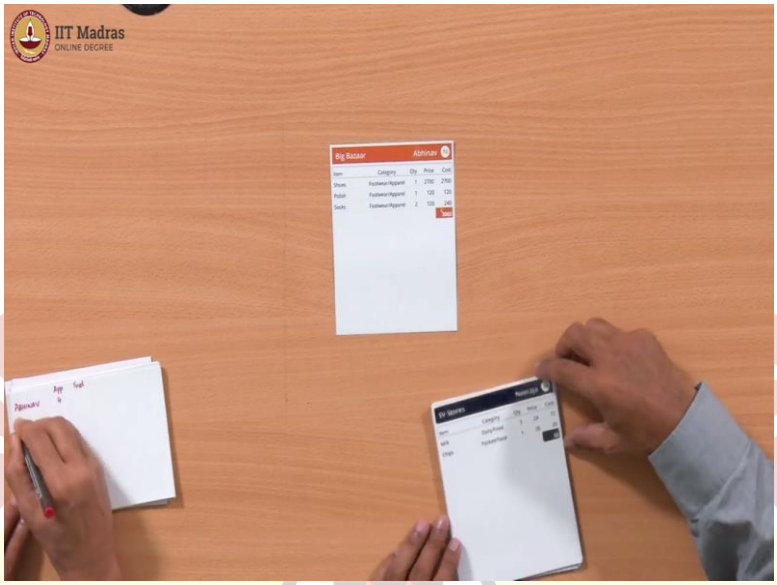
Professor G Venkatesh: I will just shuffle it. Just, just to make it, it is okay if it is cluster a little bit, helps us but…

Professor Madhavan Mukund: Be more illustrative we have a list, if it is not…

Professor G Venkatesh: And that should be all right.

Professor Madhavan Mukund: So now let us look at the next person on the list.

(Refer Slide Time: 05:13)



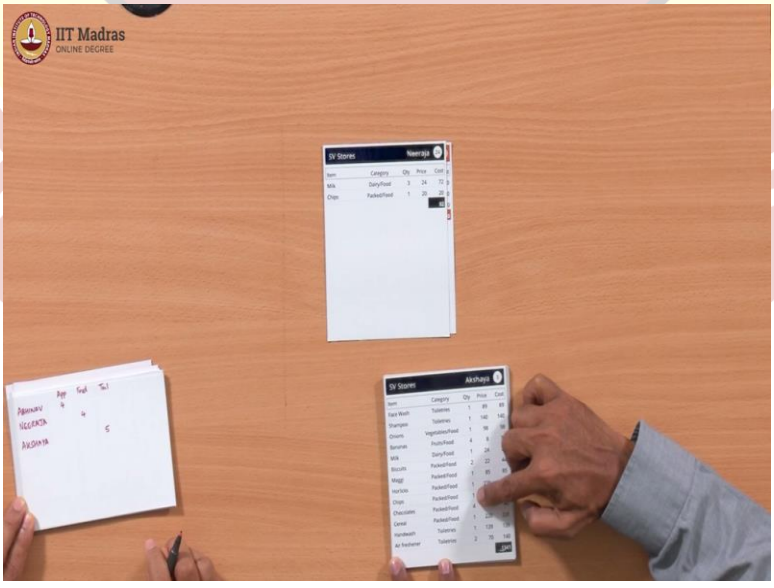
Professor G Venkatesh: Neeraja

Professor Madhavan Mukund: Next bill is Neeraja

Professor G Venkatesh: So all food in this. 4 items of food.

Professor Madhavan Mukund: So this is a new column food and there are 4 items.

(Refer Slide Time: 05:26)



Professor G Venkatesh: Akshaya.

Professor Madhavan Mukund: Akshaya.

Professor G Venkatesh: Big Bill. Akshaya has now there is something called…

Professor Madhavan Mukund: something called category called Toiletry. So I will just make Toiletry. This so toiletries is 4 items.

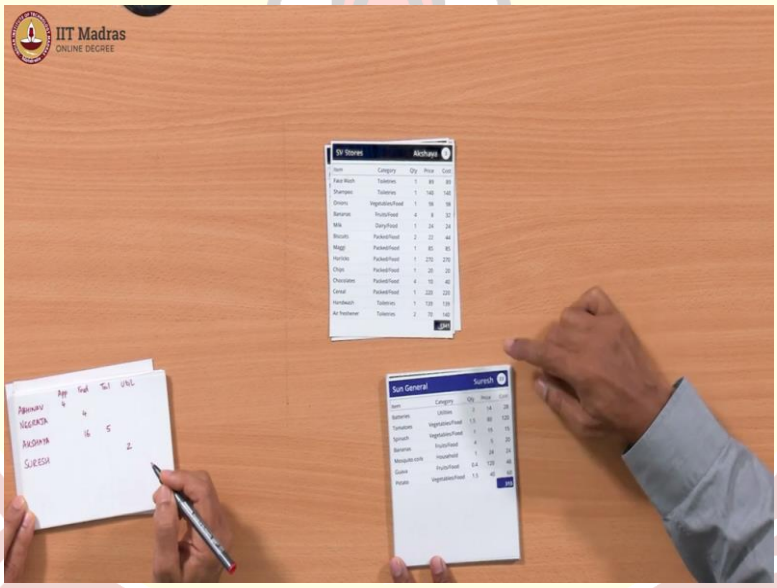
Professor G Venkatesh: 5 items.

Professor Madhavan Mukund: 5 items and Toiletries.

Professor G Venkatesh: 1 4 5 1 6 2 8 1 9 1 10 11 15 16 items and Food.

Professor Madhavan Mukund: 16 Items and Food.

(Refer Slide Time: 05:53)



Professor G Venkatesh: Then we have Suresh. Actually we should count this one item? One and a half because this smokes like KG.

Professor Madhavan Mukund: Unfortunately there is a 0.4 so we can add that.

Professor G Venkatesh: we will count it as one.

Professor Madhavan Mukund: will count it as one.

Professor G Venkatesh: One item.

Professor Madhavan Mukund: Utilities?

Professor G Venkatesh: Utilities? So, is likes like electronics utility batteries.

Professor Madhavan Mukund: we will just I think they were there are some stationary electronics we put it all in one. So, let us put it as I call it utilities but when we come to these later we will decide. Utilities there is one item.

Professor G Venkatesh: It says 2 here.

Professor Madhavan Mukund: 2 Items.

Professor G Venkatesh: Let it be 2

Professor Madhavan Mukund: And Food?

Professor G Venkatesh: 1, 2, here are 4 bananas, is one and a half kilos of tomato.

Professor Madhavan Mukund: So, should we just counted as number of items rather than quantity per item and we should go back and probably change those. Let us go back and quickly change those.

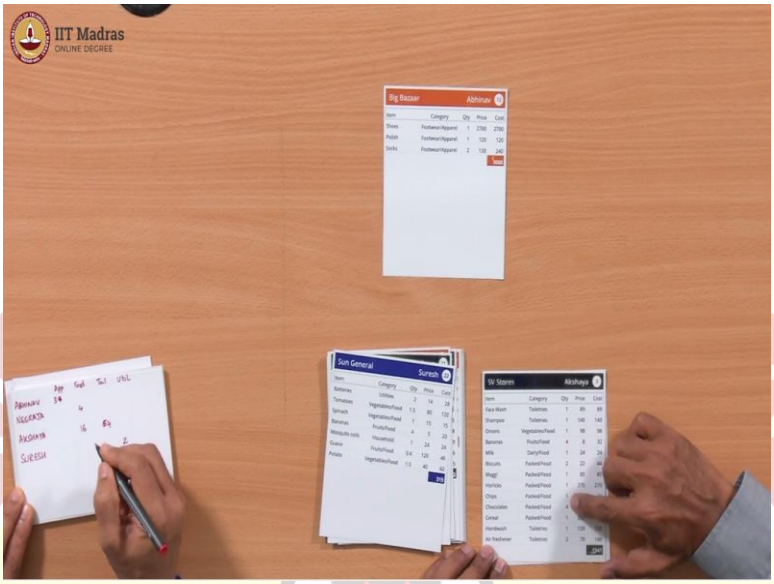
Professor G Venkatesh: So let us start again.

Professor Madhavan Mukund: Abhinav it was 4. We will make it 3, because the…

Professor G Venkatesh: Socks 2 we are looking at 2 socks let us make it 3.

Professor Madhavan Mukund: Let us make it 3.

(Refer Slide Time: 06:59)

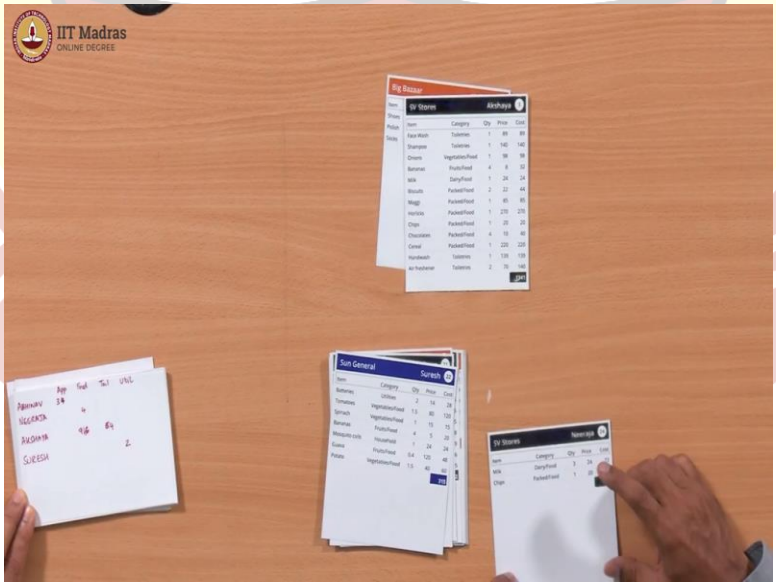


Professor G Venkatesh: Similarly Akshaya.

Professor Madhavan Mukund: So Toiletry is what instead of 5 we should make it 4

Professor G Venkatesh: 4 and 1 2 3 4 5 6 7 8 9 10 no 9 9 9.

(Refer Slide Time: 07:17)



Professor Madhavan Mukund: And Neeraja was

Professor G Venkatesh: 2

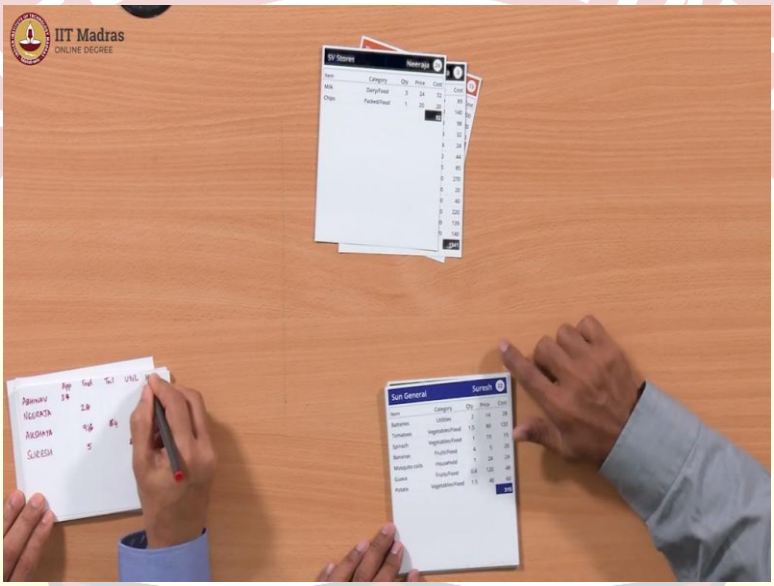
Professor Madhavan Mukund: actually 2 not 4. Now we are counting rows of items.

Professor G Venkatesh: rows of item.

Professor Madhavan Mukund: not the actual count.

Professor G Venkatesh: all right, I think this is getting better I think….

(Refer Slide Time: 07:28)



Professor Madhavan Mukund: so utilities now is only 1

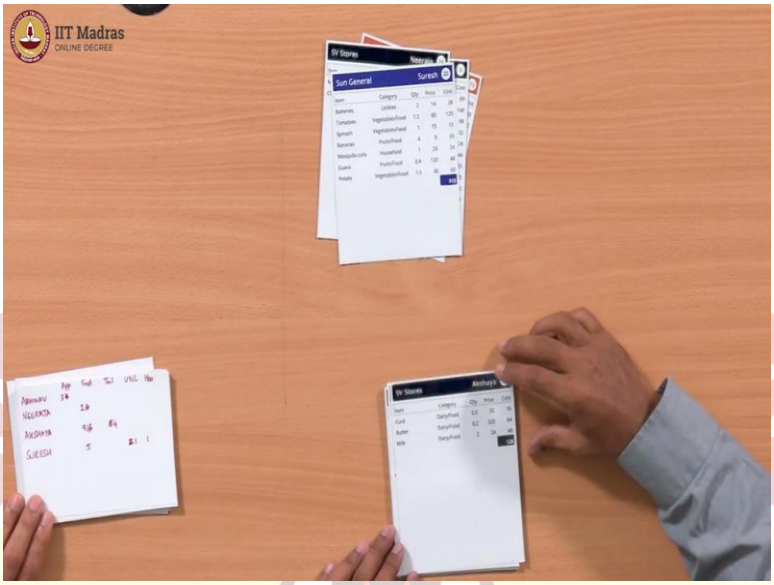
Professor G Venkatesh: Utilities is only 1 and food is 1 2 3 4 5 and there is this thing called household just general things useful for the house I think. Utensils…

Professor Madhavan Mukund: not the quite same as the batteries I think…

Professor G Venkatesh: I think no let us put it in the household 1 item.

Professor Madhavan Mukund: Household Item…

(Refer Slide Time: 07:50)



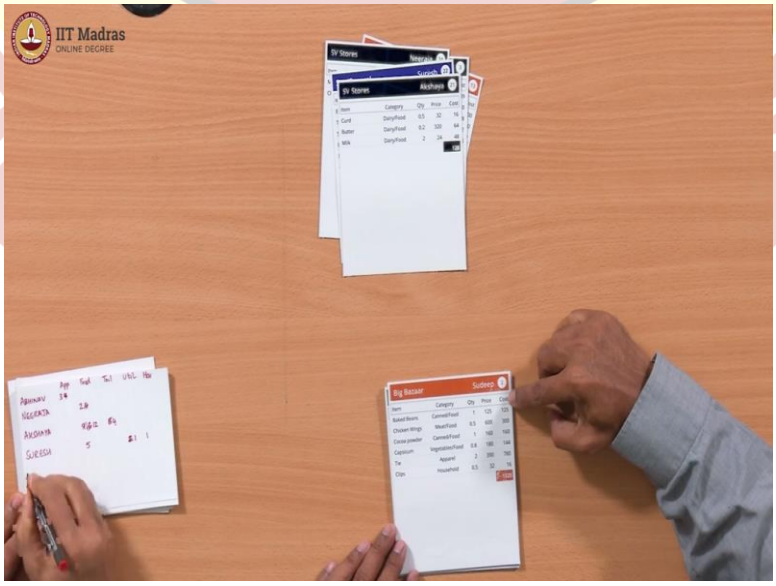
Professor G Venkatesh: Akshaya has 3 items of food.

Professor Madhavan Mukund: So, Akshaya now already had 9 items of food,

Professor G Venkatesh: Already has…

Professor Madhavan Mukund: So we have to just go back and update that. It will be discrepancy but hopefully it will be. 12 items of food.

(Refer Slide Time: 08:04)



Professor G Venkatesh: Sudeep.

Professor Madhavan Mukund: Sudeep is new.

Professor G Venkatesh: New person. 1 2 3 4 food,..

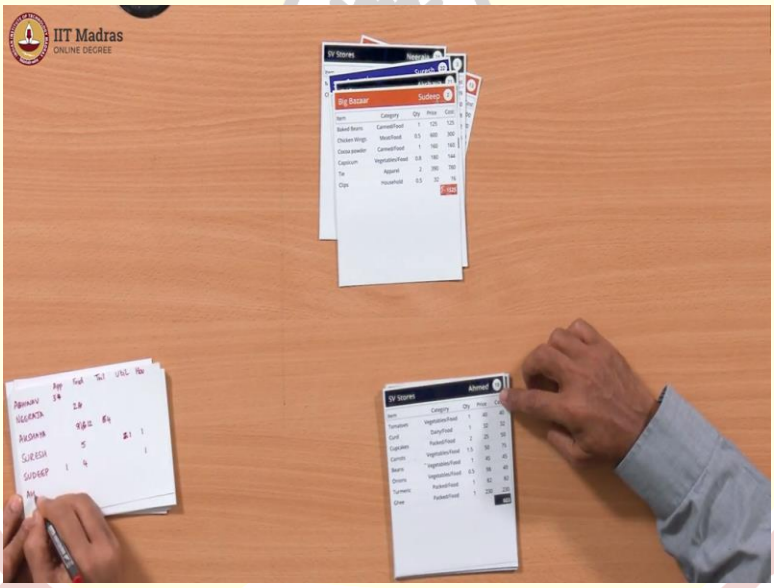
Professor Madhavan Mukund: 4 Food.

Professor G Venkatesh: 1 Apparel.

Professor Madhavan Mukund: 1 Apparel.

Professor G Venkatesh: And 1 household. Ahmed.

(Refer Slide Time: 08:21)

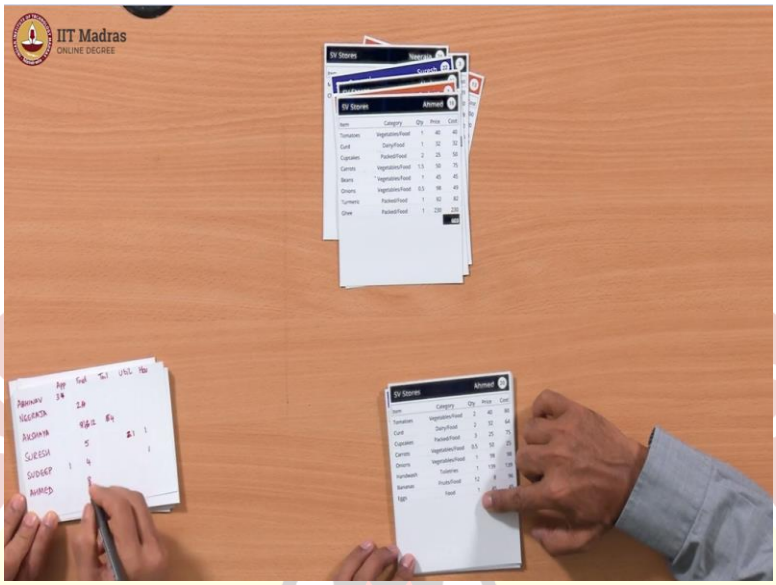


Professor Madhavan Mukund: Again new person.

Professor G Venkatesh: 1 2 3 4 5 6 7 8, food.

Professor Madhavan Mukund: All Food

(Refer Slide Time: 08:29)



Professor G Venkatesh: Another Ahmed, 1 2 3 4 5 6 7, Food.

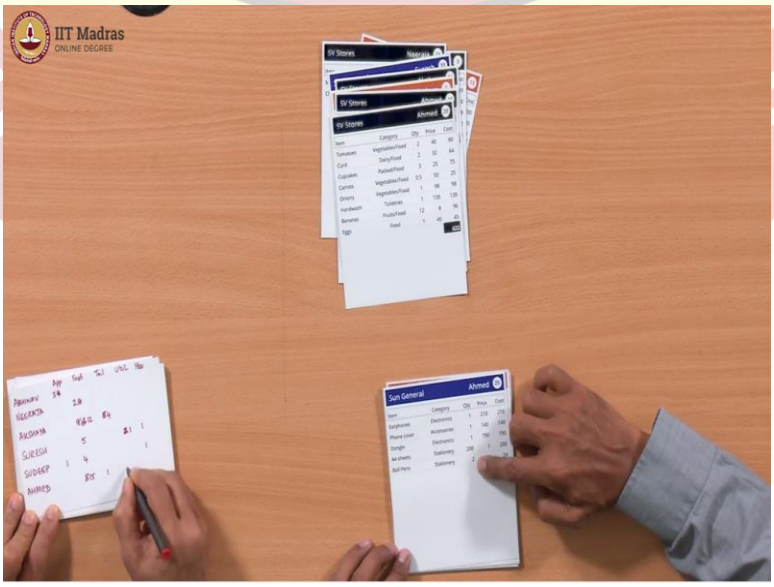
Professor Madhavan Mukund: 7 or 6? 7.

Professor G Venkatesh: 7 Food and 1 Toiletry.

Professor Madhavan Mukund: So 8 was there so 8 plus 7 15 is now the new number for Food.

Professor G Venkatesh: And 1 Toiletry. Ahmed again.

(Refer Slide Time: 08:48)



Electronics which we said will combined with utilities. Accessories also, same thing. So 1 2 3.

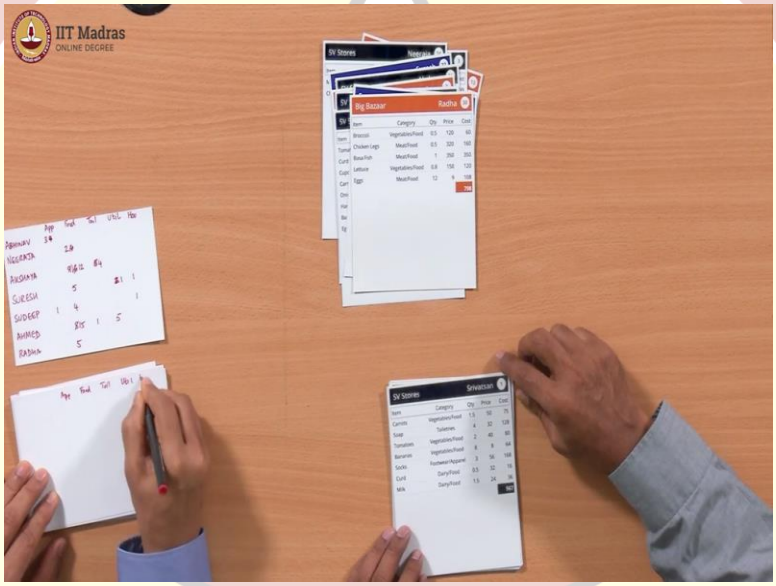
Professor Madhavan Mukund: Stationary.

Professor G Venkatesh: Stationary also we should do the same thing. So 1 2 3 4 5, 5 items.

Professor Madhavan Mukund: 5 items is. So it is a new bill but in the same row, but it is the first step. Let us put it as 5.

Professor G Venkatesh: Radha, 1 2 3 4 5 items of food. Srivatsan.

(Refer Slide Time: 09:22)



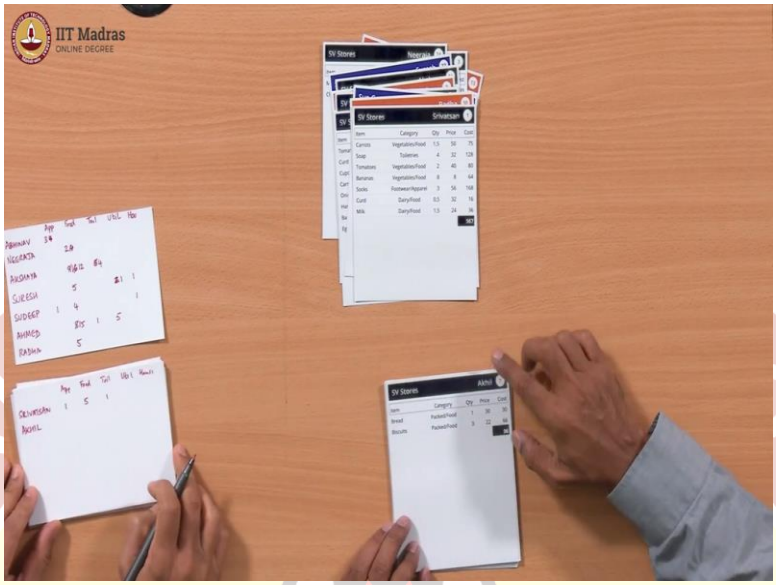
Professor Madhavan Mukund: Let me take a new card. So I will just write this Category. App, Food, Toil, Util,…

Professor G Venkatesh: And Household.

Professor Madhavan Mukund: And House. So this is Srivatsan.

Professor G Venkatesh: Srivatsan 1 2 3 4 5 Food. 1 Toiletry and 1 Apparel. Something is coming looks some pattern is there will see. Akhil.

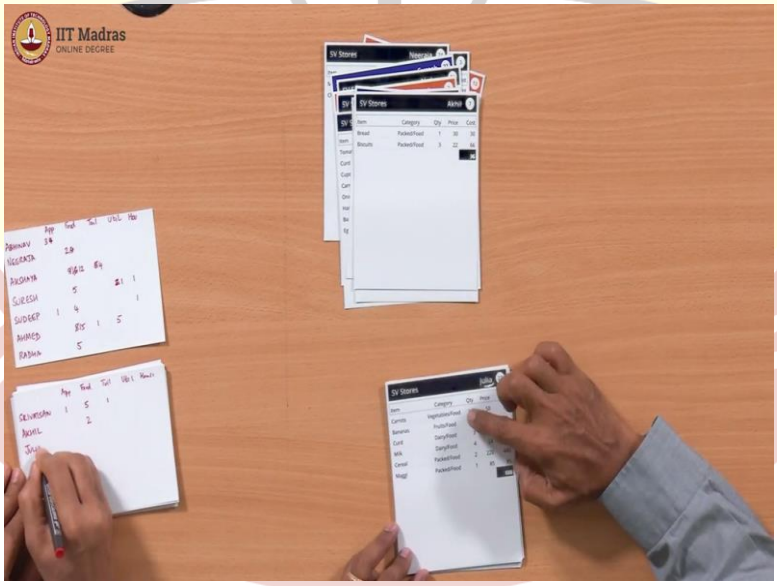
(Refer Slide Time: 09:50)



Professor Madhavan Mukund: New

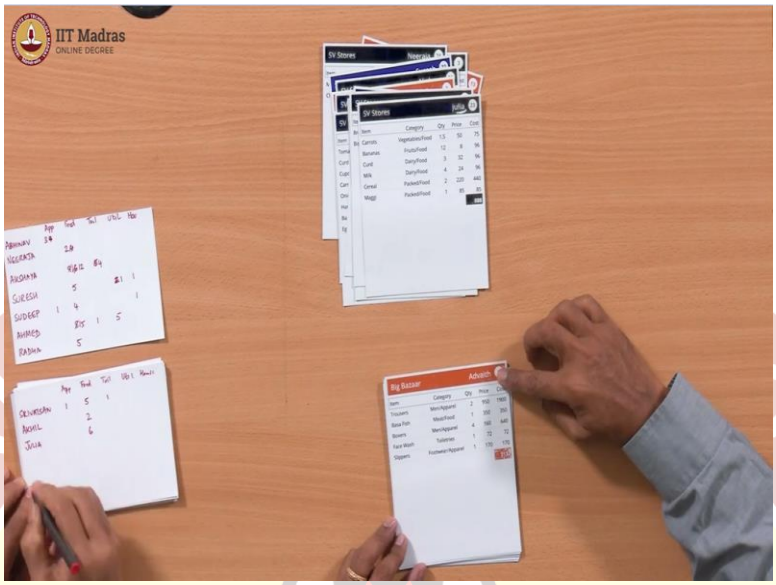
Professor G Venkatesh: Akhil is new? 2 Food. Julia.

(Refer Slide Time: 10:00)



1 2 3 4 5 6 Food. Adwaith

(Refer Slide Time: 10:08)



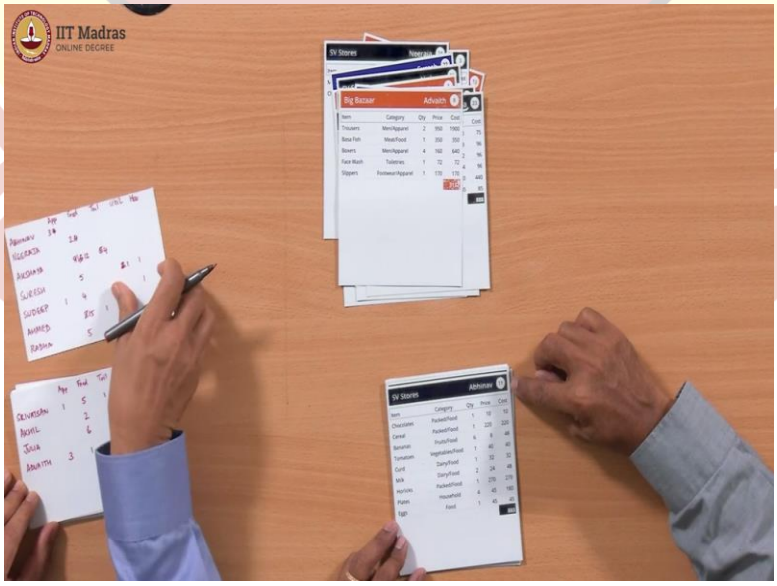
Professor Madhavan Mukund: Also new.

Professor G Venkatesh: We had Abhinav, Adwaith 1 apparel, 2 apparel 3 Apparel, 3 Apparels

Professor Madhavan Mukund: 3 Apparels.

Professor G Venkatesh: 1 Food and 1 Toiletry. Abhinav again.

(Refer Slide Time: 10:23)



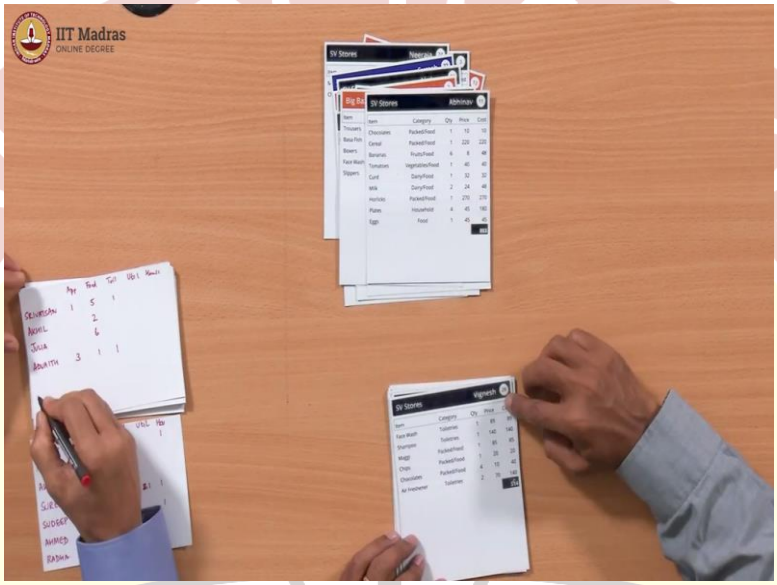
Professor Madhavan Mukund: Again Abhinav.

Professor G Venkatesh: 1 2 3 4 5 6 7 8 Food

Professor Madhavan Mukund: He did not have any Food.

Professor G Venkatesh: and 1 Household. Vignesh.

(Refer Slide Time: 10:39)



Professor Madhavan Mukund: Vignesh is new.

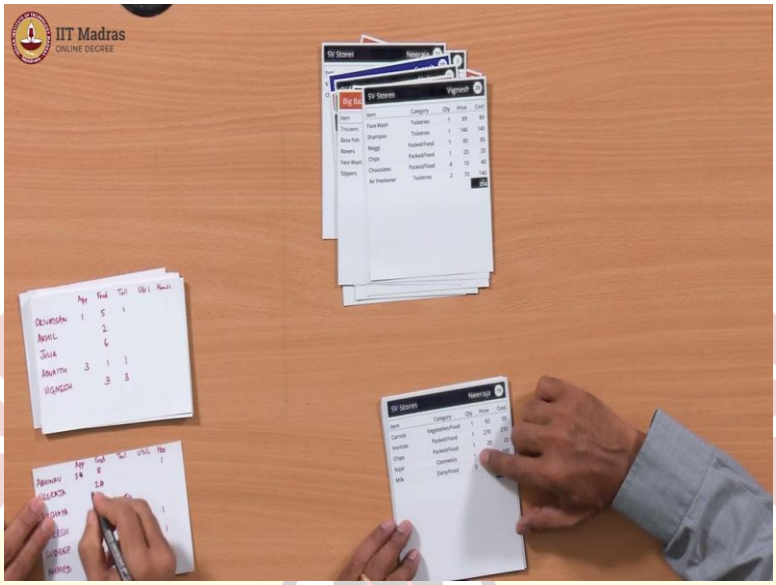
Professor G Venkatesh: 3 Food.

Professor Madhavan Mukund: 3 Food.

Professor G Venkatesh: 3 Toiletry.

Professor Madhavan Mukund: 3 Toiletry.

(Refer Slide Time: 10:50)



Professor Madhavan Mukund: Neeraja.

Professor G Venkatesh: Neeraja we have already seen. 1 2 3 4 Food.

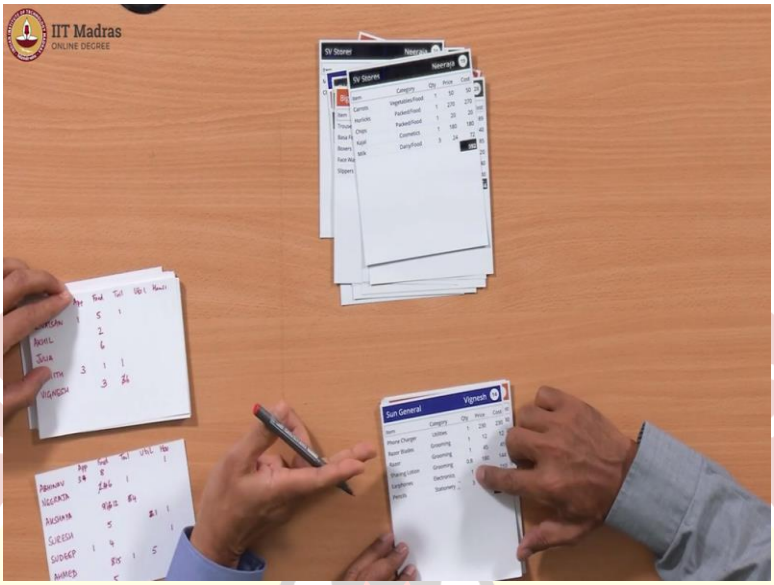
Professor Madhavan Mukund: 4 Food. So she have 4 Food.

Professor G Venkatesh: And 1 cosmetics. Cosmetics.

Professor Madhavan Mukund: Cosmetics and Toiletries.

Professor G Venkatesh: Toiletries, Toiletries and 1 Toiletries.

(Refer Slide Time: 11:04)



Professor Madhavan Mukund: Vignesh , we have an enrty.

Professor G Venkatesh: Utility.

Professor Madhavan Mukund: Utilities is,

Professor G Venkatesh: Utility 1.

Professor Madhavan Mukund: Utility Electronics

Professor G Venkatesh: What is grooming?

Professor Madhavan Mukund: Grooming I think, toiletries

Professor G Venkatesh: Cosmetics, Toiletries.

Professor Madhavan Mukund: So 3 under cosmetics.

Professor G Venkatesh: 3 under toiletries.

Professor Madhavan Mukund: That makes it 6

Professor G Venkatesh: 1 utility.

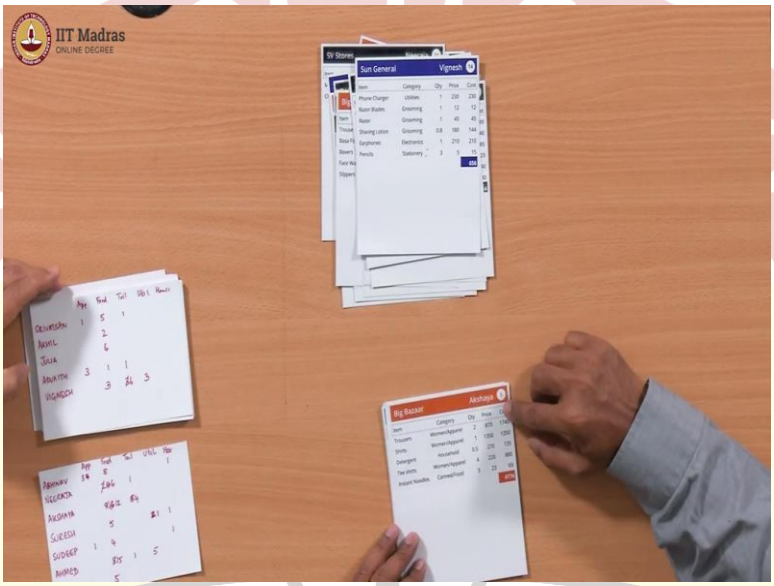
Professor Madhavan Mukund: And stationary is also Utility we say.

Professor G Venkatesh: So 123.

Professor Madhavan Mukund: 3 Utilities.

Professor G Venkatesh: 3 Utilities. Akshaya.

(Refer Slide Time: 11:29)



Professor Madhavan Mukund: Akshaya is already there.

Professor G Venkatesh: Apparel 1 2 3, 3 Apparel 1 Food,

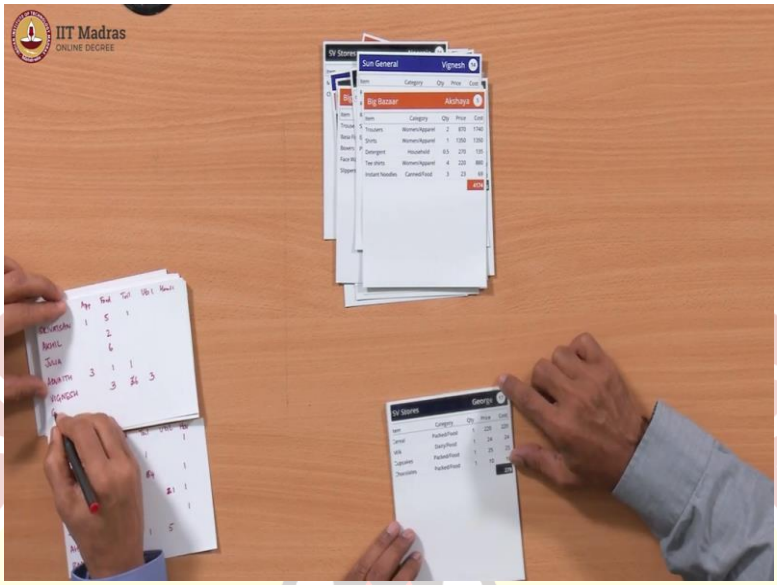
Professor Madhavan Mukund: Food is now 13.

Professor G Venkatesh: And 1 Household.

Professor Madhavan Mukund: Household is 1.

Professor G Venkatesh: I hope something is coming with,…

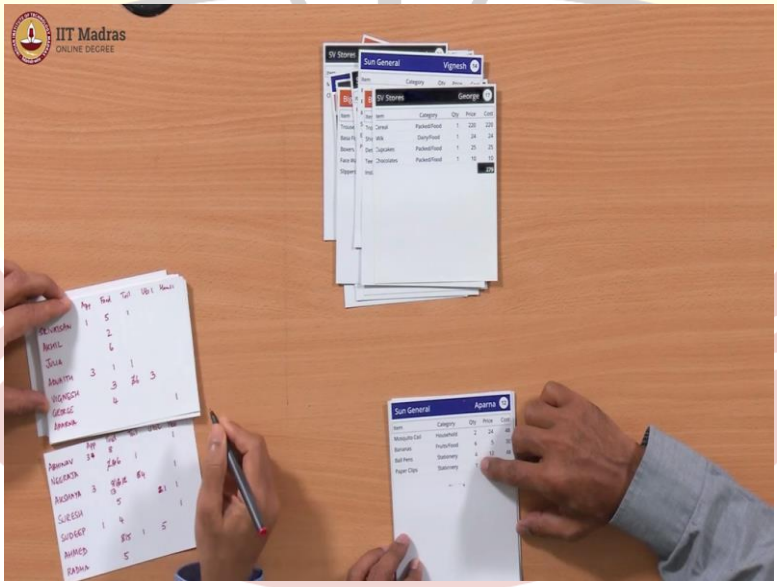
(Refer Slide Time: 11:46)



Professor Madhavan Mukund: George.

Professor G Venkatesh: George 4 Food items. George. Aparna.

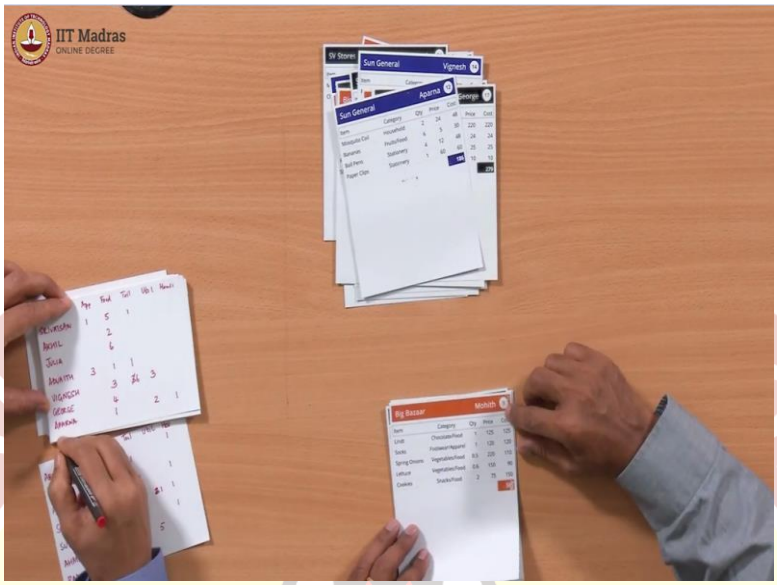
(Refer Slide Time: 11:55)



Professor Madhavan Mukund: Aparna is also new.

Professor G Venkatesh: 1 household, 2 stationary which is Utility, utilities and 1 Food. Mohit.

(Refer Slide Time: 12:09)

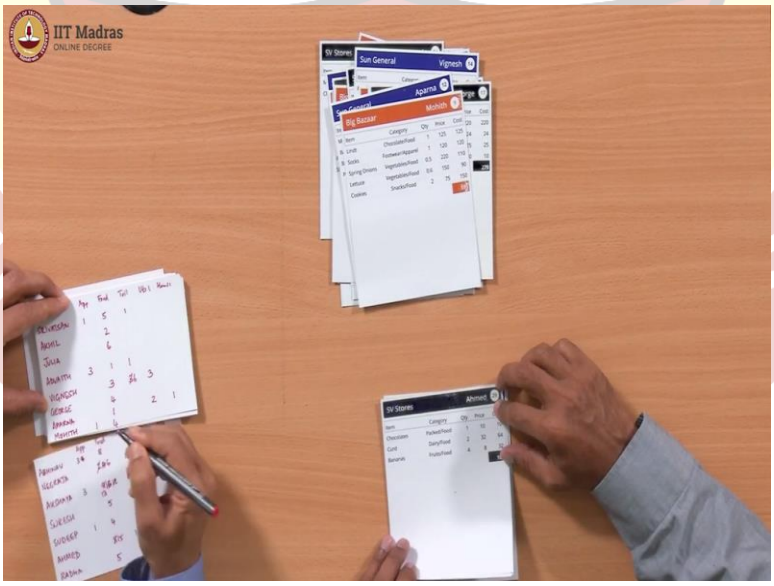


you need another card?

Professor Madhavan Mukund: Mohit can fit.

Professor G Venkatesh: 4 Food and 1 apparel. Ahmed, I think I seen Ahmed.

(Refer Slide Time: 12:19)

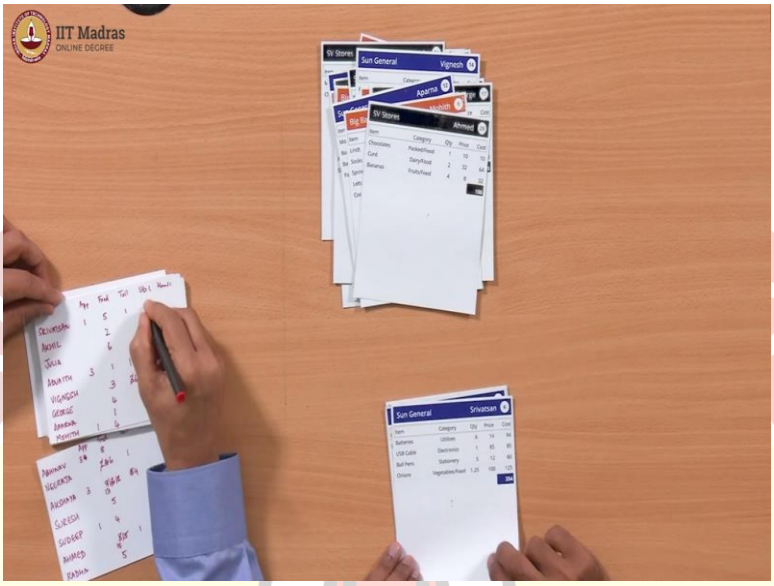


Professor Madhavan Mukund: Ahmed is there.

Professor G Venkatesh: 3 foods.

Professor Madhavan Mukund: Food is now 18.

(Refer Slide Time: 12:26)

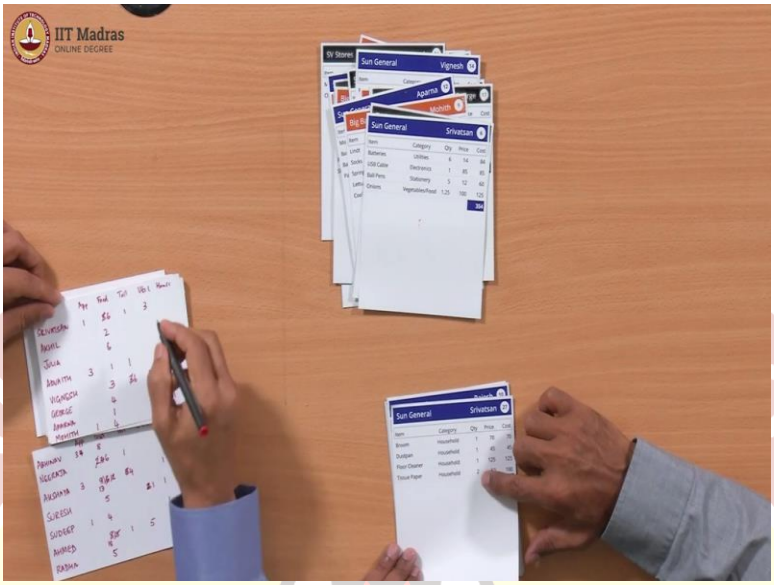


Professor G Venkatesh: Srivatsan, you have seen Srivatsan for sure. 1 utility,

Professor Madhavan Mukund: All those 3 are Utility.

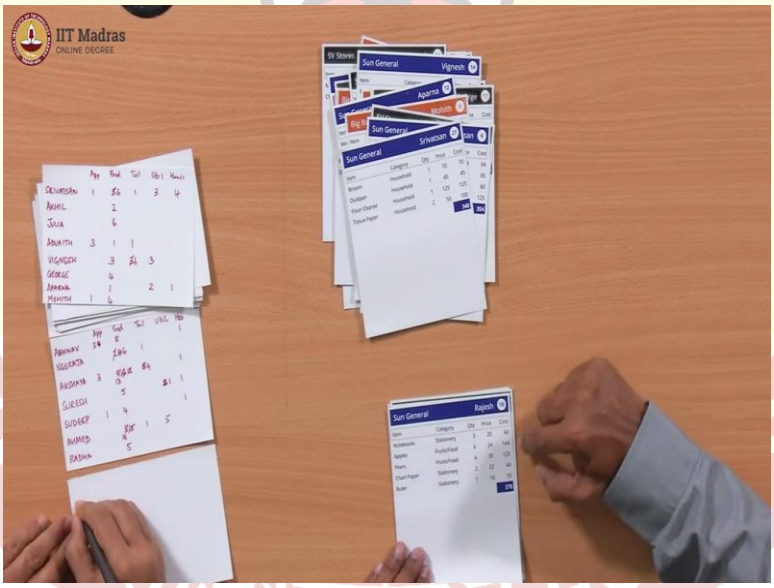
Professor G Venkatesh: 3 utilities and 1 Food. Something some up they have kind of come in we have done a lot of information simplification you have done but if it comes some pattern comes then we can try and confirm. If we get a pattern and we can confirm whether those customers are like or not. At least the first pass we can do.

(Refer Slide Time: 12:52)



Srivatsan you have seen already, 4 Households.

(Refer Slide Time: 12:58)



Professor Madhavan Mukund: Rajesh.

Professor G Venkatesh: Rajesh is 2 Food 3 stationary plus 3 utilities.

Professor Madhavan Mukund: 3 Food and?

Professor G Venkatesh: 3 utilities.

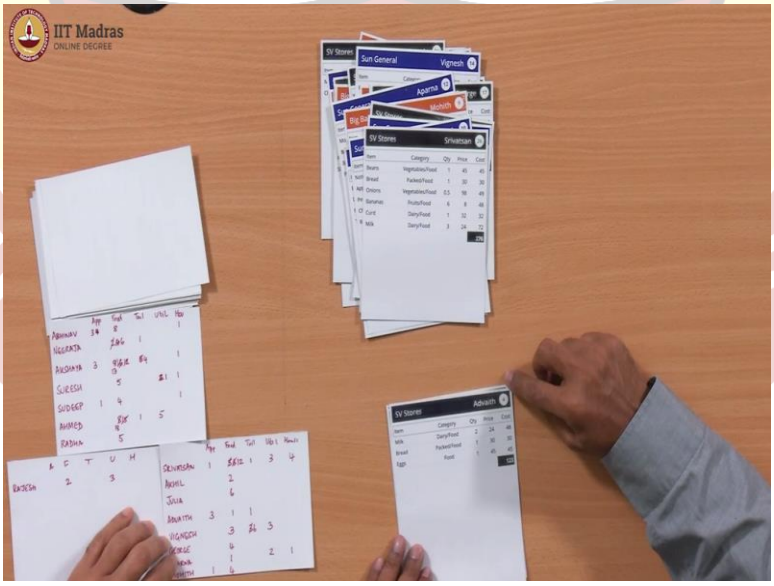
(Refer Slide Time: 13:16)

Professor Madhavan Mukund: Srivatsan again

Professor G Venkatesh: Srivatsan, all Food, 1 2 3 4 5 6 food.

Professor Madhavan Mukund: So that is 2.

(Refer Slide Time: 13:23)



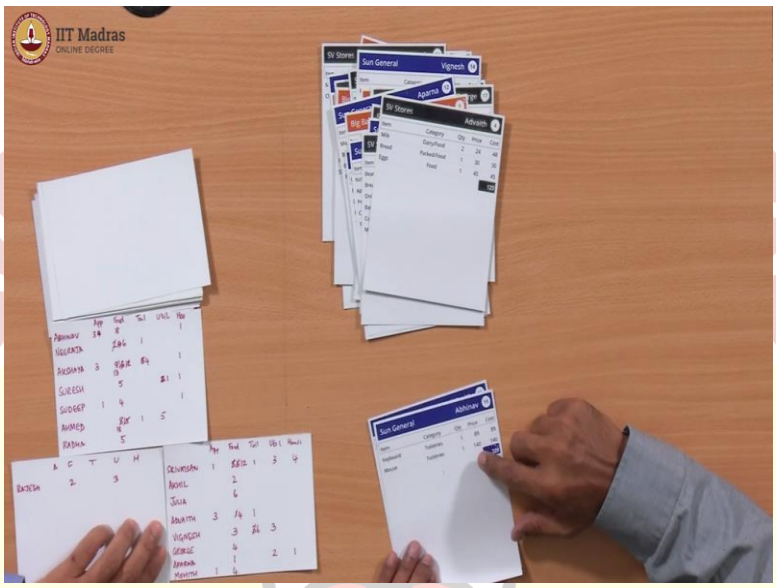
Professor G Venkatesh: Adwaith.

Professor Madhavan Mukund: Adwaith we have.

Professor G Venkatesh: 3 Foods.

Professor Madhavan Mukund: 1 4.

(Refer Slide Time: 13:30)

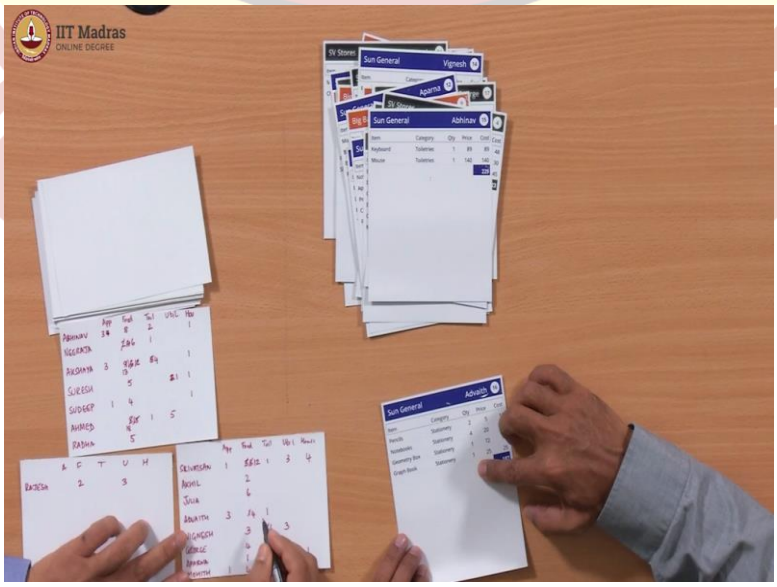


Professor G Venkatesh: Abhinav, 2 Toiletries

Professor Madhavan Mukund: I will add 2 Toiletries.

Professor G Venkatesh: Adwaith is the last card.

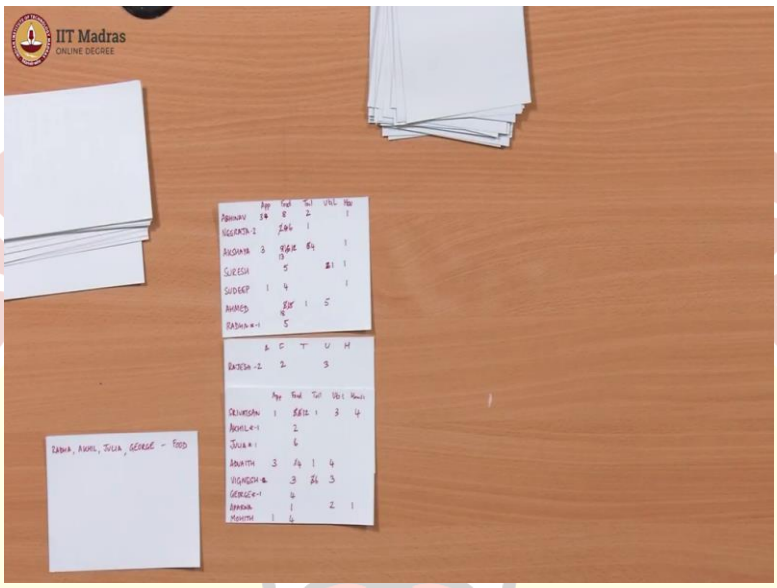
(Refer Slide Time: 13:36)



4 Stationeries.

Professor Madhavan Mukund: Stationary is Utilities.

(Refer Slide Time: 13:43)



Professor G Venkatesh: All right. So we are done with all the cards now. One Pass we did. And these are we processed it right? So we got this, we got this, this kind of, what do you mean? What do you call this a table?

Professor Madhavan Mukund: You can call this a table for sure.

Professor G Venkatesh: So it is basically a table basically is we have got for each customer, the rows are customers? Columns are categories. And the entries are…

Professor Madhavan Mukund: the number of…

Professor G Venkatesh: number of items.

Professor Madhavan Mukund: line items, Bill rows.

Professor G Venkatesh: Bill rows. We could also take the amount but we have taken the rows. It is all right I think…

Professor Madhavan Mukund: So, there are some for instance very obvious cases. So there are people like Radha Akhil, Julia, George only bought Food.

Professor G Venkatesh: So, are they similar?

Professor Madhavan Mukund: So, we could think of them as similar Of course I think in all their cases there was only one bill. So, probably to develop a pattern you need more bills, but and on basis on what we have we could consider them to be similar just by looking at this.

Professor G Venkatesh: So, should we put them in the first group? Radha, Akhil, Julia and George and then we strike it off or, what do we do? So, we know we have

Professor Madhavan Mukund: So, I will small strike it off and maybe I will mark it. So, I will say that these are all only Food. This may be the mark it to.

Professor G Venkatesh: Put a star or something this, this is keeping track is not easier. I think we have to figure out a way by which you can move down this list in a more, nice way. We will see how to do that? So, so we did one pass the remote these 4 people what else? But till now I mean you are looking like no you are looked and you are done something.

Professor Madhavan Mukund: We have to do systematically.

Professor G Venkatesh: Computer cannot just look at look at this table and do something right it has to you have to tell the computer what to do. So, you told it basically in this case you said check whether or not there is only one item and whether right?

Professor Madhavan Mukund: Across the columns only one one item is there.

Professor G Venkatesh: So, in this case Radha, all these people came out. So you would ask it to look for that,

Professor Madhavan Mukund: So maybe now logically 2 items.

Professor G Venkatesh: Let us look for 2 items.

Professor Madhavan Mukund: This is not 2 items. So let us just mark it as one. So, that we know that this is why we put this case. This are one item now.

Professor G Venkatesh: one item with the same item,

Professor Madhavan Mukund: Same item only one item across all bills. Only one type of item across all bills.

Professor G Venkatesh: All others have more than one right .we have to take that first. It could have one item or something else?

Professor Madhavan Mukund: Something else but I think everybody, is sort of. So, now if we look at 2 items, so Abhinav is 4 categories, that he has bought but Neerja has only 2.

Professor G Venkatesh: Food and Toiletry.

Professor Madhavan Mukund: She has got 2 items. Akshaya has got 4, Suresh has got 3, Sudeep is 3. So now we are doing systematically okay. Ahmed got 3, Food is in that and Radha we already has got it at. Rajesh has 2 items.

Professor G Venkatesh: Is it the same two items as Neeraja?

Professor Madhavan Mukund: For instance as not Neeraja No. Abhinav has got food and toiletries, Rajesh, Food and Utilities. They are not the same. So, it is not so easy. Srivatsan is multi item, Adwaith is multi item. Vignesh is a 2 item person,

Professor G Venkatesh: 3? 3 items.

G Venkatesh:. Aparna is, Mohit is 2

Professor G Venkatesh: is the same? No,

Professor Madhavan Mukund: Its Apparel and Food. So all of them are Food and something but not the same?

Professor G Venkatesh: Should we consider them similar? They in a sense, right the mostly buying food the maybe buying that way? The are food type people they went to the shop to buy food but they also bought an accessory. They bought a stationary item or something else. One item somewhat similar.

Professor Madhavan Mukund: Correct, but just for argument's sake. So, so for instance, if we look at Ahmed. And we look at Rajesh, Ahmed is a 3 item person, this is a 2 item person, okay,

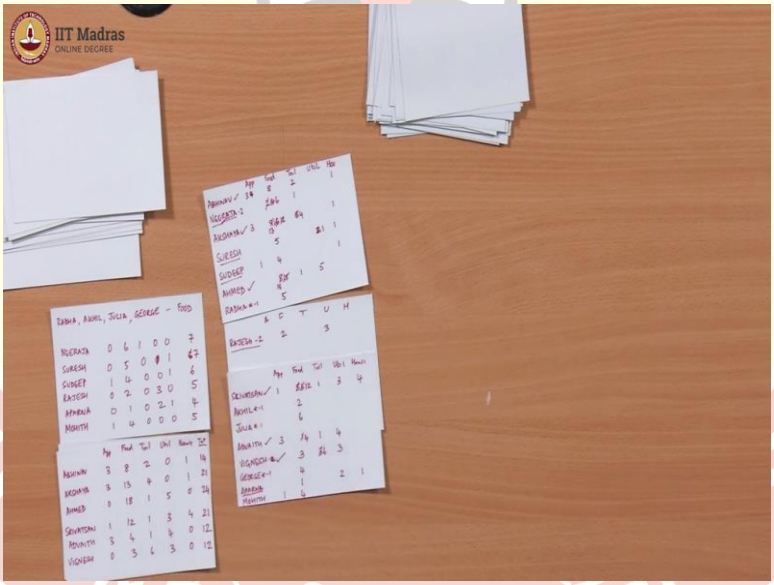
but the overlap on the 2 items and Ahmed has got 1 toiletry item which takes, maybe this is not too

Professor G Venkatesh: But Ahmed has bought so much Food. Alright, I understand what he is saying.

Professor Madhavan Mukund: So, so clearly, this is something we have to, from our perspective, understand what we want to capture. So, do we want to look for an overlap in items? What we want to look are other way of thinking about it is that people who are buying only one type of thing probably are less interesting, or one or two type of things, their patterns are less interesting for the…

Professor G Venkatesh: So should we go the other side and look at people are buying large number of items, total, total number of items,

(Refer Slide Time: 18:28)



Professor Madhavan Mukund: A large number of items. So, we will just make a separate table and just note things down. So, we have 4 columns. App, Food. So Abhinav 1. Clearly and it.

Professor G Venkatesh: 3 8 2.

Professor Madhavan Mukund: 3 8 2. I will write 0 for now so that we know.

Professor G Venkatesh: then we can keep a total so you know, that is why. So, total is the total is 3 8 11 2 13 or 14 items. So, we have basically collecting all those people which which have let us say more than 10 by more than 10.

Professor Madhavan Mukund: Neerja we will leave out for now.

Professor G Venkatesh: Neeraja is not above 10.

Professor Madhavan Mukund: Akshaya.

Professor G Venkatesh: Akshaya has got 3, 13, 4 and 3, 13, 4, none and 1..

Professor Madhavan Mukund: So again, this is 17, 20, 21.

Professor G Venkatesh: Then Suresh is small, Sudeep is small, Ahmed is,

Professor Madhavan Mukund: Ahmed is more than there. So it is a question of Totals now and…

Professor G Venkatesh: 0, 18, 1, 5, 0

Professor Madhavan Mukund: And it is 19 24.

Professor G Venkatesh: 24 items. So, these are people are buying large number of things right. There is nothing to see you that you can compare them to each other.

Professor Madhavan Mukund: Rajesh is small, Radha is small,

Professor G Venkatesh: Srnivatsan 1, 12, 1, 3, 4,

Professor Madhavan Mukund: 13, 14, 17, 21.

Professor G Venkatesh: so this is the kind of similar, kind of looks like that right there are buying large amounts of food,

Professor Madhavan Mukund: These 2 for instance are buying significantly,

Professor G Venkatesh: So. they are more similar than the others,

Professor Madhavan Mukund: But then there is nothing on but yeah let us see. Then Akhil we skip, Julia we skip, Adwaith is got certainly more than 10 items.

Professor G Venkatesh: So, let us keep up like that. 3, 4, 1, 4, 0.

Professor Madhavan Mukund: So that 12.

Professor G Venkatesh: Adwaith, Abhinav and all set to be in all into one cluster.

Professor Madhavan Mukund: Vignesh is also more than 10. Let us skip 10 as our base line. So Vignesh,

Professor G Venkatesh: 0, 3, 6, 3.

Professor Madhavan Mukund: So this 12.And now George is small upper nice 4 items with his files. So now we have these say 6 customers so,,,

Professor G Venkatesh: So, who so you know when so who are we left with? So if you leave that then we will be left with a left Suresh, Sudeep.

Professor Madhavan Mukund: Neeraja also. So the 2, so these guys we have ticked off. So Abhinav is been moved out, Akshaya is moved out, Ahmed is moved out and then Srivatsan is moved out, Adwaith is moved out, Vignesh is moved out and George this one, one item people have left out. so the people who are remaining a Neeraja, Suresh, Sudeep, Radha was a one item, Rajesh, Aparna and Mohit. So, these are the people who are, I just write them below this yeah okay. So we have Neeraja.

Professor G Venkatesh: So that we do not have to look at this cards again. Is in it, kind of done with…

Professor Madhavan Mukund: Neeraja and Suresh, Sudeep,

Professor G Venkatesh: Rajesh, and Aparna,

Professor Madhavan Mukund: And Mohit. Let us write on their things. So, it is 0 6 1 0 0 7 and Suresh is 0 5 0 1 6, 1 1. Sudeep is 1 4,

Professor G Venkatesh: 0 0.

Professor Madhavan Mukund: 0 0 1 that is 6. Rajesh is 0 2 0 3 0 that is 5. Aparna is 0 1 0 2 1 that is 4. And finally Mohit is 1 4 0 0 0 is 5.

Professor G Venkatesh: So again see what we have done we have binned.

Professor Madhavan Mukund: We have binned so basically,

Professor G Venkatesh: So, first binning we did is we took all those people who did only one item purchase. Then we did,…

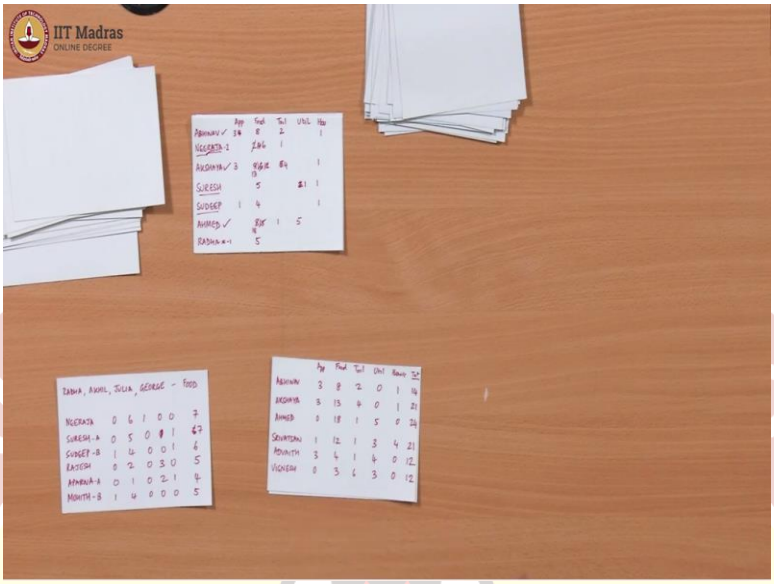
Professor Madhavan Mukund: Because there we did the numbers because there, there was some people with a large number even, we did like 5 items of food.

Professor G Venkatesh: So, we did that then the second what we did is the remaining cards we split into 2 bins, one bin which is basically people are buying large numbers of items

Professor Madhavan Mukund: So to say 10 and above,…

Professor G Venkatesh: So, 10 and above and 10 and below. Clearly there is some pattern right it's there's no point in saying this customer is like this customer. So, if you are comparing we should compare and say this is very logical it seems okay what we have done and this is very, very likely also you will go to find people who come into the shop and buy a large number of things. You are going to find some other people who come to the shop and buy a few number of things it happens actually.

(Refer Slide Time: 23:58)



There are people who come to the shop to buy only one thing that always happens right.

Professor Madhavan Mukund: very focused, very focused on wandering around looking at what is interesting.

Professor G Venkatesh: They spend time in the shop they spend an hour in the shop buy a lot of things. Some guy will want to buy only one spoon or something they will go to the shop buy that spoon and go out. So, within this we can compare all these people all alike?

Professor Madhavan Mukund: One thing is for sure everybody buys food, Food that is true does not discriminate between so then the non-food well I guess you have Neeraja is the the only one who has bought cosmetics in this group. This utility and household there are couple of people who have bought. Suresh and Aparna. So, I do not know whether we want to make such see so Suresh and Aparna have bought the same group of things.

Professor G Venkatesh: So, Suresh Aparna are similar I think. it looks like actually, except that Suresh is buy more food than Aparna and other than that they are similar in terms of purchasing.

Professor Madhavan Mukund: The others are all kind of on their own. Now Sudeep and Mohit maybe because Sudeep has got one item of household but otherwise they are identical 1 4.

Professor G Venkatesh: So, Sudeep and Mohit. So Suresh Aparna Sudeep Mohit

Professor Madhavan Mukund: This maybe this groups of gone Aparna and Suresh this category A, may be Sudeep and Mohit on Category B in this group. And now Rajesh and we will just leave them.

Professor G Venkatesh: nothing further to say. Overall we can say all of these are similar.

Professor Madhavan Mukund: I mean, of course, we are not going to find everybody goes into an in there obviously going to be groups which.

Professor G Venkatesh: So, I think all these customers I would say are more or less similar, right because they are buying food, mostly buying food, but also buying one item like that. These are people the first group which we saw this Radha Akhil Julia George went to buy just Foods nothing else.

Professor Madhavan Mukund: Now, of course, these low volume things, I mean, finding patterns, usually one needs more items, and so they have only won once to the shop. So I think there is more meaningful things will come here where they had 2 3 bills per person, then also nobody can find a pattern in this one purchase. Maybe as you said, you had a very urgent requirement, you go to a shop, you buy a light bulb and come on out.

Professor G Venkatesh: So we got two, I think we should treat this as just 2 groups. Food is one group food plus something is one group. Okay. And now we come to this third thing, which has more items, what can we say about this one?

Professor Madhavan Mukund: If we look at columns, where they differ, of course, everybody buys food, everybody seems to buy toiletries. So, these 2 columns seem to be somewhat at least represented in all, almost everybody has bought Apparel 4 out of them have got apparel. So, 4 out of them, I have got utilities and 3 out of them have got household. So, one has to distinguish based on these groups. So for instance, here, if you look at Ahmed and Vignesh, at least…

Professor G Venkatesh: they have the similar…

Professor Madhavan Mukund: zeros in the same column appear either for apparel or household. So, that is one possibility. Or you might want to look in terms of volume that, you

know, these 3 people have got significantly more food items than these other 3. So, there are many different ways one could do this, I suppose. So from the perspective of the shop, I would say for instance, it is fairly clear just looking at this information that Srivatsan and Ahmed are fairly close, because they both have a reasonably high total of food and reasonably high total utilities.

Professor G Venkatesh: So, when it looks like we are computing some kind of distance measure, yes, right? If the item, so whatever we said, if the number of items is large, similarly large, then the distance between the 2 customers on that score on that item is small. And similarly you can take the distance on each of these items, and then you add the distances. You add the squares of the distances. What thing you do? Because this is normally x and y, really x square plus y square square root of 2, find the distance or something like that. So, either you add the actual distances or…

Professor Madhavan Mukund: your column distance you add, take this some, some measure, you take for each column, add it up and add it.

Professor G Venkatesh: Add it up or take the square and add up or something you have to do.

Professor Madhavan Mukund: And the further the more the distance,…

Professor G Venkatesh: the further more the distance the further apart they are. So, ideally want to do like that. But here by examination, they are there. We are trying to figure something out.

Professor Madhavan Mukund: we are saying that these 3 Akshaya Ahmed and Srivatsan are relatively close.

Professor G Venkatesh: similar, because they have high food items.

Professor Madhavan Mukund: And in that Ahmed Srivatsan have even more closer closer because they have the same toiletry items in a high, although Srivatsan has a household thing, which is very different from Ahmed. But in that sense, Srivatsan was the only one who has a significant household thing here everybody else is 0, and 1.

Professor G Venkatesh: And Adwaith Vignesh kind of similar. They buying similar amounts of food, similar amounts of utility, but the differing in this apparel and toiletry. But when you do a distance measure, you will find the distance on this will come out low, distance on this will come out low distance on this will come out high distance on this will come out high. So, when we do whatever edition or square an ad or something we do, you will get some distance measure, it may come out less than the other pairs.

Professor Madhavan Mukund: And similarly, if you look at the first 2 rows Abhinav and Akshaya are actually actually are very similar in that sense. Because on apparael, toiletry and household, they're almost identical. And food, I mean,…

Professor G Venkatesh: utility 0, both household 1, and so the distance here is 0.

Professor Madhavan Mukund: Distance in the first column is also 0.

Professor G Venkatesh: distance here also 0, only in food, you are getting some distance.

Professor Madhavan Mukund: and actually is on the low end of the scale compared to Ahmed. So Akshaya is in the middle Abhinav a little bit below if you think of the average of this group because you have some really low once 3 4 that is a really high like 18.

Professor G Venkatesh: So actually, if you want to pair these 2 are very similar, these 2 are very similar, these 2 seem very somewhat similar, right, something like that.

Professor Madhavan Mukund: So, if you we are doing it systematically we will take each one of these and compare it with everything due to distance with everyone everything else and use that distance as a criterion for saying so here we found groups of two, we might have found groups of 3, depending on how we wanted. We said we could actually have also thought of Akshaya Ahmed and Srivastan as one group, if we emphasize the food as the most significant. That is another thing that distance need not be equal in all, all directions.

Professor G Venkatesh: So, we did find and if nothing, we can at least say that there are 3 categories, this category, and this category in this category. This category is one item buy only food item, people, these are people who buy food plus a little bit of something else. And these

are people who go and buy a number of things, many things. And within this you could do further fine grain stuff, all right.

So not bad, I mean, for, for this data, we have found something some pattern and now if shop wants to basically target these high purchase people, they could create offers for right, if more than 500 rupees you get something free. Whatever it is, so they encourage them to become buy more than 500 rupees or buy a large number of items.