Course Name - Operations and Revenue Analytics

Professor Name - Prof. Rajat Agrawal

Department Name - Department of Management Studies

Institute Name - IIT, Roorkee

Week - 07

Lecture - 35

Welcome friends. In our last session, we discussed the concept of overbooking. We know that there are chances of no-shows in various service systems. Passengers may book their tickets in advance but may not come at the time of actual service. Similarly, in many other cases, there needs to be a concept of overbooking so that you can minimize losses or, in other words, maximize the revenue generation from your limited capacity. Now, whenever we are doing overbooking, there is always a risk that all the customers who have booked the service might come. If the customers who actually come to avail the service exceed the capacity of the server, then there will be some kind of risk. Therefore, in our last session, we discussed that if more customers come than the capacity, you will not be able to provide service to some of them.

You are going to deny service. Denial may be voluntary; denial may be involuntary. But whether it is voluntary or involuntary denial, there will be some cost you are going to incur. So, there is a cost of denial, and there is a possibility of some revenue if you are going for extra booking or overbooking. So, obviously, a very simple issue is that you should book any additional seat if, by booking it, the expected revenue is more than the cost of denial of service.

If the cost of denial is higher, you should not overbook the services. Now, for that purpose, we discussed in our last session that there are different types of policies under which you can consider overbooking. The question is, if your capacity is capital C, how much overbooking should you do? Therefore, we discussed that there are four possible ways. One is the deterministic way of overbooking. We discussed this with the help of an example in our last session. Then, two very interesting policies are there.

One is a risk-based policy, and another is a service-level-based policy. And then we also have a hybrid policy, which combines the risk-based and service-level policies. Deterministic heuristic—we discussed that you have historical data, and based on that historical data, you know that, out of all the customers who are booking your service, let's say 80% are actually appearing, 90% are appearing. Whatever the historical trend is, you will decide accordingly for overbooking. Here, we discussed a parameter row, which is the show rate. So, how many customers are actually appearing will decide the overbooking amount.

So, based on your historical data, you should see what is generally the case for hotels, airlines, and various other car rental companies, etc. They have data like in the airlines—it is mentioned in various research papers that generally, the show rate is between 85 to 90 percent. So, around 10 to 15 percent of passengers do not come at the time of the flight. So, taking advantage in the simplest manner, a large number of organizations are doing overbooking using this simple deterministic heuristic. So, here the basic aspect is that we have a deterministic approach, where we have a well-defined show rate, and based on that, we discussed one example in our last session as well.

Now, in this particular session, we will focus more on this risk-based policy, where we understand that doing overbooking may involve some kind of risk because you have to estimate the cost of denied service and then weigh these things against potential revenue. So, if the potential revenue from overbooking is more than the cost of denied service, overbooking is justified. We will see how we can develop a decision-making rule in the case of risk-based policies based on these two important parameters. If the cost of denied service is more, we will not go for overbooking in such scenarios. The other important approach, which we already discussed in our previous class, is the service-level policy. Like in our inventory management classes, etc., we have also discussed that a very important aspect is we want to give maximum service to our customers. So, it involves managing to a specific target—for example, targeting that no more than one instance of denied service occurs for every 5,000 customers who come to take the service. The meaning is, if every time I have the capacity C, doing the booking B, and out of B, X are appearing.

C is the capacity, b is the booking, and x is those who are showing. Now, for example, companies generally follow this service-level policy in terms of Six Sigma, determining whether they are able to maintain a Six Sigma policy with respect to denied services. That only 3.4 customers are denied service out of a sufficient number—a large number of customers who are appearing to avail your service—or you may have a particular target based on industry standards also. So, you are a very high-level service organization. And then there is a hybrid policy where we are also looking to balance our cost of denied service and possible revenue, while at the same time providing a high service level to the customers who are actually appearing to benefit from the services.

So, these are the four different types of policies. I already mentioned that deterministic heuristic is the simplest one and the most used one also. But, in that case, we are ignoring the service-level issues, and we are ignoring the cost-benefit issues also. So, coming to this risk-based policy. So, under that risk-based policy, the booking limit is set by balancing the expected cost of denied service.

You may remember that we are calling this expected cost of denied service as capital D, which you may recall is the sum total of different types of benefits you are giving to those to whom we are denying the service. For example, either you are giving a voucher for a hotel stay, you are giving a ticket for the next flight also, or there may be a loss of goodwill also. So, all these are put together as your cost of denied service, along with the potential additional contribution from more sales. So, contribution is generally the fare, which is P, and the cost is D. So, we are going to balance these two important things: how much additional contribution you may get by overbooking, and what will be the possible cost in case you deny the service to this overbooked customer? So, the cost of denied services may have different examples for different types of cases, like a very interesting term we get—the concept of gapping—which is actually applicable in the case of broadcasters.

Now, this gapping concept is basically that you are giving a guarantee to your customers that you will get this many impressions if you put your advertisement in one of my television programs or in one of my, let us say, electronic programs. Now, maybe because of some reason, your customer is not getting that number of impressions in one program. Then, you provide this guarantee that I will continue your advertisements until you get your guaranteed impressions, and for that purpose, you have to carry forward that particular advertisement for one or two more programs until you provide sufficient or the guaranteed number of impressions to your customer. So, this is the phenomenon that when you are carrying forward the same advertisement to fulfill the requirement of the minimum number of impressions, that is known as gapping in the media industry. So, that is one example that is going in the case of the cost of denied service, and similarly, in the field of rental cars and others like airlines, hotels, etc., we have examples of the cost of denied services.

Now, how will this function be calculated to develop that balance between the benefit or the revenue and the cost you are going to incur in this particular case? Now, for that purpose, to calculate a risk-based booking limit, we need to specify the objective function that denotes the number of passengers who are showing up at departure by s. So, s is the number of customers who are actually physically available at the time of departure of a particular flight. The number of shows is a random variable that depends on both the booking limit and the total demand for the booking. Now, we are assuming that only those customers who are appearing to actually take the service will pay you. So, each customer who is appearing is paying you a price of p. So, the total revenue will be p multiplied by s.

This is the total revenue because s customers are coming, and p is the price. Now, if shows exceed capacity, if your capacity is C, then the supplier must deny services to s minus C passengers. s is more than C, so s minus C customers will not be given the services. Now, each denied customer will result in a cost of capital D. So, the total cost of denial, the total denied boarding cost, is 0 if s is less than or equal to C because, in that case, there is no overbooking, and you will not deny any customer. But, if s is more than

C in this particular case, D multiplied by the number of customers denied boarding. This is D, the cost of denial, multiplied by the number of denied customers.

So, D into s minus capital C is our cost of denied working, denied services. So, the net revenue in this case is R, which is the net revenue; ps is the total revenue, and this is the cost of denied boarding.

The risk-based objective function

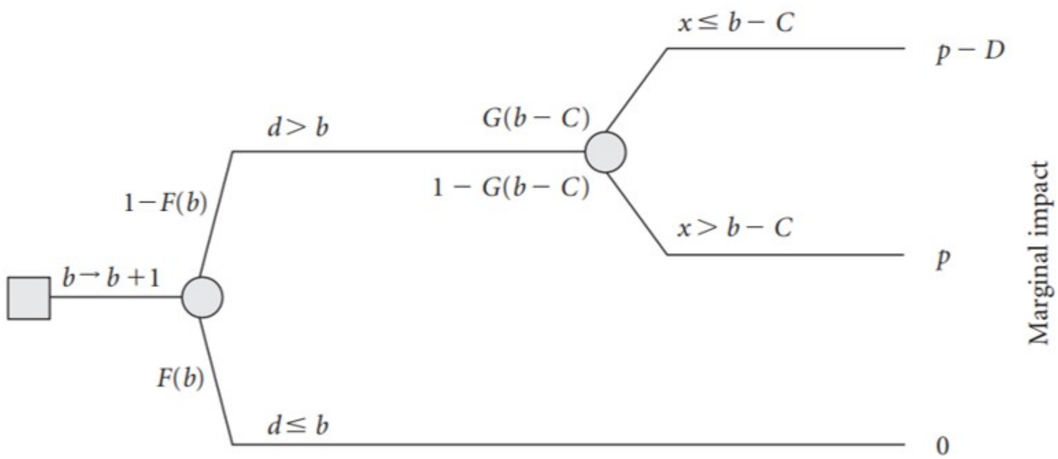
- To calculate a risk-based booking limit we need to specify the objective function. Denote the number of passengers who show up at departure by s. The number of shows is a random variable that depends on both the booking limit and the total demand for bookings.
- Since each show pays price p, total revenue is p x s. If shows exceed capacity, then the supplier must deny service to s - C passengers. Each customer denied service results in a cost of D.

The total denied-boarding cost is 0 if s <= C and is D(s - C) if s > C. The net revenue is then

$$R = ps - D(s - C)^+$$

Here, you see that we have used a plus subscript, meaning the value of D into s minus C can be either positive or 0. It cannot be negative because s minus C is the number of denied customers. So, if s is less than C, this factor will be negative, and it becomes ps plus D into that factor, but there is no meaning for a negative value here.

So, only the positive value or the 0 value will be considered for this cost of denied services. So, if it is there, it is positive; otherwise, it will not be there. Now, after understanding this net revenue function for the risk-based objective function, let us consider the case where I am going to book B seats and want to see if I should have one additional seat. Let us consider this one additional seat as an overbooked seat. Now, there are two broad possibilities.



Here, small d is the demand. Now, if small d is more than V, this additional seat will be used; but if small d is less than or equal to V, then this additional seat will not be used. So, in that case, if the additional seat is not used, there will be no marginal impact—you will not get any additional revenue, and there will be no cost of denial. So, there is no effect in this case, and the probability of this is given by F(b). If I read this F(b), it says that this is the probability that demand will be either b or less than b. Demand greater than b will have a probability of 1 minus F(b).

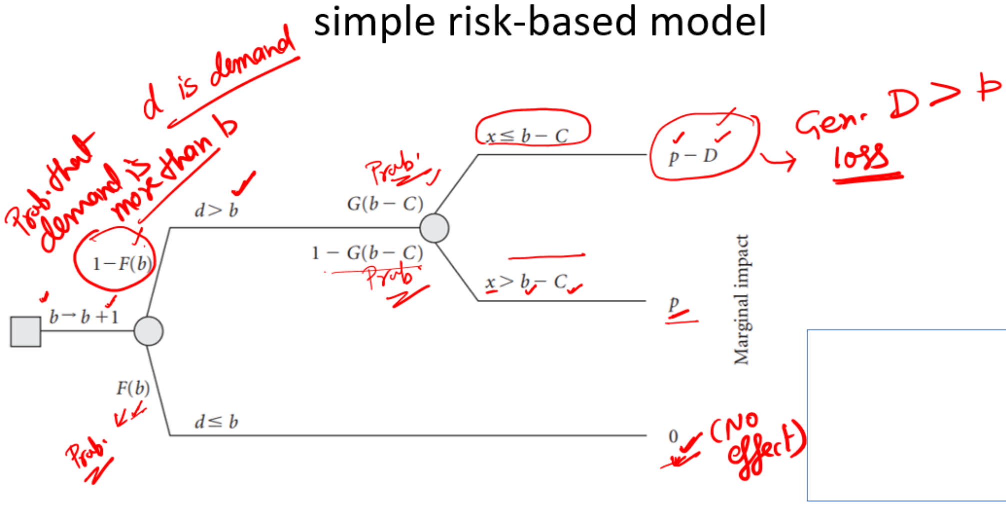
So, that is the probability that demand is more than b this statement. Demand is more than b, this is probability of 1 minus F(b). Now, when demand is more than b, what is going to happen? Then there are further two situations possible. Situation number one that x is greater than b minus C. C is the capacity of our system, b is the booking you are expecting.

So, this additional booking is going to give you a revenue of p items, the revenue of p is the price that you will get, if x is b minus c. Other case is x is less than b minus c, x is less than b minus c. Now, in this case, you have to deny somebody the booking, now in this case the revenue with for this additional customer is coming as p and then because, you are denying somebody the cost of denial the service will also come into the picture p minus d. And generally d is more, generally d is more than p. So basically this is loss, this term is representing basically the loss p minus d. Otherwise just by a visual inspection may not look that it is a loss rather you may see that the benefit is less than p but actually D is more than p. So, therefore, this outcome of this term will be negative and that indicates loss. So, if d is more than b, if d is more than b, there may be two possibilities either you have the benefit of extra revenue that is coming as p or if you are exceeding your capacity in that case it may lead to a loss of p minus d.

Now, if demand is less than equals to b minus C, the probability is G(b - C), this is the probability and demand will be greater than b minus c, this will be the probability with 1 minus G(b-C). So, these are also the probabilities. So, in this way, you have a decision tree for developing this entire scenario, that if I am going to book one extra seat, what will be the different possible impact in marginal revenue case? Either, there will be no impact or there will be a impact of extra revenue by one extra customer or there will be a possibility of loss of p minus D. So, these are the three possibilities which may happen in case of booking the additional seat. So, the total marginal impact will be the impact multiplied by corresponding probabilities that will be the total marginal revenue expected revenue for increasing the seat from b to b plus 1 and that is the decision rule. Now, for simplifying rather going deep into the mathematical calculation, infact you are welcome to do this calculation that this probability, this probability multiplied to this ,case this probability and then this probability multiplied by this.

And then, this multiplied by this is in fact 0, which is our expected revenue from this particular scenario. So, if I simplify this particular case, the practical way of implementing this whole risk-based policy is that we initialize b equals to C, the booking which we are going to accept, which is equal to the capacity of our service.

Decision tree to determine optimal booking limit in the simple risk-based model



Now, whether you are going to take an additional seat, that is C plus 1 or b plus 1, is the same thing. Now, what is the optimal number for you to decide? This becomes our decision rule. The development of this relation is coming from this decision tree only. So, rather than going into the mathematics, we are directly coming to the use of this particular decision rule for us.

So, you need to calculate these two terms: one is p by D, where p is the price revenue and D is the cost of denied services. That is one ratio, and you have to compare this ratio. For the benefit of all, this is price divided by the cost of denied service. So, this is one ratio, and you have to compare this with the probability of demand being equal or less than b minus C. So, you have to see these two comparisons: price upon cost of denied service and the probability of demand being equal or less than b minus C. Now, if the left-hand side is less than or equal to the right-hand side, you stop, and the current value of b is optimal. Whatever your current value is, the current value of b in this case is the capacity of the system. But, in case the left-hand side is more than the right-hand side.

Then, you can increase b to b plus 1 and then again go to step 2. In this iterative manner, you can reach that level where you will finally get the left-hand side less than or equal to the right-hand side, and that will be the optimal value of b. So, that will decide how much overbooking you can do in this particular case.

A Simple Risk-Based Booking Limit Algorithm

1. Initialize $b = C$.
2. If $p/D \leq G(b - C)$, stop. The current value of b is optimal.
3. If $p/D > G(b - C)$, set $b \leftarrow b + 1$ and go to step 2.

You can confirm that this algorithm will find the smallest value of b for which p/D <= G(b - C). We can summarize this as:

The optimal booking limit in the simple risk-based model is the smallest value of b for which p/D <= G(b - C)

So, after implementing this, this will get you the smallest value of b, the smallest value of b for which p by D, the left-hand side, is less than or equal to the probability of demand being equal or less than b minus C. So, in this case, we will see the implementation with the help of a very simple example, and this example gives you the implication of the rule we just discussed. We are considering a flight with 100 seats, the fare is 120 dollars, and we have given you the cost of denied boarding, which is 300 dollars. Now, in this case, P by D is 120 by 300, which is 0.4.

Now, assume that no-shows follow a binomial distribution with p equals to 0.42 and n equals to 20. This means that the expected number of no-shows is 8.4. Now, the probability G(b-C) is then as shown in the third column of the table which is there on our next page.

b	$b - C$	$G(b - C)$	Passenger revenue	Expected DBs	Expected DB cost	Expected net revenue
100	0	0.00	$10,992	0.00	$0.00	$10,992
101	1	0.00	11,112	0.00	0.01	11,112
102	2	0.00	11,232	0.00	0.09	11,232
103	3	0.01	11,352	0.00	0.73	11,351
104	4	0.03	11,472	0.01	3.78	11,468
105	5	0.09	11,592	0.05	14.25	11,578
106	6	0.20	11,712	0.14	41.91	11,670
107	7	0.35	11,832	0.34	100.68	11,731
108	**8**	**0.52**	**11,952**	**0.68**	**204.52**	**11,747**
109	9	0.69	12,072	1.20	361.39	11,711
110	10	0.83	12,192	1.90	569.46	11,623

So, these are the different levels of b. You see the capacity of the aircraft is 100. So, we are starting b equals to C. So, that is b equals to that is the starting point and you see our column number 1 that we are doing increment of 1 with every step.

So, b minus C is 0 and then the probability the difference of demand b minus C is 0 or less that is coming with the binomial distribution and in this case the passenger revenue and all other cost etc and the expected net revenue are calculated in the further columns. Now, if I increase the 100 to 101, 100 to 103 and keep on increasing one by one at 108, when b minus C is 8 and the probability is 52 percent that the difference of booking and the capacity is the probability is 8 or less than 8 that probability is 52 percent. The passenger revenue is coming 11,952 and the expected cost of derived boarding in this case is 204 and expected net revenue is 11747 that is the highest in this particular table. And this will give you a very good idea that if you see in the form of a graph that as you are booking let us say this is your capacity 100 as you are going up to 108. Your revenue is increasing, expected net revenue and after that your expected net revenue is declining.

So, as you are further increasing your capacity, your expected net revenue is declining. However, you see interesting thing is as you are continuously increasing your so, let me have this. So, this is total revenue, this is also total revenue but this is net revenue. That means your net revenue will start declining after a particular time and your total revenue may still increase, because you see that this is continuously increasing. But, because of your cost of denied boarding is started playing important role from this point onwards or let us say from this point onwards.

And therefore, your net revenue is started declining from this point onwards and this is the highest net revenue that is 11747. So, before that therefore, if you see up to this point. 11,352 and 11,351 your total revenue and net revenue are almost similar, but as you are going for more overbooking now there will be chances of some denied boardings. So, therefore, the cost of denied boarding is started appearing. It was almost negligible before this level and as you are going with higher and higher overbooking, the cost of denied boarding will also become significant over a period of time. So, therefore, in this particular scenario going by this table we see that the 11,952 is the level at which you should have the passenger revenue highest.

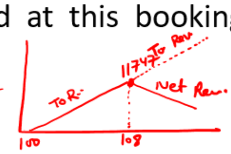
Now, without going for this entire calculation, this is not needed. This is just for showing you that how things are happening. We just need to calculate this column. This is the important column. Our probability of p by D, the ratio is 0.4.

Now, in this column where we have the probabilities of b minus C that how much overbooking we have done after 0.35, we should have the label equal to 0.4 or immediately higher label than 0.4. So, after 0.35 the next value is 0.52 which is higher than 0.4. So, this becomes our optimal level of overbooking. So, these four columns which are there to just justify that the level at which we are correcting our overbooking that the overbooking of 8 how it is leading to highest level of expected net revenue. So, in our all practical purpose on a very fast move you need not to go for all these remaining calculations, you just need to know that what are the values of Gb minus C probabilities like, in this particular case it is given with binomial distribution may be it may follow some other distribution also.

Example

Following the simple risk-based booking limit algorithm, we find that b = 108 is the first point at which G(b - c) >= 0.4 and that the optimal total booking limit is thus 108. This can be confirmed by reference to Table, which shows that the maximum expected revenue is indeed achieved at this booking limit.

b	$b - C$	$G(b - C)$	Passenger revenue	Expected DBs	Expected DB cost	Expected net revenue
100	0	0.00	$10,992	0.00	$0.00	$10,992
101	1	0.00	11,112	0.00	0.01	11,112
102	2	0.00	11,232	0.00	0.09	11,232
103	3	0.01	11,352	0.00	0.73	11,351
104	4	0.03	11,472	0.01	3.78	11,468
105	5	0.09	11,592	0.05	14.25	11,578
106	6	0.20	11,712	0.14	41.91	11,670
107	7	0.35	11,832	0.34	100.68	11,731
108	**8**	**0.52**	**11,952**	**0.68**	**204.52**	**11,747**
109	9	0.69	12,072	1.20	361.39	11,711
110	10	0.83	12,192	1.90	569.46	11,623



So, accordingly, you may check what the probability should be, and accordingly, you will see that this is the overbooking number. Now, in the same case, if I have a different aircraft, let us say a 120-seat aircraft. So, I will start this 'b' with 120 if my aircraft is a 120-seater. So, my table will start with 120, and I will increment by 121, 122, and so on, and in that way, I will find what the optimal level of my overbooking should be, same way. So, this 'b minus C' will also be the same way, because now it will be 120, 121, 122, 123, and so on.

So, this will be 128. So, the difference of 'b minus C' will always be the same because the increment is happening only by 0, 1, 2, 3, 4, 5, 6, 7, 8, and so on. So, this probability function will also give you the same value all the time, like this. So, whether it is an aircraft of 100 seats, whether it is an aircraft of 120 seats, or whether it is an aircraft of 200 seats, I will always go for an overbooking of 8 seats in this case. So, the starting capacity is not the point of deciding the overbooking.

The overbooking is basically determined by the ratio of 'p by D,' which is the most important factor for deciding this particular case. And in this way, we have seen that, yes, in this case, the passenger revenue—the expected net revenue—will change because now more passengers will be traveling in the aircraft. So, that revenue will change, but the amount of overbooking will remain the same, irrespective of the size of your aircraft. So, this is our risk-based policy. I hope we understood it very well. Then, another policy is the service-level policy.

And in the service level policy, we are going to do the overbooking in such a manner that we should give the minimum number of denied boardings to those customers who are coming to avail the service. So, here our idea is that we should provide the maximum level of service to all those customers who are coming to avail our services; that is the service level policy. In many cases, customers are preferring the service level policy over the risk-based policy. The reason is goodwill. So, when a customer has actually come to your point of service and you are denying the service to that particular customer, it creates an upset mood for the customer or a loss of goodwill, which is a bigger loss for the company.

Therefore, in the service industry, service itself is the most important criterion. So, most organizations now prefer the service level policy over the risk-based policy because the risk-based policy is more profit-oriented and less service-oriented, and in the present environment, people prefer service orientation over economic or monetary issues. So, this particular thing is more commonly used. The second important point in favor of the service level policy is that risk-based booking can lead to wide variations in booking limits and the potential number of denied boardings from flight to flight. And this will lead to bigger challenges: if more denied boardings occur, this will create many challenges for the operators and may lead to undesirable situations, losing customer confidence in those particular services.

So, organizations want to become more service-oriented. Now, in the service level policies, we generally have this issue: how do we book that level of service where we lead to fewer denied bookings? So, in this case, the airline sets a policy that a fraction of

$$\frac{E[((s|b) - C)^{+}]}{E[(s|b)]} = \frac{1}{10{,}000}$$

$$\text{or, equivalently,} E[((s|b) - C)^{+}] = 0.0001E[(s|b)].$$

bookings resulting in denied boardings should be approximately 1 in 10,000.

Generally, this rule—what should be the number of denied boardings in your particular case—is based on industry standards. So, there are benchmarks available, and based on those benchmarks, you may decide what should be the number of denied cases in your particular case.

So, it may be different for hotels, it may be different for airlines, it may be different for car rentals, and so on. Now, in this case, the term s|b refers to shows given booking limit b. So, s, as I have already explained, is the number of customers who are actually appearing, and b is the booking limit, and C is the capacity of the particular server. So, in this particular case, the way of setting b—what should be b—is the expected number of s|b minus C, and again, only zero or positive values are taken; that is the meaning of this plus superscript. with respect to the expected number of shows, particularly with respect to a booking limit. So, this is set as 1 out of 10,000; out of 10,000 customers, you are allowing only one denied service.

Or, in a simpler way, if you do the mathematical jugglery, you will find these relations. And in this way, you have a policy that specifies that the number of denied service incidents should be some specified fraction of customers served rather than the total number of bookings. This is the way airlines report denied bookings to the Department of Civil Aviation. And, as we just mentioned, Alaska Airlines experienced 0.81 involuntary denied booking boardings per 10,000 passengers in the year 2003. So, the number that 1 out of 10,000, you can say, is very close to Alaska's average achievement of 0.81 out of 10,000.

So, under this policy, the company would set b so that this left-hand side expected value equals q multiplied by E, which is the target denied service fraction—like in our case, it was 1 out of 10,000.

$$E[((s \mid b) - C)^{+}] = qE[\min((s \mid b), C)]$$

And this is the expected sales given a booking limit of b. So, in this way, you can operate your service-level policies also. Finally, organizations want to take the benefit of economies and also minimize the cases of denied services. So, therefore, nowadays, organizations are developing a combined policy known as a hybrid policy. So, here, we are trying to minimize the losses because of overbooking or the cost of denial of services.

And we are also looking at how we can minimize the cases of denied bookings in terms of the fraction of customers who are appearing to avail these services. This is known as a hybrid policy, which is more complex in nature, and it will be difficult for us to discuss those policies in the form of a course that is more focused on using modern methods of AI and machine learning to combine these two policies for getting the benefit of both systems. So, with this, we see that deterministic policies are the easiest to implement, then risk-based policies and service-level policies are the other two policies which are popular, and hybrid policies are in development where we are combining the economic benefits as well as the benefits of the service levels. With this, we come to the end of this particular session. Thank you very much.