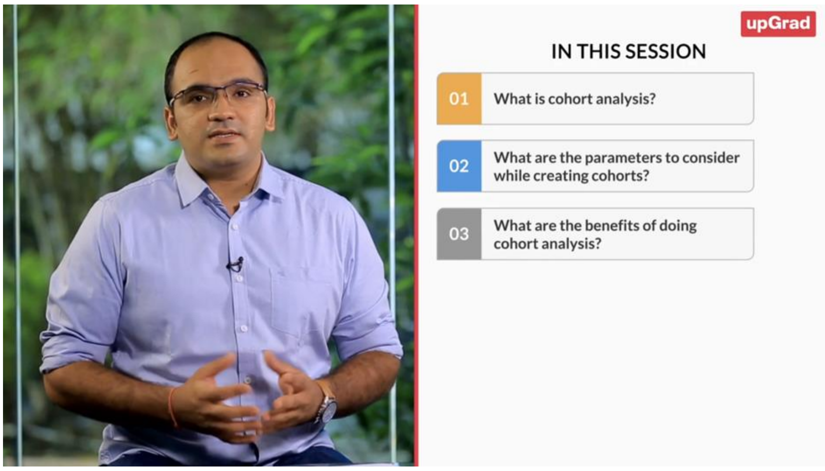
Transcription

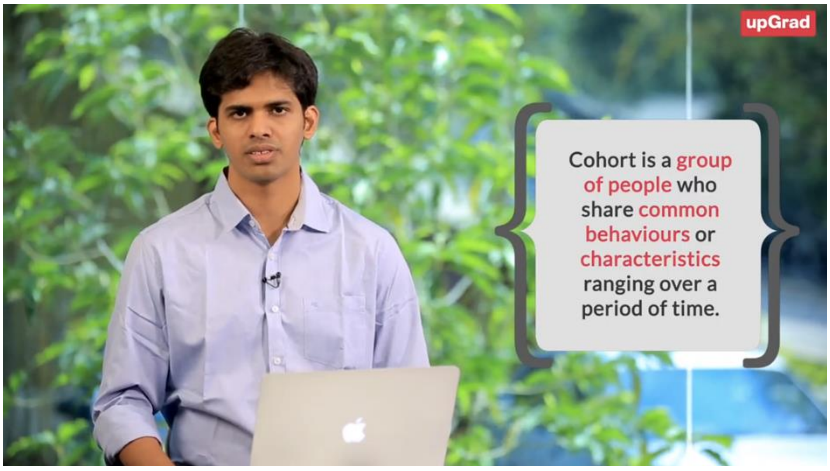
Cohort Analysis



Hi. Till now, you've gone through A/B testing, multivariate testing and segmentation and funnel analysis. Now, let’s move on to cohort analysis.

So, what is cohort analysis? What are the parameters to consider while creating cohorts? What are the benefits of doing cohort analysis?

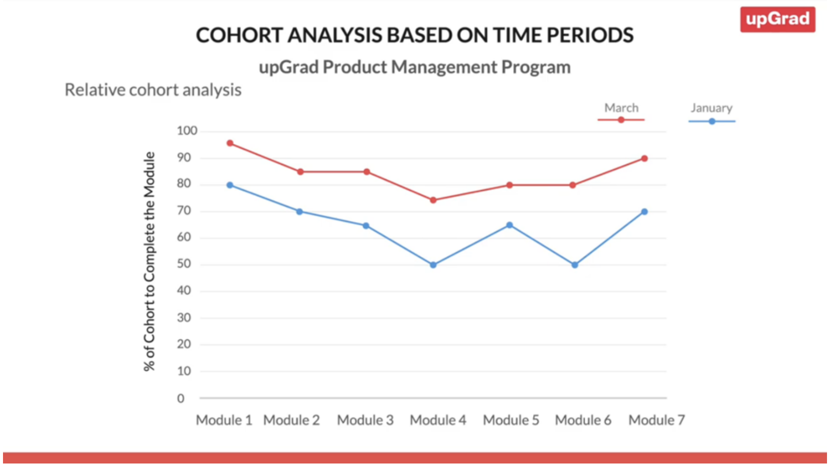
Let’s see what our experts have to say.



Before diving into cohort analysis and its significance; let’s first understand what a cohort is.

A cohort is defined as a group of people who share common behaviours or characteristics, ranging over a period of time. For example, the people enrolled in upGrad's product management program for January would form one cohort. Similarly, people who enrol for upGrad's product management course for March would form another cohort.

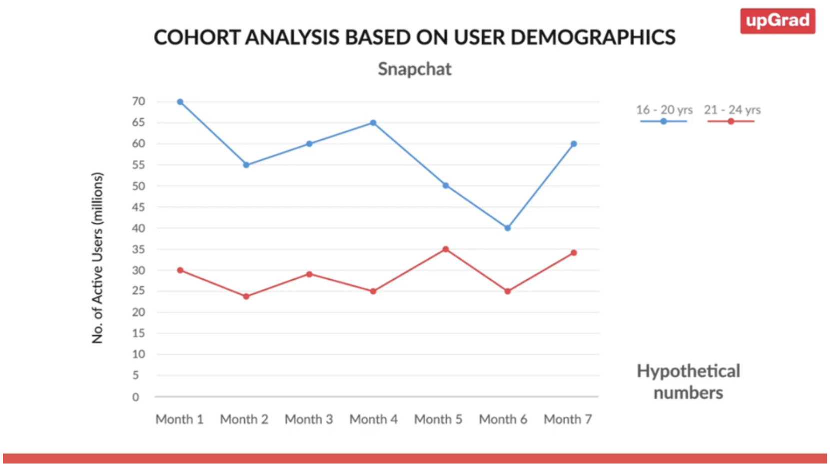
Now, cohort analysis deals with analysing the data related to such cohorts. Analysis can be absolute or relative. So if you just take a look at the January cohort and calculate the percentage of people completing any module of the course, you would be doing cohort analysis, which is absolute in nature. The real value of cohort analysis lies in studying different cohorts over a period of time, as this lets you study and compare behaviours of two or more groups of people.



For example, suppose you compare the cohort of January with the March cohort and find out that more people from March cohort were able to complete the modules. You would then figure out the reasons for this. Here, it seems that one possible reason could be changes in the program structure made for the March cohort, which could have led to higher engagement.

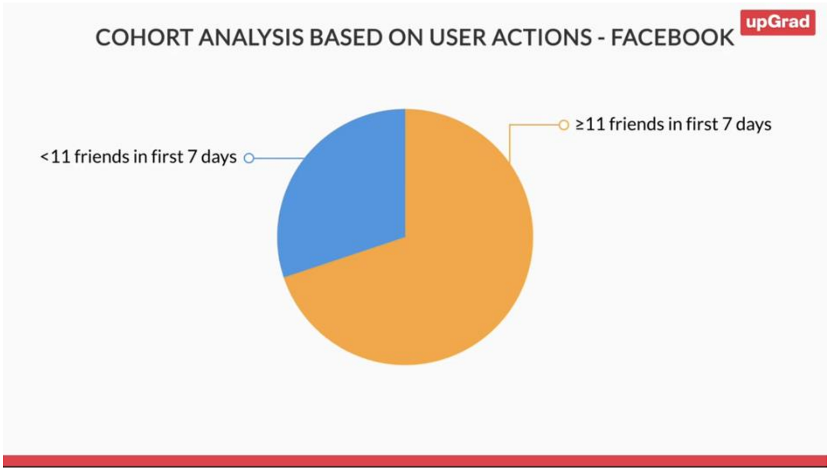
For any product, cohorts can be created based on various parameters such as user demographics, user actions and time period. In the previous example, you saw that the cohorts were created as per different time period of the program's launch, such as the January cohort and the March cohort.

Similarly, you can also create cohorts based on user demographics. For example, for Snapchat, you can create two cohorts. Cohort 1 would contain people within the age group of 16 to 20 years, and cohort 2 would contain people within the age group of 21 to 24 years.



Now, let's say that upon comparison of these two cohorts, you discover that cohort 1, which consists of people within the age group of 16 to 20 years, is more active on Snapchat, then cohort 2. This insight would be handy while designing a feature for Snapchat or in a marketing campaign.

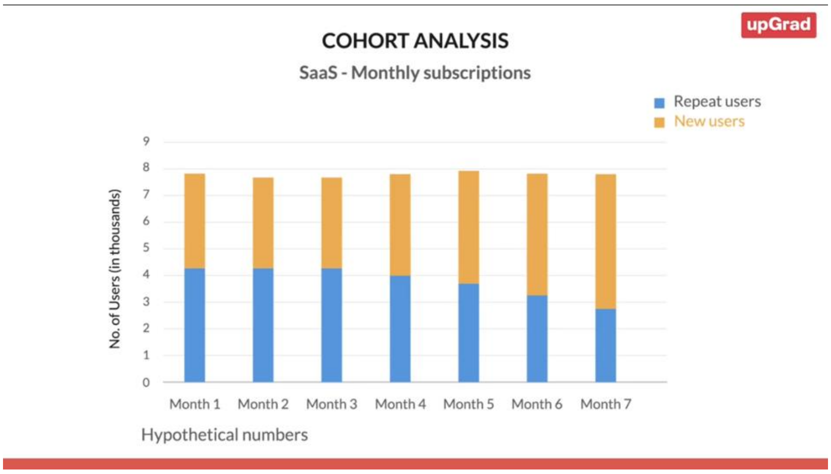
Another way to create cohorts is on the basis of user actions. Let's take Facebook example to understand this. Examples of this would be people who posted more than two times on Facebook in a day or people who liked more than twenty posts in a day. Now, in case of Facebook, some of you might not know this, but in the initial days of Facebook, what happened was they realize that people who have eleven friends in first seven days were more likely to stick to Facebook than people who didn't have any friends or, let's say less than 10 friends in first 7 days.



So there were two separate cohorts working at Facebook. One was people who formed 11 friends in first 7 days and people who didn't form 11 friends in first 7 days. And this is the agenda that Facebook walked on for years to actually, whenever a new person signs up for the Facebook, they actually make sure that the people has more than 11 friends in first 7 days and this actually helped Facebook a lot in its growth and engagement part.

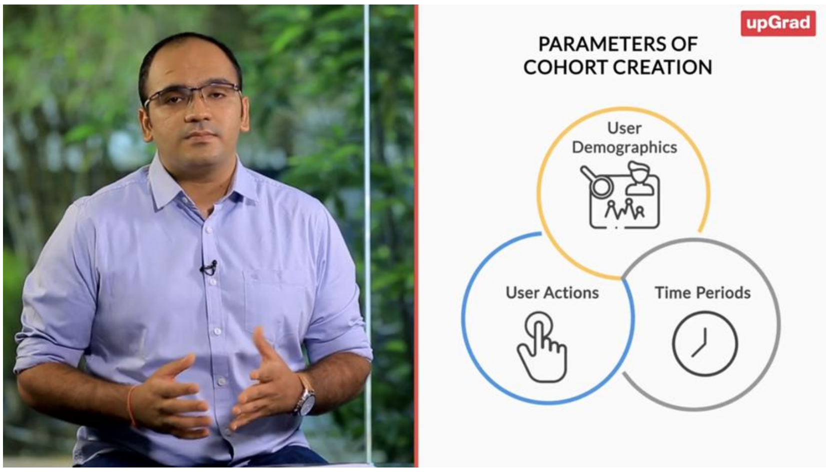
So how can cohort analysis help you, as a PM? Cohort analysis is a vital tool for a product manager. Here are the main benefits that I'm going to list that you can get from cohort analysis.

1. You can find out how user behaviour in cohorts has changed over time for your product
2. You can get insights to help you develop product roadmap and to identify what is working in your product and what is not. You can decide to add or remove features based on the data that you get from cohort analysis.
3. Cohort analysis can also be used to calculate the lifetime value or LTV of the customers. We will look into this later in sessions.
4. Cohort analysis can help you to achieve your business goals. By analysing different cohorts, you can find out what metrics are lagging behind and then further you can work upon improving the product features that directly affect these metrics.



Let's take an example of a SAS company which is selling monthly subscriptions to users. The company is growing well on a month-to-month revenue basis, but upon performing a detailed cohort analysis, the team finds out that the number of repeat users renewing their subscriptions has been falling since the last 3-4 months, and this is actually masked by the increase in new signups. Now this would not have been possible to spot while looking at general metrics like growth and revenues.

But the most important benefit of cohort analysis, apart from this is that, it helps you see the growth and engagement metrics separately. And this is very important for you as a PM because generally, the focus is on increasing the revenue of the company, but sometimes the high growth may mask the fact that you are losing repeat customers or having very low engagement for the existing customers. This is how a cohort analysis helps the SAAS Company. It is able to identify the cohort where engagement is falling rapidly. Once the team identifies these users, it gathers feedback from them and makes appropriate changes to its product offerings. So this is how cohort analysis helps you as a PM.

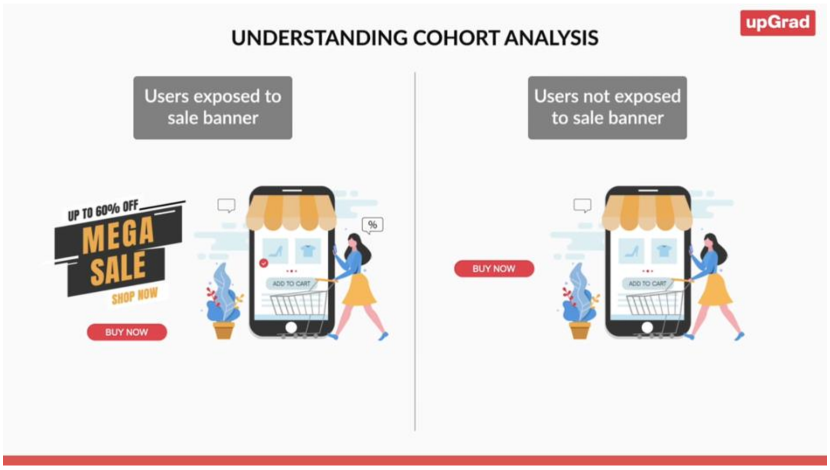


In this video, you learnt that cohort analysis means analysing the data related to cohorts which represent groups of people who share common behaviours or characteristics over a period of time. You also learned that cohorts can be created based on parameters such as user demographics, user actions and time periods. You then looked at the various benefits of doing cohort analysis.

In the next video, we will take a look at some industry examples of cohort analysis.

Cohort analysis is very useful technique that helps us understand how customer journey is impacted by some experiences or events. So, what are cohorts and what do we analyse?

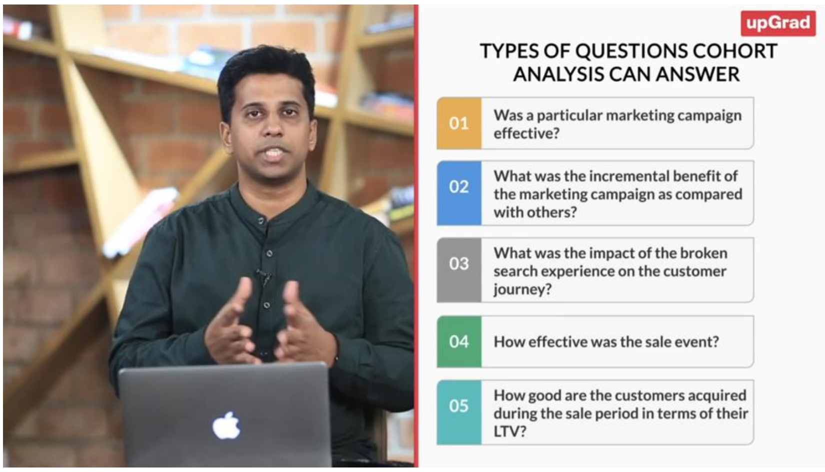
Cohorts are groups that share some common experience or characteristics within some defined timespan. More concretely, these are customers who have had some common event binding them or there was some temporal event which binds them together. So, as an example users exposed to the sale intrigue banner on the homepage, that could be one set and the users who are not exposed to this could be the other set. So, the event that binds them is that event of the banner being exposed to them.



Another example could be say there was some experience on the platform which was broken. Let's say the search was not working for a while for some customers. Then the customers for whom the search was not working fine that could be one cohort and all other customers could be other cohort. Or another simple example could be the customers who were acquired during a sale period versus the customers who were acquired during the non-sale period. So, these could be different cohorts. And by the way the last one is a very-very common use case in the industry which we will study further as well.

So, yes now that we have defined what cohorts are. What do you analyse? We study the behaviour of these cohorts overtime. And that is a key aspect which makes cohort analysis what it is and differentiates it from the other techniques. So briefly therefore a cohort analysis is:

Defining cohorts of customers or separating customers into these cohorts based on some shared experience or event in time and then analysing and following their behaviour over a time period or may be contrasting them or studying these behaviours across cohorts. And this is the preferred method when your A/B kind of testing is say not possible or not practical or may be whatever happened was not a planned event, right. For example, the search breaking was not a planned event but it just happened. And now you want to assess the impact of it.

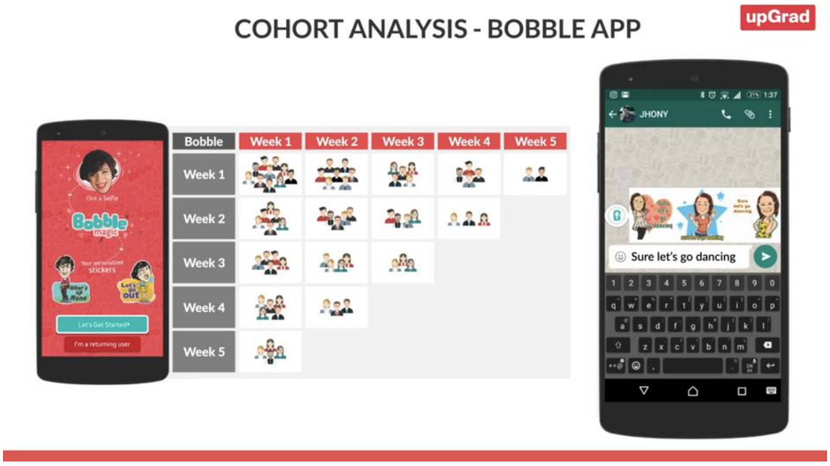


What are the kinds of question cohort analysis can help us answer?

Well you can have very common questions in the say; marketing domain like was this campaign effective or what was the incremental from this campaign versus the other and so on, or you could have something that we mentioned earlier, say the impact of a broken search experience on the customer journey. Or you could have a question like how effective was a sale in terms of acquiring new customers. Are we getting the right customers? How good are these customers that we are acquiring through our sales? These are the some of the questions and of course a lot more questions you can answer using a cohort analysis.

Again, let me re-iterate that we are not looking at an immediate conversion or clickthrough or some other metric at an instant for these cohorts. Rather we are interested in studying their behaviour over time and this is a key element which separates it from other kinds of analysis.

You now know about cohort analysis and its benefits. Let's take a look at some industry examples of cohort analysis.



An industry example of a cohort could be if you look at Bobble app, which is a keyboard which helps you express yourself with like these personal stickers, and you look at their like cohorts week after week, like all the people who sign up in a particular week, how engage they are in 10 weeks after signing up. So when you look at those improved cohorts, you can easily correlate back to like the improvements they made in the keyboard.

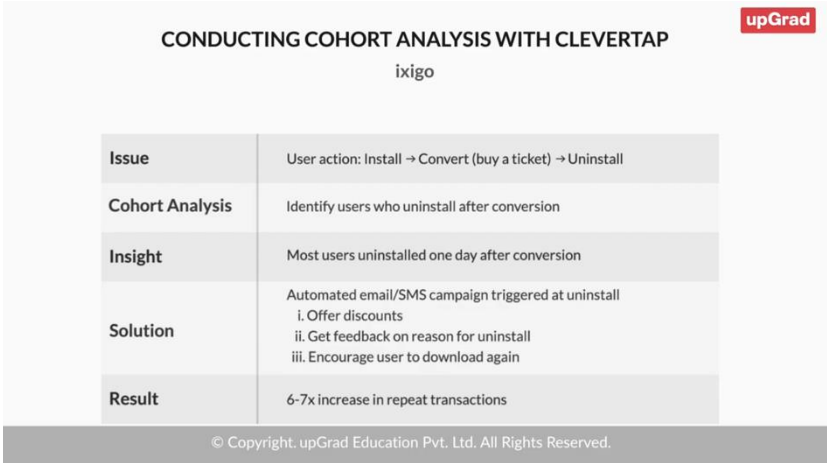
For example, one thing they changed was they allowed people to directly share stickers from the keyboard and it's very reasonable to understand how that will improve the engagement of the product.

CleverTap allows you to also do cohorts for users who do recurring actions within your app. For example, imagine a travel website, one of our clients being Ixigo. They had a major problem with a lot of people abandoning or uninstalling their app right after purchasing the ticket, right. Seems like a very obvious behaviour, especially in a market like India, a lot of users, a lot of these apps, they download the app they actually perform the conversion action and then go on to uninstall your app, right. Because once somebody books a flight on Ixigo, there is no longer kind of a need or immediate, especially for a non-frequent buyer. The app is no longer required.

In that case, how do you make sure you're not losing customers, right? Ixigo had a similar problem wherein by cohorts we were actually able to tell you how many people actually booked a flight and then went on to uninstall the very same day or the following days. One of the major days of uninstalls was the day right after the day of purchase. In this case, you were losing paid customers, people who actually booked a flight with you. You were losing them right on the next day.

How would you address such a problem because the user no longer has your app. They did do a booking, but you're not going to be in the consideration set because they don't have your app when the next time they want to book a flight, right. So how do you still control the experience when somebody is actually uninstalled your app?

The cohort shows you how many people actually booked a ticket on Ixigo and then went on to uninstall the app the very next day. In this case, what we did is enabled an email, automated email to go out as soon as the users uninstalled or maybe an SMS, to go out to only users and kind of giving them discounts on the next time they book a flight or even better getting feedback of why they uninstalled the app or even better just get them to download your app again.


upGrad

CONDUCTING COHORT ANALYSIS WITH CLEVERTAP

ixigo

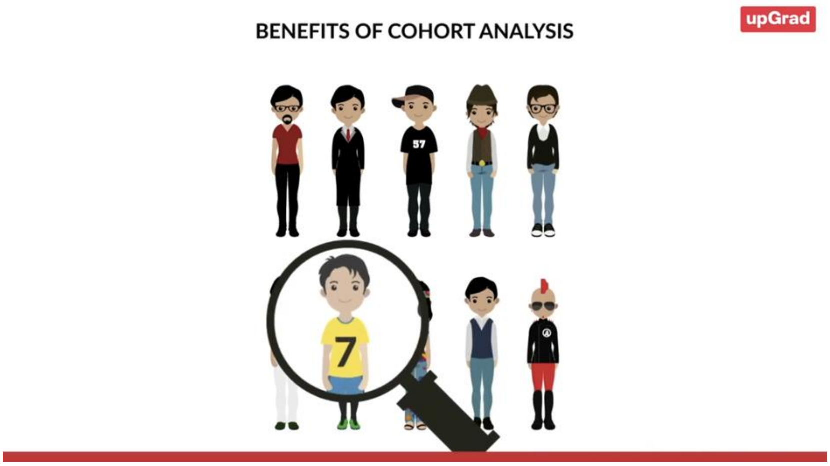
Issue	User action: Install → Convert (buy a ticket) → Uninstall
Cohort Analysis	Identify users who uninstall after conversion
Insight	Most users uninstalled one day after conversion
Solution	Automated email/SMS campaign triggered at uninstall i. Offer discounts ii. Get feedback on reason for uninstall iii. Encourage user to download again
Result	6-7x increase in repeat transactions

© Copyright. upGrad Education Pvt. Ltd. All Rights Reserved.

In this case, we were able to actually even after install - uninstall, we were able to actually talk to users and kind of still control the experience or the world view they have about Ixigo. So the next time somebody wants to book a flight, they know for sure the last email they saw was from Ixigo and they have some credits and they can go ahead and actually book.

From the Ixigo campaign, we were able to actually record transactions 7 days later or 14 days later or 3 weeks later, of people who've, actually uninstalled and reinstalled the app and come back to actually purchase. So the increase in for repeated transaction was about 6-7x.

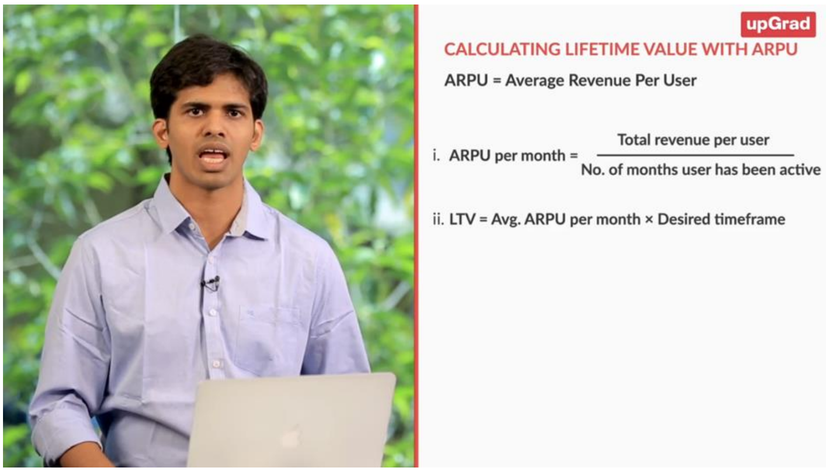
You have learned that one of the benefits of cohort analysis is calculating the LTV or lifetime value of your customer. Let’s dig a bit deeper and ask our SME about what is LTV. How do you calculate LTV?



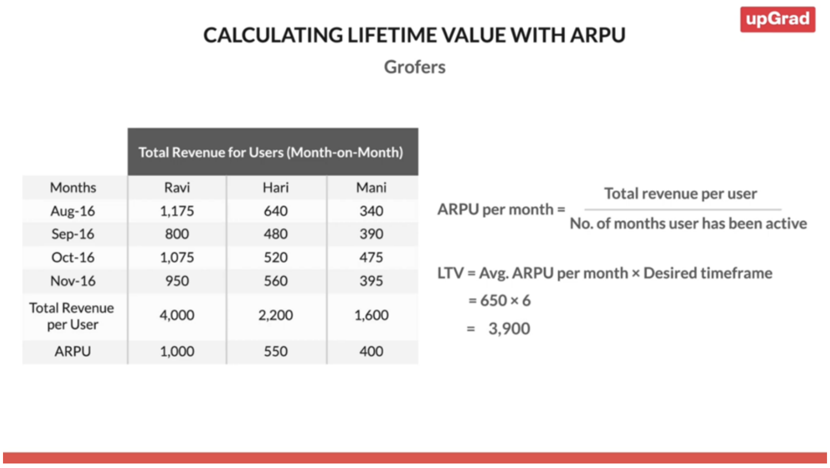
One of the most important benefits of cohort analysis is that it helps you calculate the LTV or lifetime value of your customer. Calculating LTV is important because not all customers use your product in a similar fashion. You need to identify your most valuable customers and then focus on them while working on the development of new product or feature. This is important because such customers are likely to have greater impact on your revenue figures, and hence they define the growth of the company that you are working for.

Before learning how to calculate LTV using cohort analysis, let's learn about LTV in more detail. LTV, which is short for lifetime value, predicts all the value a company will derive from its relationship with a customer, because you don't know how long a customer will use your product. You make a good estimation and state LTV as a periodic value, which means you say that customer X for 6 or 12 month value is Y rupees and hence this is the LTV.

One thing to remember, though, is LTV is never computed as total revenue divided by total number of customers. This ignores how long a customer sticks with you, which is a very important factor to consider.



One of the approaches to calculate LTV is by calculating the average revenue per user per month or ARPU in short per month. ARPU per month is the total revenue for a user divided by the number of month's that user has been with you. To calculate the LTV, you take the average of ARPU per month for each user and multiply it by the timeframe for which you want to calculate the LTV.


CALCULATING LIFETIME VALUE WITH ARPU

Grofers

	Total Revenue for Users (Month-on-Month)		
Months	Ravi	Hari	Mani
Aug-16	1,175	640	340
Sep-16	800	480	390
Oct-16	1,075	520	475
Nov-16	950	560	395
Total Revenue per User	4,000	2,200	1,600
ARPU	1,000	550	400

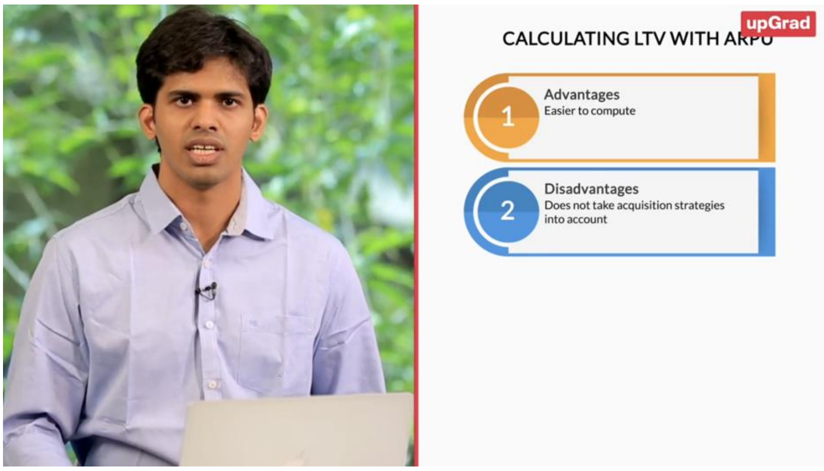
$$\text{ARPU per month} = \frac{\text{Total revenue per user}}{\text{No. of months user has been active}}$$

LTV = Avg. ARPU per month × Desired timeframe

= 650 × 6

= 3,900

To understand this better, let’s take an example. Suppose Grofers has three users: Ravi, Hari and Mani. Each users purchase on the Grofers are shown in the table on the screen. First, you calculate the monthly ARPU for all three which can be done by dividing the total revenue for a user with the number of months the user have been with you. So if you look at Ravi, you can see the purchases he has made across different months. You can add up these purchases to come up with total revenue generated from Ravi. Now dividing this with the number of months, which in this case is 4, you can get ARPU as Rs.1000 for Ravi. Similarly, you can estimate ARPU for Hari and Mani, which comes out to be 550 rupees and 400 rupees respectively. Now you take the average of all three which comes out to be 650 rupees to calculate the LTV for 6 months. You multiply this number by 6 to get the total LTV as 3900 rupees. Using this approach, you can calculate LTV as the ARPU is relatively easier to compute.

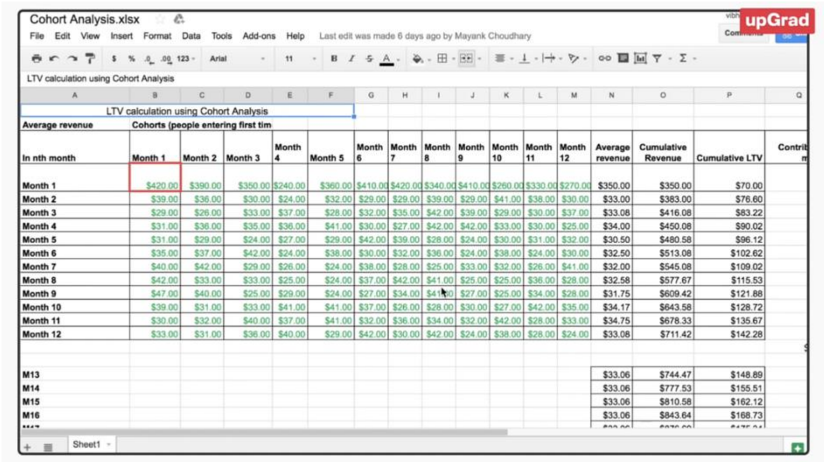


One of the disadvantages of this approach being that it does not take into account the changes in your customer’s behaviours. So if your company is growing at a fast pace and you are having lots of new customers, this number might be misleading. It does not tell you if new users are spending less or more or if a user made only one big purchase in a month and then did not return at all. Also, if you have acquired new customers using a different strategy, it will not help you see if they are spending more than the user acquired from other strategies.

You learnt about LTV and how you can calculate it using the ARPU approach. Next up, we'll look at how you can calculate LTV using cohort analysis.

So, far we have covered how LTV is calculated using ARPU approach. Now, let's hear it from our SME on how you can calculate LTV using cohort approach.

Let's now look at how you can calculate LTV using cohort analysis. One of the advantages of using this approach is that the cohort analysis does not assume all months to be the same when it comes to revenue. Let's calculate LTV for an E-commerce business using cohort analysis.



LTV calculation using Cohort Analysis

Average revenue	Cohorts (people entering first tim													
In nth month	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10	Month 11	Month 12	Average revenue	Cumulative Revenue	Cumulative LTV
Month 1	$420.00	$390.00	$350.00	$240.00	$360.00	$410.00	$420.00	$340.00	$410.00	$260.00	$330.00	$270.00	$350.00	$350.00	$70.00
Month 2	$39.00	$36.00	$30.00	$24.00	$32.00	$29.00	$29.00	$39.00	$29.00	$41.00	$38.00	$30.00	$33.00	$383.00	$76.60
Month 3	$29.00	$26.00	$33.00	$37.00	$28.00	$32.00	$35.00	$42.00	$39.00	$29.00	$30.00	$37.00	$33.08	$416.08	$83.22
Month 4	$31.00	$36.00	$35.00	$36.00	$41.00	$30.00	$27.00	$42.00	$42.00	$33.00	$30.00	$25.00	$34.00	$450.08	$90.02
Month 5	$31.00	$29.00	$24.00	$27.00	$29.00	$42.00	$39.00	$28.00	$24.00	$30.00	$31.00	$32.00	$30.50	$480.58	$96.12
Month 6	$35.00	$37.00	$42.00	$24.00	$38.00	$30.00	$32.00	$36.00	$24.00	$38.00	$24.00	$30.00	$32.50	$513.08	$102.62
Month 7	$40.00	$42.00	$29.00	$26.00	$24.00	$38.00	$28.00	$25.00	$33.00	$32.00	$26.00	$41.00	$32.00	$545.08	$109.02
Month 8	$42.00	$33.00	$33.00	$25.00	$24.00	$37.00	$42.00	$41.00	$25.00	$25.00	$36.00	$28.00	$32.58	$577.67	$115.53
Month 9	$47.00	$40.00	$25.00	$29.00	$24.00	$27.00	$34.00	$41[illegible]	$27.00	$25.00	$34.00	$28.00	$31.75	$609.42	$121.88
Month 10	$39.00	$31.00	$33.00	$41.00	$41.00	$37.00	$26.00	$28.00	$30.00	$27.00	$42.00	$35.00	$34.17	$643.58	$128.72
Month 11	$30.00	$32.00	$40.00	$37.00	$41.00	$32.00	$36.00	$34.00	$32.00	$42.00	$28.00	$33.00	$34.75	$678.33	$135.67
Month 12	$33.00	$31.00	$36.00	$40.00	$29.00	$42.00	$30.00	$42.00	$24.00	$38.00	$28.00	$24.00	$33.08	$711.42	$142.28
M13													$33.06	$744.47	$148.89
M14													$33.06	$777.53	$155.51
M15													$33.06	$810.58	$162.12
M16													$33.06	$843.64	$168.73

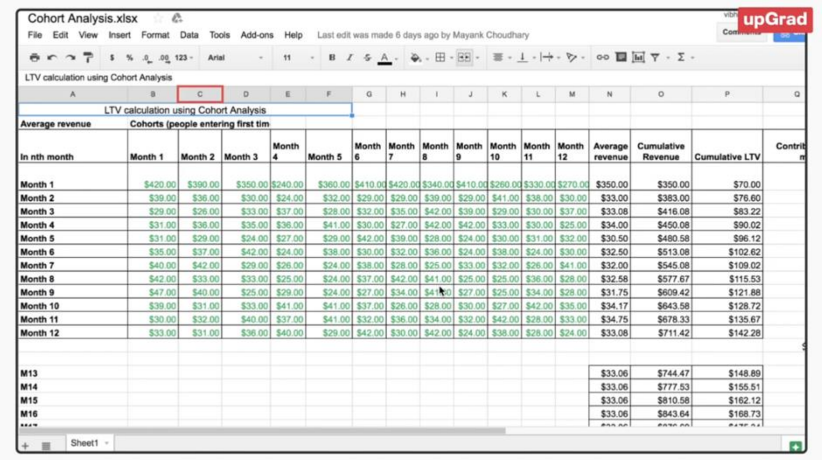
Assume that business has been into operations for couple of years now and we are taking a year's data. So, I want you to think that suppose there is a user who has come into the first month. He will be on the platform for next 24 months. And we have all 24 months data but here for the sake of simplicity, we are considering only first 12 months data. And the reason we are doing this because the person who came in the 12th month, we will have the data for 12th to 23rd month. So, we will have complete 12-month data for that user as well.

So in essence, we have two years of data but we are just seeing the first 12 months data and in this example, we are going to show you, how we calculate LTV for an E-commerce business. So, what you have to do here is first look at the different cohorts month-wise. So this is month 1, month 2 and month 3 and month 4 up to month 12.

Now what it means is people who came in these months. So suppose Ravi came in month 1, so Ravi belongs to that cohort month 1. But if somebody else came in month 2 that person belongs to month 2 cohort. Now Ravi came in month 1 and Ravi spent these many in month 1. In month 2 Ravi spent another amount, in month 3 another amount and similarly for next 12 months Ravi spent some particular amount in a particular month.

Now this table illustrates the average spent by that particular cohort in that month. And as you can see people coming in different months, they are spending different amount in their first month, in their second month, in their third month. For some cohort this might be same but not for all cohorts. And this is how an E-commerce business operates like people are getting acquired at different point of the business and depending on the product availability for that particular e-commerce business some people might purchase more than others.

So, at a particular point in let's say 6 months, we didn't have electronics but in 7 month they launch electronics. Because of which the first-month purchase for the next cohort which is month 7 will increase, right. So that's why you don't see all the numbers same for different cohorts.



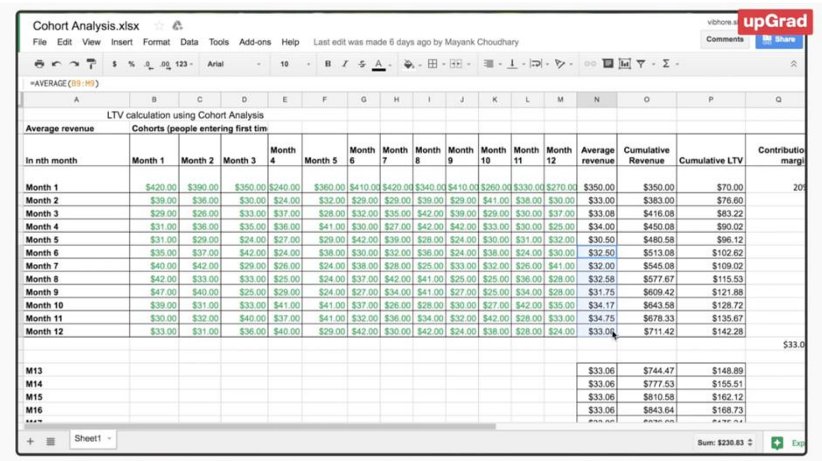
Cohort Analysis.xlsx

LTV calculation using Cohort Analysis

Average revenue	Cohorts (people entering first tim														
In nth month	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10	Month 11	Month 12	Average revenue	Cumulative Revenue	Cumulative LTV
Month 1	$420.00	$390.00	$350.00	$240.00	$360.00	$410.00	$420.00	$340.00	$410.00	$260.00	$330.00	$270.00	$350.00	$350.00	$70.00
Month 2	$39.00	$36.00	$30.00	$24.00	$32.00	$29.00	$29.00	$39.00	$29.00	$41.00	$38.00	$30.00	$33.00	$383.00	$76.60
Month 3	$29.00	$26.00	$33.00	$37.00	$28.00	$32.00	$35.00	$42.00	$39.00	$29.00	$30.00	$37.00	$33.08	$416.08	$83.22
Month 4	$31.00	$36.00	$35.00	$36.00	$41.00	$30.00	$27.00	$42.00	$42.00	$33.00	$30.00	$25.00	$34.00	$450.08	$90.02
Month 5	$31.00	$29.00	$24.00	$27.00	$29.00	$42.00	$39.00	$28.00	$24.00	$30.00	$31.00	$32.00	$30.50	$480.58	$96.12
Month 6	$35.00	$37.00	$42.00	$24.00	$38.00	$30.00	$32.00	$36.00	$24.00	$38.00	$24.00	$30.00	$32.50	$513.08	$102.62
Month 7	$40.00	$42.00	$29.00	$26.00	$24.00	$38.00	$28.00	$25.00	$33.00	$32.00	$26.00	$41.00	$32.00	$545.08	$109.02
Month 8	$42.00	$33.00	$33.00	$25.00	$24.00	$37.00	$42.00	$41.00	$25.00	$25.00	$36.00	$28.00	$32.58	$577.67	$115.53
Month 9	$47.00	$40.00	$25.00	$29.00	$24.00	$27.00	$34.00	$41.00	$27.00	$25.00	$34.00	$28.00	$31.75	$609.42	$121.88
Month 10	$39.00	$31.00	$33.00	$41.00	$41.00	$37.00	$26.00	$28.00	$30.00	$27.00	$42.00	$35.00	$34.17	$643.58	$128.72
Month 11	$30.00	$32.00	$40.00	$37.00	$41.00	$32.00	$36.00	$34.00	$32.00	$42.00	$28.00	$33.00	$34.75	$678.33	$135.67
Month 12	$33.00	$31.00	$36.00	$40.00	$29.00	$42.00	$30.00	$42.00	$24.00	$38.00	$28.00	$24.00	$33.08	$711.42	$142.28
M13													$33.06	$744.47	$148.89
M14													$33.06	$777.53	$155.51
M15													$33.06	$810.58	$162.12
M16													$33.06	$843.64	$168.73

So different columns b, c, d, e are different cohorts and then in rows like 4, 5, 6, 7 there is month 1, month 2, month 3 in first column. This is basically their spend in these months. So, as you can see like in different cohorts in month 1. Let's say Ravi spend $420, in subsequent months the spend is much lesser and such is the behavior for all cohorts like people who are starting in month 2 are doing the same, in month 3 the same and in month 4 the same.

So, this is one behavior which will be very useful for you as a product manager because what you can do is, you can actually engage people who moved from month 1 to month 2, like people who are in their first month. You know they will purchase something but in their second month, their purchase will drop down. So you can maybe send some coupons or something to actually increase the revenue for the second month of the user. So, this is one way in which this cohort analysis will help you as a PM.



Cohort Analysis.xlsx

=AVERAGE(B9:M9)

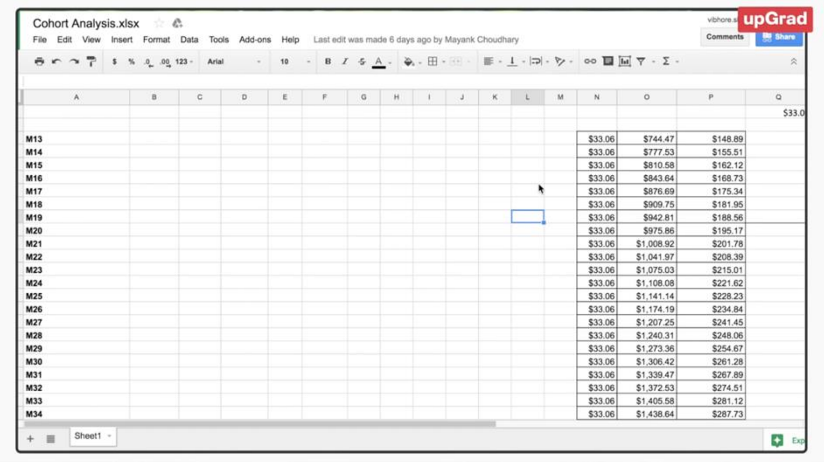
LTV calculation using Cohort Analysis

Average revenue	Cohorts (people entering first tim															
In nth month	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10	Month 11	Month 12	Average revenue	Cumulative Revenue	Cumulative LTV	Contributio margi
Month 1	$420.00	$390.00	$350.00	$240.00	$360.00	$410.00	$420.00	$340.00	$410.00	$260.00	$330.00	$270.00	$350.00	$350.00	$70.00	20
Month 2	$39.00	$36.00	$30.00	$24.00	$32.00	$29.00	$29.00	$39.00	$29.00	$41.00	$38.00	$30.00	$33.00	$383.00	$76.60	
Month 3	$29.00	$26.00	$33.00	$37.00	$28.00	$32.00	$35.00	$42.00	$39.00	$29.00	$30.00	$37.00	$33.08	$416.08	$83.22	
Month 4	$31.00	$36.00	$35.00	$36.00	$41.00	$30.00	$27.00	$42.00	$42.00	$33.00	$30.00	$25.00	$34.00	$450.08	$90.02	
Month 5	$31.00	$29.00	$24.00	$27.00	$29.00	$42.00	$39.00	$28.00	$24.00	$30.00	$31.00	$32.00	$30.50	$480.58	$96.12	
Month 6	$35.00	$37.00	$42.00	$24.00	$38.00	$30.00	$32.00	$36.00	$24.00	$38.00	$24.00	$30.00	$32.50	$513.08	$102.62	
Month 7	$40.00	$42.00	$29.00	$26.00	$24.00	$38.00	$28.00	$25.00	$33.00	$32.00	$26.00	$41.00	$32.00	$545.08	$109.02	
Month 8	$42.00	$33.00	$33.00	$25.00	$24.00	$37.00	$42.00	$41.00	$25.00	$25.00	$36.00	$28.00	$32.58	$577.67	$115.53	
Month 9	$47.00	$40.00	$25.00	$29.00	$24.00	$27.00	$34.00	$41.00	$27.00	$25.00	$34.00	$28.00	$31.75	$609.42	$121.88	
Month 10	$39.00	$31.00	$33.00	$41.00	$41.00	$37.00	$26.00	$28.00	$30.00	$27.00	$42.00	$35.00	$34.17	$643.58	$128.72	
Month 11	$30.00	$32.00	$40.00	$37.00	$41.00	$32.00	$36.00	$34.00	$32.00	$42.00	$28.00	$33.00	$34.75	$678.33	$135.67	
Month 12	$33.00	$31.00	$36.00	$40.00	$29.00	$42.00	$30.00	$42.00	$24.00	$38.00	$28.00	$24.00	$33.08	$711.42	$142.28	
																$33.0
M13													$33.06	$744.47	$148.89	
M14													$33.06	$777.53	$155.51	
M15													$33.06	$810.58	$162.12	
M16													$33.06	$843.64	$168.73	

Sum: $230.83

Now coming to the LTV calculation what we can see is we compute the average revenue which is in column N and we calculate the average revenue for every month like month 1 to month 12 and what we get is as you can see the revenues are getting stabilized towards the end. So what we are doing is, we are calculating the average revenue of last 6 months which is month 6, month 7 to month 12 and month 7 to month 12, if you calculate the average, it is around $33.06.

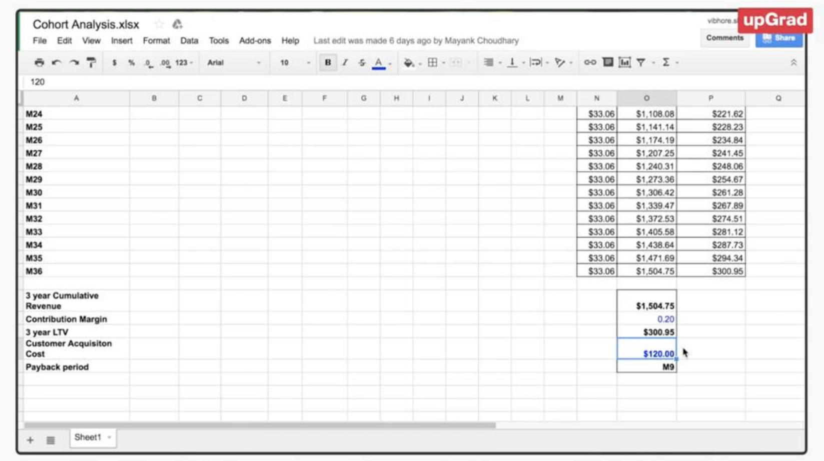
Now, this average amount we can assume that in coming months which is let's say 13th month or 14th month or lets' say 15th month user will spend the same amount. And therefore if you have to compute the LTV for a particular period which is let's say 3 years what you can do is you can assume this amount to be constant for next 2 years. Now beware that as you get more and more of data you will change this number and you will be more efficient or more accurate in predicting the LTV but for now, we have 12 months data so we will go by that. So this way what you get is the average revenue from month 1 to months 36.



	N	O	P
M13	$33.06	$744.47	$148.89
M14	$33.06	$777.53	$155.51
M15	$33.06	$810.58	$162.12
M16	$33.06	$843.64	$168.73
M17	$33.06	$876.69	$175.34
M18	$33.06	$909.75	$181.95
M19	$33.06	$942.81	$188.56
M20	$33.06	$975.86	$195.17
M21	$33.06	$1,008.92	$201.78
M22	$33.06	$1,041.97	$208.39
M23	$33.06	$1,075.03	$215.01
M24	$33.06	$1,108.08	$221.62
M25	$33.06	$1,141.14	$228.23
M26	$33.06	$1,174.19	$234.84
M27	$33.06	$1,207.25	$241.45
M28	$33.06	$1,240.31	$248.06
M29	$33.06	$1,273.36	$254.67
M30	$33.06	$1,306.42	$261.28
M31	$33.06	$1,339.47	$267.89
M32	$33.06	$1,372.53	$274.51
M33	$33.06	$1,405.58	$281.12
M34	$33.06	$1,438.64	$287.73

Now what we calculate next is cumulative revenue which is in column 0 and here what I am doing is I am adding month 1 and month 2, to put the total cumulative amount in month 2 and so on. So, this way I can say that when I see this number in month 12, I can say that in last 12 months the user has spent $711. So, in month 12 what we see is the user has actually given us $711 approximately in revenue.

Next, we go on calculating cumulative LTV. Now LTV in this case will be dependent on the contribution margin for the E-commerce website or E-commerce company. And the contribution margin we have assumed here is around 20%. So, for everything that you sell on Rs. 100 you get Rs. 20. So what we do is we multiply cumulative revenue by 0.2, which is 20% and then we get cumulative LTV. And this way if you move down what you see is the cumulative LTV for first 3 years for a particular user for a average user is around $300.



	N	O	P
M24	$33.06	$1,108.08	$221.62
M25	$33.06	$1,141.14	$228.23
M26	$33.06	$1,174.19	$234.84
M27	$33.06	$1,207.25	$241.45
M28	$33.06	$1,240.31	$248.06
M29	$33.06	$1,273.36	$254.67
M30	$33.06	$1,306.42	$261.28
M31	$33.06	$1,339.47	$267.89
M32	$33.06	$1,372.53	$274.51
M33	$33.06	$1,405.58	$281.12
M34	$33.06	$1,438.64	$287.73
M35	$33.06	$1,471.69	$294.34
M36	$33.06	$1,504.75	$300.95
3 year Cumulative Revenue		$1,504.75	
Contribution Margin		0.20	
3 year LTV		$300.95	
Customer Acquisiton Cost		$120.00	
Payback period		M9	

So here is a final table that we have made. Now the 3 year cumulative revenue is around $1500, contribution margin is around 20%, 3-year LTV is around $300, and customer acquisition cost which we can get from marketing teams because they are ultimately responsible for it is around $120. What it means is for every customer that we acquire, on an average we are spending around $120.

So you have LTV, cumulative LTV and you have customer acquisition cost, so you can also calculate payback period which is, in which month do I get these $120 back. And, if you move back and you look at the table you see that in month 9, I have gotten $121 in cumulative LTV. So, my 9th month is the actually payback period. In 9 months, I am getting whatever I have spent in acquiring this customer. So, this is the example where we calculated the LTV and finally payback period for the E-commerce platform.

Now, this is helpful in many ways. One this is helping you understand what is the behaviour of the cohort. So, as you can see, you found out that people are spending in month one much higher than they do in the subsequent months. And this you can use to promote coupons in their second month, third month or fourth month to increase the revenue. The second thing that we found was the payback period in our case is 9 months which means that we need some cash before for acquiring the customers because after 9 months only we would be able to break even on customer acquisition cost.

upGrad

CALCULATING PAYBACK PERIOD

Uber

Referral	= ₹50 off on first ride
Customer acquisition cost (CAC)	= ₹50
No. of rides per customer	= 10/month
Revenue per customer	= ₹50/ride = ₹500/month
Uber's commission per customer	= 10% per ride = ₹50
Payback period	= 1 month

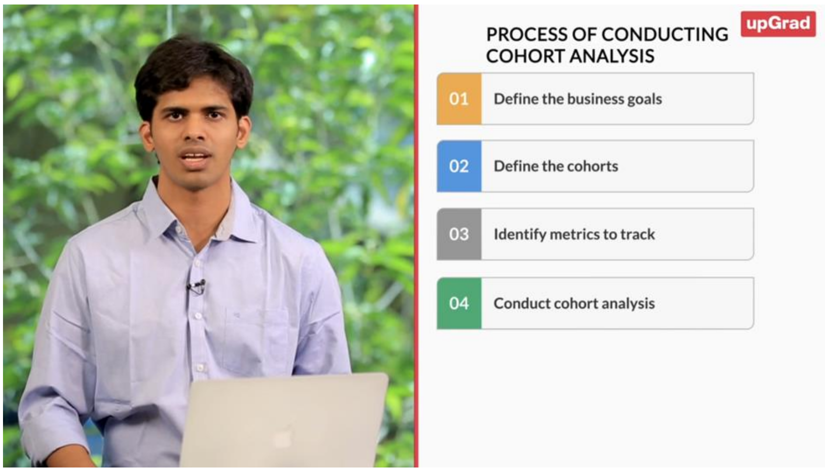
Now, the same analysis we can use in different scenarios. So, I will take a very small case. Let's say Uber is running referral programs and it is giving Rs.50 off for the first ride for a new customer. So, essentially Uber is losing Rs.50 as a customer acquisition cost. We will assume that it's not spending any money in advertising about the referrals. It is just relying upon the customers it already has. It is not doing any TV commercials or any newspaper ads or anything.

So, Rs.50 is the customer acquisition cost. And what you realize, what Uber realized or as a PM you realizes at Uber is people are taking around 10 rides per month, just be with me. So 10 rides per month, on an average Rs.50 per ride, so that is Rs.500 in a month and Uber gets 10% as the commission.

So, of Rs.500, 10% is Rs.50. So, in a month the user, average revenue per user per month is around Rs.500 and out of that UBER gets 10% which is Rs.50. So, this Rs.50 was again the customer acquisition cost. So, in a month Uber has got in its payback. So, in that way we see that one month is a payback period for Uber for its referral acquisition. And, that explains why Uber focuses so much on referrals because it's a very good source of acquisition for users for which it is not spending much but it is getting good customer's, people who actually ride on a day-to-day basis.

You now know about cohort analysis and how it can be used to calculate the LTV of your customer. But is there a process to conduct cohort analysis? Let's hear more on this from our subject matter expert.

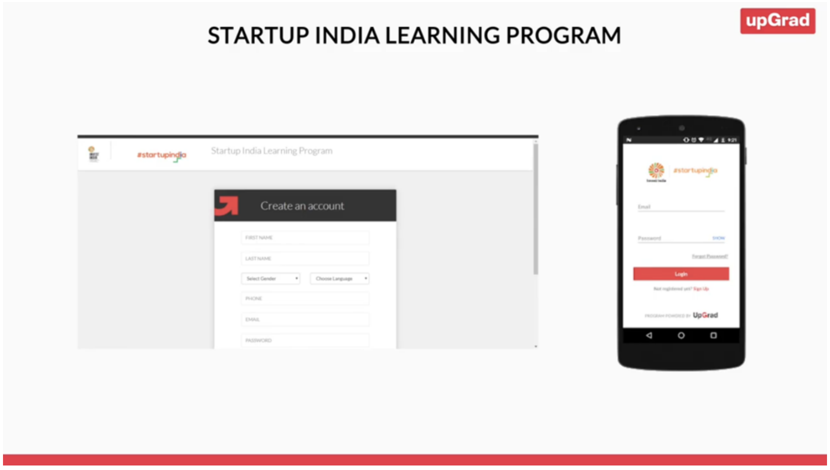
With the advancement in Analytics tool such as Google Analytics, Mixpanel and others it has become quite easy to perform cohort analysis. We will discuss how to use these tools later. But first let's look at the steps involved in conducting cohort analysis.



1. So, first thing first, you need to define your business goals for which you want to do the cohort analysis.

2. Then you would need to define the cohorts. This can be done based on the parameters such as user demographics, user actions, time periods etc. which we have already discussed.
3. Next thing would be you would list down the metrics through which you can track your progress towards achieving these defined business goals.
4. And, final thing is you would do the cohort analysis using Analytics tool like Google Analytics or Mixpanel.

Let us take an example to understand this better. So, just a brief background here. We launched the Startup India Learning Program in January this year. We first started with the website but after a couple of months we launched the app as well in the March. Now we had to decide whether to focus more on app or web. Because everything has a development cost and because we wanted people to complete the course we would focus on which ever platform; app or web is helping people engage more in completing the program.



So, let me break it down for you to understand.

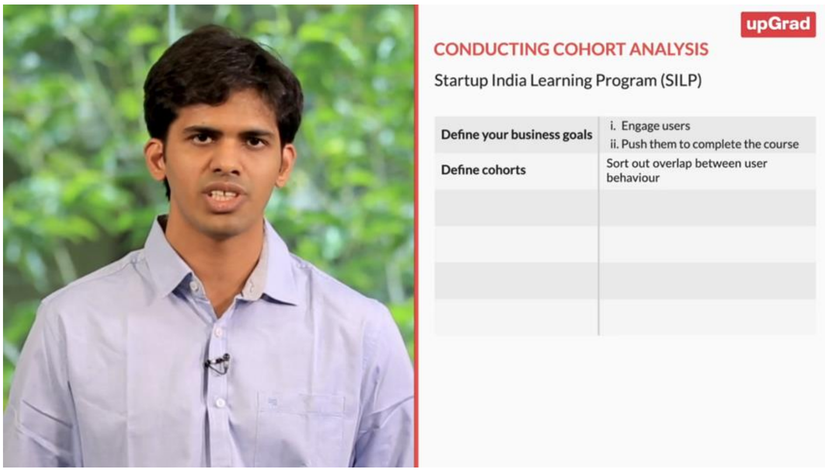
1. So, the first step as we discussed is defining the business goal. So, in this case it was engaging people and pushing them to complete the course. Now let me give you some context regarding that.

So SILP is 4-week-long program and for people to complete this program they have to come on this platform for at least 4 or 5 times. So even if they are consuming content for a week in one go, they would have to come at least 4 times which means we have to re-engage users again and again. So that they can complete the programs.

Once they complete the program they also come to know about other programs for upGrad for which they might come to website and they might explore those programs. And, then they will finally take it. So, that we can get revenue out of the traffic that is coming out of Startup India Learning Program. So that's the business goal that we have to engage these users and finally make them complete the program.

And that is where we found that we have to distinguish between app and web because development on app is a different thing than development on web and each has its own cost. So, where we should focus on? So this was the first step where we said that we have to see which investment is better, app or web, right.

And it was tied to a business goal which is which people are getting more engaged or actually completing the program in Startup India Learning Program. So this was the first step.



2. The second step was actually breaking it into cohorts.

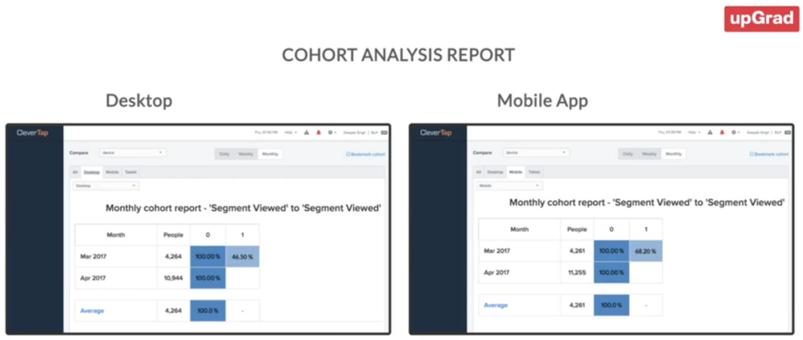
So, here it was based on device which people were using to complete the program or engage with the program because it was app versus web. We could easily see where people were first going. Are they downloading the app or are they going to the website and registering themselves there and then what are they using subsequently. Are they moving from app to web or web to app? So these cohorts were very tricky because there was an overlap between user behaviors.

So, to remove that dependency we just focused on two sort of people. One who were just using app and two who were just using web to engage with the program or to complete the program. So we had two very different cohorts, people who are using app as a mode to complete and engage with the program and people who are using web as a mode to complete and engage with the program.

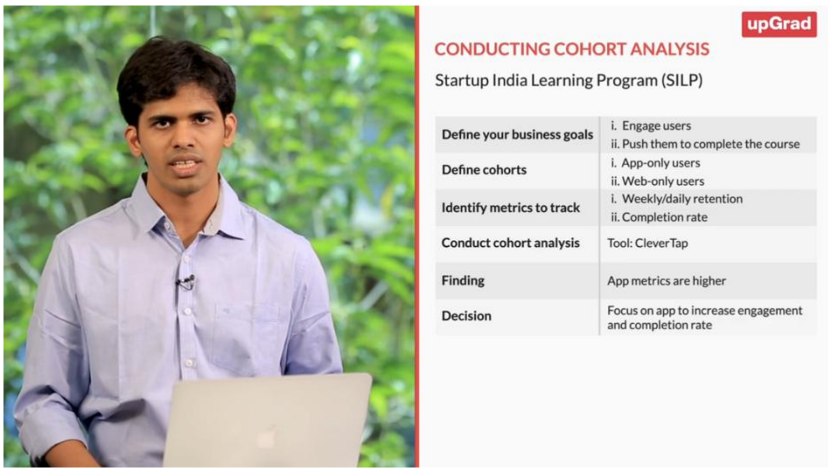
3. So, once we have these two cohorts defined the third step was defining the metrics.

So, the metrics in this case once we are clear of the business logic is very clear. The first metric is engagement where we have to see whether people are coming to the platform every day or every week. In short, weekly retention is that metrics or daily retention is the metrics that we are looking at. The second metric which was obvious from the business goal is the completion rate. So, people should complete the program. So that they can get to see the value that upGrad offers and they can move to other programs, other paid programs. So, these are the two metrics that we defined in this case which is the first was weekly retention or daily retention and the second was actually the completion rate.

And finally, once we had done this we did cohort analysis using CleverTap as an analytics tool. So CleverTap is another tool which we can use to do this sort of web analytics. You can also use Mixpanel or Google Analytics for the same thing but since we were pushing our data into CleverTap, we looked at the CleverTap numbers.



Now in CleverTap when we looked at the numbers what we found was, the retention in case of app was much higher and it is actually intuitive if you think about it, because people on app tend to be, tend to have that app available to them 24/7 whereas people on web need to find their laptop, they have to sit somewhere, they have to plug-in earphones and then they have to start. On app it’s just opening your mobile app, opening your app and then final plug-in.

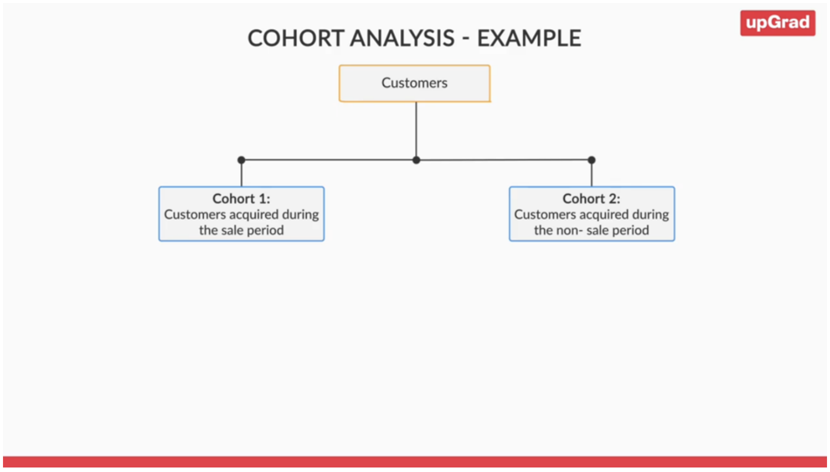


So the engagement was much higher, people were actually completing more as well. So both the business metrics or both the metrics that we were trying to achieve are the numbers the numbers were supporting that we should actually focus on app more to increase the engagement or completion rate. So this is the demonstration of how cohort analysis helped us make decisions which would actually improve the business.

Next, we will see how can we do cohort analytics using tools like Mixpanel and Google analytics.

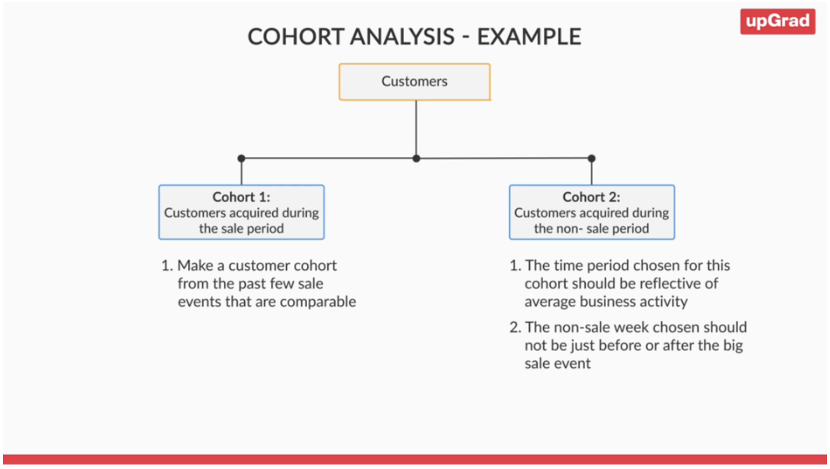
So now, let's look at a concrete example of where cohort analysis can help us and how. And let's try to answer a broader question right.

Let's say you were working at an e-commerce company and e-commerce company has lot of sale event. And a question is what is the value of the customers acquired during sales or you know are there any differences, rather between the customers acquired during sales versus non-sales? And you could have some ideas as to why this question is coming in the first place. Some people may suspect that you know the customers acquired during sales are not so sticky in the sense that they don't have such a strong repeat behaviour. Maybe they came in because of some pricing or the offers which you had at that point in time and they purchase less often and maybe their general lifetime value is not so much right. So to answer this kind of a question or to let's say, study this behaviour, we can use cohort analysis.



We can divide into two cohorts. Cohort 1 could be the customers who are acquired during a sales period. Cohort 2 could be: customers acquired during the usual non-sale period. And then what you will do is, after defining these cohorts, follow the behaviour of these customers. The behaviours you are interested in this case could be say, repeat behaviours. So the customers acquired during sales or non-sales, how many of them repeat in the subsequent visits? How many of them buy in the subsequent visits? You could look at their purchase behaviour like how many of them or what is, the general revenue per order for these customers, the revenue per customer or the number of orders they make in the subsequent weeks. You can study these behaviours.

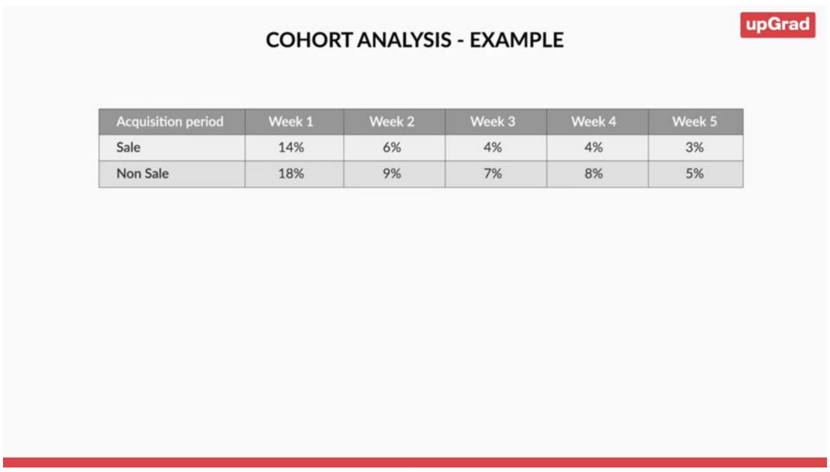
So let's take one example where we will make these cohorts by acquisition period and we will study the repeat rates. We will study how much they repeat over the subsequent weeks. First off, let me clarify that we don't typically do this based on just one week of data, so, for example, we have mentioned acquisition period as sales. It doesn't mean that we just had one sale event.



What is a better approach is maybe look at the past, few sale events which are comparable, not one big grand sale but more common sale events which are comparable and take all the customers over there and make a cohort out of them. So, that you get a better sense of it. And in the non-sale time you pick some weeks, which are your average business as usual days and those weeks, some N weeks become your non-sale cohort. And off course, you have to be careful that you're choosing the weeks accordingly, for example, if you chose non-sale week and that this week was right before the big sales week, then, of course the behaviour is not worth studying. It's not really reflective of what a general non-sale acquired customer is. So, just be careful in choosing these weeks right, they should be clean data.

Once we have these right, let's say we have enough customers in the sales cohort, enough customers in the non-sales cohort, and now what we are interested is in looking at in each subsequent week how many of these customers were coming back to the platform?

So, the first view you want to look at is for each cohort, sale and non-sale, over the subsequent weeks. So week 1 is one week after acquisition, week 2 is two weeks after and so on. Each subsequent week for the next five weeks, how many of those customers came back? What proportion of those customers has come back? And if you see you have these numbers for sale and non-sale and right here, you can see that for the non-sale customer's right, the repeat rates are a little higher in general.



Acquisition period	Week 1	Week 2	Week 3	Week 4	Week 5
Sale	14%	6%	4%	4%	3%
Non Sale	18%	9%	7%	8%	5%

You can see that non-sale customers week, 18% come back compared to only 14 for the sale customers. In the week 2, non-sale customers have a 9% repeat, whereas sale is 6 and so on you see that there is consistently a higher stickiness in the non-sale customers. So you do get a sense from the data as well that the customers acquired during sale don't really repeat so much over the next weeks and therefore have a lower retention. So you could stop here. Cohort analysis here itself gave you a very good insight, very good view and confirm your hypothesis that indeed sale acquired customers are less sticky. The value could be a little lower than your non-sale customers. But let's take this one step further.

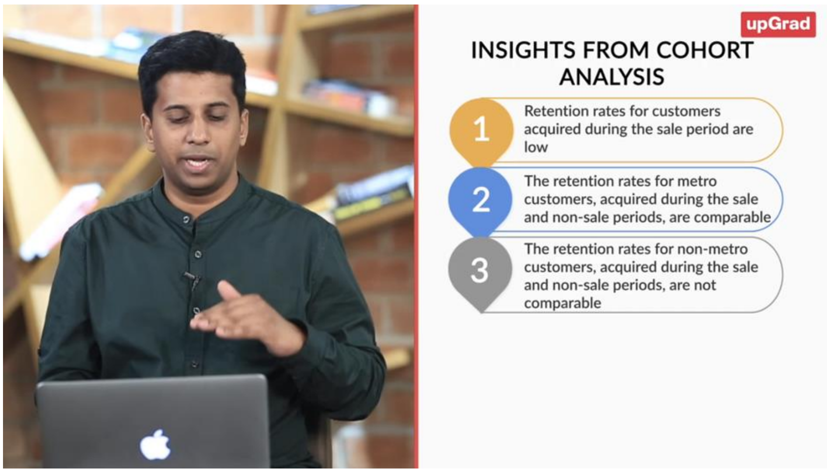
Let's now divide these cohorts further by some relevant segment. So the idea over here is that the cohort itself may not be a homogeneous cohort. The cohort could have different kinds of customers and maybe if we broke it down by some customer segments, we may find some other insight as well. Let's do that.

COHORT ANALYSIS - EXAMPLE

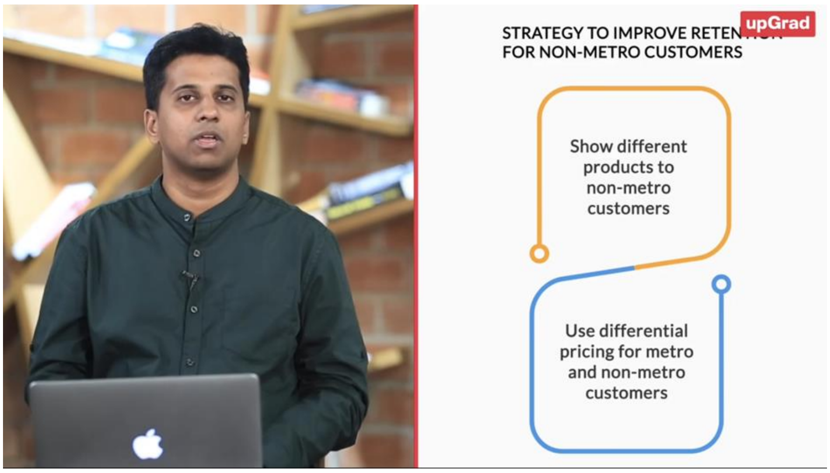
Acquisition period	Week 1	Week 2	Week 3	Week 4	Week 5
Sale	14%	6%	4%	4%	3%
Non Sale	18%	9%	7%	8%	5%
Sale					
Metro	16%	8%	7%	4%	2%
Non metro	9%	4%	3%	2%	1%
Non - Sale					
Metro	20%	10%	8%	4%	3%
Non metro	17%	8%	7%	4%	2%

In this case, let's take a very simple segmentation. Let's simply divide the customers into metro and non-metro. So what will we do? We have the sales cohort. We will break the sales cohort into metro and non-metro and look at these retention numbers for each of them. So if I come to the sales cohort here is something very interesting.

The metro cohort right, the retention rates for the week 1 and 2 are not very far away from the overall non-sale numbers. So maybe the customers acquired during sales from metro's are still not too different from the general non-sale customers. They could be of similar value. But when you look at non-metro, you have something very interesting: you see that the retention rates for non-metros are far lower right. You have just 9% non-metro guys coming back after acquisition during a sale. And the next week the number is 4 and you can see that these numbers are far lower than the non-sale and they're far lower than the metro numbers as well.



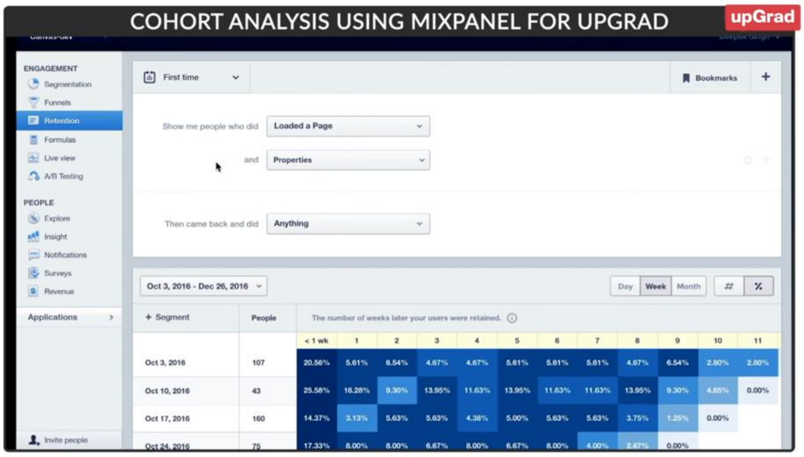
So what's coming out, is that although sale customers, sale acquired customers generally have a slightly lower retention, but not all the customers acquired during the sale are equal. If you just broke it down by metro or non-metro as a segment, we see that metro customers could still be of similar value to general non-sale customers, but non-metro customers definitely are very, very different in terms of retention and stickiness and their repeat rates are very low.



This is the observation you get from the data and from this observation you can form your own hypothesis and a hypothesis over here could be that maybe your non-metro customers are very price sensitive. Maybe the offers are what attract them to the platform in the first place to begin and when, in the usual day you don't have those offers or the pricing is back to your normal level. Maybe you don't have much for those non-metro customers. So maybe you have to identify a better strategy for them. You could use your domain understanding over here, maybe when they are on the platform in usual days, maybe surface different products for them. Maybe try out some differential pricing or maybe ease some of the cost they incur around shipping and so on. So again, there could be a lot of ideas to improve the situation.

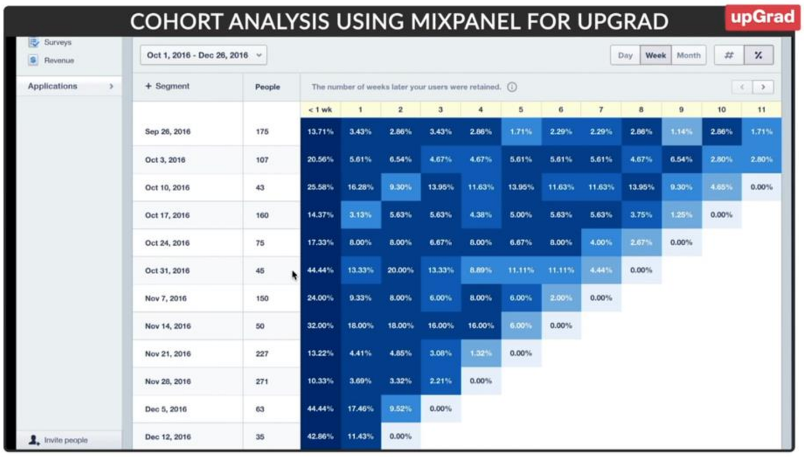
So cohort analysis when cut by different segments, can give you very revealing insights around your customer behaviour and can help you come up with hypotheses to improve your overall platform, your offerings and your product.

You have learnt that cohort analysis can be carried out using tools such as Mixpanel and Google Analytics. So now let's look at how you can conduct cohort analysis using Mixpanel.



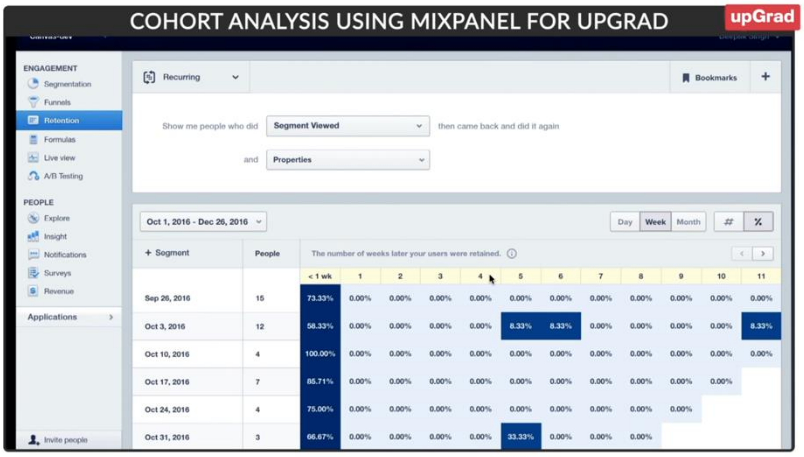
So, let me take you through the Mixpanel cohort capabilities first. As you might have noticed there is a retention tab in Mixpanel where what you can see is the first time recurring and addiction retention. Let me take through the first-time tab first. So, the first time is essentially for the users who are coming to your website for the first time and what you can see is let's say these are the people who came on October 3rd, 2016. These are 107 people who came to your website and then they loaded a page and then came back and did anything. So, this is the retention tab but if you see, this is the cohort of 107 people who came on October 3rd, 2016.

Now if I want to see how many of these people from this cohort actually came back a week later or two weeks later or three weeks later, I can see that in this tab. I can also do it day-wise or I can do it month-wise. I can also change it from percentage to numbers and numbers to percentage. So, this is essentially a cohort for you because what you are seeing here is these are the first-time users who are coming on a specific date. Let's say in this particular month or in this particular week or in this particular date and then you are analysing their behaviour.



For example, you might notice that people who came on let's say October 3rd, their retention was 20%, whereas people who came on October 31st, their retention was 45%. So, both of these cohorts have a different behaviour. And since this is the week, that means in this particular week, that is the October 3rd to 10th. The people who came were less retained compared to the people who came at October 31st week. So, this is the first-time cohort. Then you can also go into the recurring and you can say that okay, let me see people who viewed segment: who came and viewed segment. And I will switch this view to days. So, people who viewed segments on October 1st is zero, October 3rd its 5. So, there are five people in this particular cohort: People who viewed segment and came on October 3rd.

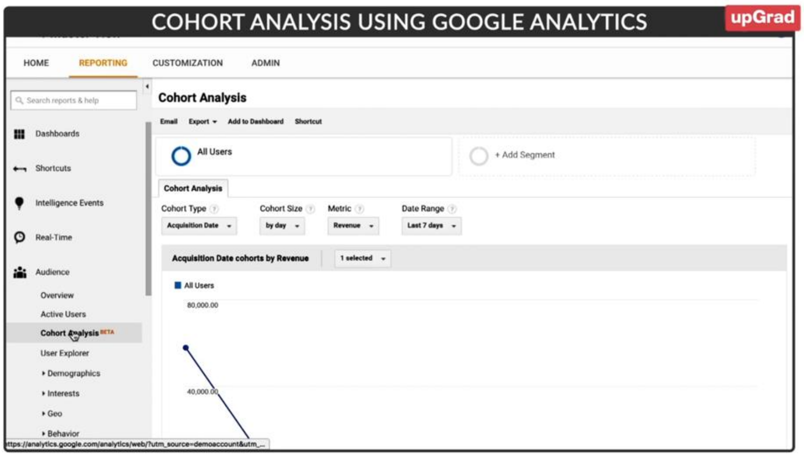
And, less than 1-day retention is 20%, 75%, 100% and 0% for the rest of the days. If I convert it into week, what I get is, in one particular case where November 7th to November 14th, there is week 1 retention off 12.5%. Further, what we can notice is that users in some cases are coming after 4 or 5 weeks. So, retention is non-zero after 4-5 weeks.



Now, you should notice that this is sample data. So, don't go by the exact numbers in this data. But this is how you analyse the cohorts in Mixpanel which is you see that on this particular date, or in this particular week or in this particular month, this is the number of people who came and then what those people do. Did they come back and viewed some segment? Did they come back and just loaded the page and went away? Did they come back, viewed a video, used chrome browser and went back?

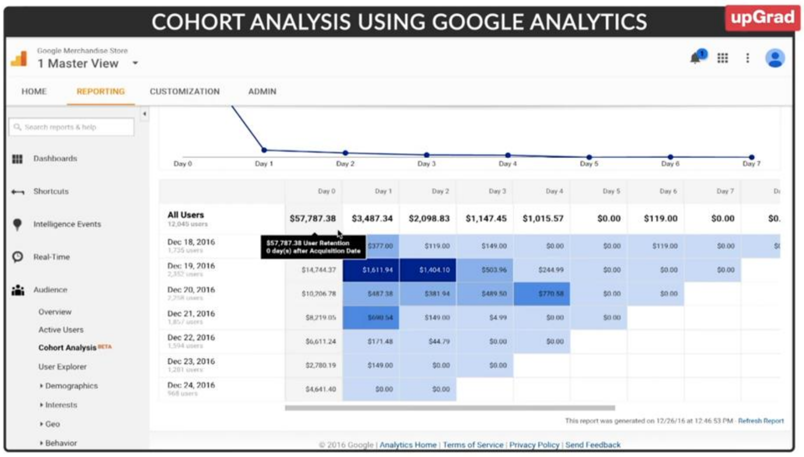
So you can create a segment for these users who come into different time span and these are your cohorts and then you can you analyse which cohorts are making more sense than others and that would tell you in what direction you should move or which kind of marketing is working for you or what kind of cities should you target so that you can retain users in a better way. So, that is cohort analysis in Mixpanel.

You learned about how to conduct cohort analysis using Mixpanel. Now, let's look at how you can do this using Google Analytics.



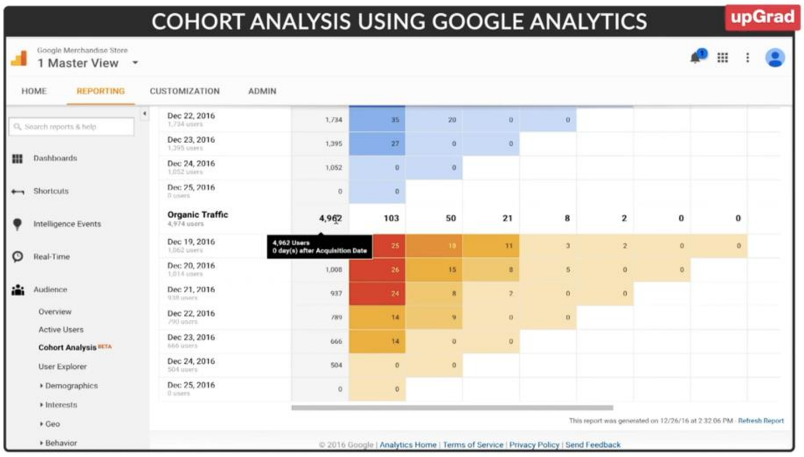
Now, let's move to the cohort analysis in Google Analytics. So, Google Analytics as you can see in the audience tab. There is a cohort analysis, which is a beta version of it. It's not something that Google has actually rolled out fully, but Google is making progress in this direction. Now, when you click on the cohort analysis tab, what you get is the cohort type, cohort size, metric and date range. Now, let me explain for a moment what these different things mean.
So cohort type is:

1. What kind of cohort are we looking at like as you can see, there is just acquisition date, which means that these are the people who came, whose first date was this, who came on this particular date on your website? This was the acquisition date, so that is a cohort type.
2. Cohort size is something which is by day, by week, by month, like how are you sizing the cohort? Are you looking at this on a day to day basis or on a weekly cohorts or a monthly cohort, just like we saw in Mixpanel?
3. The third is: what is a metric is, on which this cohort analysis is based on. For example, you can choose a metric like revenue, you can choose a metric like sessions, and you can choose a metric like user.



Now, what does a metric like revenue means in terms of cohort. So, as you can see, I can choose by day revenue and date ranges - last seven days. So, what I get is that on December 19th, 2016; day 0 people spent around 14,744. Now, the same cohort of 2,352 users came on day one and spent $1611 and then on day 2, $1404 and, as you might notice, some of these cohorts are spending much more than the other ones. Now, this sort of cohort analysis will give you a confidence like on what particular days are these weekdays or weekends, the people coming are spending more.

Similarly, what you can see is the metric you can change to, let's say users. Now, what it means is on December 18th, whatever number of users came, which is 1735, how many of those came to day 1 like 57 of those, then 29 on day 2, then 27 on day 3. Now, the thing to notice is again, on December 19th, the 2300 users that came out of them 104 came back on day 1 and 70 came back on day 2. So, this way, what you can see is you can calculate a percentage and you can just see which of the cohorts are behaving in what way.

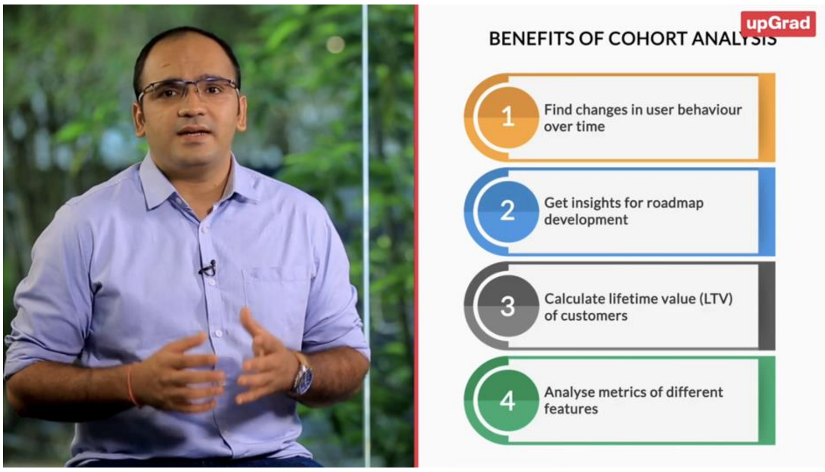


The interesting thing to do here would be adding a particular segment. For example, here if I add the organic traffic right, if I apply the organic traffic in this particular acquisition date cohort, what I get is the organic traffic is 4,962 out of which 103 are coming, a day after acquisition date right and you can compare this organic traffic with the all user traffic. So, as you can see here, the difference is not significant like, this is approximately 50% of total users. But day 1 retention of organic traffic is actually much-much less than the all users, which mean that organic traffic users are not getting retained that much.

So this is how you do cohort analysis in Google Analytics. The best thing about Google Analytics is, you can create a lot of segments and then you can analyse those segments. And these segments can be by age, by city, by source, medium or campaign, anything and you can do the same things in Mixpanel as well. You can use different dimensions of Mixpanel to segment users and based on the activities that they did on a particular day. You can put them into a particular cohort and then you can see the behaviours of the cohort.

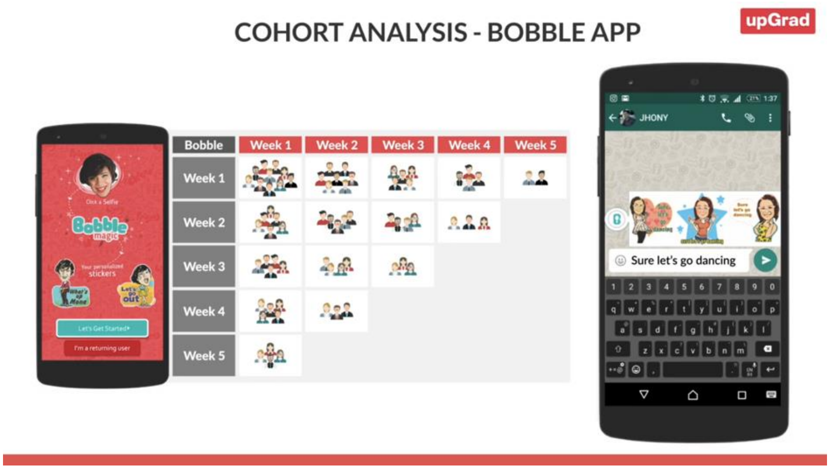
This is all for now. In the next video, we'll do a quick recall of what you've learned in this session. See you there.

Hi there. Let us do a quick recap of what you learned in the session. We started the session by understanding what a cohort means in a products context, which is a group of people who share similar characteristics over a period of time. For example, the people enrolled in upGrad's product management program for January would form one cohort. Similarly, people who enrolled for upGrad’s product management program for March would form another cohort.

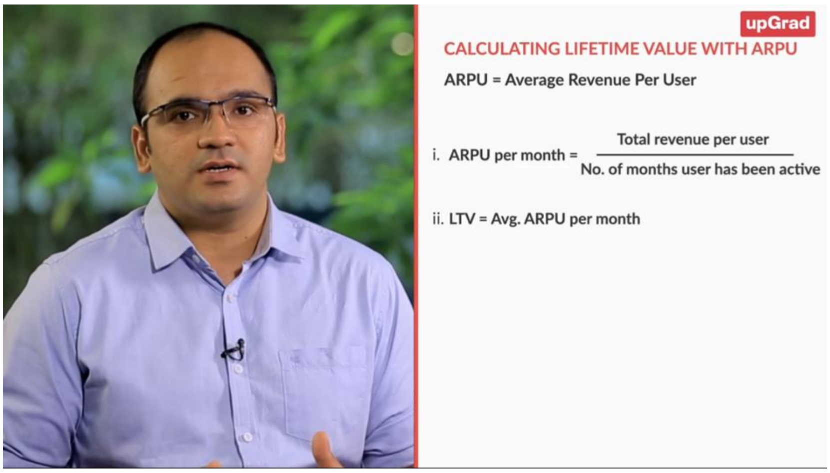


For any product, cohorts can be created based on various parameters, such as user demographics, user actions, and time periods. Further, you learnt about the benefits of conducting cohort analysis. These were, finding changes in user behaviour over time, get insights for roadmap development, analyse performance of features, calculate lifetime value of customers, analyse metrics of different features.

Another important thing which cohort analysis helps you with is that it enables you to see growth and engagement metrics separately. And this is very important for you as a product manager, because the focus is generally on increasing the revenue of the company. But sometimes high growth may mask the fact that you're losing repeat customers or have very low engagement for your existing customers. We went through an example of a SAS company to understand this.



Then you saw how the Bobble app analysed its cohorts week after week, identified that usage increased because of the improvements made in the keyboard. For example, one thing the team changed was to allow users to share stickers directly from the keyboard. You also looked at the CleverTap example where CleverTap used cohort analysis and help Ixigo increase its repeat transactions by about 6x to 7x. Next you learnt about LTV or lifetime value, which predicts all the value a company will derive from his relationship with the customer.

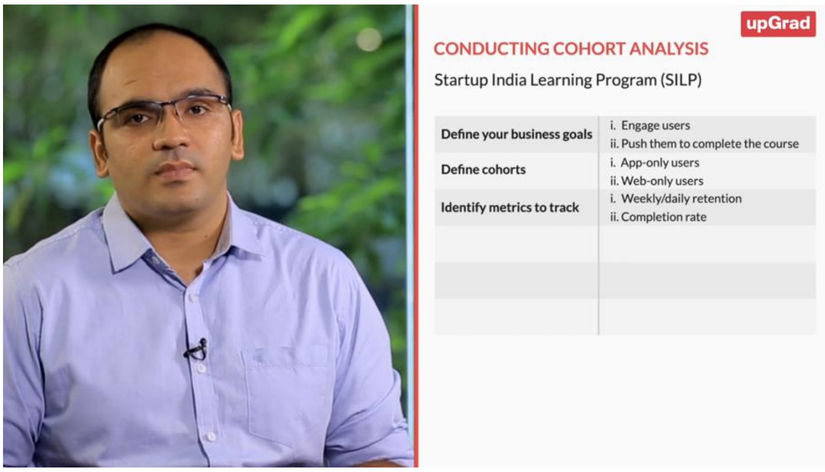


One of the approaches to calculate LTV is by calculating the average revenue per customer per month or ARPU per month. ARPU per month is the total revenue for a user divided by the number of months that user has been with you. To calculate the LTV, you take the average of the ARPU per month for each customer and multiply it by the timeframe for which you want to calculate the LTV. We calculated the LTV for Grofers using the ARPU approach. But one of the disadvantages of this approach is that it does not take into account the changes in customers' behaviours.

Then you learnt how to calculate LTV using cohort analysis for an ecommerce business. One of the advantages of using this approach is that the cohort analysis does not assume all months to be the same when it comes to the revenue.

You also looked at the steps involved in conducting cohort analysis.

1. First, you need to define your business goals for which you want to do the cohort analysis.
2. Then you have to define the cohorts. This can be done on parameters, such as user demographics, user actions and time periods, which we already discussed.
3. Next, list down the metrics through which you can track your progress towards achieving the defined business goals.



Finally, you would do the cohort analysis using analytics tools, such as Google analytics or Mixpanel. You went through the example of start-up india learning program to understand this process. Here, the business goal was to engage people and push them to complete the course. The cohorts were created based on which platform people used to do the program, which was either website or app. The metrics tracked were weekly retention rate and completion rate.

Finally, upon looking at the numbers from CleverTap for both weekly retention and completion rate, the team found the app to be more engaging and even the completion rate for the app was higher than the website. So, we started focusing on the app more. At the end of this session, you saw how to use Mixpanel and Google Analytics to conduct cohort analysis.

And with that, we've come to the end of the session. See you next time.