Marketing Analytics

Professor Swagato Chatterjee

Vinod Gupta School of Management

Indian Institute of Technology, Kharagpur

Lecture 07

What Consumers Want

Hello, everybody, welcome to Marketing Analytics course. This is week two and I am Dr. Swagato Chatterjee from VGSoM, IIT, Kharagpur who will be taking this course. So, in the week two we will talk about What Consumers Want. So, this will be the overall topic of this particular week and there will be as usual six sessions and we will discuss that how we can find out what consumers want.

So, to start with the first basic marketing thing that we have to understand is that there are certain concepts called Need, Want and Desire or Demand which is there in marketing 101 and there, these three things are different and for a marketer, these three things are very important that why these things are different and what I should focus on and what I should not focus on and so on.

(Refer Slide Time: 1:11)

So, need, want and demand. So, the first basic thing that I want to focus on is that what is a need? So, need is all the basic necessities of a human being, which is required to survive. For example, the physical health, the food, the shelter, safety, financial support. So, initially, we used to talk about food, shelter and probably the clothing that we wear, but instead of that, there will be much more basic things that are required, probably a little bit of connectivity, as it has been written here, mental well being, quality education.

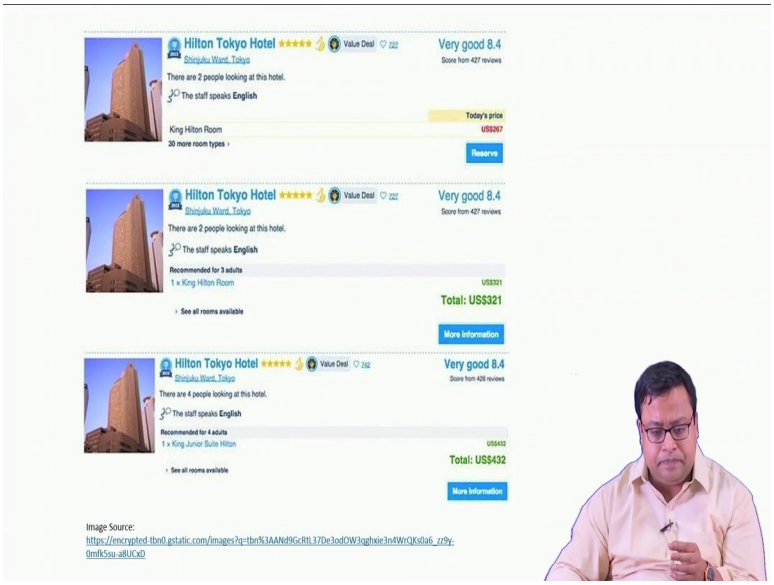
Now, all of these things have become a part of our basic needs. Now, this is our basic need, and we need this to survive, but there is another concept called want. So, basic need is something without which you will not survive. On the other hand, there is something called want which is actually a little bit higher version of need and it is probably sometimes opposite to need, it is not something that you will not survive without of it.

So, you will actually survive without your want as well. But still you probably love to have that, love to consume that or love to have those kind of services with you. So, that is called want. For example, let us say I have given a picture of a baby here, a kid, for him food, quality food, good food is something that is his need. But baby wants candy, that candy that he is asking for, he or she is asking for is something that is his want.

So, without that candy he will survive, but still, he wants that because that is something he founds pleasurable, or probably something that attracts him more or her more and etcetera etcetera. So, that is the basic difference between need and want. For example, in the case of us, when we travel, let us say a safety, a safe place to sleep at night is something that is a basic need. On the other hand, a hotel with a comfortable bed and a good bath and under running hot water and Wi-Fi facility and etcetera etcetera is your want.

So, without a comfortable bed, without the running water, you will still survive, you will still be able to travel, but your travel experience will not be that good. So, majorly marketers focus on this want part, and they want to convert this want to a desire or demand part. So, desire with the price tag is something called demand. What is desire? So, when a want is targeted towards a specific thing, then that becomes a desire, and that has a certain price tag with it that becomes a desire.

(Refer Slide Time: 4:06)



For example, let us say I am, as I told that I would be traveling, and I am looking for various kinds of product options and etcetera or probably, in this case, hotel options and etcetera, where I will stay, and this Hilton Tokyo Hotel has rooms of different quality. And you will see that the room's quality has slowly improved over one step to other steps. Now, all of them are provided by this Hilton Tokyo Hotel. Now, the one is at 432 dollars, the another is 321 dollars and another is almost 267 dollars which is very low, the value deal is 267 dollars.

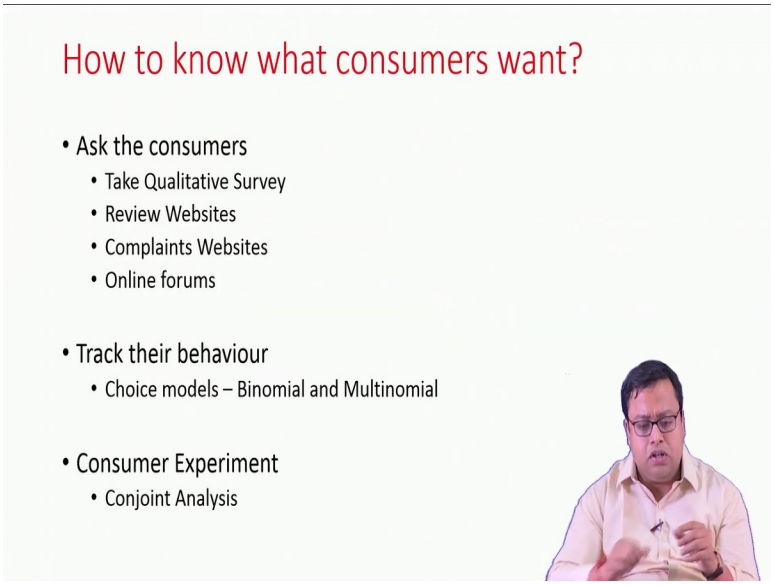
Now, what they are trying to say is that which one of this particular products that they have, and associated prices that they have, which one will be having higher desire for the customer, customer will want to have which one and if the price and the desire matches, if the willingness to pay of the customer is matching with the desire, matching with a price, then the desire will convert to a demand.

So, he wants to stay in this, the last one which is let us say King Janitor Suite Hilton room, he wants to stay in that, but that is his desire that 'I would want to' so good hotel is something that is a need, lots of facilities in the hotel is the want, when that want is

targeted toward a particular solution, which is in this case, this King Junior Suite Hilton, that is a probably, that is a desire, and when the desire converts to willingness to pay of the consumer, that willingness to pay if that matches with the price, that is quoted here 432 dollars, then that becomes a demand.

If the willingness to pay is lower or equal to that particular price that you are quoting, then that becomes a demand. Now, why this story is important because this story is important to know that what converts a need to a want, what converts want to a desire and how to make that desire to a demand by giving the right price for that particular desire, all of these things is important for a company to know and that is where the what customers want, what customer values, this kind of story comes up.

(Refer Slide Time: 6:38)



So, what are the various sources of information from which you can know what customers want. So, there are lots of places where customers themselves say what I want, what I do not want, and etcetera. One of the major sources of that is probably the qualitative survey that you take; oftentimes you take surveys, qualitative, sometimes quantitative as well, all form of surveys that you take to say that whether you are happy with us, if you are not happy, then what is the place where you are not happy, what kind

of features you are asking for, what are you looking for, what are you not looking for, all of these details you can ask the customer.

So, that is the first way basically, simply straightforward ask the customers what you want, but you do not want what you like what you dislike. Sometimes you do not ask somebody else asks for you. For example, let us say the review websites like Amazon has all the reviews listed there. Sometimes there are verified customer, and Amazon asks those customers that okay, ‘ do you want to put a review?’ Sometimes a customer who is not verified customer has bought from some other source can also go and put their review there.

So that is an accumulation of lots of reviews, sometimes by verified customers, sometimes may not so verified customers. So, that is also a major source of information. A product company can actually go through all those product reviews that have been posted there and can get an idea that what are the things that they want to, they should improve, what are the things that they should not improve or they are okay with, and what has higher importance, what has lower importance and etcetera.

And we will do a small case study, we have started doing a small case study in the last week also, we will continue on that. And we will do later point of time more case studies on the text part of the review also to find out what matters? What is something that customers want? Review websites can be Tripadvisor.com also. So, Tripadvisor.com as I was telling or a lot of other hotel reviews have both quantitative data and qualitative data, both types of data is collected.

So, if you have both types of data, you can actually mix and match quantitative techniques and qualitative techniques, probably text mining techniques like let us say probably natural language processing as well along with your quantitative techniques, all form of quantitative techniques, regression, ergonometric techniques or machine learning techniques and you can all of them you can combine them together and create insight.

So, in the latter part of this particular course, I will try to do that, how that is done. But, here also we will see some of the easy ways to do the same thing. It can be the complaints

website also so, Consumercomplaints.com or Consumercomplaints.in, if you go to those kinds of websites, you will find out how people are giving, proposing complaints. Now complaints are generally negative reviews, there will be nothing positive there in most of the cases, but you also will know that what are the major reasons for which customers are complaining.

And there can be online forums also where people discuss with each other, it is not only, for example, what consumers want in that space, online forums comes very handy when you are doing a product improvement. For example, let us say Salesforce.com there is a let us say forum, online forum, users forum and users forums are there for all other various kinds of tools that are there.

For example, Microsoft's Power BI, or let us say IBM's Watson all of these guys will have their own online forums, where the users will actually talk about what kind of improvements has to be done, what kind of improvements is required, what they are facing difficulties and etcetera. So, one user posts, another user gives the reply of that and etcetera. Now, sometimes one user gives a product recommendation or improvement recommendation; other users come and vote there. So, from those kinds of information, you can get an idea of what customers are asking for.

Now, the second part, this is what the customer themselves either by you asking or they are on their own motivation, they are expressing their requirements, their needs, but not always is something that happens. So, sometimes you have to track customer's behavior, what they are doing, what they are not doing, to find out that what customers want. For example, one of the basic things that we do here, in this case, is choice models. And we will discuss in length about choice model, choice model is nothing but quantitative modelling about customer's choice, as simple as that.

Now, customers can choose many things, customer can choose probably, whether I will continue with this service or not. So, let us say you have taken a Netflix subscription, and Netflix wants to know that out of all my customers, let us say 1 million customers, which of the customers have more chance of reneging, that means they have more chance of not renewing their subscription in the next year and what are the drivers of that, what is the

reason, what customer are asking for, which they do not have that is why they will churn out. If they can find out that then they can actually create designs, they can create products, they can create offerings, which is targeted towards those customers who are trying to churn out.

So, here it is a binomial model, binomial because there are two options available either customer stays or customer goes out, one or zero. Now, it can be a multinomial model also, for example, let us say the customer has three choices let us say, customer can, or more than three choices let us say, customer can stay where he is, customer can upgrade, customer can downgrade or customer can churn out. So, customer has four choices, today I have a subscription of the medium plan, I can upgrade to a premium plan, or I can downgrade myself to a basic plan or I can stop using the product at all.

So, there are four options available to this customer in this kind of a situation and it is available for many other, for WordPress let us say, WordPress has lots of plans, if you are a user of Wordpress.com, where you can create all your websites in the platform rather than downloading, I would say codes and etcetera, you want to just do a drag and drop kind of designing of a website, you can use Wordpress.com and Wordpress.com has multiple features along with multiple pricing options, and as more and more features come out, the pricing options are also there, and the prices go up.

Now you can upgrade, you can downgrade, you can unsubscribe to certain products, you can subscribe to certain products. All these options are there. When all those options are there in your hand, and you want to model consumer's behavior and you want to know from that information that what consumers are asking for, you have to do a multinomial choice modeling that is also part of our, of whatever we will cover and I will come to that.

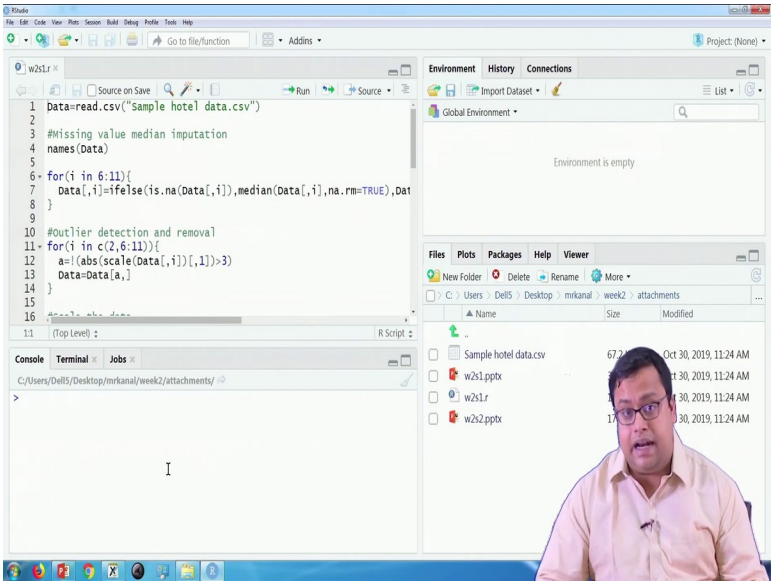
And the third way which is a combination of the previous things that I told is something where I am doing an experiment, experiment means I am creating the situation, it is not a natural occurrence that is happening, I am creating situations for the consumers, and then I am asking the consumer that in this situation how you will behave. So, I am tracking his behavioral data only or his preference data only, but I am not directly asking that what

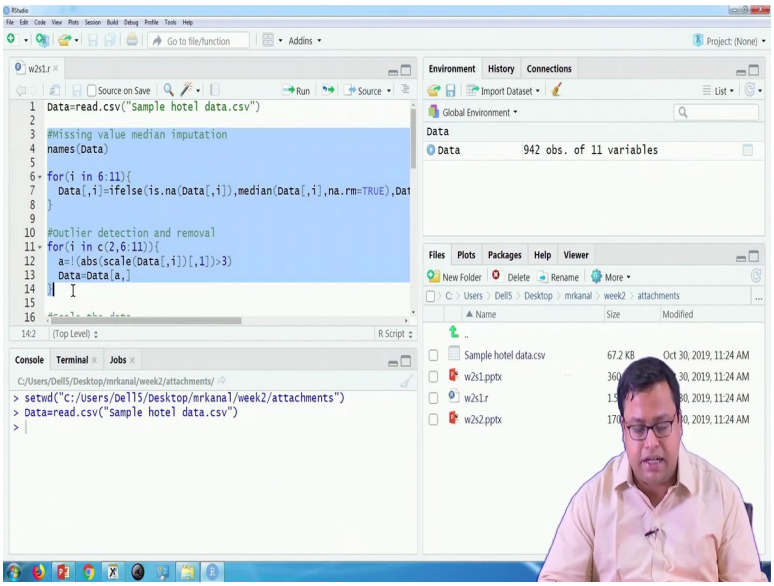
you want or what you do not want or I am not directly tracking his behavior in an actual situation, sometimes we try to create certain situations which are not available in the market till now.

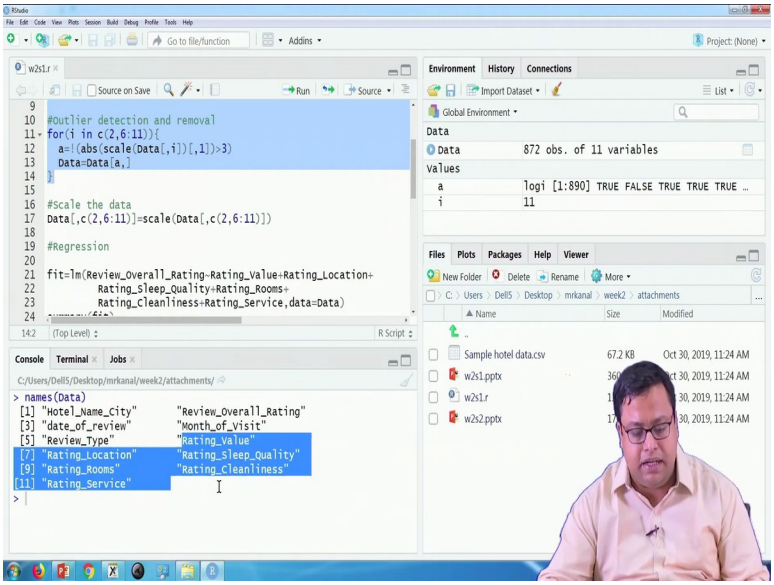
Let us say you are trying to introduce a new product. So, if the customer will ask or has a need or they will like this new product or not, some new features you are trying add on, that new features, whether customers will appreciate that particular feature or not, you cannot get that from this behavioral data, the option number two, you cannot get that, why? Because that offering is not there in the market. If that offering is not there in the market, you have no idea whether the customers will like it or not.

So, then you have to create situations for the customers for some sample customers that okay, I have created the situation, now you tell me whether you like this or not. So, that part is called consumer experiments. One of them is conjoint analysis where we create multiple options for the consumer and then ask the consumer to either rate or rank or give a choice about those particular options. So these are various things that are there.

In today's session, session number one of week two, we will talk about the first part and that too, with quantitative data, not with qualitative data, and then we will in the next sessions talk about the rest of the parts. (Refer Slide Time: 16:02)







So, I will come back, in your files, you will find out that there is a week two S1, w2s1.r file, I will ask you to open that. So, it is opening. So, as usual, when we start, I would ask you close everything other than this one, clean your console, clean your global environment, everything is clean. So, we will work with the same data set that we were working before, the same hotel review data.

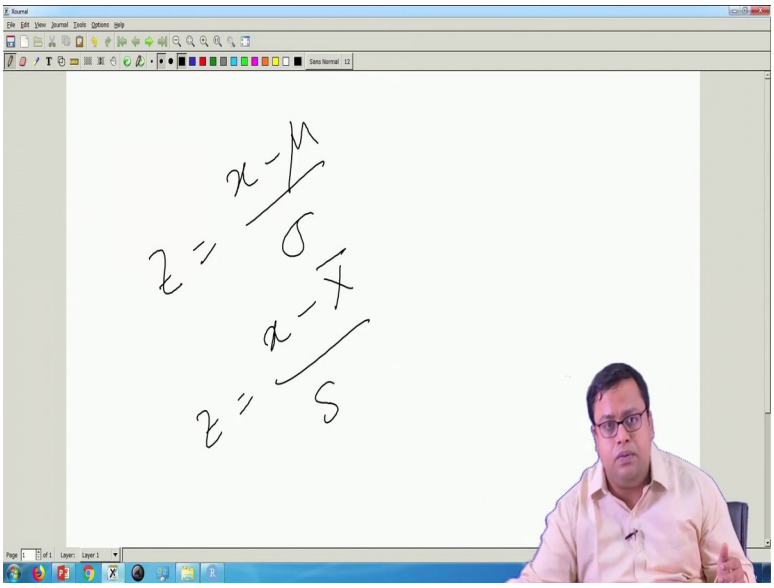
So, if I just put the session directory to the source file location, make sure that here is the sample hotel data file, and here is my W2S1 file they are holding up in the same folder. So, I can now read my data set and my data set has 942 observations of 11 variables. This

is what we have done in the last class. We will also find out the missing value; we will also find out the outlier and remove them.

So these codes are actually taken from the last class, so we will select this and run it directly. And then I will try to compare the attributes, remember there are six attributes, if you remember that there were six attributes and those attributes where rating value, locations, sleep quality, value means value for money, location, sleep quality rooms, cleanliness, and service.

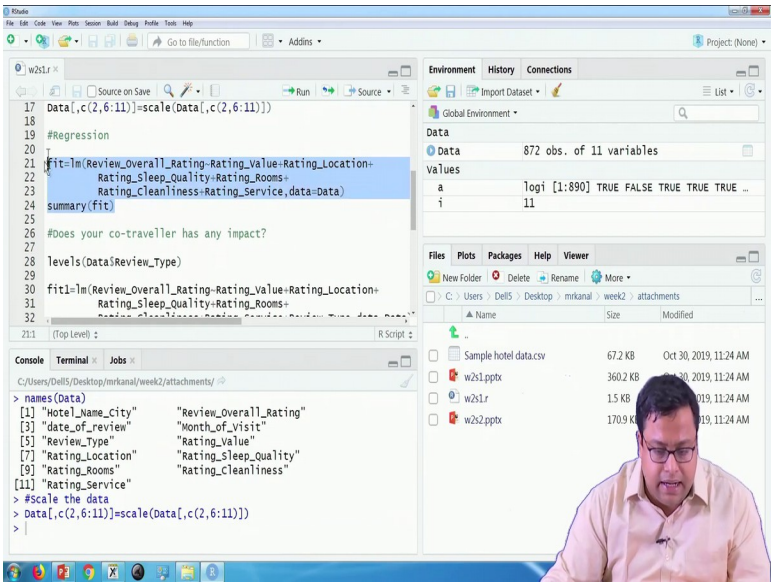
These are the six aspects, and I want to know their relative importance. So, one thing is they are all one to five point scale so I can directly run a review, but sometimes the means are different, no? So, to make sure and probably the standard deviation of these variables are also different, to make sure they are compatible, we change it to their scaled form. So, I change both the y variable and the six x variables to their scale form. So, scale means standardized form, standardized means the $z= (x- \mu)/\sigma$.

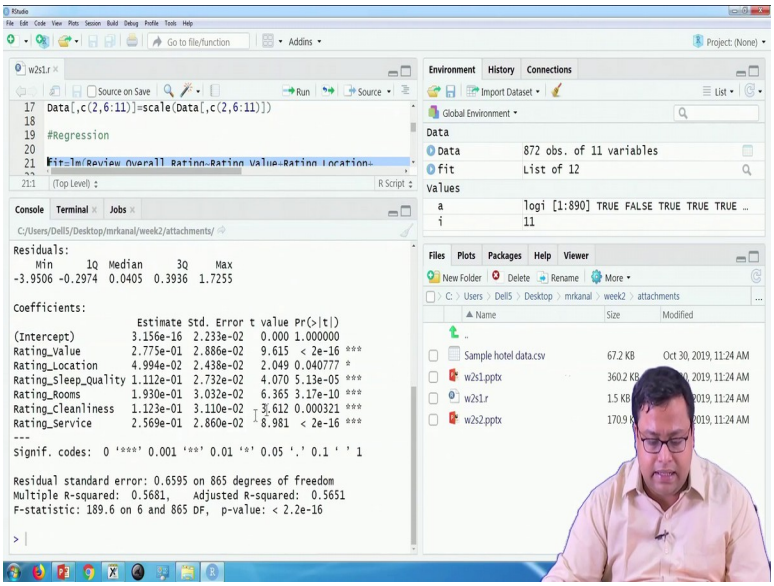
(Refer Slide Time: 18:28)



So, that is the standardized form, so where $z= (x- \mu)/\sigma$, if it is a population and $z= (x- \mu)/S$,if it is a sample. So, whatever so you find out that and you find out the corresponding z value.

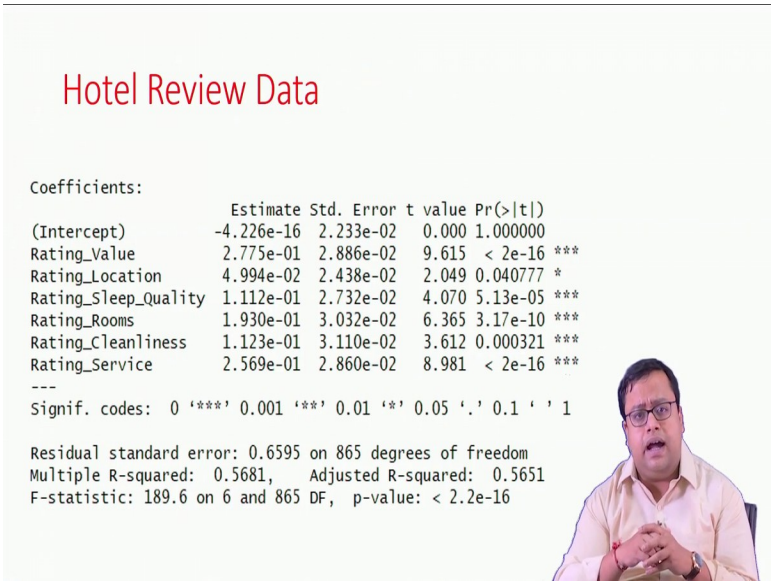
(Refer Slide Time: 18:47)





So, once I find out the z value, what I will do is, I will be doing a regression. So, this is the regression we have done before where I have taken a review overall rating as my y axis and all the six things one by one, so location plus sleep quality plus rooms plus cleanliness plus service and so on as my x variables. So, once I do that and run that I get a data, I get a result ,and the result is given here and that is the same result I actually got out here and now, I want you to see this result carefully.

(Refer Slide Time: 19:24)



So, I have just run the regression and the F-statistic is 189.6, degrees of freedom is 6,865 and p value is much lower than 0.05. So, we are okay with this regression results. Previously we have also discussed about the correlation, so correlation is fine and then our adjusted R squared is 0.5651, in a marketing context it is good enough. The data size is 900, so it is not very big data that I can say that all these results are meaningless or we have to further probe in a greater depth that whether it is coming up to be in the making sense or not.

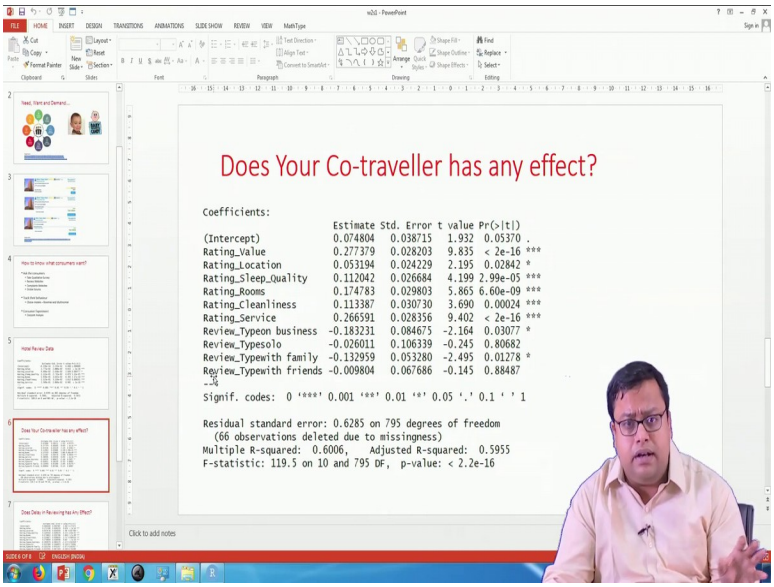
On the other hand, I can see here carefully that the most important factor is value for money. So, all the six factors are significant first of all, location is less, I will probably say that even the t value is lower than the other ones. So, the p value is also lower, and the coefficient value, which is 4.994 into 10 to the power minus 2, means almost 0.05, so that is 0.05 means that is the lowest out of all of them.

So, another but the highest one is the value for money, which is 0.28 almost 0.28, so 0.28 is the highest one, so here it is 0.28. So, this one is the highest one and then actually 0.05 this one is the lowest one, and then you can find out which one is more important than the

other one. So, this much basic information I can gain that what customers want? The most important thing that the customer wants is value for money.

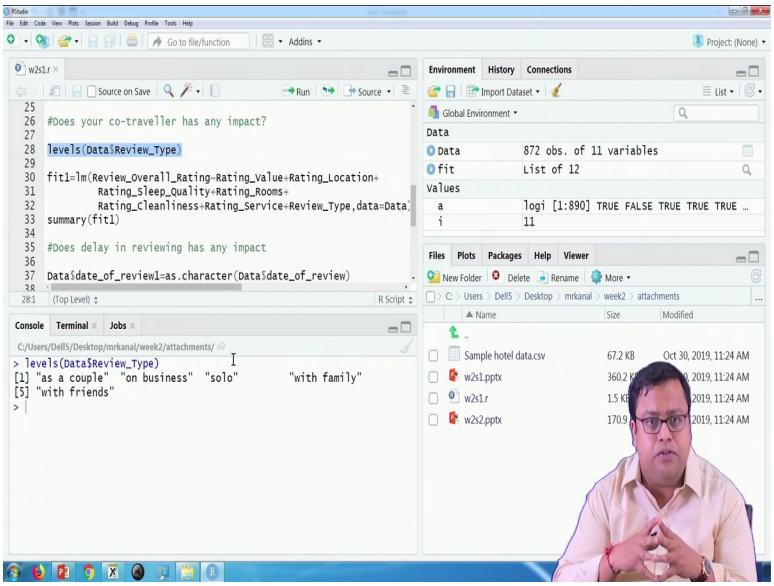
And then the next important thing is service, the next important thing is rooms, room quality, then the next important thing is cleanliness, sleep quality and so on. So, I will get an idea about what customers are asking for. Now, the question, the next question that comes up is that whether these, anything else matters other than the six attributes. So, these are let us say these are reviews, like these are actually something that customers have experienced before, they come back and talk about this. So, we have to know whether there is something else also that customers are asking for or not.

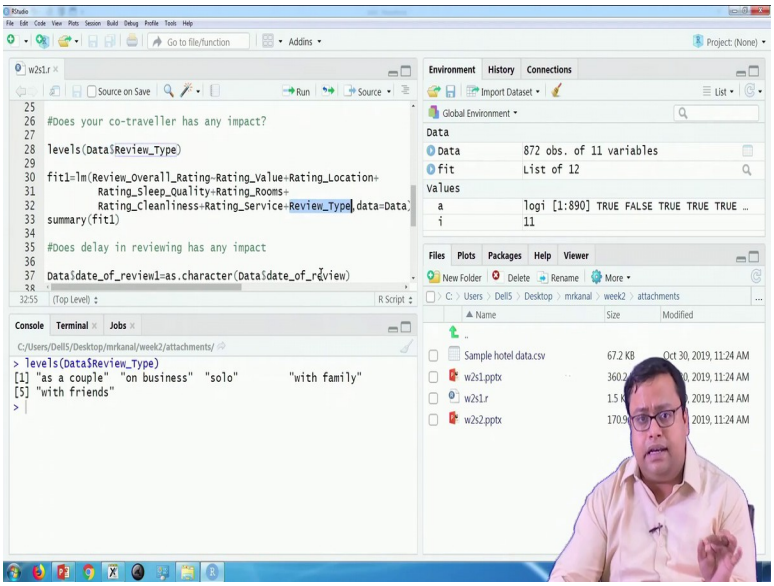
(Refer Slide Time: 21:58)



So, the second thing that we are asking is, ‘Does your co-traveler has any effect on your overall satisfaction?’ So, whom you are traveling with now, this is not something that the company can control but company should have an idea that whether the existence of a co-traveler affects customer satisfaction or not or what customers is asking for not.

(Refer Slide Time: 22:20)



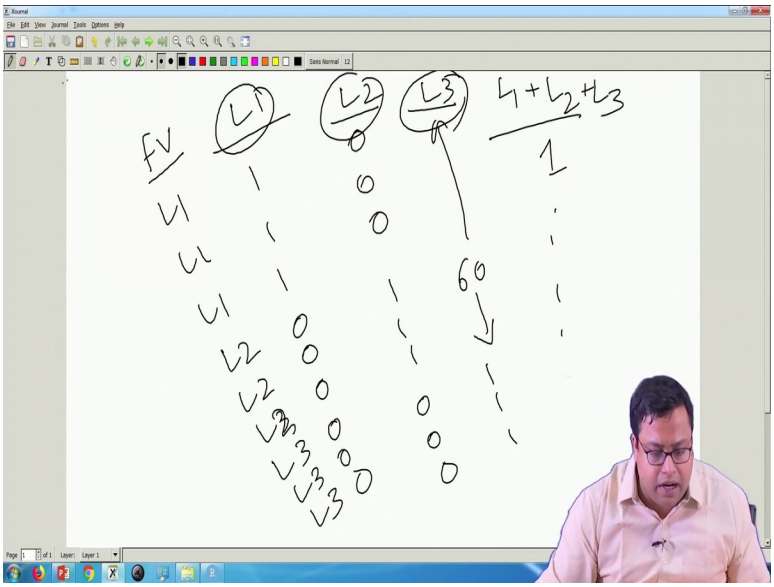


To run that we run another regression, if you just come down in this file 'Does your co-traveler has any impact?' so what are the co-travelers? I find out the levels of travel type. So, line number 28, I run and these are the levels that are coming up so if you see carefully it is as a couple so you are traveling as a couple or you are traveling on business or you are traveling solo or with family or with friends.

So, except business travel, all other four travels are probably leisure travels and then in the leisure also you can travel alone. You can travel as a couple means with your spouse or you can travel with your family, if you have kids and parents and etcetera together or you can travel with your friends. So these are the five possible options are there.

And I just add that the review type, I just add here in the regression. Now, regression in this particular case, you know, in many other softwares you have to actually create dummy variables. And you have to create it, if you know, you have to create four dummy variables out of this five, why four dummy variables? Because you have to take care of the multi co-linearity issue.

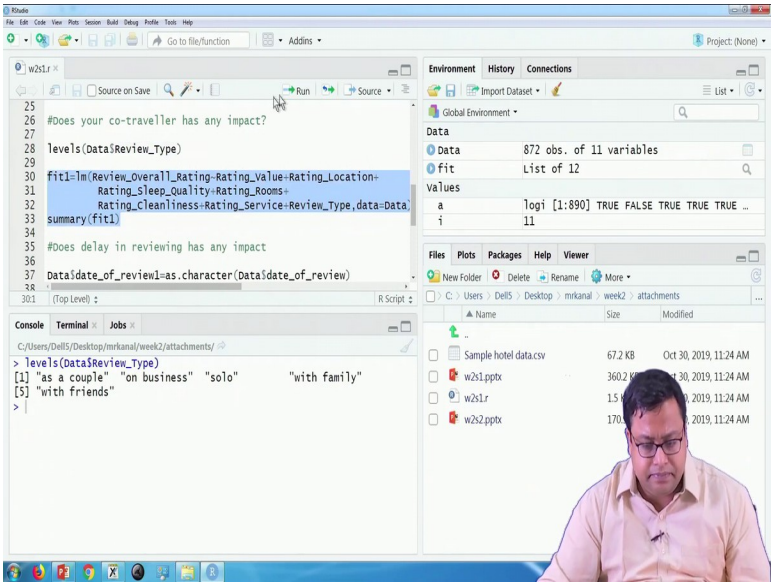
(Refer Slide Time: 23:35)

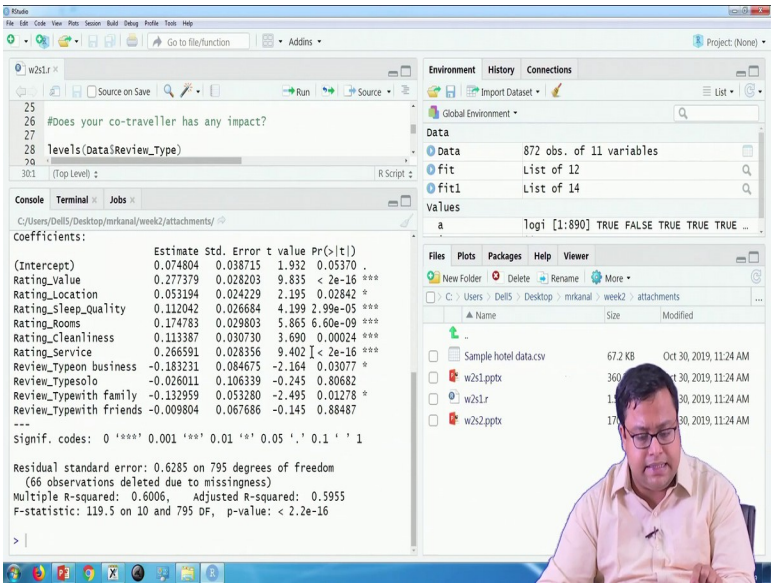


So, let us say I have, x1, so one factor variable fv, with just three levels L1, L2, L3. So, L1, L1, L1, L2, L2, L2, and L3, L3, L3, something like that. And I create L1, L2 and L3. So the first one will be 1, 1, 1, 0, 0, 0, 0, 0, 0 the second one will be something like this and third one will be six 0 and then 1, 1, 1. Now, if you just add them up in L1 plus L2 plus L3, this is always 1, this is always 1, for all of these guys.

So, you cannot take L1, L2, L3 together in the model, you have to drop any one of them. So, you choose which one you want to drop and based on that you create a dummy variable to find out that which one has an effect and which one does not have an effect.

(Refer Slide Time: 24:41)

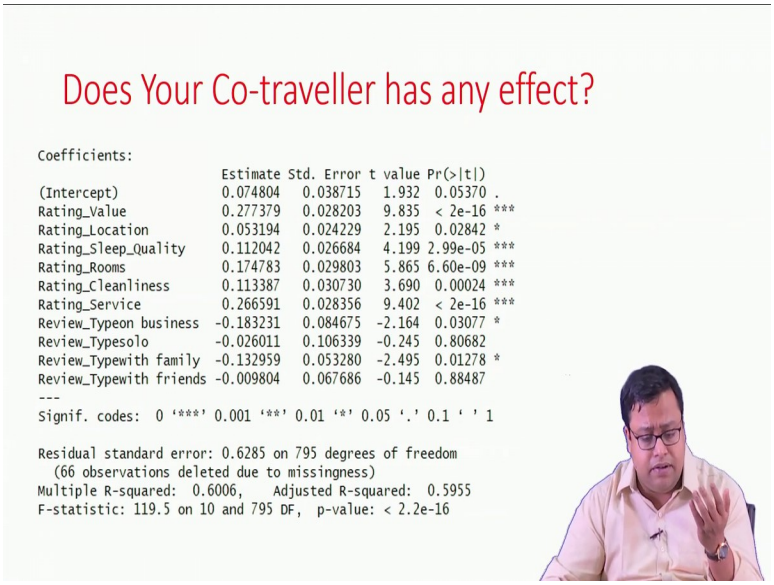




Now, here in this particular case, this guy will actually create the dummy variable on its own, you do not have to do that, that is, I would say is advantage of using a factor

variable in your model. So, once I run this, I get certain results again and this is the same result I have kept here.

(Refer Slide Time: 25:00)



Now, understand the result carefully. So, the previous ones are still there; even here rating value for money is more important than other things. But the two extra information that I got is that four new factors are coming up. So, remember there are five factors, five types of review type, one has been dropped, and other four have been reported here.

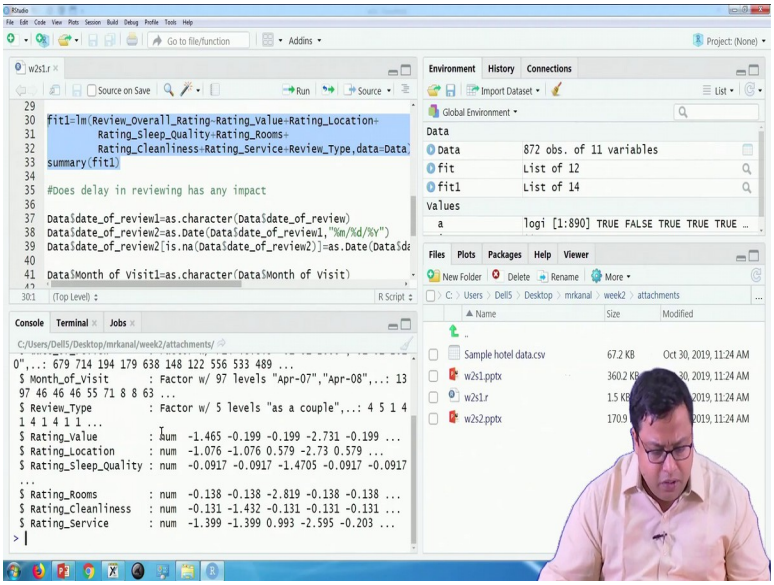
So, business, solo, with family and with friends, and probably the one that he has been dropped is with as a couple. So, as a couple is the level which is the first level, which is probably even in a nomenclature wise, alphabetically it is the first one, so that is why it gets dropped. So, this is something that R does, alphabetically whichever is the lowest level, so the first level that will get dropped.

If you want to change that, what gets dropped, you have to change that particular level alphabetical order, or you have to interpret it properly. So, that interpretation is important, that is what we are going to do. So, if you have an idea about how to interpret dummy variables, it is not something that is new. So, here what I am trying to say, the first thing that we are trying to say is in comparison to as a couple which got dropped, this solo guys and with friend guys are insignificant means that they are similar to them, they are not different from them.

But business and family will have lower satisfaction score -0.18 which is significant and -0.13, which is also significant. So, people who are traveling for business or people who are traveling with family, they might have lower rating, lower overall satisfaction level than people who are traveling solo or with friends or as a couple. So, you can say that people who are traveling in smaller group are most satisfied guys, people who are traveling with their family is the second best and the least satisfied guys will be the business travelers.

So, you have to pamper more towards the business travelers then towards the guys who are family travelers, and for other guys who are traveling with friends or solo and etcetera, they find out their own corner, and you might not have to focus on them much. Now, this is a very interesting finding, and this is something that you, if the customers, the hotel companies, can actually learn from the behavior of the customers.

(Refer Slide Time: 28:13)

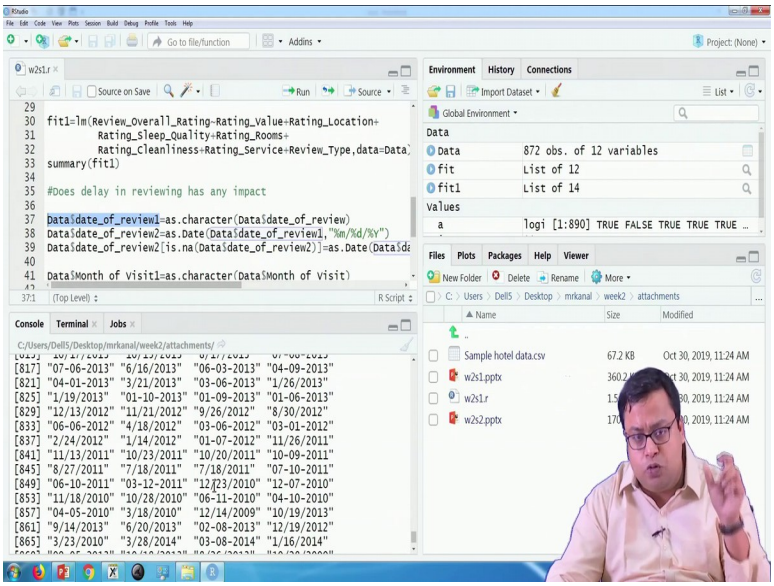


Now, another thing that becomes important is that whether the distances, remember these reviews are posted long back, so the memory will play a role. So, whether the distance, the time distance between the date of the travel and date of the review. So, you traveled in January, and you wrote a review on June, the six months whether that has any impact

or not because that will actually impact whether you have anything, in you, what kind of things you remember, or what kind of reviews you post and etcetera.

So, in a easy way, we will say that, so, let us see first of all to do that you will have to see the structure of the data, in the structure of the data you will see that there are two variables that we have, one is the review, month of visit, which is a factor variable and another is date of review, which is also a factor variable. So, we have to convert these to in date variables.

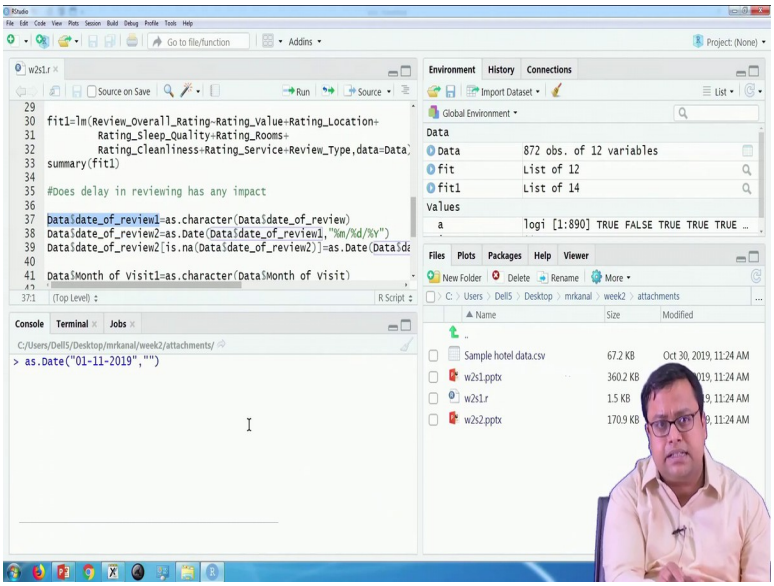
(Refer Slide Time: 28:40)



I discussed previously that there is another form of a vector, which is called date. So, the first thing that I do is change it to character variables. So, I write data dollar, date of review one; I am creating a new variable is as character date of review. So, whatever was date of review, I change it to its character form. So, if I run these and then show you what these guys are, these guys are nothing but the character forms.

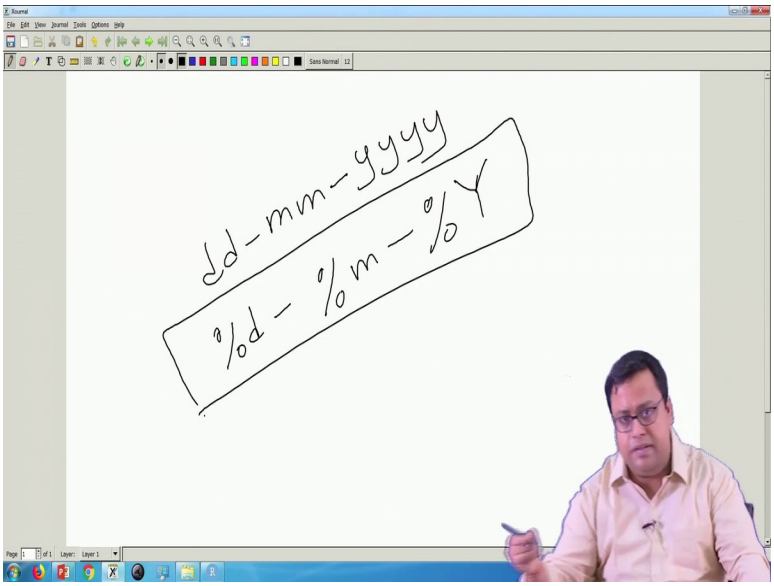
Now, here there are two types of dates, two formats of dates that have been written, one is the month, day and year written as dash. And then another is probably month, day and year written as slash, so I have to create this thing.

(Refer Slide Time: 29:24)



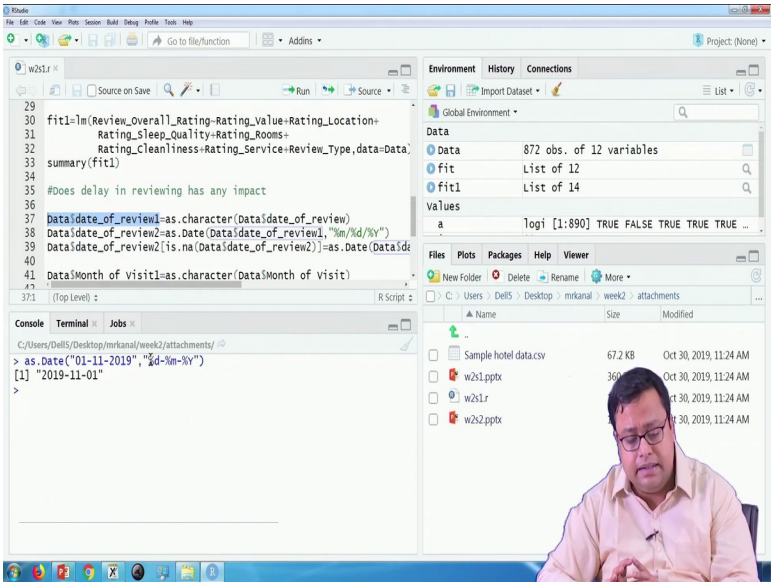
So, there is a function called as.date, you can go to help and you can go to Google and find out how as.date works. So, as.date, if I write let us say that 01-01-2019 and I want them so, 01 let us say one one 2019. So, first November 2019, I want this guy to read it as a date, I have to write the format exactly whatever the text is corresponding format I have to write. So, here what is the format?

(Refer Slide Time: 30:03)



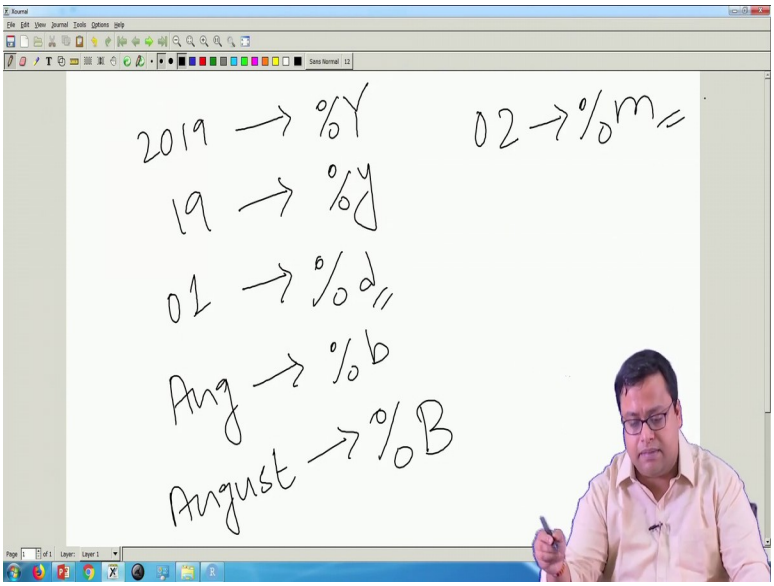
The format is actually if you check carefully, the format is, dd/ mm/yyyy this is the format, right in common sense. So, in here what we write is % d - % m - %Y. So, this is the syntax. Why it has been written like this?

(Refer Slide Time: 30:42)



So, I will go and do it like that. So, % d - % m - %Y, if I write that, now it has been read as, see the format has changed and this is the format. Now, this guy is actually reading this part as a date.

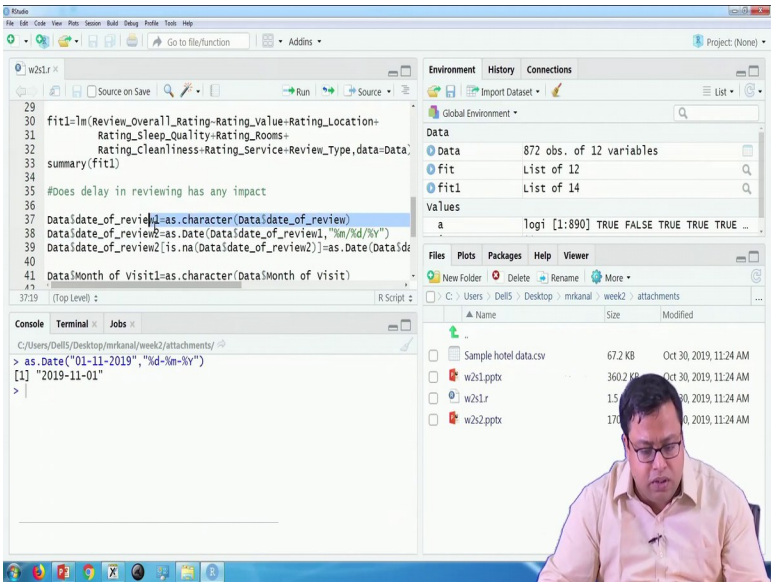
(Refer Slide Time: 31:01)

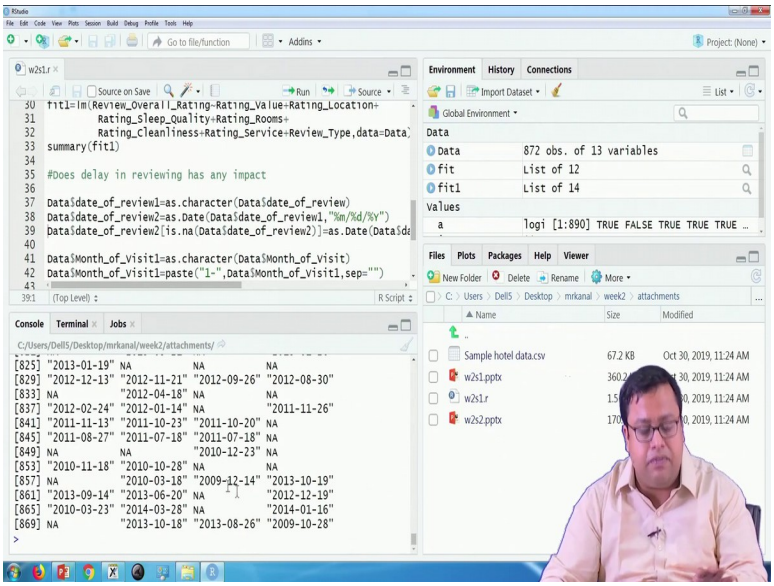


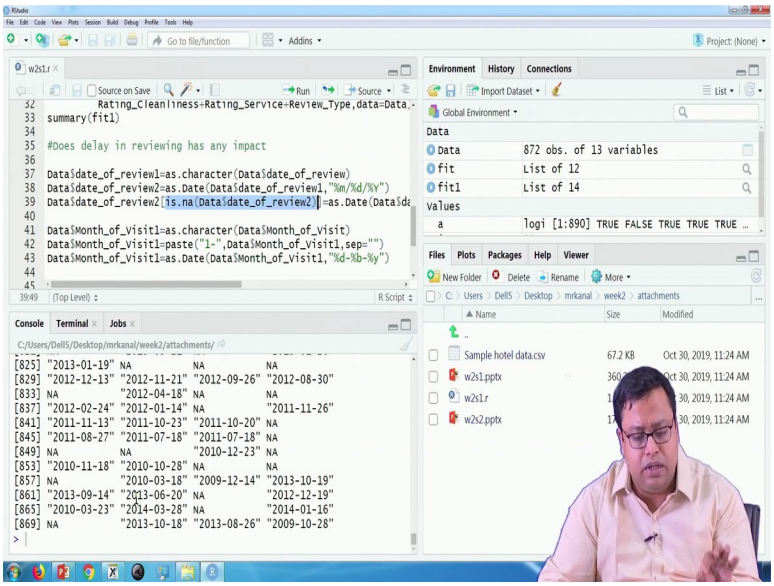
Similarly, so why %Y, why is not something else. To learn that we have to go to the help file of as date. So, some basic thing is like that as date will ask you, if you have written 2019, correspondingly you have to write % Y, if it is only 19 you have to write % y. Similarly, if it is date, date is generally double digit, so you have to write percentage d for date, for a date, let us say 0 1 or 3 1 or so on.

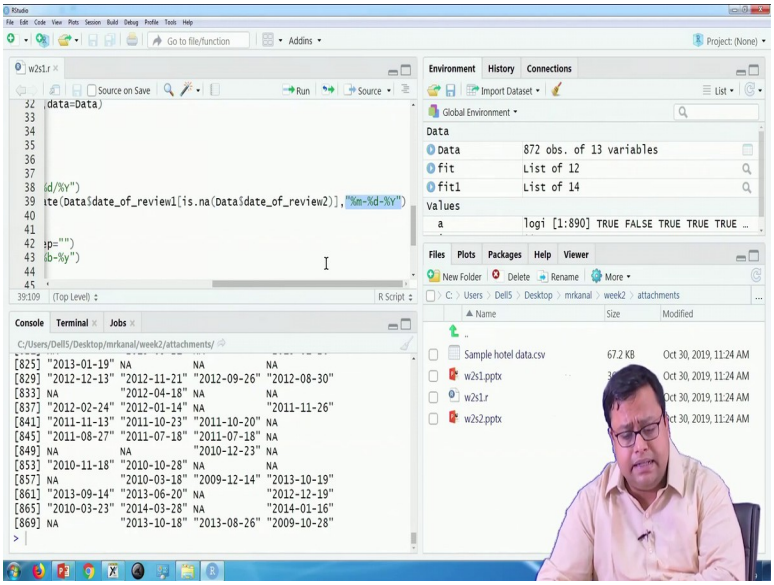
And then for month if it is AUG , you have to write % b and if it is AUGUST, full thing you have to write %B, it is let us say month, second month, 0 2 you have to write %m, so m for month, d for date. So, that is some, so these are some basic norms, you can go to the time date formatting and you can find out the other ones as well. And the other things like slash or dot and etcetera, has to be absolutely what is written in the text.

(Refer Slide Time: 32:08)







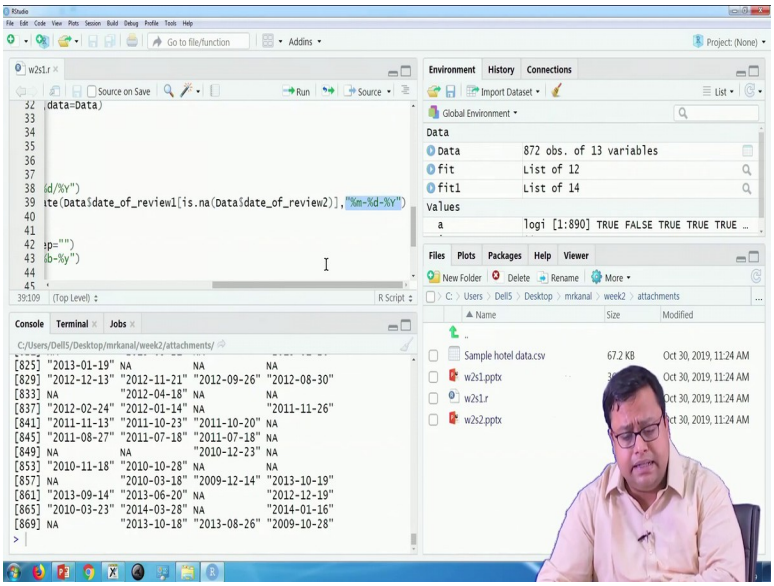


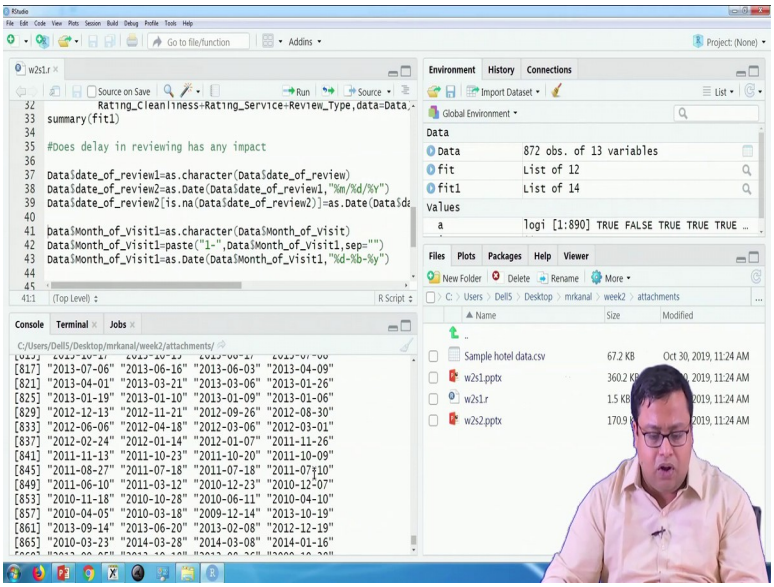
So, here if I have two forms, I first change the slash ones and then the dot ones. So, I write data review2 is equal to as.date, data review1 and what is the format I am asking percentage m means month slash date slash year, year in full form. So, if I run that and now want to see data review2, you will see there are lots, some of them are dates coming properly and then for the others lots NA is coming.

Why NA is coming? Because some of them are in the dash form, they are not in slash form. So, this guy is giving wrong result that is why. So, for that what I have to do is? I will be changing this thing in, so I am writing for those guys why it is NA, where is data,

data review2, data date of review2 is coming as NA, change it to its date form and the date form is corresponding to this.

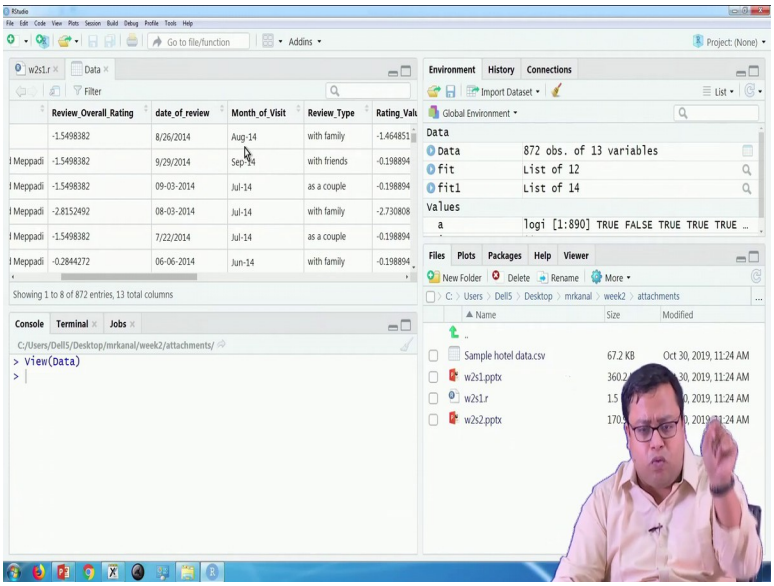
(Refer Slide Time: 33:11)

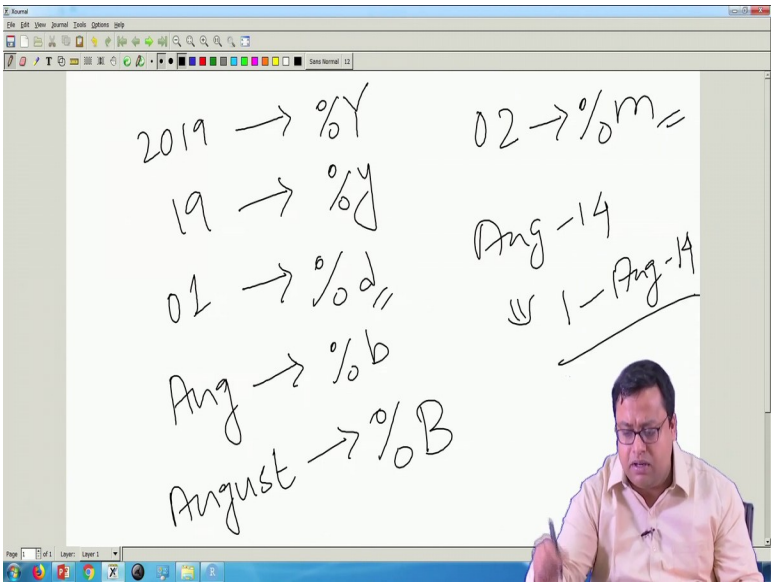


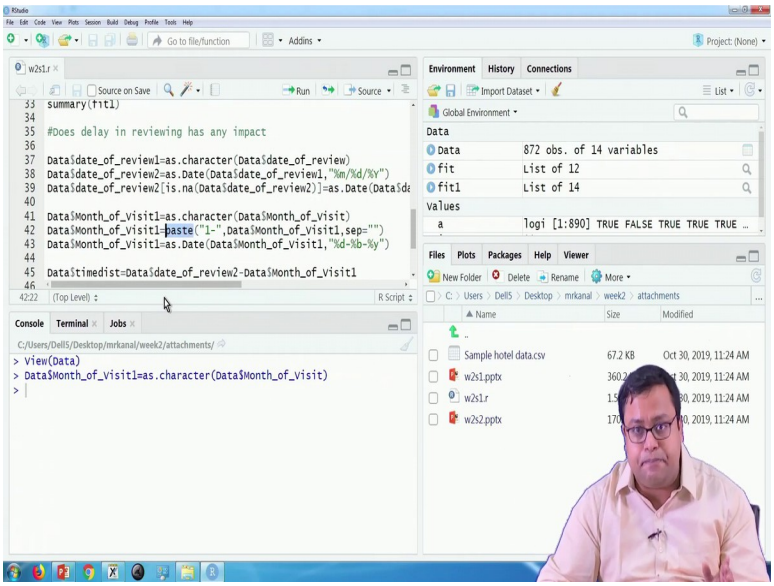


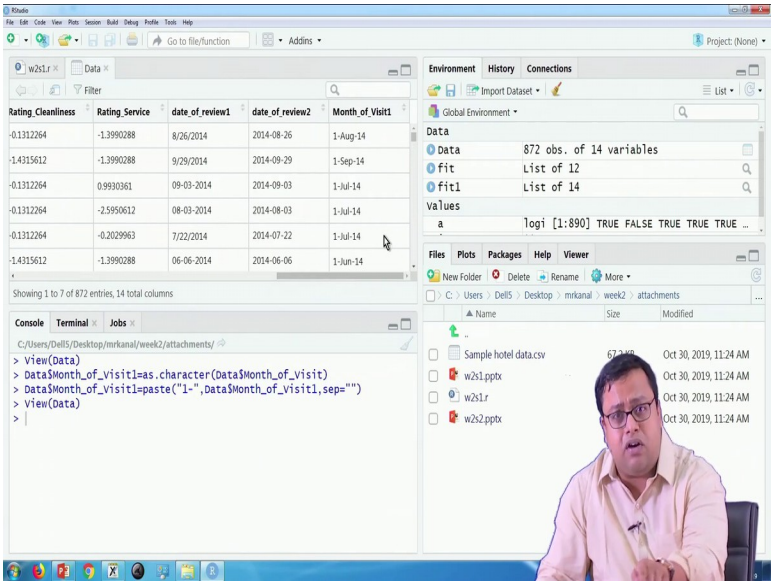
So, if you run it carefully if you see it carefully. Now, if I run this and then if I want to see data review2 all the dates are coming properly, so the NA ones are replaced with the new format. So, first I format the slash ones, then I format the dash ones and then I put it together. So, this is the date of review, the same thing I do it for the month of review.

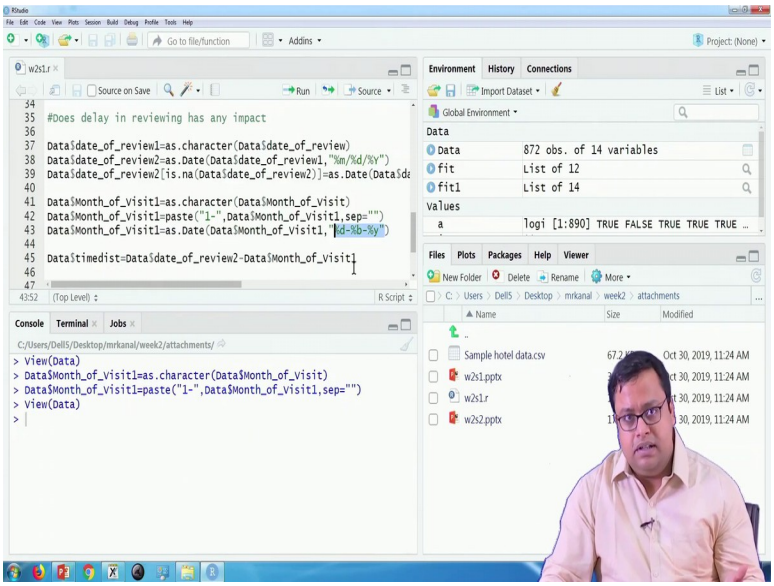
(Refer Slide Time: 33:35)






summary(fit1)
#Does delay in reviewing has any impact
Data$date_of_review1=as.character(Data$date_of_review)
Data$date_of_review2=as.Date(Data$date_of_review1,"%m/%d/%y")
Data$date_of_review2[is.na(Data$date_of_review2)]=as.Date(Data$da
Data$Month_of_Visit1=as.character(Data$Month_of_Visit)
Data$Month_of_Visit1=paste("1-",Data$Month_of_Visit1,sep="")
Data$Month_of_Visit1=as.Date(Data$Month_of_Visit1,"%d-%b-%y")
Data$timedist=Data$date_of_review2-Data$Month_of_Visit1
> View(Data)
> Data$Month_of_Visit1=as.character(Data$Month_of_Visit)
Data 872 obs. of 14 variables
fit List of 12
fit1 List of 14
a logi [1:890] TRUE FALSE TRUE TRUE TRUE ...
Sample hotel data.csv
w2s1.pptx
w2s1.r
w2s2.pptx


Rating_Cleanliness Rating_Service date_of_review1 date_of_review2 Month_of_Visit1
-0.1312264 -1.3990288 8/26/2014 2014-08-26 1-Aug-14
-1.4315612 -1.3990288 9/29/2014 2014-09-29 1-Sep-14
-0.1312264 0.9930361 09-03-2014 2014-09-03 1-Jul-14
-0.1312264 -2.5950612 08-03-2014 2014-08-03 1-Jul-14
-0.1312264 -0.2029963 7/22/2014 2014-07-22 1-Jul-14
-1.4315612 -1.3990288 06-06-2014 2014-06-06 1-Jun-14
Showing 1 to 7 of 872 entries, 14 total columns
> View(Data)
> Data$Month_of_Visit1=as.character(Data$Month_of_Visit)
> Data$Month_of_Visit1=paste("1-",Data$Month_of_Visit1,sep="")
> View(Data)
Data 872 obs. of 14 variables
fit List of 12
fit1 List of 14
a logi [1:890] TRUE FALSE TRUE TRUE TRUE ...
Sample hotel data.csv
w2s1.pptx
w2s1.r
w2s2.pptx



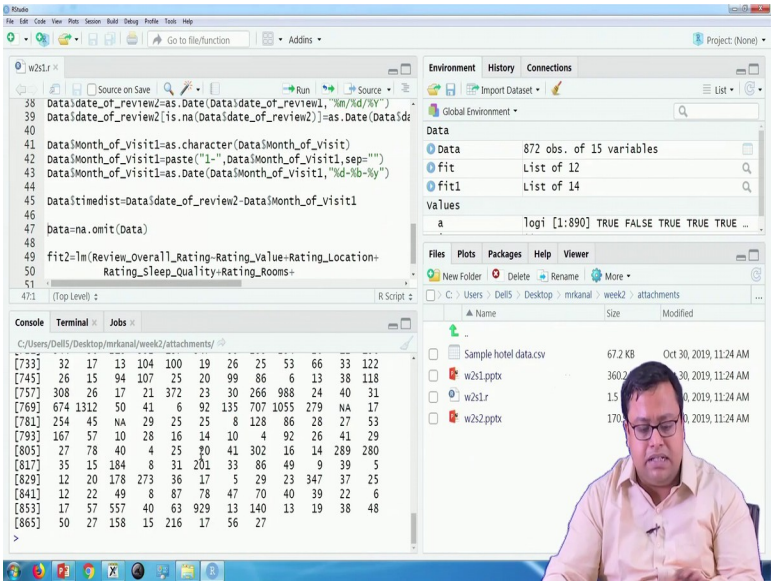
Now in case of month of review, month of travel, if you see the month of travel only has the month, it does not show which date, to be safe, we put one dash for August 2014, for August 2014 I change it to 1 August 2014, first date of that particular month I change it so that it, I make sure that everything has some value. Now, it will create certain bias, but we do not have information so we have to do our best.

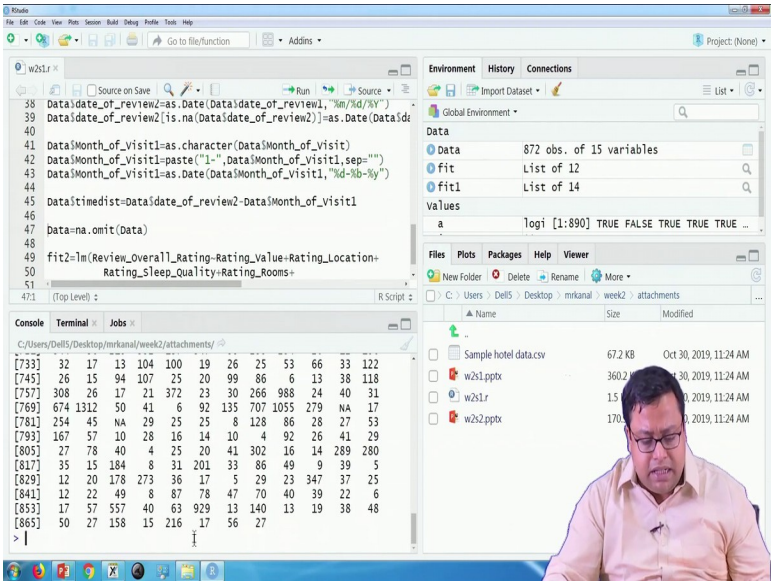
So, I have done that, so how did I do that? I have first changed it to character as usual, and then I use a paste formula. So, what is paste, it joints two texts, the first text is 1 dash, the second text is whatever was there in that particular column, which is data dollar month visit1 and separator means whether there will be certain gap. So, separator is equal to nothing, double quotes nothing written under the double quotes means that you just paste them together, you do not put anything, if I do not give that it will write 1 dash and then space and then the rest.

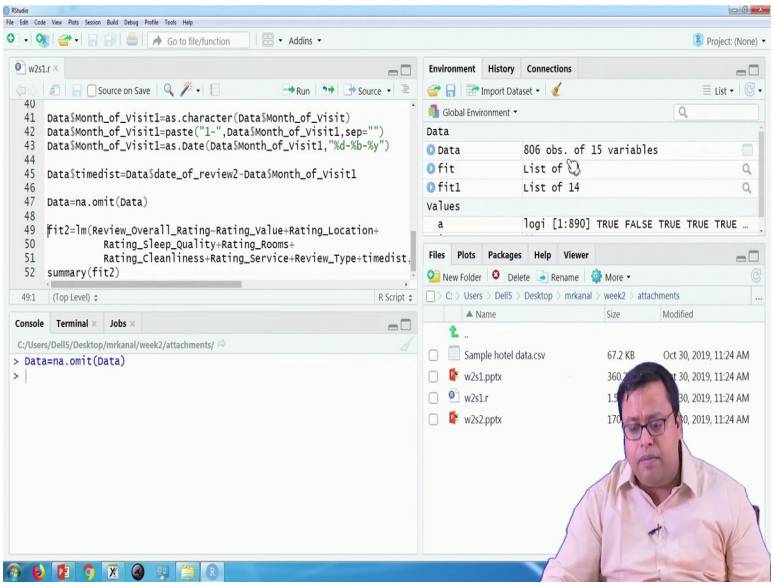
So, that is something that is not happening here. So, data dollar month visit1 if I just run this and then if I want to see what is, what I am getting. You will see at the end of the day it is 1 August, 14; 1, September 14; 1 July, 14; so 1 has been added. Now, it is still a character; I have to change it to its date form by writing this formula. So, %d dash %b, because AUG is written and this %y because dash 14 is written, dash 14 is written not 2014.

So, I run that so, if I run that what happens is I get two date columns and now the time distance between these two is the date of review, which happens closer today minus date of visit, so visit happens first and then the review. So review minus visit, so that gives me the time distance.

(Refer Slide Time: 35:40)

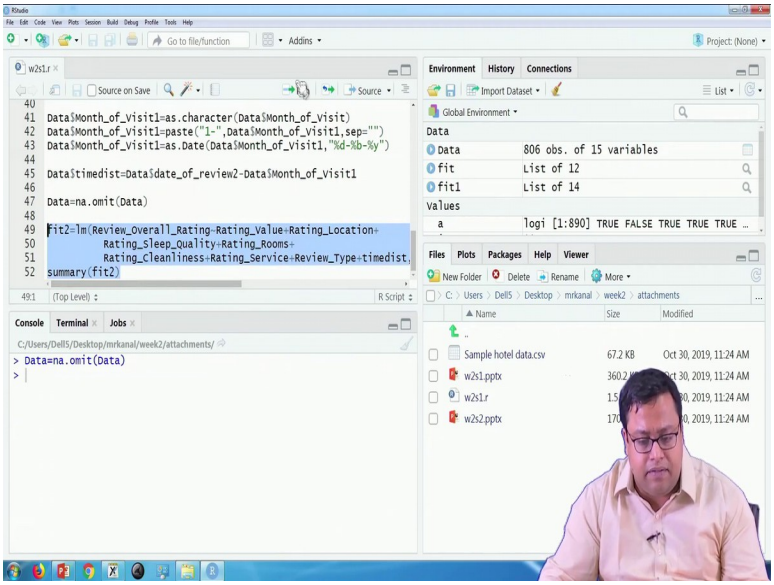


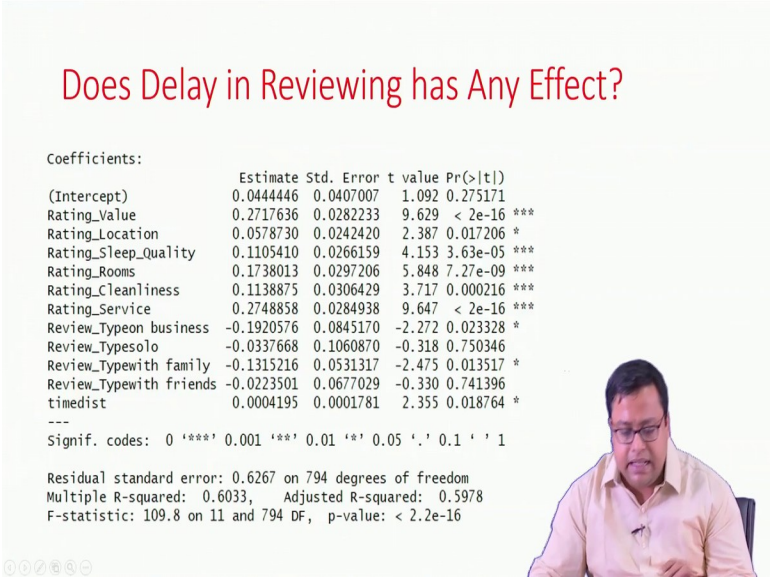




So, if I just plot the time distance for some of the values it is coming NA because the date of travel is not given, but for many of them the time distance is available. So, for the, I will just blindly remove the NA ones, the not available ones, so from 872 we got 806, which is around 66 observation got dropped, still we have enough data to handle and only thing that I do is I introduce a timedist variable here in the model.

(Refer Slide Time: 36:13)





And when I introduce this timedist variable and run the regression, I get the result which looks like this, you see the result, the result looks like this. The additional information by doing all of these things, the additional information that I am getting is this timedist. So, I am running a regression keeping all the other variables that were already there in the model. I am not disturbing them. I am introducing one more variable, which is the date of your travel minus the date of your review, or the other way review minus travel.

At that time distance I am bringing in, and that is giving us a small but significant result. What does that mean? A time distance increases, your satisfaction increases that means the later you give the review, the more satisfying results are coming up. What is a suggestion? If that is something that happens? Should we actually delay the review seeking, if you seek the review should you seek the review very close to the travel date, or should you seek the review very much away from the travel date, that is something that is an important question.

And this is something you can actually go and search in Google Scholar by my name, we have a paper on this, me and Professor Aruna Divya Tatavarthy from IIM Ahmedabad and Professor Piyush Sharma from Curtin University, we actually have a paper on similar kind of data and similar kind of analysis we have done to find out that we are trying to question what customers want, first of all, whether they are focusing on outcome oriented factors or probably service or process oriented factors when they go to a service.

And whether this preference of process versus outcome changes depending on whom you are traveling with or depending on what is the time distance after which you are giving your review or depending on what kind of expertise level you have, whether you are a probably naive traveler, all of these things will have an impact on what consumers want and what consumers do not want.

So, you can go and actually read that paper, I can actually share probably the older version of that particular paper, I can make it available, you can read the paper and you can know more about what we are trying to discuss. So, that is all for session one. We will come back in session two with some new information. Thank you.