

Foundations of R Software
Prof. Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture - 52
Graphics: Bivariate and Three-dimensional Plots

Hello friends. Welcome to the course Foundations of R Software and you can recall that in the last couple of lectures, we had talked about different types of graphics and then and I took a some examples very simple example. And I explained you how you can make them beautiful and impressive.

And all those graphics were in the univariate direction; that means, you had only one variable for which you wanted to create the graphics. Similarly, one when you try to increase the number of variables then you, then the way the graphic graphics have to be prepared that also changes. For example, if you have two variables, then means you would like to see what is the joint behavior of the two variable.

For example, if I say height and weight for the small children as the height increases weight also increases, as the weight increase the height also increases. So, in this case if you have the peer data that is on height of child 1, weight of child 1, height of child 2, weight of child 2 then you would like to have a bivariate plot so that you can see what is the joint variation of the two variables and how these variables are going to going to affect each other.

And similarly, if you have more than two variable, then also you have the similar type of thing, you would like to create a graphic in the three-dimension, but certainly in the three-dimension you cannot plot anything we can see in the three-dimension.

But we cannot plot a graphic in the three-dimension, but the way we try to plot we try to plot them in the two dimension in such a way. So, that they are trying to give us the effect of three-dimension. So, these are the things which are very easily available in the very nice way in the R software.

Surely covering everything in graphics also that is very difficult for me, possibly we can have a full course on the graphics in the R software. So, once again I will try to take here

some nice representative graphics in this lecture today and then I will stop with the graphic also. And now we are coming very close to the end of the course also.

So, after that I would expect that you can look into the help, you can look into the course material, books etc. and then try to see what are the different possibilities and then you can practice yourself. So, with this request let us begin our this lecture and try to understand about two-dimensional and three-dimensional graphics in the R software, right.

(Refer Slide Time: 02:47)

Association of two variables:
Example

- Number of hours of study affect the marks obtained in an examination.
- Electricity/power consumption increases when the weather temperature increases.
- Weight of infants and small children increases as their height increases under normal circumstances.

So, you see whenever we have two variables from the statistics point of view, we are always interested in finding out their association. For example, the number of hours of study affect the marks obtained in the examination. You know that as a teacher we always say, if you study more you will get more mark. Similarly, when the temperature of the other increases, then the power consumption will also increases.

For example, in summers people try to use air conditioner, cooler etc. whereas, during the nice weather when it is not too cold, not too hot people do not use these equipments and the consumption of this power or this electricity decreases. Like similarly the weight of small children increases as their height increases under the normal circumstances and vice versa also.

(Refer Slide Time: 03:30)

Association of two variables:
The observations on both the variables are related to each other.
How to know the variables are related?
How to know the degree of relationship between the two variables?
Graphical procedures – Two dimensional plots, three dimensional plots etc.
Quantitative procedures – Correlation coefficients, contingency tables, Chi-square statistic, linear regression, nonlinear regression etc.

So, now you are always interested in learning about this that how you can study the association of two variables. And in the statistics the this is a very big job for us, right to that how to get it done and for that we have graphical procedures, we have quantitative procedures both actually and in this case the observation on both the variables are related to each other.

So, the first question comes here how to know whether the variables are related or not? And sometimes you also want to know the degree of relationship between the two variables whether the degree is a strong or say not so strong or weak. In order to understand or get the answers of this thing, we have two types of procedures: one is the graphical and another is the quantitative procedures in a statistic.

For quantitative procedure, we have the concepts of correlation coefficient, contingency tables, chi square statistic, linear regression, non-linear regression etc., but certainly we are not going to do it here. But I am just telling you that how these two graphical and quantitative procedure analytical procedures go together.

Similarly, in graphics also we have two-dimensional plot, three-dimensional plots else through which we always try to get an idea that how the variables are related to each other, right.

(Refer Slide Time: 04:34)

Association of two variables:

How to judge or graphically summarize the association of two variables?

X, Y : Two variables

n pairs of observations are available as $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$

Handwritten notes: 5 children, H, W

So, for example, our main question is now how to judge or graphically summarize the association of two variables. So, suppose we have here two variables, which I am trying to denote here as say X and Y and we have one pairs small number of observations of the paired observations on say X and Y which are denoted as $x_1, y_1, x_2, y_2, x_n, y_n$ for example, if I have data of say 5 children on their height on their weight. So, x_1 is the height of the child 1 and w and y_1 is the weight of the child 1.

Similarly, for the child 2, x_2 is the height y_2 is the weight and so on these are called a paired observation, right. So, we have such an observation and we want to plot them.

(Refer Slide Time: 05:14)

Scatter plots:

Plot the paired observations in a single graph, called as scatter plot.

Scatter plot reveals the nature and trend of possible relationship.

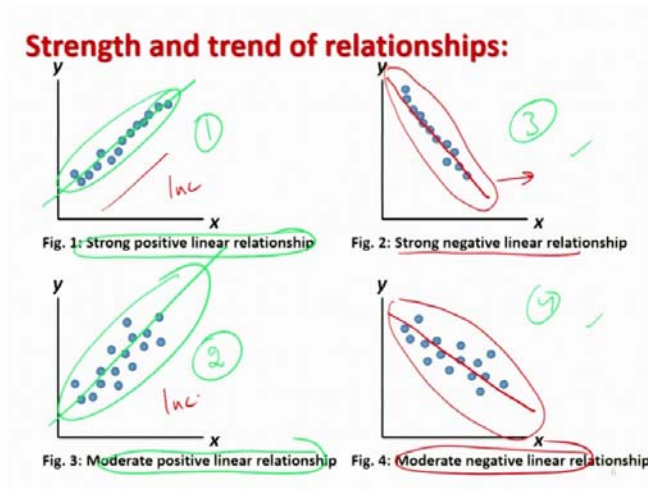
Relationships : Linear or nonlinear.

Linear relationship between X and Y Nonlinear relationship between X and Y

So, when we try to plot the paired observations, they are called as a scatter plot and scatter plot you also have learned in the earlier lectures that you can create using the command plot. But that was for the single variable, right. And if you try to make here such a scatter plot they can reveal the nature entrant of the possible relationship and this relationship can be linear or non-linear.

For example, if you try to create a scatter plot and if it is look like this or like this, then you can very clearly conclude from this one plot number 1, that the relationship between X and Y is quite approximately linear, but in the case of plot number 2 certainly you can say that the relationship between X and Y is non-linear, right.

(Refer Slide Time: 05:56)



Similarly, the concept of strength and trend of relationship that can also be studied from the scatter diagrams. For example, if you try to see at let me call this picture says 1, 2, 3 and here 4, right or yeah. So, if you try to look at here this picture and this picture. So, what you see here? Here, the points are very close to the line and in the picture number 2 the points are quite away from the line which I have drawn.

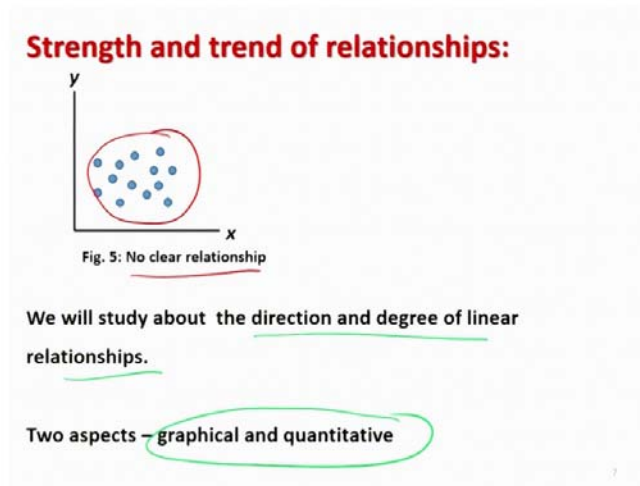
So, in this case you can say that the relationship between X and Y is quite strong. Whereas, in the case number 2 we can say that the relationship between X and Y is say not so strong. And similar is the case if you try to see here in this case also picture 3 and 4 also, you can see here we have a similar type of relationship, but there is a difference.

The difference here is this that here the line was like this, but now the line is here like this, right.

The line is passing through like this and here like this. So, the relationship between X and Y is decreasing whereas, in the case of 1 and 2 this was increasing, but the degree of relationship if you try to see here the points are very close to the line, which is drawn here and here the points are not so close compared to the points in picture number 3.

So, we can say that here in the case number 3 you have strong negative relationship and in the case number 4, you have here moderate to moderate negative linear relationship. So, this is how we can judge by looking into the scattered diagrams about the strength and trends of the relationship.

(Refer Slide Time: 07:22)



And then if there is a no pattern like this one between X and Y, then you can say that ok there is no clear relationship. So, now, I have explained you that how you can take a conclusion about the relationship from the scattered diagram that you are trying to plot here. So, people try to actually look for the direction as well the degree of linear relationship and then we can combine both the graphical and quantitative aspects together to take a final call.

(Refer Slide Time: 07:49)

Bivariate plots:
Provide first hand visual information about the nature and degree of relationship between two variables.
Relationship can be linear or nonlinear.
We discuss several types of plots through examples.

So, that is the job which we try to do as a statistician, but here now our main objective is that how we can create such plots, how we can create such bivariate plots which can help us in taking this type of call. So, this bivariate plots provide us the first hand visual information about the nature and degree of relationship between the two variables, right.

And this relationship can be linear or say non-linear and now we will try to discuss this type of relationship through example and we will try to create the bivariate plots.

(Refer Slide Time: 08:14)

Bivariate scatter plots:
Plot command:
`x, y`: Two data vectors
`plot(x, y)`
`plot(x, y, type)`

type	
"p" for points ✓	"l" for lines ✓
"b" for both ✓	"c" for the lines part alone of "b"
"o" for both 'overplotted' ✓	"s" for stair steps.
"h" for 'histogram' like (or 'high-density') vertical lines	

So, earlier you have command plot. So, now, this plot command can also be used for making the bivariate plots. So, inside the plot you had the length couple of commands. So, that those commands can be used here very easily without any problem.

And the way it is going to be used here it is you write plot, plot and within the parenthesis you try to write down the data vector on x and y separated by comma and after that you have many options.

So, for example, if you there is an option here type. So, if you try to give here type is equal to p, that is the default actually you will get here points. If you give here the type is equal to l, then you will get the lines. If you take a type equal to b that the natives were point and line both. Then similarly if you try to take care c; then if you try to see in the both option whatever is there only the remaining part is there.

If you try to take it here o then it is over plotted as is for the stair steps, h is for histogram types or say high density vertical lines ok. What are these thing? How the graphs will look like? Under these types of option I will try to show you on the same data, right.

(Refer Slide Time: 09:22)

Bivariate scatter plots:

Plot command

x, y: Two data vectors

```
plot(x, y)
plot(x, y, type)
```

Get more details from help: `help("type")`

Other options:

<code>main</code>	an overall title for the plot.
<code>suba</code>	sub title for the plot.
<code>xlab</code>	title for the x axis.
<code>ylab</code>	title for the y axis.
<code>aspthe</code>	y/x aspect ratio.

And then beside those things, you have other options also like I said main for giving an overall title, sub a for giving a subtitle xlab title for the x axis ylab title for the y axis aspect for to control the aspect ratio and so on.

(Refer Slide Time: 09:36)

Bivariate scatter plots:

Example

Data on marks obtained by 20 students out of 500 marks and the number of hours they studied per week are recorded as follows:

We know from experience that marks obtained by students increase as the number of hours increase.

Marks →	337	316	327	340	374	330	352	353	370	380
Number of hours per week →	23	25	26	27	30	26	29	32	33	34

Marks	384	398	413	428	430	438	439	479	460	450
Number of hours per week	35	38	39	42	43	44	45	46	44	41

But anyway, let me try to take here a very simple example and I will try to show you these graphics on this data so that you can understand them very easily.

So, suppose there are 20 students and they have appeared in an examination and their marks are obtained out of 500 and then it was asked from those 20 students that how many hours they have studied in a week. And then we try to see that whether the number of hours of study in a week and the marks obtained in the examination are they interrelated or not. Well, that is a very common statement by our teachers, by our family members that if you study more you will get more marks.

And we want to see here on the basis of this data, whether this holds true or not. So, you can see here we have a data here like has on the marks and the number of hours of study in a week. So, it is like that, if a student has a study 23 hours he got 337 marks, 2nd student is studied 25 hours in a week and that is true and got 316 marks and so on, right.

(Refer Slide Time: 10:34)

Bivariate scatter plots:
Example

```
marks =  
c(337, 316, 327, 340, 374, 330, 352, 353, 370, 380, 384, 39  
8, 413, 428, 430, 438, 439, 479, 460, 450)
```

```
hours =  
c(23, 25, 26, 27, 30, 26, 29, 32, 33, 34, 35, 38, 39, 42, 43, 4  
4, 45, 46, 44, 41)
```

12

So, all this data is here and I have compiled this data into two different data vectors, one for the marks and one for the hour. So, the relationship is very simple, the value at the first position in the mass corresponds to the number of hours studied by the student number 1 and so on.

(Refer Slide Time: 10:49)

Bivariate scatter plots:
Example

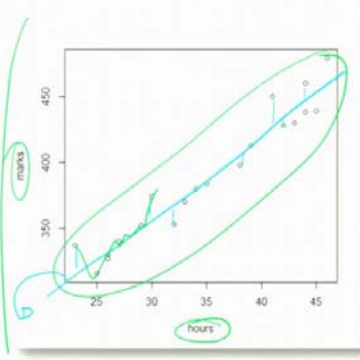
plot command:

x, y: Two data vectors

Various type of plots are possible to draw.

```
plot(x, y)
```

```
plot(hours, marks)
```



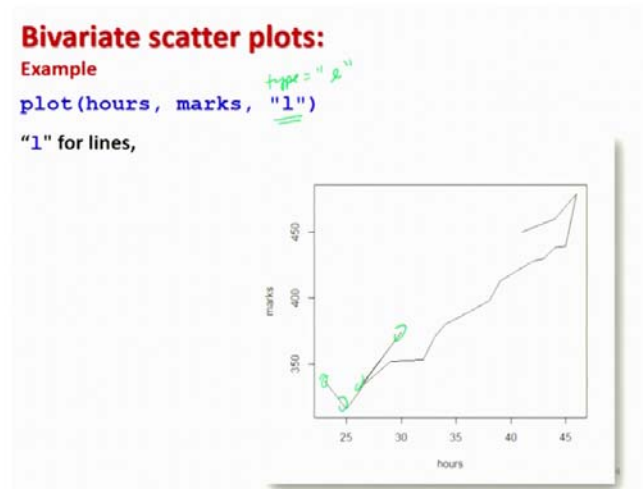
13

So, now I have here two data vectors and I try to use here a command plot. So, I simply use here plot x, y, right. So, I try to use here plot, hours, and marks. So, you can see here

you will get here this type of graph, this is here hour this is here marks and this is here this dot, right. And you know that if you want to change the color of this dots etc. etc. you can do very easily, but now I try to show you that what you can conclude. If you hypothetically think that ok.

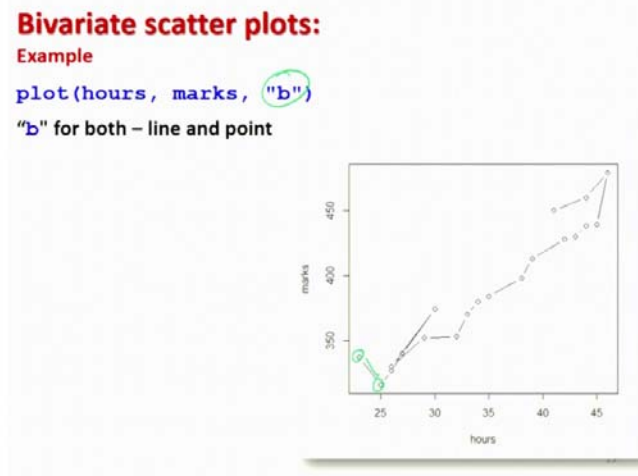
You can create here a line somewhere here passing through here like this, then you can actually judge that how close are this point to this line and that will possibly give you the degree of linear relationship or degree of association between the hours and marks, on the basis of the given cut of data, right. So, later on I will try to show you that how you can create such hypothetical equation here and how you can plot such a line.

(Refer Slide Time: 11:40)



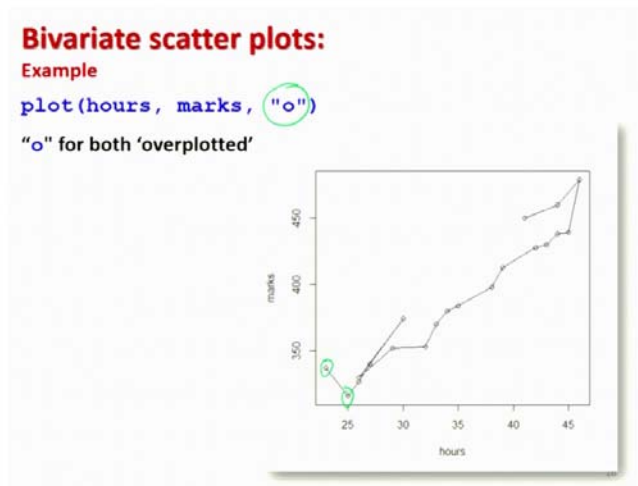
But anyway, let me try to change my here these options, right. If you try to see here I try to take here the option here say l which is actually written here like this type is equal to l like this, so l just for line. So, you can see here all these points which are here is in this graph you can see here they are the points, but now they are joined here like this and so on, right. So, this is happening in the case of a line.

(Refer Slide Time: 12:04)



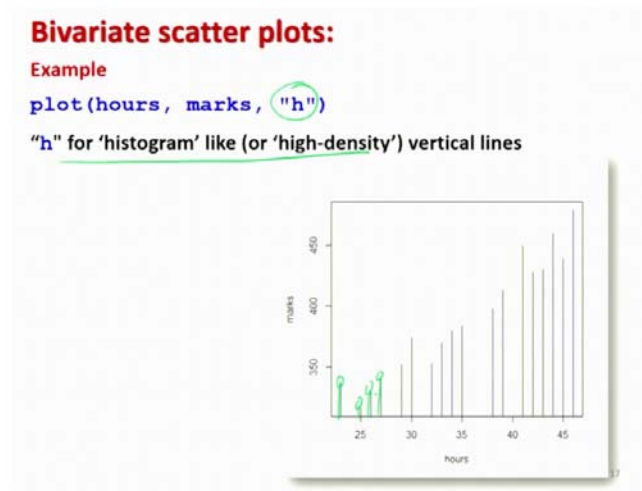
Similarly, if you try to take here the option here b, type is equal to b, then you can see here this points and lines both are occurring together this is the meaning.

(Refer Slide Time: 12:13)



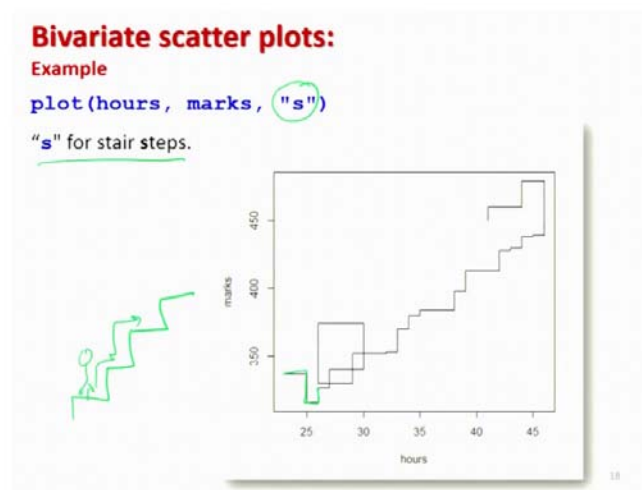
And similarly, if you try to change the option to here say type is equal to o, then you can see here that this lines and this points the line is over plotting over the points. So, this is the meaning of here o.

(Refer Slide Time: 12:25)



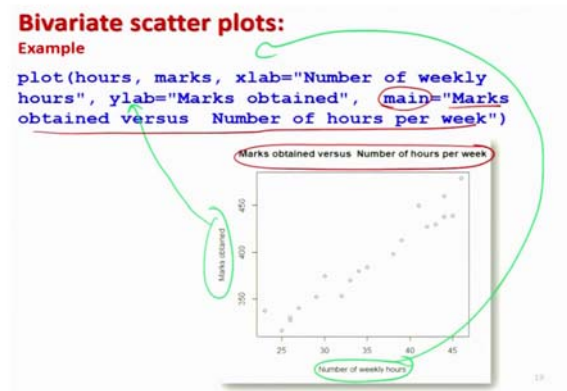
And similarly, if you try to take here the type is equal to here h, which means like histogram or say high density vertical lines. So, every point is giving you here a line which is falling vertically on the x axis, right.

(Refer Slide Time: 12:38)



Similarly, if you try to take here type equal to here s. So, all the points here they are joined like a stair case, right. So, s means a stair steps, you have seen the steps in your stairs and there are steps in your home on which you actually walk, right like this.

(Refer Slide Time: 12:51)

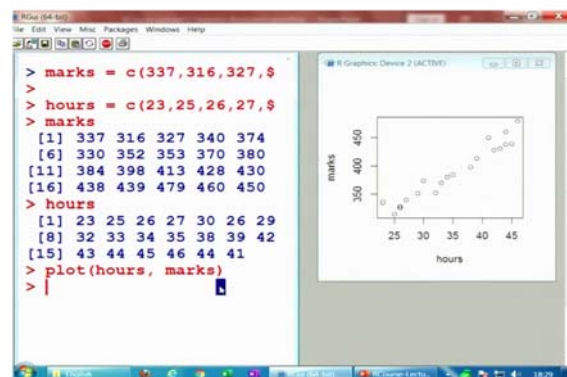


So, these are the different types actually I will try to show it on the R console and if you want to add here some options. For example, now you know that if you want to add here number of weekly hours here, you have to give here the option here xlab and if you want to obtain here the option on the y axis you can use here the option here y lab.

And if you want to add here the title of this plot, which is the marks obtained versus number of hours of per week you can use option here main and then you can give this title inside the double quotes and so on. That you can also change the color etc. So, these are very simple things which you can do very easily after learning so much of graphics, right.

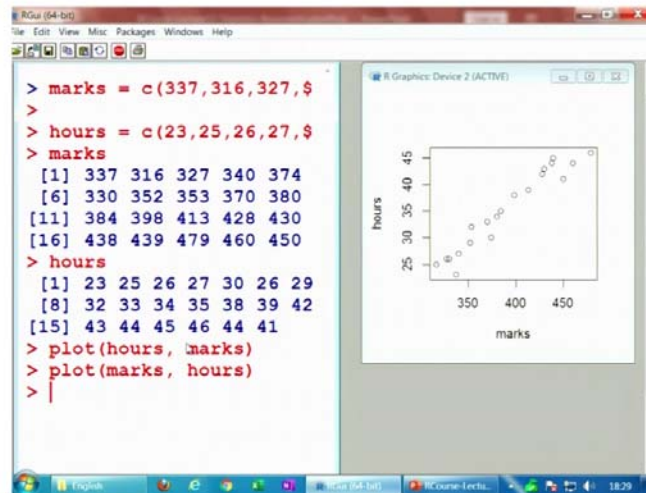
So, let me try to first show you these operations on the R console itself. So, let me try to first get here the data on here marks and hours, right.

(Refer Slide Time: 13:41)



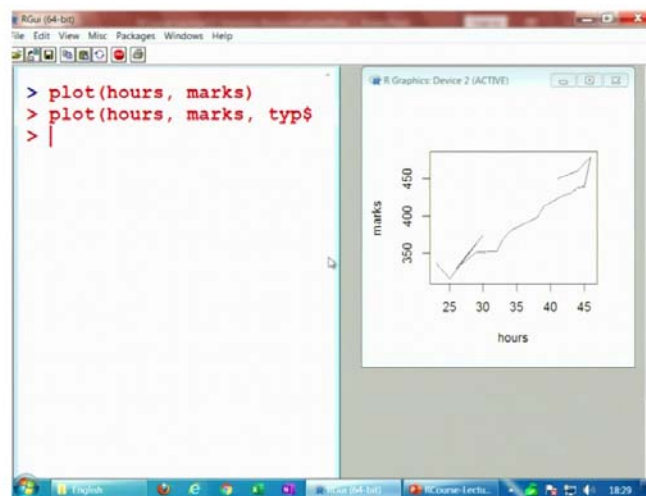
So, if you try to see here, I have this is the data here marks and this is the data here hours, right. And now I see here plot the hours and here marks. So, you can see here you get here this type of air plot, right.

(Refer Slide Time: 13:58)



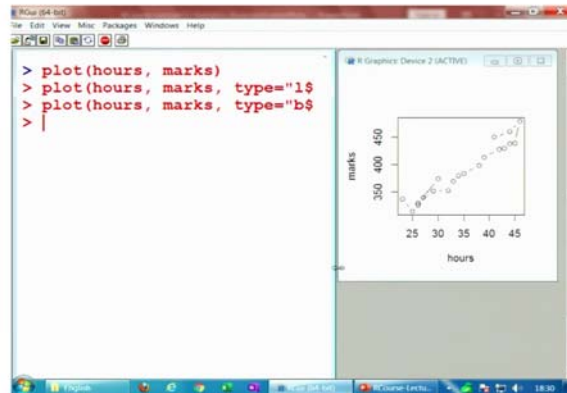
And if you try to see here hours on the x axis and marks on the y axis. But if you try to plot it here see here marks comma hours, then you will see that their locations are going to be changed yeah like this.

(Refer Slide Time: 14:07)



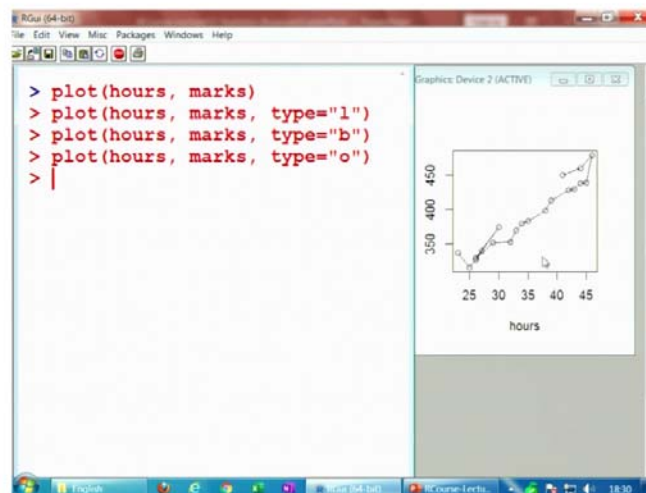
But anyway, let me try to clear the screen and let me try to create the graph here plot. And then now, I will try to add here the option for the here type. So, type if I try to say here this is here line. So, this you can see here just see here how the things are changing.

(Refer Slide Time: 14:23)



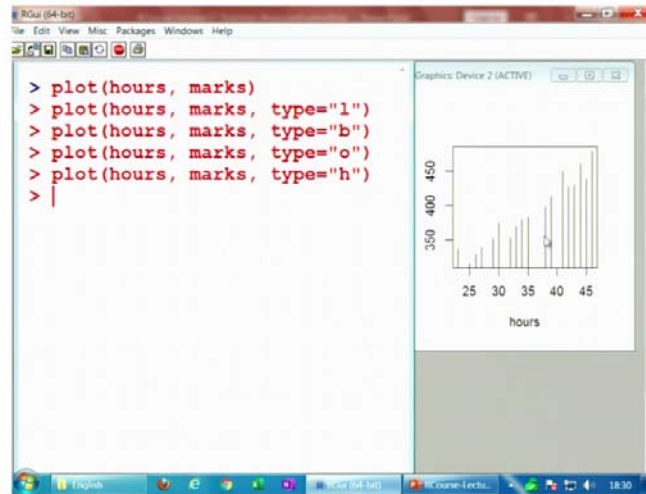
So, if I try to make it here this option of the line here, I can reduce here the size of the graphic so that I can show you here everything, right. So, you can see you can adjust all these things on the graphics without any problem. Now, instead of line if you try to take here both, we can see here now both are there, right.

(Refer Slide Time: 14:42)

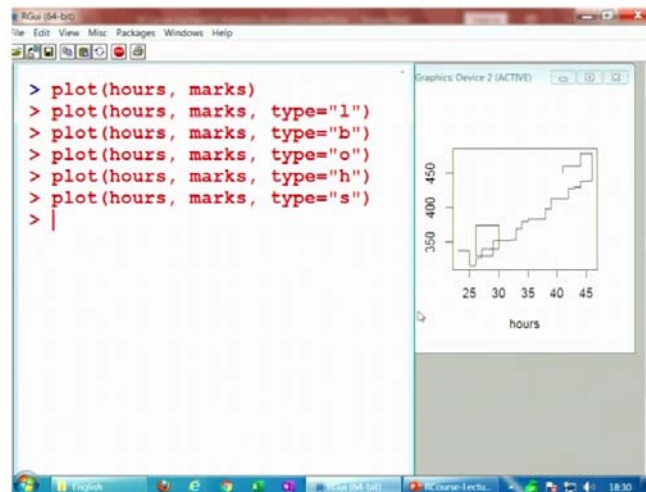


And similarly, if you try to take here instead of here both, if you try to take here over plotted type is equal to o. You can see here now this graph is change here like this.

(Refer Slide Time: 14:51)

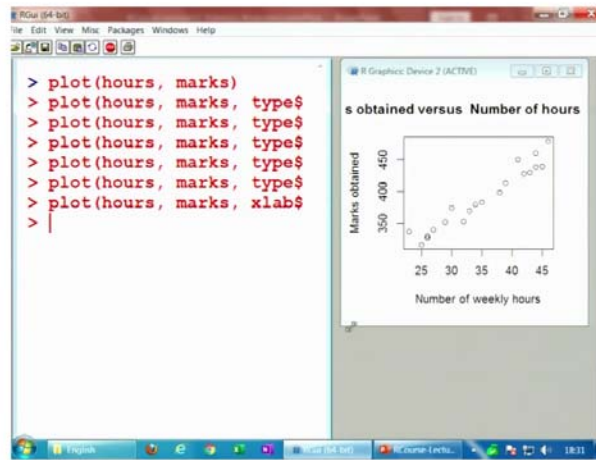


(Refer Slide Time: 14:56)



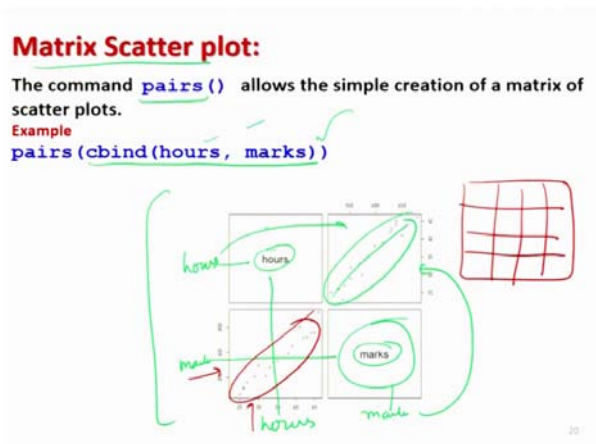
Now, if you try to take the option here h, which means high density lines you will get here this type of graph. And if you try to take here the option here s, that is steer steps you get here this type of graphic, right. So, you can see here that is not a very difficult thing for you to understand and after this if you try to just try to add here some titles etc.

(Refer Slide Time: 15:14)



So, you can just see here if you simply try to execute it, you are getting here this type of offshore option, right yeah. I am trying to manage both the screen on the same computer, but you can see it very clearly when you try to make it large.

(Refer Slide Time: 15:28)



And anyway, I come back to my slice and try to give you here one more command here. So, sometime you see you have these types of variable and you would like to see their joint variation, but you would like to create all possible combination of the plots in a single plot.

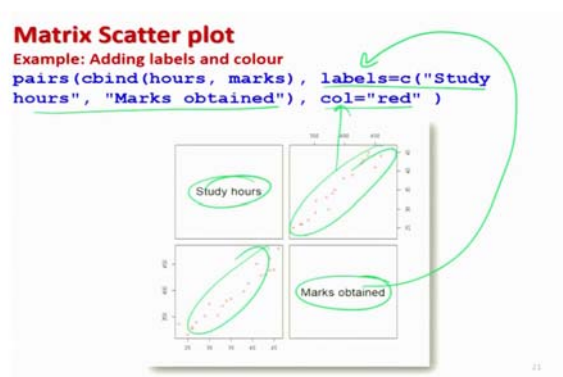
So, that is called here as a matrix scatter plot and this command can be obtained by using the command `pairs`, `pairs`, right. So, the rule is very simple, inside the `pairs` you have to write down here `c` bind and you. For example, here you try to write hours comma marks and if you have more than two variables that you can join here separated by commas.

Right, say `x` comma `y` comma `z` etc. So, now since you have taken here two variables. So, it will give you here a plot like this one which is a 2 by 2 plot and it looks like a matrix. So, that is why it is called as a matrix scatter plot. So, this will be like on is here it is hovers and it is here marks and on the `x` axis also this is here hours and this is here mark. You can see from here, right.

Now, this is a plot between hours and hours. So, which has no value and this is the plot here between marks and marks which has no value, marks from here, marks from here, hours from here, hours from here. Now, if you try to see this is a plot between say here hours and here mark. So, you can see here this type of trend and this is another graphic here which is between say hours and here marks. So, the only difference is that the roles of the `x` and `y` axis they are changing.

But if you are trying to take here say 4 variables possibly it will look like this. All the graphics will be there. Now, that will be a challenge and that will be your job to conclude what you really want to conclude at the end, right. But anyway my job was to explain you that you can do these things on the R software also.

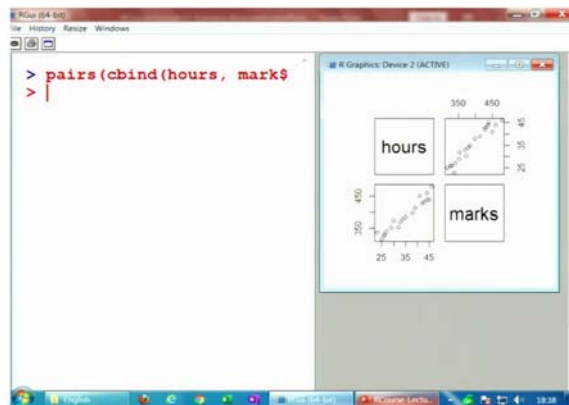
(Refer Slide Time: 17:07)



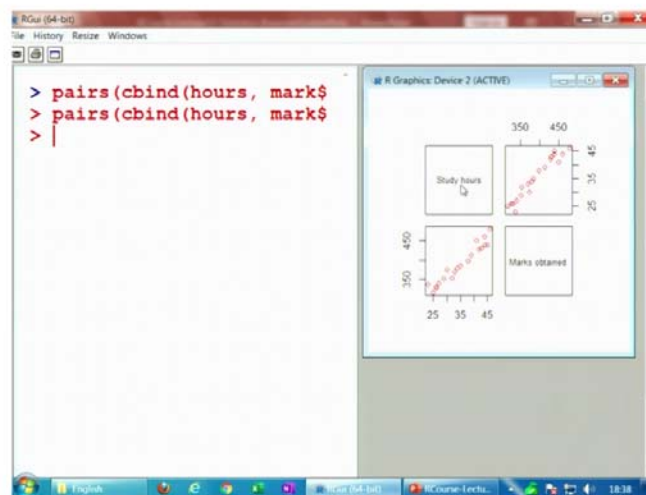
For example, if you try to see here, when I try to obtain matrix scatter plot and if I want to make here the labels of my choice.

I can use the command here labels, labels say study hours and marks obtain and I can make it with the color to be here red. So, you can see here these labels are controlled by the this option here labels and then this color of this dots, that is changed by this ul option. I do not think it will be a very difficult thing for you so, but let me try to show you that how you can obtain it on the R console.

(Refer Slide Time: 17:37)

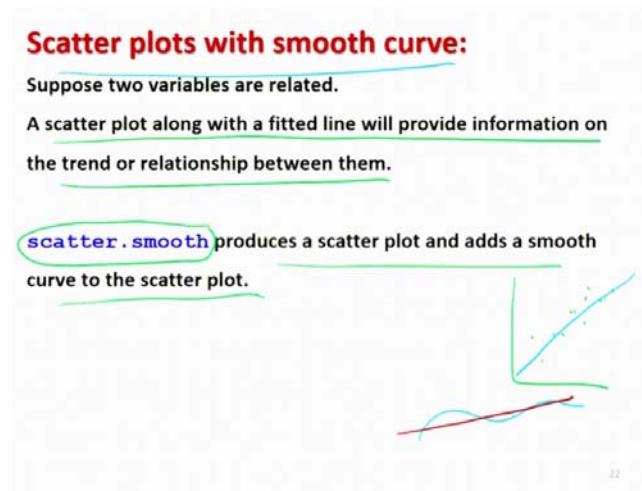


(Refer Slide Time: 17:49)



So, if you try to see here this is your here matrix plot, right. And similarly, if you try to take here the other option, where you are trying to change the color and you are trying to add the labels you can see here this is here like this. You can see here now labels are added here study hours and marks obtain and this color is now changed to red. So, you can see here that is not a very difficult thing, right, and I am sure that you must be confident at how well you can do it very easily.

(Refer Slide Time: 18:02)



Now, in this scatter plots, if you try to see here whenever you are trying to make a scatter diagram like this one I always ask you that you try to create here a hypothetical line like this one and try to see whether these points are close to the line or not. Well we have statistical methods by which we can find the equation of such a line. So, what is happening?

That when we want to plot such a line all the statistical tools are used inside the program, the equation of the line is obtained and then this line is plotted over the scatter plot, right. So, now, we try to obtain such scatter plot which have a smooth curve. Smooth curve means it may be like this one, but if you are trying to fit only a straight line possibly for the same data, it may look like this, right. So, now, we are looking for a smooth curve. So, suppose there are two variables which are related.

So, what we try to do? A scatter plot is created with a fitted line and that will provide other information on the trend and the type of relationship between them. And in order to do it we have a command here `scatter dot smooth`, `s c a t t e r dot smooth` and all in lowercase alphabets and this will produce a scatter plot and this will add a smooth curve to the scatter plot, right.

(Refer Slide Time: 19:16)

Scatter plots with smooth curve:

`scatter.smooth` is based on the concept of LOESS which is a locally weighted scatterplot smoothing method.

LOESS is used for local polynomial regression fitting.

Fit a polynomial surface determined by one or more numerical predictors, using local fitting.

Use `help("scatter.smooth")` to get more details.

23

How that I will try to show you. So, actually just for the information this is based on the concept of LOESS, which is a locally weighted scatter plot smoothing method. And it is used for local polynomial regression fitting and it fits a polynomial surface determined by one or more numerical predictors using the local fitting.

Well, I am not going into these details these are the things which we try to study in statistics, right, but we are simply using here. So, if you want to have more information about that scatter dot smooth you can just look into the help, right and try to see.

(Refer Slide Time: 19:47)

Scatter plots with smooth curve:

```
scatter.smooth(x, y = NULL, span = 2/3, degree = 1, family = c("symmetric", "gaussian"), xlab = NULL, ylab = NULL, ylim = range(y, pred$y, na.rm = TRUE), ...)
```

x, y	x and y arguments provide the x and y coordinates for the plot.
span	smoothness parameter for LOESS.
degree	degree of local polynomial used.
family	if "gaussian" fitting is by least-squares, and if family = "symmetric" a re-descending M estimator is used.
xlab	label for x axis.
ylab	label for y axis.
ylim	the y limits of the plot.

And then you will see here there are many many options, the most simple option here is that you try to write down the scatter dot is smooth and then try to give here this data on say x and y. And then you can control the span, the family by which it is going to be estimated, right degree xlab, ylab etc. is ylimb.

But and there are many many commands here, which you are going to use for example, x and y they are the arguments that will provide the data span will control the smoothness parameter, degree is the degree of the local polynomial, which is used family is the we try to use in stats is Gaussian fitting or like that. xlab is the label for the x axis, ylab is the label for the y axis y limit the limits on the y axis.

(Refer Slide Time: 20:27)

Scatter plots with smooth curve:

Example

Data on marks obtained by 20 students out of 500 marks and the number of hours they studied per week are recorded as follows:

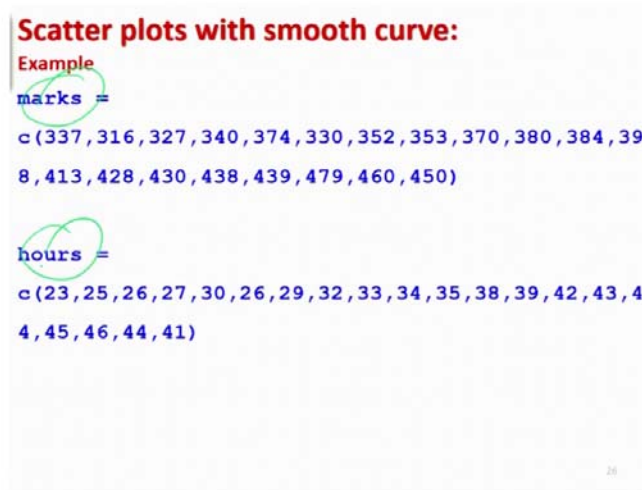
We know from experience that marks obtained by students increase as the number of hours increase.

Marks	337	316	327	340	374	330	352	353	370	380
Number of hours per week	23	25	26	27	30	26	29	32	33	34
Marks	384	398	413	428	430	438	439	479	460	450
Number of hours per week	35	38	39	42	43	44	45	46	44	41

So, I try to take here thus, I have just given you here just repeated this so that you can understand this is the same example.

(Refer Slide Time: 20:36)

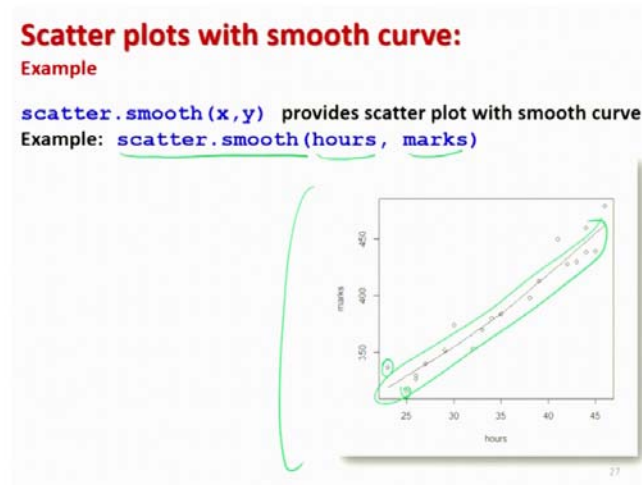
Scatter plots with smooth curve:
Example
marks =
`c(337, 316, 327, 340, 374, 330, 352, 353, 370, 380, 384, 398, 413, 428, 430, 438, 439, 479, 460, 450)`
hours =
`c(23, 25, 26, 27, 30, 26, 29, 32, 33, 34, 35, 38, 39, 42, 43, 44, 45, 46, 44, 41)`



So, from this example, where we have the marks and number of hours which I have stored here in these two data vectors, marks and hours.

(Refer Slide Time: 20:42)

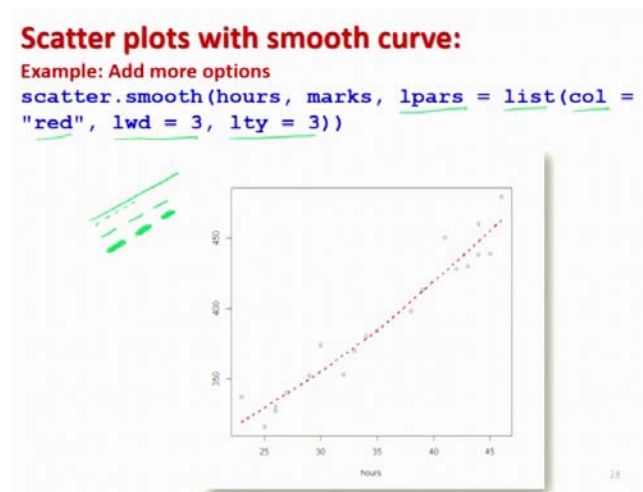
Scatter plots with smooth curve:
Example
`scatter.smooth(x, y)` provides scatter plot with smooth curve
Example: `scatter.smooth(hours, marks)`



So, I will try to use the same data set and I will try to create here a scattered smooth plot. So, you can see here I just try to use here the command here scatter dot is smooth and inside the parenthesis hours and mark.

And you can see here you get here this type of graph. You can see here earlier you had obtained only this dot, but now there is a line here also. So, now, looking at this line you can make a better judgment whether the points are close to the line or not. So, that is the advantage, right.

(Refer Slide Time: 21:08)

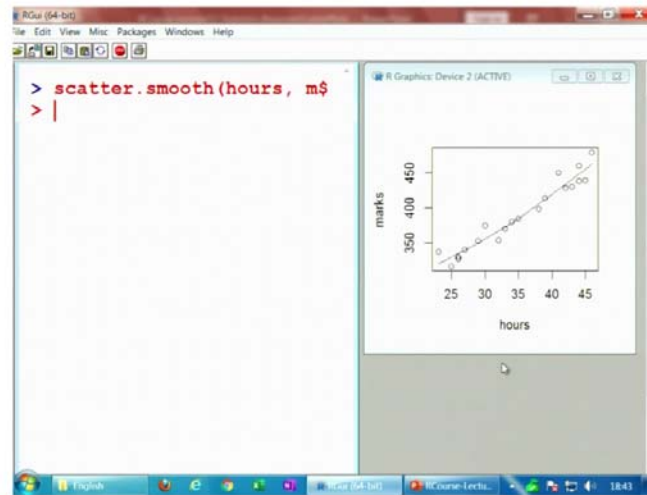


And similarly, if you want to make some modifications in the line you can just use here different types of options here.

Like lpars which is actually a list now you know, what is the list where color is equal to red then lwd is equal to 3 lty is equal to 3. So, these things are trying to control the this spacing of the line like this, this one, this one, this one and the width like this one, right because this line can also be like this like this.

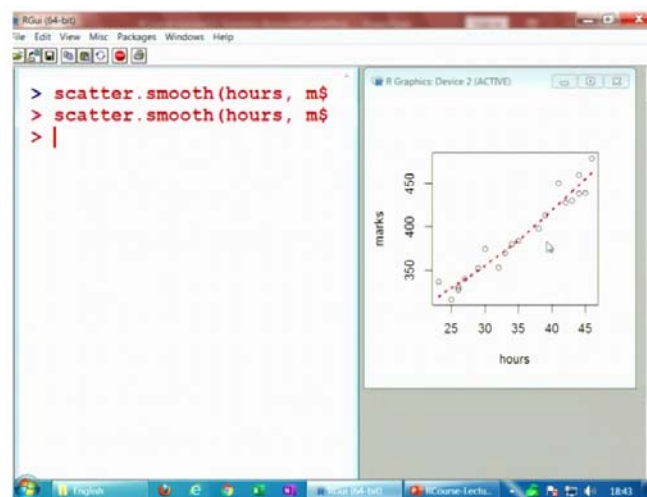
So, these things are, but in order to understand this thing I will say simply you try to take different values and try to plot them in the R software and try to see that how do they affect. And then you will have an idea that how these things are going to be changed.

(Refer Slide Time: 21:51)



So, let us try to first make this scatter smooth plot on the R software so that you get here more confident. So, I try to you can see here this is here the graph, which you wanted to have, right. So, you can see now here it is very convenient to look at what is that hypothetical line and try to see how close are the points to this line, right.

(Refer Slide Time: 22:13)



And similarly, if you want to make here these type of changes here, you can use here this command here you can see here that this line is changing. But anyway, these are the

cosmetic changes and what it depends on the on your need that how you want to move forward ok.

(Refer Slide Time: 22:21)

Three dimensional scatter plot:

```
scatterplot3d(x,y,z)
```

Plots a three dimensional (3D) point cloud of the data in **x**, **y** and **z**

Need a package **scatterplot3d**

```
install.packages("scatterplot3d")  
library(scatterplot3d)
```

29

So, now I try to take here the one more topic that is about the three-dimensional scatter plots, right. Three-dimensional scatter plots you know that they cannot be done in the three-dimension, but what we try to do here that we try to make the two-dimensional plot which will look like a three-dimensional plot.

So, this command here is scatter plot 3D which is here like this `scatterplot3d` and inside the parenthesis `x` comma `y` comma `z` this will plot a three-dimensional plot and, but it is based on a package here `scatterplot3d`, right. So, for that you need to install this package and then you have to upload it. But I will try to take here one very simple example to show you that how these things can be done very easily.

(Refer Slide Time: 23:02)

Three dimensional scatter plot:

Example

The data on height (in cms.), weight (in kg.) and age (in years) of 5 persons are recorded as follows. We would like to create a 3 dimensional plot for this data.

Person No.	Height (Cms.)	Weight (Kg.)	Age (Years)
1	100	30	10
2	125	35	15
3	145	50	20
4	160	65	30
5	170	70	35

30

So, yeah just this is a very simple hypothetical example, which I have created to explain you that I have considered 5 persons and then you know that height, weight and age they are inter connected, inter related. As the age increases usually under normal circumstances the height and weights also increase.

So, now, we would like to create here a three-dimensional figure such that there are three axis x, y and z and where we are trying to take the variables height weight and h for these three axis and then if we would like to explore more that how we can move ahead. So, you can see here this is person number 1, the height is 100 centimeters, then weight is 30 kgs, age is 10 years. This is person number 2, the height is 125 centimeters, weight is 35 kg and age is 15 years and so on.

So, what I try to do here, that I try to create the data vectors for the height, for the weight, and for the age, right. Like as here you can see.

(Refer Slide Time: 23:57)

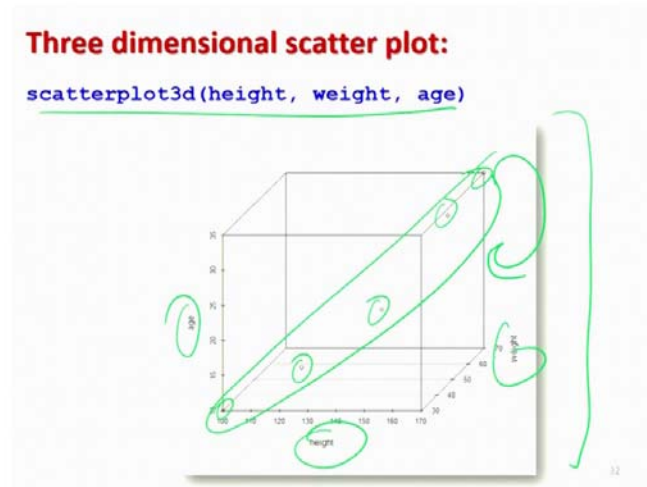
```
Three dimensional scatter plot:  
scatterplot3d() Plots a three dimensional (3D) point cloud  
install.packages("scatterplot3d")  
library(scatterplot3d)  
height = c(100, 125, 145, 160, 170)  
weight = c(30, 35, 50, 65, 70)  
age = c(10, 15, 20, 30, 35)
```

So, now I can explain you that you need to use the command scatterplot3d, where you have to be very careful s c a t t e r plot. That is in the lower case alphabets, but now there is here a number 3 and then it is here lower case alphabet d.

And inside this parenthesis, you have to give this values, right and then you have to first install a package scatterplot3d and you have to upload it. I already have done it on my

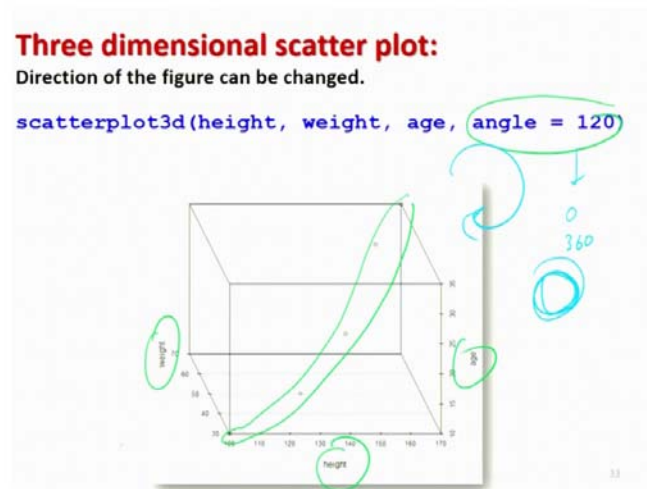
computer so but you also please do it. And then you see I have created here the data vectors on height weight and age, right.

(Refer Slide Time: 24:29)



As soon as you try to plot it, it will look like this. So, you can see here these are here the data vectors, right. So, I have taken it suitable way so that you can see it very clearly. And now there are many many options in this scatterplot3d.

(Refer Slide Time: 24:43)

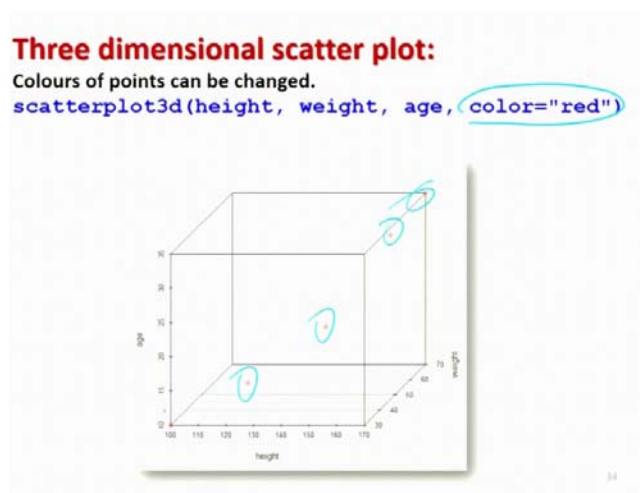


For example, if you try to see here there is an option here angle, angle is equal to here 120. So, if you try to see here this was the inclination. Now, this inclination is changed by 120 degrees and this inclination becomes here like this, right. It is changed and you can see here the axis here are age, weight and height.

But now it becomes here weight, height and age. So, now, this entire box is being rotated like this clockwise or anticlockwise that you also can control. So, now, if you have this type of option and if you write a simple loop and you try to change the angle, say from 0 to 360 degree and then you try to play the program. Do not you think that this picture will start moving like this.

And you can move this picture in any direction and can see what is the location of this point. So, if you try to see this is a very simple logic to view the two-dimensional picture or a three-dimensional picture in a two-dimensional way, right. So, it depends on your capability what you can do.

(Refer Slide Time: 25:40)



And similarly, if you want to change here the color. Now, this is a very simple thing here you can use here color; color that the same thing you can see here now the color is changed to red, right.

(Refer Slide Time: 25:51)

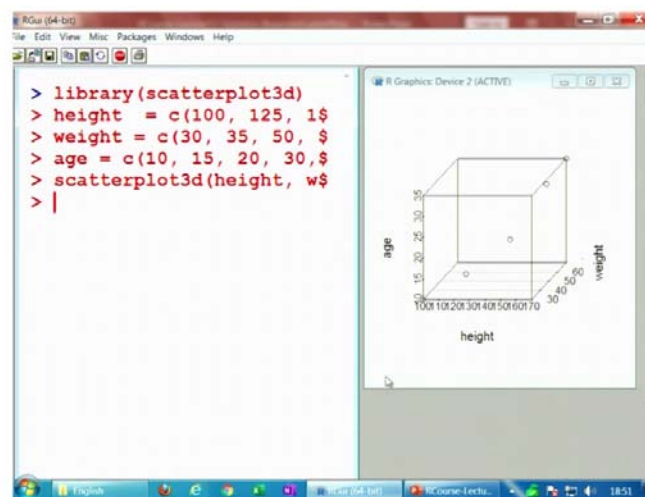
More functions:

- `contour()` for contour lines
- `dotchart()` for dot charts (replacement for bar charts)
- `image()` pictures with colors as third dimension
- `mosaicplot()` mosaic plot for (multidimensional) diagrams of categorical variables (contingency tables)
- `persp()` perspective surfaces over the x-y plane

35

So, these are the things which you can do here and beside those things we have many more plots, contoured plot, dotchart, image, mosaicplot, perspective loss etc. and yeah it is very difficult for me to consider all these plots. But surely, I will try to show you these commands.

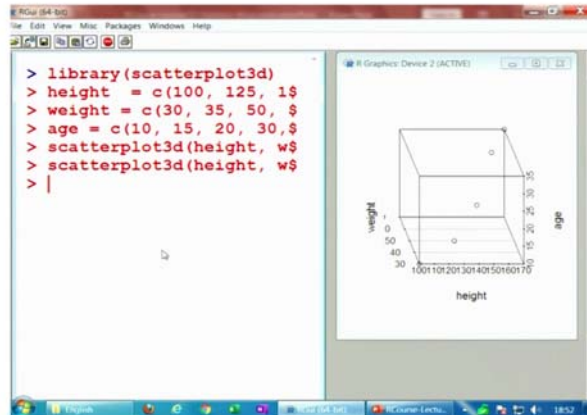
(Refer Slide Time: 26:09)



So, I try to use them I client screen and I first I try to load the library and then I have stored the data on say height weight and age and then I try to create here this

scatterplot3d and as soon as I do it, you can see here this is created here. And if you try to yeah make this picture larger you can do a you can see a better view, right. So, I hope you will not mind, right.

(Refer Slide Time: 26:33)



And then, if you try to change here the angle also that also. You can do here like this you can see now the angle is actually change. If you try to see here earlier one that was like this and now the new one it is here like this. So, you can see here because change and similarly if you want to add here a color also you can see here that this color can be changed here like this, right. So, you can see here it is not a very difficult thing that you cannot actually do.

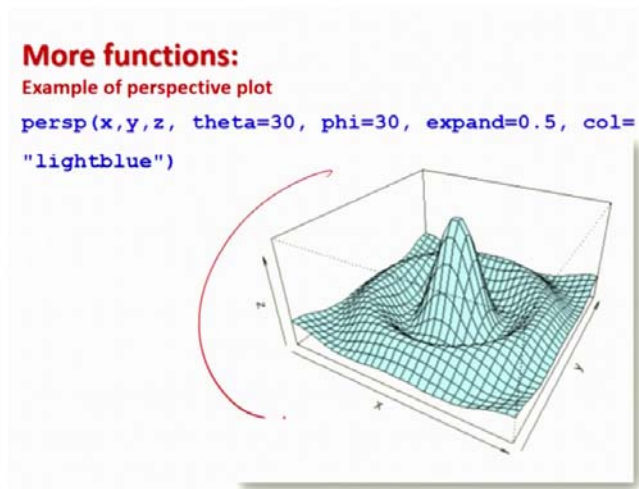
(Refer Slide Time: 26:56)

More functions:
 Example of perspective plot

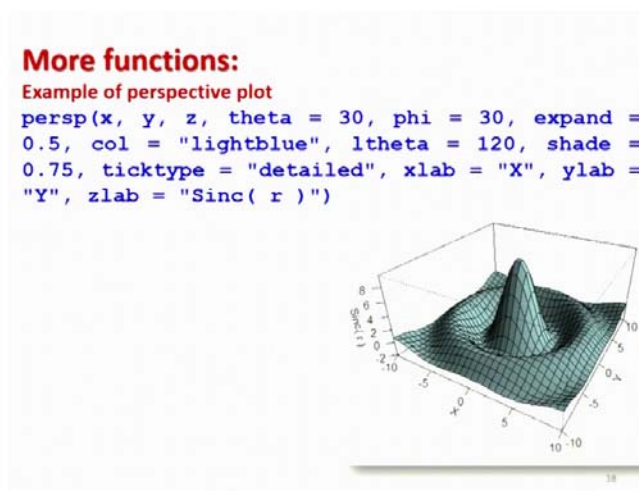
```
persp() perspective surfaces over the x-y plane
x = seq(-10, 10, length= 30)
y = x
f =function(x,y){r = sqrt(x^2+y^2);10*sin(r)/r}
z = outer(x, y, f)
z[is.na(z)] = 1
op = par(bg = "white")
```


So, after this actually I just want to give you a very quick example that how beautiful graphics you can make, this is the example of the perspective plus. I simply try to take this values, I am not going into the detail that what are these values I just want to give you a demonstration.

(Refer Slide Time: 27:08)



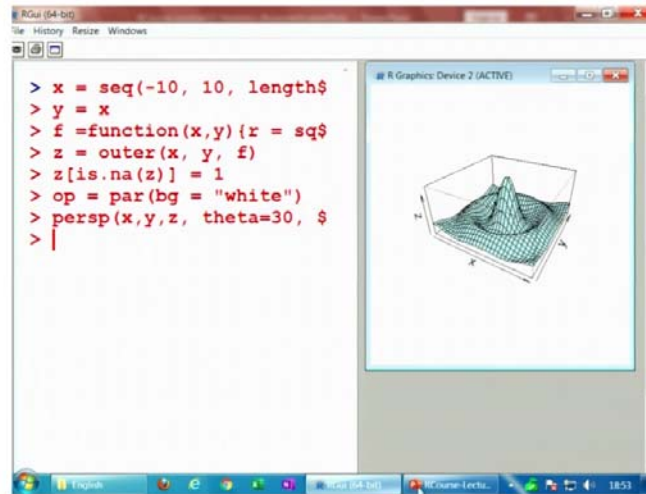
(Refer Slide Time: 27:13)



And then after that I use this command and you will get here this type of graphic. And if you try to change some parameter values you will get here this type of graphic, right. So,

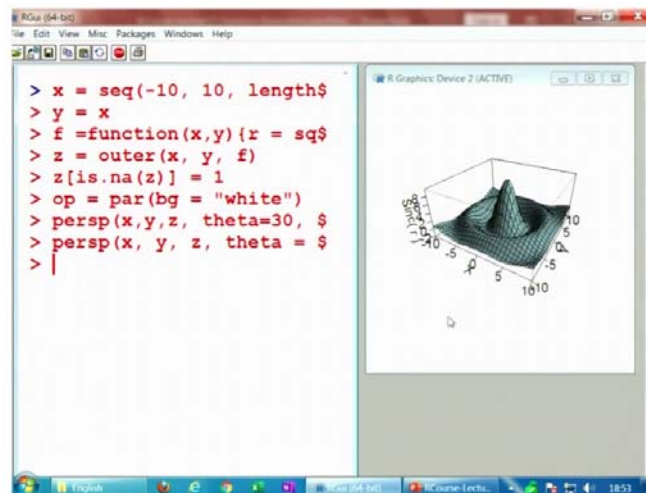
just as an advertisement of the R software, I would like to show you these things on the R console, right and then I will stop in this lecture, right.

(Refer Slide Time: 27:28)



So, if you try to see here, I have taken just these value. Surely I am not going to explain it in more detail, but I will leave it up to you that how far you want to go, but if you try to see here this curve is here, right.

(Refer Slide Time: 27:46)



And if you try to make it here like this perspective here you can see here you can see here now this curve is changed and this curve becomes here like this, right. You can increase the size minus the size, ok. So, now, I come to an end to this lecture and in this lecture I have tried my best to give you a quick review of the bivariate plot and a small glimpse of the three-dimensional plot.

But surely, as I said earlier, this is the beginning of the graphics that you want to learn in the R software. Now, there are many many commands, many many types of graphic, there are different packages like ggplot, they can create a very beautiful very impressive very informative graphics, but surely as I say always that if I try to take all of them in the same course possibly this course will become only on the graphics.

So, and that was not my objective also, my objective was simply to take the fear from your heart, do not forget, right. And possibly I feel that I have been successful and if you have followed the course and I am sure that you have followed the lectures, I am sure that you will not have fear to create any type of complicated graphics also in the R software. And I will see you in the next lecture, till then good bye.