HR Analytics

Prof. Santosh Rangnekar

Department of Human Resource Management

Indian Institute of Technology, Roorkee

Dr. Abhishek Singh, Assistant Professor

Department of Human Resource Management

Indian Institute of Management, Rohtak

Week:12

Lecture:58

Dear participant, in this session, we will learn about Tableau, how we can visualize the data by using the Tableau software, right. So, in this course, as you are already aware, we have used three softwares or three tools to visualize the data. One is Tableau, second one is Power BI and third one is the excel. So, in this session, we will learn about the Tableau. So, before starting the hands-on exercise on Tableau, just I would like to tell you how to prepare your data about the, for the Tableau exercise, right? What is the process that you are supposed to do, right? So, what do you need to do whenever you are preparing the data? So, if you remember the data science framework that we have learned in the first week right. So that says first you need to identify, first you need to define the problem, first you need to define the problem right.

So first you will define the problem. After defining the problem what you will do? You will define the variables, define the variables which you are needed to describe the problem or to find out the solution, define the variable. So, variable like age, gender, work experience. department, in which department you are working.

So that is like this data you need to prepare a one excel sheet. So in that employee related data that you will see the problem that you want we want to address. So what I have done I have created one dummy data for this tableau exercise. So first I will explain you the dummy data which I have developed for this exercise in excel. So, in that this age, gender, work experience, department and then we will upload on a tableau and then we will use it.

So, let us start the use of tableau, let us understand how we can upload this data on

tableau and we can visualize this data. So, dear participants as I discuss how to prepare the excel sheet. So I have prepared one excel sheet for this exercise. So I will discuss that excel sheet first with you and then I will open the software and upload this excel sheet. So in this excel sheet that is what you can see.

This is the first I want to make you understand. This is the completely dummy data. This data just I have prepared only for this exercise purpose. This data I have not collected from any company, I have not downloaded from any source. This just variables that I have created here and by using these variables I will show you how to analyze the data on Tableau.

or how to visualize the data on Tableau. So let us understand the variables which what are the variables are there. So here we can see the first column in which serial number is there. So you can write employee ID, you can write serial number just to understand how many data point that we have collected. So here you can see the serial number is there.

So you can see how many data point that dummy data that I have created. So, 516 data point that we have collected in this excel that is what you can see, right. So, this is the number. So, it could be the number, employee ID, employee number or if you are downloading the data from HR IS, from your HR information system, so that may generate one unique ID. So, that unique ID will help you to count how many numbers are there.

Second thing that you can see, second variable that we have, that is the gender, right. So what is their gender, though 516 employee data is there, so what is their gender, how many males are there and how many females are there, right. Third thing that you can see that department, so it is the recruitment related data, so for which department they have applied, right. Fourth one that you can see the date. So date, so interview date on which date their interview is scheduled.

On which date their interview is scheduled. Next variable that you can see from which city they are coming. Next you can say status of for interview or status of interview like whether they are coming or not coming. Source of application from which source they have applied. how much work experience that they have like for how many years they have worked before this job so that is what is mentioned and then education level that you can see what is their qualification next you can see the age what is their age category Next that you can see the salary which is in 1000.

So how much salary that they are expecting from this particular job. So just I have written 12, 13, 14 but it is 1000. So K, how many thousand salary that one particular

individual is expecting from the organization. So this is the data, this dummy data that I have created. So, as you remember I already said first you need to define the problem.

So, based on data if you want to know how many people are coming from which city. So, it could be one of the problem. Or you want to understand from which source how many people have applied, for which department maximum number of people have applied. So, such kind of data, such kind of questions that we will try to answer through in this data through the tableau. So, this is the excel sheet that I have generated dummy data.

So, I hope these all variables are clear to you. So, now let me open the tableau. So, I have downloaded the public version of the Tableau right. So, it is opening I have clicked on it. So, you can see you can also download you can go to this Tableau website and that free version of Tableau is there that you can download and for the practice purpose and one trial version that you can download and it will work for 14 or 15 days and this public Tableau public is available this also you can use.

right and you can use and you can try and you can work on this also so this i am using this tableau public right so here you can see on connect various types of data files are there which you can use like text data right microsoft access pdf file microsoft excel so for this exercise i have developed or i have created the excel data so i have i will click on this excel so I have clicked, I will click on this Excel and this is the file that I have, I am going to use to analyze this data. So, here I am going to use this data. So, I have clicked on this Excel and the moment I will open it, then you will see, this sheet has opened. and you can see the seat is there right seat number one that is what you can see and here the need more data then you can click on this that is what you can see and all variables that you will see which i had discussed with you that are there so if you want to check you can check it here like first serial number is there second second that you can see a gender is there then department is there, then interview date is there, city is there, from which city they are coming, status of interview is there, source of application is there, work experience is there, education is there, age is there, salary is there. So these all variables that you had created in the excel sheet, all these variables are here.

right and that is how you will see the tableau sheet which is there in front of your screen. In the screen that is the screen that is what you can see. Now you can see at the left hand side you can see the sheet number 1 And above that this go to the work sheet, right. This is the option that is coming. So you need to click on go to the work sheet.

So the moment you will click on this, go to the work sheet, the moment you will click on this, so here you can see at the bottom sheet number 1 will come, right. and you will see

the all variables list that has come here. So, that variable that we had created city, department, education, gender, source of application, status of interview. age, interview date, salary, serial number, work experience and some variables that you can say that automatically they are generated. So if you want to understand what is the understanding about this variable, so what you can see here, you can click on this and this dialog box will open and if you want to know whether these variables are clearly uploaded or not.

So you can click on this describe and after clicking on describe you can see this is the variable. So discrete dimension that is what you can see here. City that you can see. So the detail about the variable that you will get. So here you can see the city.

So how many cities are there? 6 cities are there like Bangalore, Chennai, Delhi, Hyderabad, Calcutta and Mumbai. So, if you want to know about the variable like for example the city that I have taken. So, you can click on this and then you can understand which type of variable it is, whether right information is uploaded here or not. Similarly, you can click on this department and if you want to make some changes you can click on this transform and then you can make changes, split, custom split, mark date, So whatever change that you want to do it that you can and in the same way if you want to know what is the meaning of this department, so you will click on this you will understand. So discrete dimension that is what you can see and all departments names are there.

So if you need to know the more information about the variable if you are unable to understand what this variable is. like in a excel sheet sometime it may be possible you are getting the huge data you are unable to understand these variables. Now simple variables that I have created here so you can understand easily but sometime it happens you are getting the huge data through your HR IMS. So here you will on the left hand side you will get the list of all these variables and if you want to know which variable means what. So what you can do you can click on that variable and then you can understand whether that variable is on nominal scale, ordinal scale, interval scale or ratio scale that is what you can understand right.

And in this side if you left hand side if you will see clearly here two things one line is there right after this major names you can see this line is there. So, some variables are above the line and some variables are below the line. So, you will see all variables which are below the line they are quantitative in nature. like age, interview, salary, serial number, work experience, count, so these all are variables are in quantitative in number. But above the line if you will see the variables they all are in you can say that text format, city, so name is there, text format that is what you can say or you can say descriptive format that is what you can say.

So number or text, there is only two type of data. So below the line that you will see all variables are presenting the number and above the line all variables are presenting the text. That is what you can see. So that is how you can see here. So now let us start with making some of the graph and then we will understand how to make it.

Before starting making the graph I want to tell you that the way we have used this pivot table, if you are good at with the pivot table you won't face much problem with the tableau even with the Power BI also. So here you can see that on the top this rows and columns are there, rows and columns are there so that you will see and pages are there and here filters are there. So pages, if you need to the sheet 1 on which we are working, if you want to add it, you can click next to the sheet number 1 and then you will get the next sheet. That is what will happen and next to the pages you can see one is data view is there and next one is analytics is there. So here if you want to the after putting the data here if you want to make some line graph, bar graph, median, quartile.

So analytics like basic statistical tool that you want to apply so you can apply through here also. So, now what we will see? We will see how to use this filter, rows and column. So, if you understand what is row and what is column, then you will be able to work very quickly on this tableau. One suggestion that I have, once you are, you have excel, after having the excel, what you do? You have got this list of variables. Now you decide for which variable you want to make which type of graph.

For example, this city is there, right? If you want to know how many people are coming for the interview from which city, right? So what do you need to do here? First you need to identify how many people are there. So how many people are there? So that you will get it from the count. right, number. So, you got this number, right, that I have put in a rows. Now, you want to know how many people are there in a city, right, city wise.

So, now you can see the number. Highest number of people are coming from the Chennai for the interview, right. I have applied from the Chennai. Now, I want to know how many people are coming from the different different cities. So, here you can see this status of interview. So, along with count I will put it this also.

So, here you can see from here. So, you can see other from Bangalore, right? So, here how many people are coming? and not coming, Chennai coming and not coming. So, 131 are coming and 131, 129 is not coming. So, in the case of Bangalore, two people are there and both of them are coming. In the case of Delhi, 10 are coming and in the case of Hyderabad, 37 are coming. So, status of interview that is what you can see here coming, not coming, right.

So, from Kolkata 55 people are not coming, Mumbai that is what you can see one is coming and 150 people are not coming. So, still that visualization is not very much clear, right. So, what you can do? You can put this into the filters, right. So, status of application you put it in filter.

So, it happens. when you will try to put it sometime it will happen right you will not be able to put. You select from here and put into the filter right it is not coming right. So, sometime it happens, sometime it happens you will try to check right. Now you want to know from which city how many male and how many female are coming right. Just you write it gender here So, male and female wise data that you will get it.

So, here you can see male, female, male, female. So, that data is divided. If you want to convert the same thing into the table, you can see here, you have converted into the table. So, just you need to, you can see here all types of graphs are there, you need to click on that. So, basically what I am trying, in the first session of this tableau, I am trying to make you understand only one thing. You got the list of variables and you got this tool rows and column.

So here you have to identify the problems. Problems like for example if you want to know what is the average age of people who are coming from the different different cities. right? If that is what you want to know, right? What is the average age of people who are coming from the different different cities? So, average age, so first you write the count, so that column that you have and now you have put it into the age. So, now it is showing the sum. So, do you think the sum of age is makes any sense here? I hope no.

So, sum is not making any sense here. So, what so how you will change it? So, you need to see here on this green age is there and bottom triangle is there just click on this you will see the list, one dialog box will be open, in dialog box you can see the sum, average, median, count, maximum, standard deviation, so whatever you want to count here. So, here what is the, we want to know the average, so just click on the average. So average that you want, so this graph that you got it, the count and then average. So now you can see this graph is not giving any information to you. Why it is not giving the information? Because you do not have the clarity about the where to put this rows and where to, what to put in rows and what to put into the count.

So at the beginning level such kind of mistakes will take place with you, it will happen because you have not decided which type of graph that you have to make and in which in column which variable you need to put, in rows which variable you have to put. So, when you are making the graph on Tableau or any data visualization tool, first make the graph

in the advance in your mind. If you are a beginner, then I would suggest make the graph on pen and paper and define everything like what would be the title, what would be x axis, what you will write on y. how you will present it, everything that you write on, everything that you write on x axis, y axis, title, everything that you mention in the rough and then the same thing that you repeat here. So here we want to know the averages but this graph is not making any sense.

So what you can do, now let me reverse the things. So how we can reverse? So let me reverse the thing. So I will put age. into the column and then I will put count into the rows right again that we are saying getting the same right same thing that we have done. So, let me put this count into the here and sum into the rows again you got this right.

So, I have clicked here. So, I have changed the graph because that graph was not giving the right thing. So, I have changed, so here you can see the sum major value is there, that is what you can see here. So, I have removed, this also I removed, right. Again I will try, so that is how you have to keep trying it. So, what do you have to do? Just think again, what do you need to write? You need to count how many male and females are there, right.

Let us start with this. So, how many male and female? right? You got this, right? You got the count and here you can see the male and female. In the same way, I want to know the average age, right? So, I have written here average. in the case of male and female. So, this is the sum is there, sum of age.

So, I need to convert into the average. So, from major, major to average. So, I have clicked on the average. So, average age in the case of female is close to 40. right and that number is there and here you can see the total number more than male or more.

So, averages is close between 45 to 50. So, here you can see 47.5. this is the averages for the male and this is the averages for the female. So I want to make this from that graph to this bar chart that is what I want to make it. So here you can see the male and female. Male and female is there and this is the average age is there, right.

So, that is how you can just any keep on. So, in this in this case also you will get the same, but the presentation has changed. So, female average is 44 and male age is 55, right and in the same way you can just keep checking which graph is more suitable. So, in this you can see this graph is not suitable, is not making any sense here, right. that I am trying to make you understand here, like that is how you will get the variable, text variable and number variable.

What you have to put in column and rows. So at the beginning stage, just you keep playing like this. You write one variable and then see what you are getting on this sheet. and then you write one numerical variable in the rows, column, keep changing, keep doing it because if step by step if you will go I will tell you the step in the first step does this, second step does this then if you want to explore more then you will not be. So what is my suggestion is in this session you develop a create a dummy excel sheet understand ask question to yourself based on this variable and then make a graph accordingly whatever graphs are there. So Simple that I gave you the example, so let me understand how many people are there from the various cities.

How many people are there from the various cities? That is what I want to know. So, this is the graph that is there. So, here male and female also there. So, here gender that I have put and city I have put, gender. So, if you want, I want to remove this, I want only that number like from which city how many people are coming.

So, that number I want, so I have written count here, right? So, this table, so I want to make it a table, so this table that I got it. If I want to make it pie chart, pie chart would be there, but you need to check here and if you want to write something, so you can write here. You can click on this pen and then you can write what are the things that you want to add here. So, this table that you have made bar graph that you can make like here this count how many people are coming from which city. So, here you can see the maximum number people are coming from the Chennai 260 because this bar graph is telling to you.

So, that is how you can just keep asking the question to you, identify for this question which variable that you need to use and pick up that variable, put it into the rows and column and see which graph that you are getting. If you are not getting any information from that, see in the table which graph is suggested, click on that and then you will get the particular graph. So, I hope you would have understood how to prepare the excel sheet, how to upload the data on the tableau and then how to initiate this tableau and how to frame the questions related to your data set and find out the answers. So some of the questions we will discuss in the next session. So dear participant, I hope you would have learnt how to prepare the excel sheet.

and how to upload on the tableau, right, how to upload this excel sheet on a tableau, how to understand these variables, so text variables and numeric variables. And the moment you have understood these all variables then ask question related to it and then select the variable, put into the column and rows and see which type of graph that you are getting. If you are not getting the answers in your way the way you wanted then change the graph, then change the graph and then see whether that presentation is appropriate or not, right. So, this practice you keep doing it and I hope you will understand how to use this

tableau, right. So, in coming session we will go in little bit more in depth and we will ask some more question related to this particular data and then we will focus on.

So, few variables like age and work experience and salary. So, in this session just we simply discuss about the a few variables, right, city, gender like that. So, in coming session we will discuss the question related to these work experience, related to the age and expected salary. Thank you.