**Video Script: Section 3 Video 1 Using group and color**

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Description | Action on screen | Narration |
| 1 | Introduction  (Outcome and why it is desirable)  This should give the viewer an idea of the outcome of the task at the beginning of the videos and set the stage and expectations of the viewer. | Refer to PPT | In this section, we’ll see how to use more aesthetics to add extra information to a graph.  In this video, you’ll understand about the ‘group’ and ‘color’ aesthetics. |
| 2 | Context(Problem/Solution)  Present the viewer with a real-world solution and how the situation would pose as a challenge. It always helps to draw the viewer's attention using a use-case. Metadata template can be used here. |  | When we have two variables to visualize, we can use the x and y coordinates to represent the data. For a third variable, we can use color. Remember that in ggplot terminology, these are all called aesthetics. For visualizing more than three variables, we just need to add more aesthetics.” |
| 3 | Guidance (How to do it and how it works): |  | So far we have only used ‘colour’ as an aesthetic to differentiate between data points. We can use more aesthetics simultaneously, for example group and color to convey more information. |
| 4 |  | Highlight and run:  library(ggplot2)  ?ChickWeight  # How many chicks per diet?  aggregate(Chick ~ Diet, data = ChickWeight, FUN = function(x) length(unique(x)))  On the screen:  Diet Chick  1 1 20  2 2 10  3 3 10  4 4 10 | Open activity\_03\_01.R  Run the first 10 lines to load ggplot2 and get a first view of the data.  The Chickweight data contains information about the growth of chicks under different diets.  As you can see, we have from 10 to 20 chicks assigned to each of the 4 diets. |
| 5 |  |  | A line connecting the weights over time for each chick is appropriate, so we’ll be using geom\_line,  Butwe also want to differentiate between diets. |
| 6 |  |  | For this, we map the diet to colour and an individual chick to group. The aesthetics ‘group’ is used for connecting the data points, when used with geom\_line. |
| 7 |  | Highlight and run:  ggplot(ChickWeight) +  geom\_line(aes(x=Time, y=weight, colour=Diet, group=Chick)) +  ggtitle("Weight versus age of chicks on different diets")  A description... | Let’s run the ggplot command. |
| 8 |  |  | This way, each chick has its own line and each line is coloured according to the chick’s diet. As usual, we also get the default legend for free. |
| 9 |  |  |  |
| 10 |  |  |  |
| 11 |  |  |  |
| 12 |  |  |  |
| 13 |  |  |  |
| 14 |  |  |  |
| 15 |  |  |  |
| 16 | Conclusion:The video concludes by showing the viewer that the goal has been achieved, and reminding them why they should be happy about that. A PowerPoint summary slide with the key points emphasized would make it easier for the viewer to remember what was covered in the video | Back to the PPT | You can convey more information by mapping multiple aesthetics at once. As you’ve seen, the aesthetics ‘group’ is useful for collecting data points in different subsets of the data.  In the next video, we’ll look at using another aesthetics: ‘size’. |