**Video Script: Section 5 Video 4 – Customizing the color palette for continuous data**

|  |  |  |  |
| --- | --- | --- | --- |
| 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 video, we’ll see how to customize the colour palette for continuous data. |
| 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. |  | By default, ggplot2 uses a gradient from black to blue for colouring continuous variables. You might want to use a different gradient or even more than two colours, to visualize different value bands, for example. |
| 3 | Guidance (How to do it and how it works): | Switch to activity\_05\_04.R in RStudio | Open activity\_05\_04.R and run the first example. |
| 4 |  | Highlight and run:  # default  p <- ggplot(diamonds)+geom\_point(aes(x=carat,y=depth,colour=price))+scale\_x\_log10()  p + ggtitle("Default palette") | We’re plotting the diamonds dataset, with carat vs depth (on a log scale). The colour is a function of the price. |
| 5 |  | A description... | Instead of the usual black to blue gradient, we might want to use, for example, yellow to red. To change the gradient, use scale\_colour\_gradient() |
| 6 |  | Highlight low='yellow', high='red' | The parameters *low* and *high* control which colours you want to use. |
| 7 |  | Highlight limits=c(5000,10000) | You can also use the parameter ‘limit’ to focus on a specific band of value in your data.  For example here, we decide to only look at prices between 5000$ and 10'000$. We can do that without having to subset and modifying the original data. |
| 8 |  | Highlight and run:  p + scale\_color\_gradient(limits = c(5000,10000), low = 'yellow', high = 'red') +  ggtitle("Gradient with custom gradient")  A description... | The gradient now goes from yellow to red and for prices from 5000 to 10'000, as we wanted. |
| 9 |  | Highlight and run:  p + scale\_colour\_gradient(limits=c(5000,10000), low='yellow', high='red', na.value=NA) +  ggtitle("Gradient with custom gradient\nwithout unneccessary datapoints") | Datapoints not in this range are shown in gray.  Use the parameter na.value to change their colour to another, or set it to NA to make them invisible.  It only affects the plots, not the data. |
| 10 |  | A description... | You can also use multiple gradients. For example, scale\_colour\_gradient2 asks for an intermediate colour to go through. |
| 11 |  |  | Suppose we want to visualize whether a diamond has a price below or above average.  We can use a first gradient from red to white for diamonds below the average price and a second gradient, from white to blue for more expensive diamonds. |
| 12 |  | Highlight  midpoint = mean(log10(diamonds$price)) | For this, we set the parameter *midpoint* at the average price. We could change the colours by using low, mid and high but they’re already at red, white and blue by default. |
| 13 |  | Highlight trans = ‘log10’ | As an extra, we also use the parameter ‘trans’ to define the gradient on a log scale. It is more appropriate because a 100$ difference between cheap diamonds is not the same as a 100$ difference between expensive diamonds. |
| 14 |  | Highlight and run:  p + scale\_color\_gradient2(midpoint = mean(log10(diamonds$price)), trans = 'log10') +  ggtitle("Two gradients using the average price as midpoint")  A description... | We see which diamonds are below or above average very easily (e.g. from 1 carat onwards, diamonds are more expensive than average)  We also see how the log scale automatically appears on the colour guide. |
| 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 PPT | In this video, we saw how to override the default palette for continuous variables and achieve interesting effects.  In the next video, we’ll see how to customize non-data elements, like legend title and labels. |