**Video Script: Section 5 Video 3 – Customizing the color palette for categorical 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 override the default colour palette and make our own. |
| 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. |  | The default colours for a categorical variables are chosen by picking up evenly-spaced hues and keeping luminance and chroma constant. |
| 3 | Guidance (How to do it and how it works): | Switch to activity05\_03.R on RStudio | Let’s have a look at a simple plot. |
| 4 |  | Highlight and run:  p <- ggplot(diamonds) + geom\_histogram(aes(x = cut,fill = clarity))  p + ggtitle("default palette")  A description... | Eight colours, for eight categories, are automatically chosen by default. You can see the color progression from orange to brown to eventually pink expected by going round a colour wheel.  This is a simple enough system that gives good results. |
| 5 |  |  | If you have a good idea of the color scheme you’re after, you can use scale\_something\_manual, with which you spell out the colour for each value of your categorical variables.  Scale\_colour\_manual will override the palette for the ‘colour ‘ aesthetics and scale\_fill\_manual will override the palette for the ‘fill’ aesthetic. |
| 6 |  | Highlight and run:  p + scale\_fill\_manual(  values = c( "I1" = "red", SI2 = "red", SI1 = "blue", "VS2" = "blue", "VS1" = "green", "VVS2" = "green", "VVS1" = "black", "IF" = "#FEAC12")) | Run the next example. |
| 7 |  | A description... | Colours can be defined by name (‘red’, blue’ etc.) or by their RGB code. |
| 8 |  |  | Choosing a good palette is actually fairly difficult. Scale\_colour\_brewer and scale\_fill\_brewer offer a good collection of palettes that have been carefully designed to give a good contrast on geographical maps. |
| 9 |  | Switch to colorbrewer2.org | You can see them in action on the website colorbrewer2.org |
| 10 |  | Play with the interface while talking. | There you can choose between sequential, diverging and qualitative palettes and see their effect.  The website also tells you whether it is suitable a colour-blind audience. |
| 11 |  | Back to RStudio.  Highlight and run:  p + scale\_fill\_brewer(type='qual',palette="Set1")  A description... | We used the palette ‘Set1’ of type ‘qualitative’ from colorbrewer2.org for our plot.  The colours chosen have a very sharp contrast and are very distinctive. |
| 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 PPT | In this video, we’ve seen how to control the colour palette for categorical variables.  In the next section, we’ll see how to override the default palette for continuous variables. |