**Video Script: Section 1 Video 2 – Understanding the structure of a plot**

|  |  |  |  |
| --- | --- | --- | --- |
| 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. | Opening slide | **In this video**, we are going to look at the structure of a ggplot2 object.  This object contains the necessary information for producing a graphic, when it is rendered. |
| 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 gg in ggplot is short for grammar of graphics. The main idea is that instead of having a single complicated function that draws a plot, you build up a plot from many smaller functions, just like combining words to form a sentence.  Although the syntax differs from many other plotting languages, it makes your plotting code more readable and easier to debug.  This can be disconcerting at first, but you’ll understand it better as you follow along this video. |
| 3 | Guidance (How to do it and how it works): |  | Let's see a complete ggplot2 command. Most of the terms are probably obscure for the time being but let's ignore the details for now and concentrate on the general structure of the plot. |
| 4 | Consider a multi-layered ggplot2 plot (provided) | In Rstudio, open activity01\_02.R  The code is run and a plot appears.  01_02_image.png | Open Activity01\_02.R in RStudio and run the code by highlighting the whole script (you can use ctr+a to select all lines) and pressing ctrl+enter to copy it to the console. |
| 5 | Decompose the object in its layer. | Highlight:  library(ggplot2)  set.seed(42)  small<-diamonds[sample(nrow(diamonds),1000),]  head(small) | The first lines of the script call the library and prepare some data to be plotted.  Here we randomly sample 1000 rows of the built-in diamonds dataset.  Head(small) shows the first rows of our small dataset. Diamonds are described in terms of carat, price, shape etc. |
| 6 |  | Highlight:  ggplot(small) | Let's decompose the ggplot2 command into a few elements.  The first line, ggplot(small) says that we want to create a ggplot of the small dataset that we just created.  It's important that this argument is always a data frame.  At this point, nothing will be plotted: only the data has been set up, no drawing directives have been used yet.  If you run this on its own, you'll get the following message:  'Error: No layers in plot' |
| 7 |  | Highlight:  + geom\_point(aes(x=carat,y=price,color=cut)) | geom\_point tells ggplot that we want to add points to our plot. That is, we are creating a scatterplot. Geom is short for geometric object.  We want to plot the carat variable on the x axis, the price variable on the y axis, and change the colour depending on the cut variable. Notice that these instructions are wrapped inside the aes function. aes is short for aesthetics.Note the '+' sign joins all the commands together: this is how each layer is added to the object. |
| 8 |  | Highlight:  scale\_y\_log10() | Next, scale\_y\_log10 tells ggplot to make the y-axis log-scaled.  Scales are an important part of ggplot2 and have applications not only for axes but also color schemes, sizes and shapes. |
| 9 |  | Highlight:  facet\_wrap(~cut) | facet\_wrap splits the plot into several facets. That means several panels. In other graphical packages, this is sometimes called trellising or latticing. |
| 10 |  | Highlight:  ggtitle("First example") | ggtitle adds a custom title to the plot. |
| 11 |  | Highlight the whole ggplot2 command. | By putting these 5 commands together, we obtained the plot you see on the screen.  This is the general structure of a plot with ggplot2: layers are added one at the time, each with their own specific logic. This will allow you to build very complex plots in a principled manner. |
| 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 | Last slide of the PPT | The seemingly complex code for the plot is now clearer. Now we'll move onto the details of the commands and the grammar behind ggplot2, in particular the concepts of 'aesthetics' and 'parameters'. |