**Plotting with ggplot2  
Post-session activities**

1. **Loading packages and reading in data**
   1. Load the *ggplot2* and *dplyr* packages  
      **Hint:** use the library() function separately for each package
   2. Read the “winemag-data\_first150k.csv” file and store its contents in an object called “wine”
2. **Creating scatter plots**
   1. Use the filter() function to keep only rows whose variety is “Malbec” **OR** “Tempranillo”, **AND** whose price is less than 100. Store this as an object called malbec  
      **Hint:** filter(wine, variety == “x” | variety == “y”, price < z) will retain rows whose variety is equal to (==) “x” or “y”, and whose price is less than (<) z
   2. Create a scatter plot of malbec with price mapped to the x axis and points mapped to the y axis
   3. Create the same plot again, but place a 2d density contours *underneath* the points  
      **Hint:** use geom\_density\_2d() for this
   4. Create the same scatter plot as in (b), but this time map the variety variable to the colour of the points
   5. Create the same plot as in (d), but create separate subplots for each different country  
      **Hint:** use facet\_wrap() for this
3. **Creating line plots**
   1. Execute the function data() to see all the datasets built into R (and the packages you have loaded) available for you to play with.
   2. Execute data(economics\_long) to read in the built-in economics dataset and explore it (run ?economics\_long to find out more about the variables)
   3. Create a line plot of economics\_long with date mapped to the x axis, value mapped to the y axis, and variable mapped to the colour aesthetic. Why is this plot so difficult to read?
   4. Create the same plot, but this time instead of colouring by variable, create separate subplots instead.
   5. Edit the plot above by adding the argument scales = “free\_y” inside facet\_wrap(). How does this change the plot and why?
   6. Add a geom\_smooth(span = 1) layer *on top* of the lines. What happens to the line when you set span equal to 0.2?
4. **Creating bar plots**
   1. Use data(presidential) to load data of the terms of US presidents from 1953 to 2009, explore it
   2. Use mutate() to compute a new variable called term that represents the length of each president’s time in office. Overwrite the original data object with this dataset  
      **Hint: you can subtract dates**
   3. Create a bar plot that represents the frequencies of the name variable. Why does “Bush” appear twice?
   4. Change the value in the 8th row 1st column to be “HW Bush” and change the value in the 10th row 1st column to be “W Bush”
   5. Create a bar plot with name mapped to the x aesthetic and term mapped to the y aesthetic  
      **Hint:** for geom\_bar() to plot the y value, you need to add the argument stat = “identity” to geom\_bar()
   6. Edit the plot in (e) by replacing name with reorder(name, start) inside the aes() function. This reorders the presidents in order of their presidency.
   7. Map the party variable to the fill aesthetic, and set the colour of the bars equal to “black”
5. **Creating boxplots / violin plots**
   1. Create a box plot using the malbec dataset you created in part (2a), mapping country to the x axis, and points to the y variable. Which country ranks lowest for these wines?
   2. Map the variety variable to the fill aesthetic. Why do some countries only have a single boxplot?
   3. Copy and run the following code:  
        
      ggplot(malbec, aes(country, fill = variety)) +  
      geom\_bar(position = "dodge")  
        
      What happens when you change the position argument to “stack” and then to “fill”?
   4. Apply a theme of your choice to a plot of your choice
   5. Use the labs() function to give your plot a title, subtitle, and more accurate x and y axis labels  
      **Hint:** add labs(title = “this is a title”, subtitle = “this is a subtitle”) as a layer with “+”
   6. Save a png of your plot using the ggsave() function

**If you have any questions or want feedback, please contact** [**hefin.rhys@ucb.com**](mailto:hefin.rhys@ucb.com)