Tidy Data and Visualisation with R Exercises

Before you start, make sure you've read the document that describes the zebrafish dataset we're using in this practical. And make sure you've put the required file (`Amp.counts.tsv`) in your home directory.

Load the example dataset (`Amp.counts.tsv`) using `read_tsv`.

```
library(tidyverse)
results <- read_tsv('Amp.counts.tsv',
                    col_types = cols(Chr = col_character(),
                                     Strand = col_character()))
```

Tidy Data

1. Use `filter` to find out how many genes have an adjusted p-value less than 0.05.

2. Find out which gene has the smallest p-value by sorting the data using `arrange`.

3. Make a new column in the data that is -log10 of the adjusted p-value column. You can use the `log10()` function to calculate this.

4. Make a new data.frame that contains the Gene, Name and all the normalised count columns.

5. Make the new data.frame data tidy using `pivot_longer`.

Plotting

Load the penguins dataset with `library(palmerpenguins)`.

1. Using the `penguins` dataset, make a plot of body mass (body_mass_g) against bill length (bill_length_mm).

2. Now colour the points by flipper length (flipper_length_mm) and use the `viridis` colour scale using `scale_colour_viridis_c`.

3. Change the shape of the points to a hollow shape (one of 21-25).
   Make species the fill colour, and pick 3 colours to use with `scale_fill_manual`.