**Data Viewer Instructions**

Opening the program

Open Google Chrome and go to the URL: https://paleo13.shinyapps.io/data-viewer-biol2015/

**NOTE: Please do not use a different browser.**

Loading data

Data for the mangrove herbivory project collected during week 1 in 2014 has been provided. No ther data has been provided. **Please do not attempt to load other datasets**. This may crash the program.

**To load all the formatted data,**

1) click on the 'load data' tab on the sidebar;

2) in the drop down menu under 'Project', select the 'Mangrove herbivory' option;

3) leave the 'load data for specific group?' checkbox unticked;

4) click the 'Load Data' button

**To load data for a specific group,**

1) click on the 'load data' tab on the sidebar;

2) in the drop down menu under 'Project', select the 'Mangrove herbivory' option;

3) click on the 'load data for specific group?' checkbox to make sure it is ticked;

4) in the drop down menu under 'Week Number', select 'Week 1';

5) in the drop down menu under 'Group Colour', select a color (eg. Green),

6) in the drop down menu under 'Group Name', select data for the group names you wish to load.

**NOTE:** if the drop down menu under 'Group Name' does not contain any values, try selecting a different group color, and then selecting the group color you want to force the program to reload the data.

Exploring data

*'Variables' sidebar tab*

The options in the 'Variables' tab on the sidebar are used to select variables to visualise and perform statistical analyses. **If any of the drop down menus in this sidebar tab are empty, then the data has not been loaded correctly, please refer to the above instructions on how to do this.**

The 'Response variable' drop down menu is used to specify the response (aka dependent) variable for the statistical analyses, and the variable that will appear on the y-axis in the plots. This drop-down menu will contain the names of columns in the formatted dataset that have continuous values.

The two 'Predictor variable' drop down menus are used to specify predictor (aka independnet) variables, and will appear in the x-axis or legends in the plots. The first drop down menu contains the names of columns in the dataset that contain either continuous or categorical values, and the second menu contains names of columns that contain categorical values.

To investigate variation in a single variable, set the 'Response variable' menu to the desired variable, and the 'Predictor variable' menus to '----'. To investigate relationships between two variables, set the 'Response variable' and the first 'Predictor variable' menu accordingly, and set the second 'Predictor variable' menu to '----'. To investigate relationships between three variables, set all menus accordingly.

The 'Error distribution' drop down menu contains the names of various error structures that can be used in the statistical analyses. For all columns in the mangrove herbivory dataset, the only valid value is 'Gaussian' (the default).

*'Plot' tab*

This tab contains a plot to visualise the data and model predictions. Depending on the nature of the data, the plot may be a :

- **histogram:** the blue bars indicate relative frequency (density) of data. The black line denotes the density of data using a kernel. The blue dashed line indicates the mean of the data.

- **scatter plot:** open circles denote data. The lines indicates model predictions, and the gray shape indicates standard error of model predictions.

- **bar plot:** bars indicate predictions and standard errors according to the specified model (either one-way or two-way ANOVA).

*'Diagnostics' tab*

This tab contains the standard plots to assess the standard assumptions of general linear models in R.

*'Results' tab*

This tab contains various summary statistics and results from statistical analyses.

The 'Data Summary' section contains a table with the mean, variance, standard deviation, and standard error for the response variable, grouped by any categorical predictor variables. This information will be useful for answers questions in the booklet on the island.

The 'Modal Summary' section contains output from the 'summary' function in R.

The 'Backwards step-wise term deletion routine' section contains output from the 'anova' function in R. This section is useful for investigating the significance of relationships.

The 'Model Post-Hoc' section contains the results from a generalised pairwise Tukey's Honest Significant Differences analysis with a Bonferroni correction. This section will only appear categorical predictor variables are specified. This information is useful for determining which factors are different to each other.

*'Data' tab*

This tab contains the complete data set loadded into the program. This can be useful if users wish to see what the raw values are.