# Assignment 2

**Due date: 16 October 2022**

**This assignment is worth 14% of final mark** for STAT 2402.

---

This assignment involves analysing one dataset. The task is described below. Page 2 provides information on the requirements of the assignment and page 3 contains details of the dataset.

**Task 1.** The data frame `bumpus` is available in the file `bumpus.txt` in the folder Computer Labs/Data.

*Aims of Analysis*: To determine how the survival of the sparrows depends on the morphomological measurements and other characteristics in the data.

---

Please turn over

# Assignment Submission Information

Submission: Assignment 1 *must be submitted the LMS submission link* by *11:59 pm on Sunday 16 October 2022*. A pdf document is preferred, but Word document or html will also be accepted. You are strongly encouraged to use `R, knitr` and LaTeX to prepare your submission.

Please be sure to include your name and student number on the assignment. Late submissions will attract a penalty (5% per day late) unless the unit lecturer is advised as soon as possible of any extenuating circumstances and an extension is granted.

---

Requirements of the assignment:

The assignment task involves the analysis of data. If you have difficulty accessing the data set you should contact the unit co-ordinator *immediately*.

You should hand in a report based on your analysis. The report should be *no longer than four single A4 pages* (excluding any relevant Figures, Tables). *Submissions will be marked down for exceeding this page limit*. The aim of the report is to convey the aims, methodology and results of your data analysis in a concise, readable fashion.

Style of report

It is strongly recommended that you structure your report into sections, along the following lines.

This is a *technical* report, so you should include model equations. However, no `R` code is required.

- *Executive summary:* Briefly introduce the data, the context, the aims of the analysis and the main findings. This should not exceed half a page.
- *Introduction:* Briefly describe the context, the data and the aims of the analysis.
- *Methodology:* Describe the statistical methods that you use. Roughly speaking, you should give enough information to enable a professional statistician to reconstruct what you did.
- *Results:* Describe the results of your analysis. You will include appropriate exploration of data (tables and graphs), and describe your findings. You will also describe the model fitting procedure (such as steps in arriving at the final model). No `R` outputs or code should be given here. Note that this section contains results only, not interpretation.

  Select your graphs and tables carefully. You will only include graphs and tables that highlight a specific aspect of the data or analysis that cannot be equally simply described in words. Any graphs or tables that are included in the report *must be discussed*.
- *Discussion:* Draw conclusions **based on your results**. Discuss any issues in the analysis that may affect the conclusions (e.g. potential weaknesses of your analysis, alternative analyses).
- *References:* A list of references that you have used and cited in your report. Any standard form of referencing and citing may be used, but be consistent in the format that you use.

---

Assessment:

Your report will be assessed as follows.

- Exposition (40%): The report should be well organised, readable, and concise. The report should be readable English, not computer language. The writing should be concise and to-the-point. Graphics should be selected for appropriateness and relevance. Marks will be deducted for exceeding the page limit. Marks will be deducted for writing that is vague, circumlocutious, illogical or self-contradictory.

- Statistical content (60%): marks will be awarded for the correct use of appropriate statistical techniques, and for the correct interpretation of results from these techniques. Marks may also be awarded for the 'completeness' or thoroughness of the analysis (e.g. careful inspection of the data at the beginning). Marks may be deducted for grossly inappropriate selection of technique, gross misinterpretation of the results, failure to notice important features of the data, and grossly unjustifiable conclusions (e.g. speculation, exaggeration).

Warning: The work that you submit must be your sole effort (i.e. not copied from anywhere or anyone else). Penalties for plagiarism will apply according to faculty policy.

## Information on the Data

On 1 February 1898, after a severe storm some English sparrows were brought in to the Anatomy Laboratory of Brown University, Rhode Island. Several morphological measurements were made on the birds. Some of the birds survived while others perished.

Analyse the data to determine the relationship between survival and the other variables in the data. Interpret your results appropriately.

The variables are as follows.

- `ID`: an identifier of bird
- `Sex`: sex of the bird (m = Male, f = Female)
- `Survival`: a binary variable indicating survival status of the bird (T = Survived, F = Died)
- `TotalLength`: measured from tip of the beak to the tip of the tail (mm)
- `AlarExtent`: measured from tip to tip of the extended wings (mm)
- `Weight`: weight of the bird (g)
- `BeakHead`: length of beak and head, measured from tip of the beak to the occiput (mm)
- `Humerus`: length of humerus (inches)
- `Femur`: length of femur (inches)
- `Tibiotarsus`: length of tibiotarsus (inches)
- `SkullWidth`: width of skull measured from the postorbital bone of one side to the postorbital bone of the other (inches)
- `Sternum`: length of keel of sternum (inches)