

## Q1 - Visualization and analysis of the Palmer dataset

The Palmer penguin dataset consists of 344 records of the physical attributes of three species of penguin living on three islands in Antarctica (Table 1) [1].

Attribute	Type	Values in the dataset
species	categorical	Adelie, Chinstrap, Gentoo
island	categorical	Torgersen, Biscoe, Dream
bill length	numerical	32.1mm - 59.6mm
bill depth	numerical	13.1mm - 21.5mm
flipper length	numerical	172mm - 231mm
body mass	numerical	2700g - 6300g
sex	categorical	Male, Female

In this report, consideration is given

Table 1: Attributes of the Palmer penguin dataset

to data cleaning and preparation, the dataset is explored through visualization and analysis is carried out to compare the accuracy performances of a small number of AI approaches.

There is a table at Table 1 and a figure at Fig. ??.

Here is an example of an equation:

$$\pi = 4 \left( 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \dots \right) \quad (1)$$

or

$$\pi = 4 \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \quad (2)$$

where  $\pi$  can be written in line by using \$'s. Here is a vector:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (3)$$

You can write in **bold**, or *italics* or **true type**, often the latter is used for specific commands or libraries in a programming language, as in ‘I used **numpy** v1.23.4 to...’. Notice the use of the left quote symbol found in the top left of the keyboard to get the left quote. There is also blackboard bold often used for things like  $\mathbb{R}$  for real numbers and there is calligraphic for fancy things like  $\mathcal{L}$  but this is becoming increasingly irrelevant to what you are likely to need!