**Data Citizenship Unit, Level 4**

**Practice / Exemplar Assessment on Zoo and Aquarium animals**

**In this assessment you will need to:**

* **download a dataset from a link given to you**
* **carry out simple summaries of the data**
* **interpret the data to answer some questions or solve a problem with the dataset.**
* **choose the best types of visualisation to use to show your interpretation**
* **create two visualisations.**
* **make recommendations based on your interpretation of the dataset.**

**Access the Dataset (1 mark):**

1. Download and open the North American zoo and aquarium species dataset

**Excel Dataset (with data dictionary):** <http://dataed.in/DCL4Eexcel>

**Original Dataset (CSV without data dictionary):** <http://dataed.in/DCL4Edataset>

**Understanding the dataset (2 marks):**

2. What does each column contain?

3. Do a data quality check and find out if there any data quality issues

**Manipulations (10 marks):**

4. You will not need data on confidence limits (the CI columns). Remove all of the columns that are not required for your analysis.

5. Remove the rows of poor quality data where there is a ‘yes’ in the Male Data Deficient or Female Data Deficient columns. Then delete these two columns as they should be no data stored here now.

6. Find out if males or females of these species live longer on average. Calculate the difference between the Male and Female MLE (median life expectancies) columns. Use conditional formatting to colour the cells differently depending on the values.

7. The TaxonClass variable shows the type of species. Create another data table that summarises the sample size and MLE data about each unique Taxonclass value.

**Visualisations (9 marks):**

8. What is the most popular species? What are the least popular? Create a chart with the most popular or the least popular creatures in zoos and aquariums.

9. Create a chart showing which TaxonClass species type lives longest on average?

10. Create a graph showing the species that have the greatest difference between the male and female life expectancies

**Interpretation (3 marks):**

11. What type of graph did you choose to show the most common species? Why did you choose this?

12. What type of graph did you choose to show the average life expectancies of different types of species (taxonomical classes)? Why did you choose this type of graph?

13. What type of graph did you choose to show the differences between male and female average life expectancies? Why did you choose this?

Assessment Total = 25 marks