

Assignment 2

Due at 11:59pm on Sunday 21st of November through Moodle.

Provide the Python script, and the final dataset you produced. Code should be well commented throughout.

An Excel file (WING.xlsx) containing data on the shape of butterfly wings is on Moodle. The information is spread over multiple sheets in the Excel file. There are details on the data at the bottom of each sheet. Details on the experiment are given here:

<https://datadryad.org/stash/dataset/doi:10.5061/dryad.1s15q>.

Using Python, organise the data into a single dataset containing any information you feel is relevant. The final dataset produced should follow best practices for data management. Write your final dataset to a csv file, and upload this file as part of your submission.

(70 marks)

Complete the first two chapters from the 'Cleaning Data in Python' course on Data Camp. The chapters are called 'Common data problems' and 'Text and categorical data problems'. You will be given 10 marks for completing the chapters. The remaining 20 marks will be given based on the number of XP points achieved out of 1750 for the two chapters.

(30 marks)