**Computational Biology, Programming and Practice**

**Name:** Areeba Salman

**Student ID:** 20697113

**Course Code:** LIFE4138

**Date:** 28/01/25

**Coursework Part 2:** Python

**Identify the heaviest pumpkin grown in any of the competitions – what variety was it? Where was it from and when was it grown?**

* Heaviest Pumpkin Weight: 2705.614652628638
* Variety: 'Golden Hubbard’
* Location: Houston, Florida, Mexico
* Year: 2454

**Plot the relationship between the estimated weight and actual weight of the pumpkins. This can be in either lbs or kg but remember to ensure that both axes are in the same units. Colour the points in your plot based on the weight\_class column that you created in question 4. Ensure the plot has appropriate axis labels and is clear and well presented. Save this plot to your computer.**

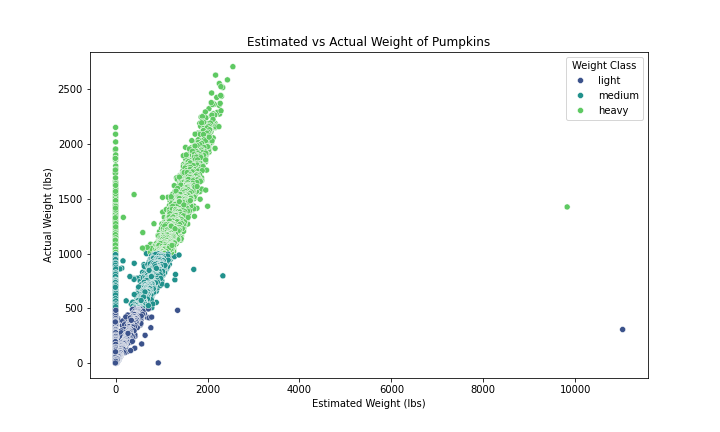


Figure 1 shows the relationship between the estimated weight and the actual weight of pumpkins.

**Summarise your filtered dataset from question 6:**

* 1. **Identify the mean weight of pumpkins for each of your three countries. Which country had the highest mean pumpkin weight?**
  2. **Identify the mean weight for each variety of pumpkin for each of your countries. Which variety in which country had the lowest mean weight?**

|  |  |
| --- | --- |
| **Country** | **Mean Weight** |
| China | 774.161283 |
| UK | 785.512086 |
| USA | 774.187512 |

UK had the highest mean pumpkin weight.



Kakai in USA had the lowest mean weight.

**Using your filtered data set from question 6, plot pumpkin weight distributions (in either lbs or kg) for your three countries as a boxplot. Ensure the plot has appropriate axis labels and is clear and well presented. Save this boxplot to your computer.**

A screenshot of a computer screen

Description automatically generated

Figure 2 explains the boxplot of pumpkin weight distributions for China, UK and USA.

**Redraw your plot from question 8 as a facet plot showing the data from each variety of pumpkin as a separate sub-plot. Save this plot to your computer.**

A diagram of different colored squares

Description automatically generated

Figure 3 shows the redrawn boxplot for China, UK and USA with their four different varieties.