

Statistics assignment 2024

Sean van der Merwe

March 2024

Instructions

In this stinky experiment, the shelf life of brands of sausage are being compared across Gauteng. The researchers think that the proportion of sodium affects the shelf life, and are curious as to whether there are brand differences after controlling for sodium proportion and source. They take samples from two shops on each of two sides of Gauteng. At each of these 4 shops they buy two packets of each brand, a month apart. For each sausage they take a sample from each end and place all the pieces in separate sealed containers. On the best-before date of each sausage, those sausage pieces are analysed and given contamination scores (from 0 to 100).

1. Find the sheet in the given Excel file that corresponds with your student number or email stub and load the data in a statistical analysis program. [5]
2. Illustrate the full data set any way you prefer that allows you to see and understand the patterns in the data. Key variables to include in the plots are contamination score, sodium, and brand. [10]
3. Describe the patterns you see. Does it look like the effect of sodium might differ between brands (interaction)? [10]
4. Draw a box plot of the scores, split by brand and month. Which brand has the lowest (best) median, and when? [10]
5. Do a Kruskal-Wallis test for whether there is a statistically significant difference between the brands, ignoring all other effects. Does the result invalidate or support (or neither) the conclusion from the box plot? [10]
6. Perform a formal regression and/or ANCOVA of score on the explanatory variables you think need to be studied or controlled for, to allow the researchers to test the null hypothesis: that there is no systematic effect of sodium proportion or brand on the shelf life. Justify your choices and interpret your results in detail. [30]
7. Looking now at all the plots, tables, and statistics, make a statement as to what you think the results of the experiment are (if anything). [7]
8. Also explain what you personally learned from this assignment. [8]

Please submit a Word or PDF file with the code, output, and explanations neatly linked and clearly identified. Also attach supporting files as appropriate (*e.g. Rmd or Excel files that you created*). These must be submitted by emailing to vandermerwes@ufs.ac.za at least 24 hours prior to the follow up discussion session. [10]