Statistical Consulting For Biology

Nathan Hughes

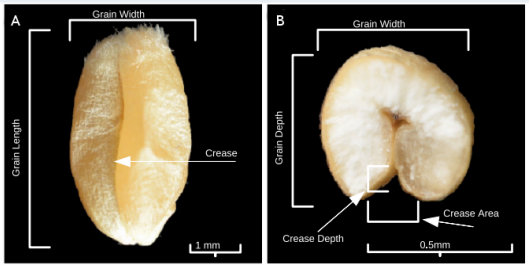
April 30, 2018

Outline

What was the data set

We focused on looking at a CT image dataset

- Its output (morphometric) data.
- This was part of a Wheat (*Triticum aestivum*) drought experiment.



Experiment Details

Drought Experiment Parameters

Normal Watering	Droughted
25C	25C
30C	30C
35C	35C

Working as a group

Insight

- I was already very familiar with the data
- Which meant I was too close to it to be "open-minded"

Benefits of discussing the data

Explaining teaches

- Going through and explaining data makes you re-evaluate it
- Stops you taking data for granted!
- Helps you recognise your own mistakes and misunderstandings

Analysis is a living and evolving process

- "Individuals and interactions over processes and tools"
  - Kent Beck *et al.*

Group discussion

Suggestions

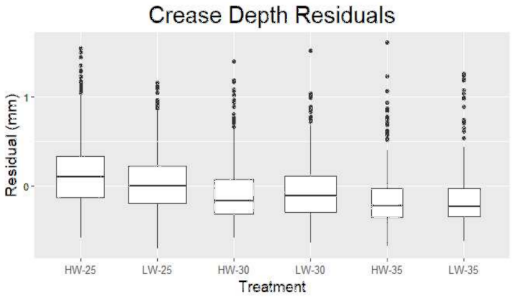
- Our group suggested that Principal Component Analysis would be well suited to the data
  - It may provide a method of looking at which traits are most effected by the experiment treatments
- Another idea was using MANOVA to explore mean differences in treatments

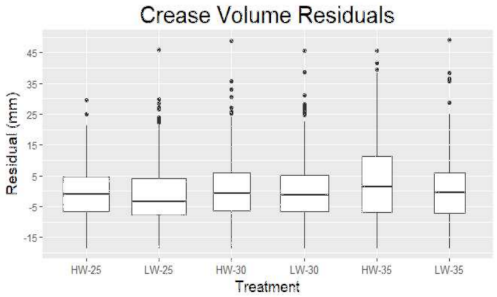
Presenting the data

Interest in how best you could convey information

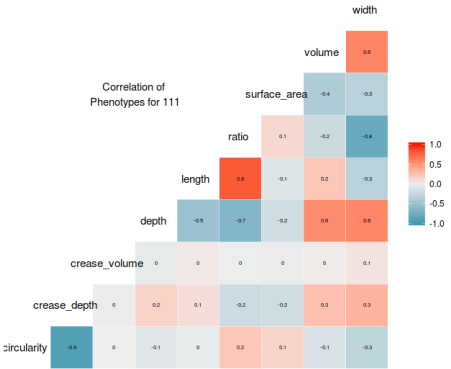
- We looked at the data from a visualisation point of view
- How you could show details on morphometric phenotypes in concise charts, plots or diagrams
- We knew that some of the data couldn't be used, but not *why*

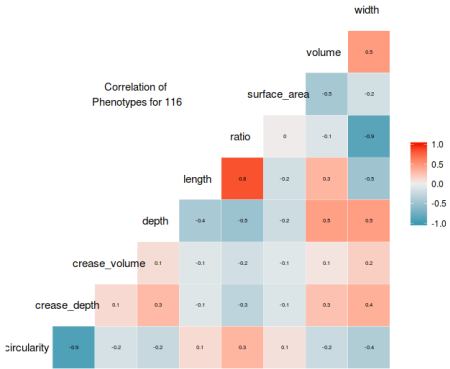
Showing the errors in the data





Showing relationships of traits





Feedback from audience

Suggested Models

- Multiple Regressions
- Simple linear models
- More suggestions to use MANOVA

Similar with the group's initial ideas

- Which shows how different fields instinctively tackle/solve problems

Thanks for listening!

Any Questions?