Capstone Data Analysis Project – Preliminary Data Analysis
Biostatistics

Name: Sara St. Clair Score = 22.5 /25

Submitted on time? (●) Y ( ) N

GENERAL REQUIREMENTS (10 POINTS):

Element	Points	Score	Feedback
Effective git/GitHub	1	1	
Well-organized	1	1	
Strong commentary outside of code chunks	3	3	
Effective use of comments within code chunks	2	2	
Code provides correct values and reduces "human intervention"	2	2	
Link on Canvas	1	1	

STATISTICAL ANALYSES (15 POINTS):

☐ Took initiative to learn new methods as appropriate

☑ Generally followed the our workflow:

Plot -> Guess -> Create model -> Check assumptions -> Interpret -> Final plot

Statistical analysis 1:

Question: Are weed lifeform and biovigalence status independent from each other?

Workflow checklist

☑ 1. Plot data

☐ 2. Guess relationships

☑ 3. Create model: Chi-square test of independence

☑ Correct model?

☑ 4. Check model assumptions, if needed

☑ 5. Interpret model

☑ Interpretation is correct

☑ 6. Replot

☑ 7. Clear results statement

☑ In prose

☑ Outside of code chunk

Statistical analysis 2:

Question: Are weeds in different habitats occuring independent of biogeographic realm?

Workflow checklist

- [x] 1. Plot data
- [ ] 2. Guess relationships
- [x] 3. Create model: chi-square
  - [x] Correct model?
- [x] 4. Check model assumptions, if needed
- [x] 5. Interpret model
  - [x] Interpretation is correct
- [x] 6. Replot
- [x] 7. Clear results statement
  - [x] In prose
  - [x] Outside of code chunk

Statistical analysis 3:

Question: Is there an association between lifeform and conservation status (Red_UK)?

Workflow checklist

- [x] 1. Plot data
- [ ] 2. Guess relationships
- [x] 3. Create model: chi square
  - [x] Correct model?
- [x] 4. Check model assumptions, if needed
- [x] 5. Interpret model
  - [x] Interpretation is correct
- [x] 6. Replot
- [x] 7. Clear results statement
  - [x] In prose
  - [x] Outside of code chunk

Additional feedback

Overall, good, though as you know, rather repetitive. To extend your learning, I'd like you to learn how to conduct post-hoc tests with this sort of analysis and I've provided a link with information.

Please carefully read and apply the feedback in the prelim-analysis-feedback file.

You'll note that I didn't check any of the boxes for "guess relationships" - I'd like to see more biolgical thinking a priori about which groups you might expect to see have certain relationships as part of your alternative hypothesis statements.