Midterm Project

1 Overview

In the first half of this course, we have explored methods for cleaning and analyzing data to extract meaningful insights. This midterm project challenges you to apply these techniques in an open-ended context by analyzing a provided dataset. Unlike structured labs, this project requires you to independently determine the best approach for addressing the research questions. However, you are encouraged to ask for clarification when needed.

2 Background

Conservation ecologists monitor species populations to assess ecosystem health and the effectiveness of conservation strategies. Population trends can reveal environmental changes, emerging threats, or successful interventions.

There are several methods for estimating population sizes. Traditional techniques include mark-recapture, where animals are tagged and re-identified, and camera-trap surveys, which record sightings of individuals. However, Dr. Rachel Reid is really rolling up her sleeves and investigating a third method: analyzing biological remnants such as scat and carcasses.

In a 2015 study, Dr. Reid investigated whether morphological (e.g., size, shape), biogeochemical (e.g., chemical composition), and contextual (e.g., location) traits could distinguish between bobcat, coyote, and fox scat. Morphological traits are the most cost-effective for field identification, while biogeochemical traits require laboratory analysis.

You are provided with Dr. Reid's dataset. Your goal is to analyze this data to evaluate whether specific morphological or biogeochemical traits can reliably differentiate species. Additionally, you will research biological and ecological factors that may explain observed patterns.

1. Which (if any) morphological and biogeochemical traits distinguish between originating species of the scat samples?

2. Why do you think those traits differ across species?

3 Instructions

You will organize your results as a paper with the following sections. Please also see section description for tasks that should be completed.

Part I: Introduction and Background

1. Read the provided article by Dr. Reid.
2. Research the three species (bobcat, coyote, and fox), focusing on their diet, habitat, and physical characteristics. Summarize your sources in a bibliography.
3. Find an article reviewing population estimation methods for (conservation) ecology and their respective strengths and weaknesses.
4. Write an introduction. You should specifically discuss the following:
    - The ecological significance of tracking these species.
    - Why morphological and biogeochemical traits might serve as useful identifiers. Please make sure to discuss each species separatly.
    - Your hypotheses based on the provided research questions.

Part II: Methods

1. Data Preprocessing:
    - Describe how you cleaned the dataset, including handling of missing values, data type conversions, and any transformations applied.
    - Justify any decisions made about outlier treatment (removal, transformation, or retention).
2. Statistical and Visualization Techniques::
    - Describe which statistical tests you used to assess differences between species and why those tests were appropriate.
    - Discuss the types of visualizations you used and reason for their use.

Part III: Analysis

1. You should load and clean the data set. All columns must be converted to the correct types.
2. Create a table that categorizes each variable in the data set as morphological, biogeochemical, contextual, or not a trait (e.g., species). Include this table in your results section

3. Use the visualization and statistical testing techniques you've learned to evaluate the relationship between the morphological and biogeochemical traits and species. (Do not analyze contextual traits.) Determine which traits can be used to distinguish at least one species from the rest. Provide appropriate visualization to support your work and findings.

Part VI: Discussion and Interpretation

1. Make a table of all traits that separate at least one species from the rest. Use descriptive statistics (e.g., mean, mode, etc.) to characterize the values for each species. E.g., how is bobcat scat differ from grey wolf scat?

2. Using what you learned from your research, explain how the traits with differences relate to differences in the biology of the three species.

3. Explain why you think that predictive morphological and biogeochemical traits might be more useful than contextual traits to ecologists.

4 Submission and grading

- **You are not allowed to work with or discuss the midterm project with anyone.**

- Write up your report in a separate document. Include any plots you've generated as appropriate – remember that your goal is to make an argument addressing the research question and hypotheses, so include plots for that purpose. Aim for roughly three pages of text plus figures. Use complete sentences. Convert this to pdf and submit it. Submit your .ipynb file as well.

- You will be evaluated on choosing the appropriate techniques for analyses, accuracy of the analyses, effort and completeness in interpreting and discussion the data, and presentation and formatting of plots (e.g., axis labels). A full rubric will be available in canvas.

- NO OTHER SUBMISSION TYPES WILL BE ACCEPTED.

- **Late policy**: 5% of total points deducted per day for three days – after that no submissions allowed.