Individual Assignment Overview and Proposed Workflow

Noah Holland

2023-03-27

Background

The data set I'm using is data concerning chigger mites collected from North Carolina state parks in the last few years and their possible role as a vector of the Spotted Fever family of diseases caused by Rickettsia bacteria. The data was collected to test the hypothesis that free living chigger mites can act as hosts for Rickettsia bacteria. To do that free living chiggers were collected from different North Carolina state parks and tested for Rickettsia bacteria using PCR. Data was recorded on number of chigger mites collected from each site, Rickettsia infection status, and habitat collected from (grassland, grassland edge, woodland).

Workflow/Pseudo-code

- Tidy data, removing "NA"s and inconclusive test results
- Calculate incidence rate of Rickettsia infection in the mites by site
- Make heat maps mapping what sites have a greater population of chiggers and calculated incidence rate
- Make a bar chart comparing number of different species of Rickettsia infecting chigger mites
- Make a bar chart comparing Rickettsia incidence rate in different habitats
- Make histogram comparing number of Rickettsia positive mites by the month they were collected