REPORT

Exploratory Data Analysis of data-"fungal epidemics in Daphnia populations".

BACKGROUND: Data on algal resources, Daphnia traits, and epidemic size in 12 lakes.

Data on algal resources (log-transformed particulate carbon, nitrogen, and phosphorus), Daphnia traits (size, reproduction, and parasite load), and epidemic size (integrated prevalence) calculated from weekly sampling of 12 lakes in Indiana, USA.

Insights:

- This dataset has 12 observations of 10 variables as Lake, ln. N, ln. C, ln. P (resources), Edible C.P molar ratio, SizeUninf, AERUninf, SporeYield, IntPrev, SizeInf.

- Bar Plots: Distribution of 9 variables across 12 lakes:
  1. Epidemic Size
     The lake Goodman has highest Integrated Prevalence (Epidemic Size) of value around 20. Fig 1(a)
  2. Spore Yield: It is more in the Lake Goodman with a value more than 15000. Fig 1(b)
  3. Infected Host Size
     Almost same in all the 12 lakes. Fig 1(c)
  4. Aerial host density(uninfected) is highest in the Scott Lake. Fig 1(d)

- Scatter plots:
  5. Epidemic size vs Resource Availability:
     Epidemics grew larger in lakes with more carbon and and phosphorus. Fig 2(a)

6. Resource availability vs Size of Uninfected Host: With increase in resource availability, the size of uninfected host increased in all the lakes. Fig 2(b)
7. Resource availability vs Spore yield: The production of fungal spores within infected *Daphnia* increased with increased availability of resources. Fig 2(c)
8. Epidemics vs Aerial density of Uninfected Hosts:
   The epidemics grew largest in lakes with greater host density. Fig 2(d)
9. Epidemics vs Size of Uninfected Host: Fig 2(e)

GGplot Visualization:

1. Epidemic size vs Resource availability:
   The lake Goodman has higher ln.C and higher Integrated Prevalence, also with higher ln.P and Integrated prevalence. Fig 3(a,b)
2. Epidemic size vs Spore yield: The lake with higher Integrated prevalence and higher Spore yield is Goodman Lake. Fig 3(c).

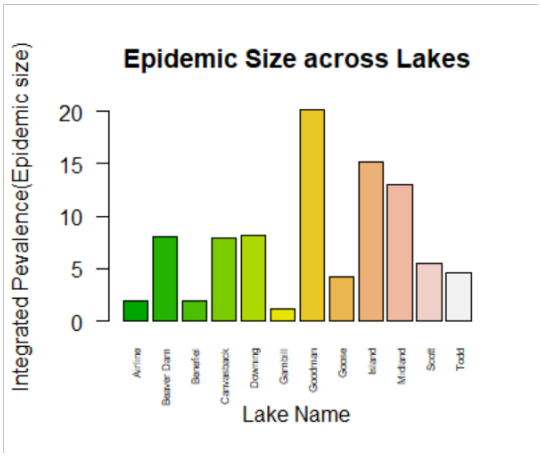


Fig 1(a)

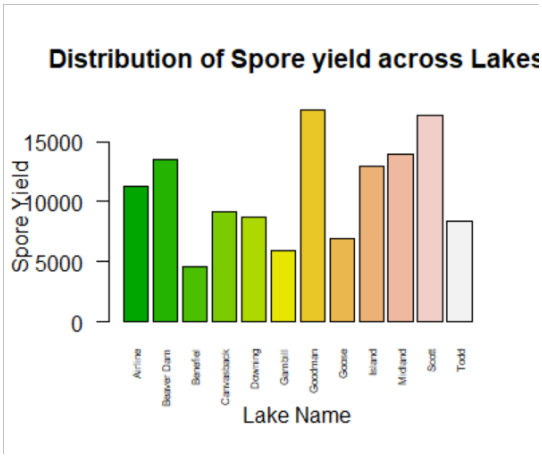


Fig 1(b)

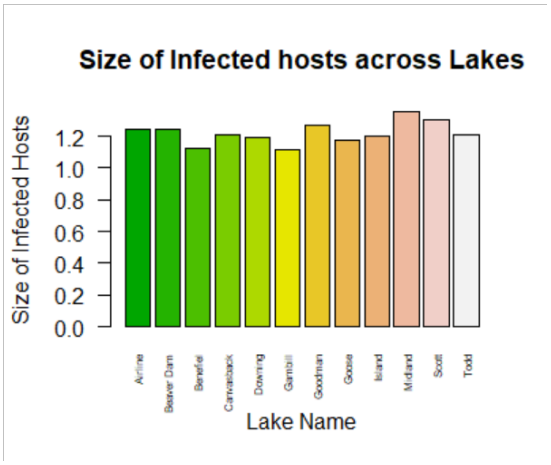


Fig 1(c)

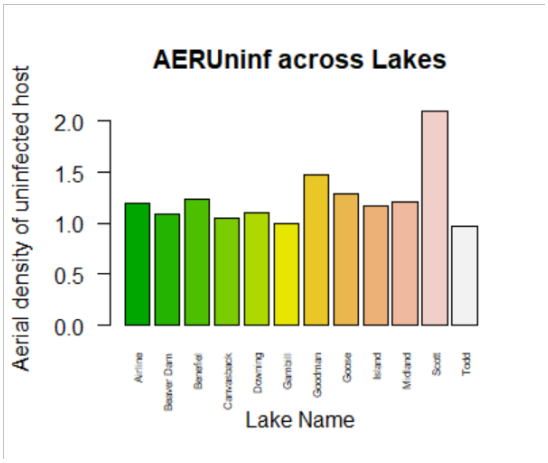


Fig 1(d)

- Scatter plots:

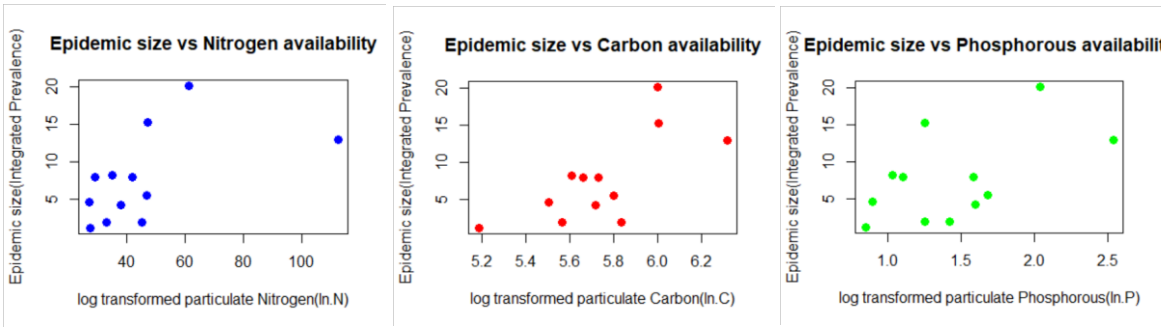


Fig 2 (a)

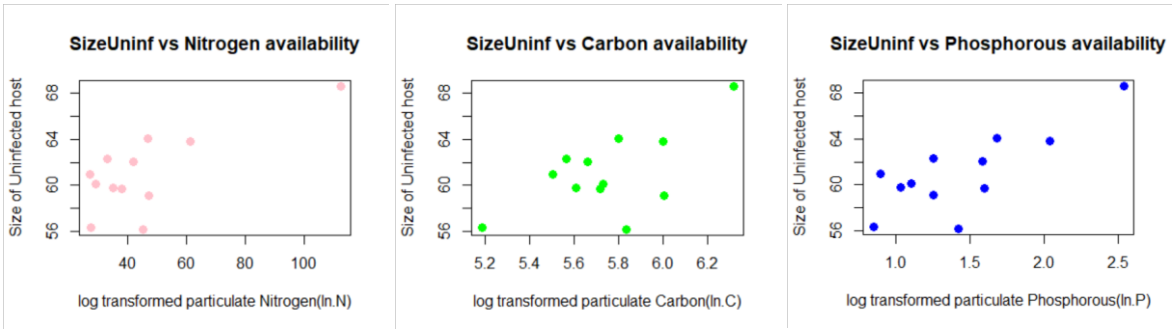


Fig 2(b)

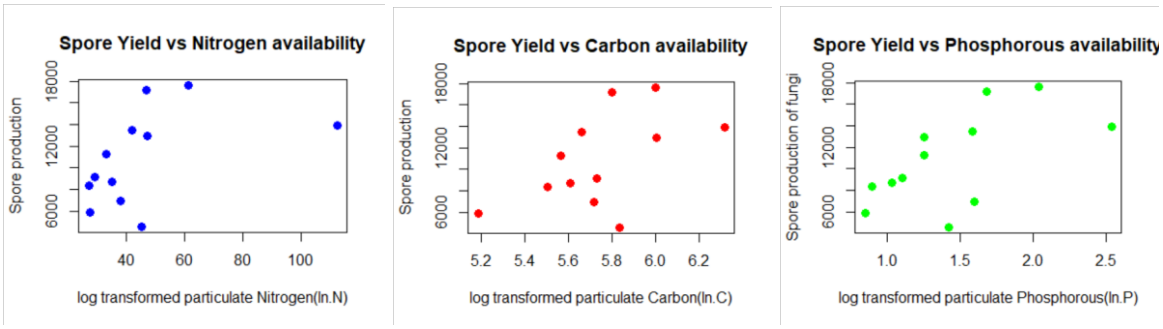


Fig 2(c)

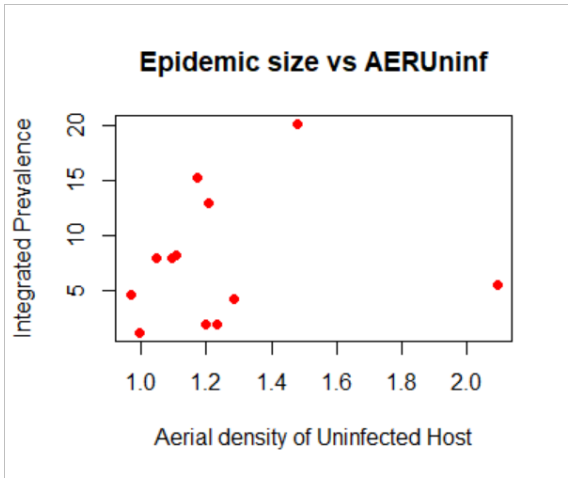


Fig 2(d)

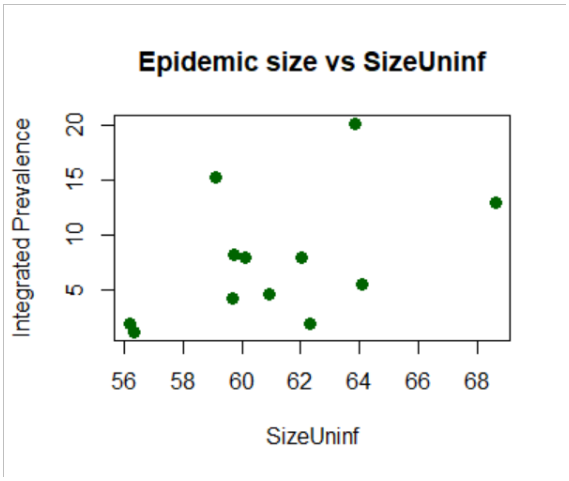
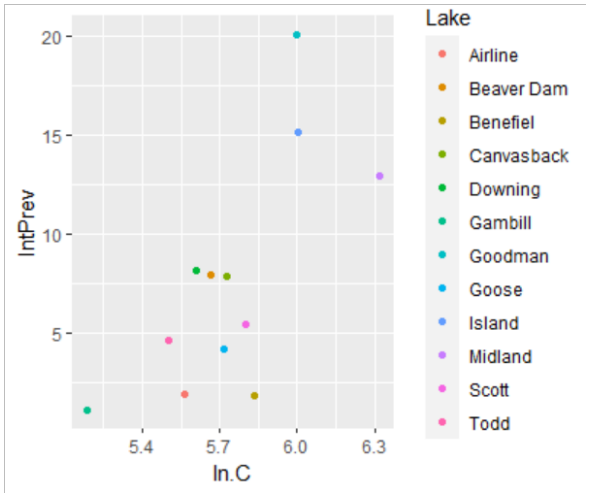
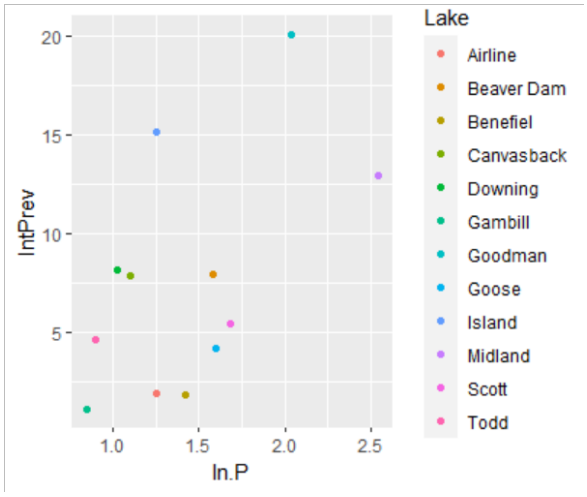
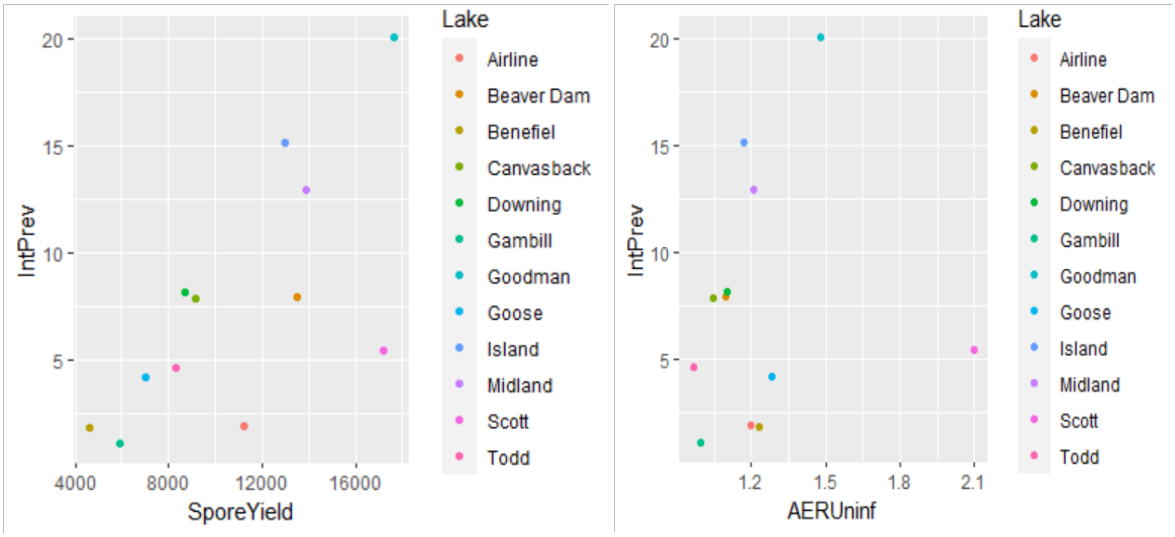


Fig 2(e)

Ggplot Visualization Plots:






IntPrev
SporeYield
20
15
10
5
4000
8000
12000
16000
Lake
Airline
Beaver Dam
Benefiel
Canvasback
Downing
Gambill
Goodman
Goose
Island
Midland
Scott
Todd
IntPrev
AERUninf
20
15
10
5
1.2
1.5
1.8
2.1
Lake
Airline
Beaver Dam
Benefiel
Canvasback
Downing
Gambill
Goodman
Goose
Island
Midland
Scott
Todd