AP Computer Science Principles

Species versus Area Project

One of the most fundamental ecological relationships is that as the area of a region increases, so does the number of different species encountered. While this makes sense and may even seem obvious, this observation seems to have first occurred late in the eighteenth century and slowly taken hold in the nineteenth century. During this period, naturalists such as Alfred Wallace and Charles Darwin accompanied sailing expeditions to islands around the globe. In the process of recording and collecting what were new and exotic species to Europeans, certain ecological patterns and trends slowly became apparent. Johann Rheinhold Forster, the naturalist on Captain Cook's second voyage to the South Pacific (1772), seems to be the first to have noticed this particular point.

Islands only produce a greater or less number of species as their circumference is more or less extensive.

Simply put, the number of species increases with area. A less obvious insight would occur later to others making careful collections of data: **the increase in species occurs at a decreasing rate.** This species-area relation may be the oldest ecological pattern to be recognized; H. C. Watson described a species-area curve for plant species in Britain in 1859 and deCandolle produced a similar study in 1855. Beginning with a small plot in county Surrey, Watson identified the plant species present in ever increasing areas of Great Britain. The general pattern is one of increasing species diversity with increasing area sampled. The first mathematical description of the species-area relationship was proposed by Arrhenius in 1920 and modified by Gleason in 1922.

The association of increased area with an increasing number of species at a declining rate has been tested numerous times. It persists over areas both small and large and with animals as well as plants.
http://math.hws.edu/~mitchell/SpeciesArea/speciesAreaText.html

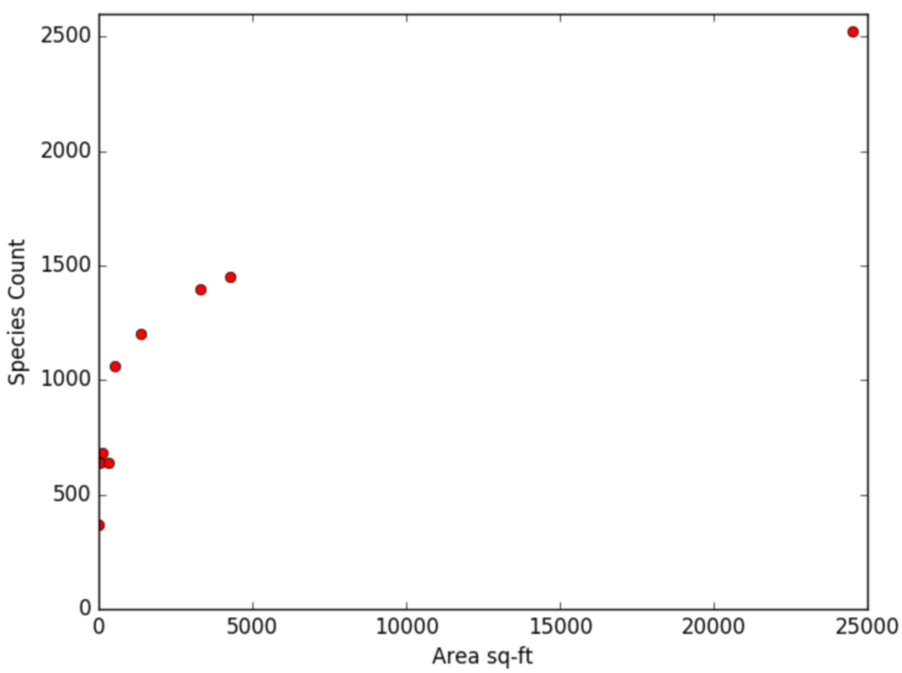
The data below give the number of endemic vascular plant species in mainland coastal areas (mi^2) of California at or above 33 degrees latitude.

Table 1: Data from [Johnson, Mason, and Raven 1968].

Location	Area	Species
Tiburon Peninsula	5.9	370
San Francisco	45	640
Santa Barbara area	110	680
Santa Monica Mountains	320	640
Marin County	529	1060
Santa Cruz Mountains	1386	1200
Monterey County	3324	1400
San Diego County	4260	1450
California Coast	24520	2525

Plot this data using the Python program below:

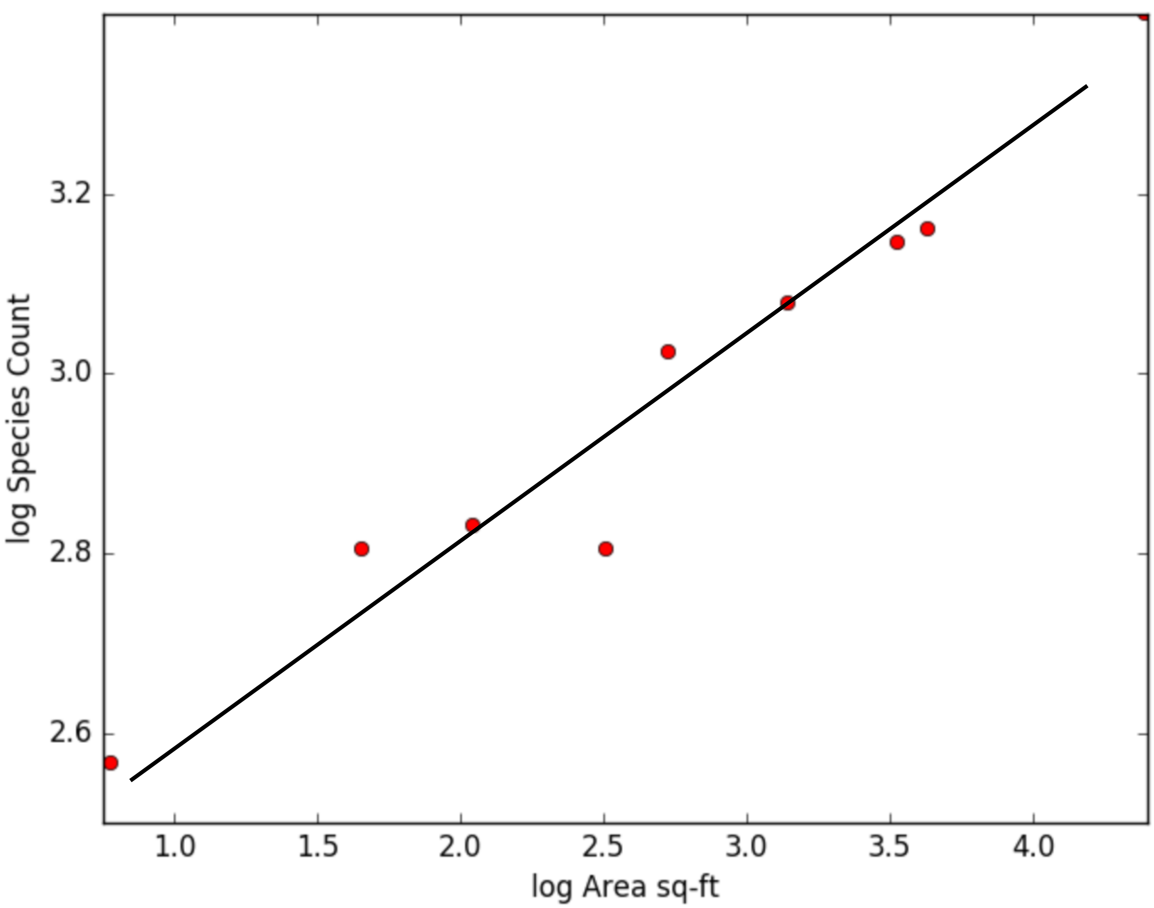
```
import matplotlib.pyplot as plt

plt.xlabel('Area sq-ft')

plt.ylabel('Species Count')

xdata = [5.9,45,110,320,529,1386,3324,4260,24520]

ydata = [370,640,680,640,1060, 1200,1400,1450,2525]

plt.plot(xdata,ydata ,'ro')

plt.axis([0, 25000, 0, 2600])

plt.show()
```



The data seems to lie on a curve. Let's find this curve. To do we need to linearize the data by taking the log of the x and y data. Replot the data using the Python program below:

```
import matplotlib.pyplot as plt

import math

plt.xlabel('log Area sq-ft')

plt.ylabel('log Species Count')

xdata = [5.9,45,110,320,529,1386,3324,4260,24520]

ydata = [370,640,680,640,1060, 1200,1400,1450,2525]
```

```
log_x = []
log_y = []
for item in xdata:
  log_x.append(math.log10(item))
for item in ydata:
  log_y.append(math.log10(item))
plt.plot(log_x,log_y ,'ro')
plt.axis([.75, 4.4, 2.5, 3.4])
plt.show()
```



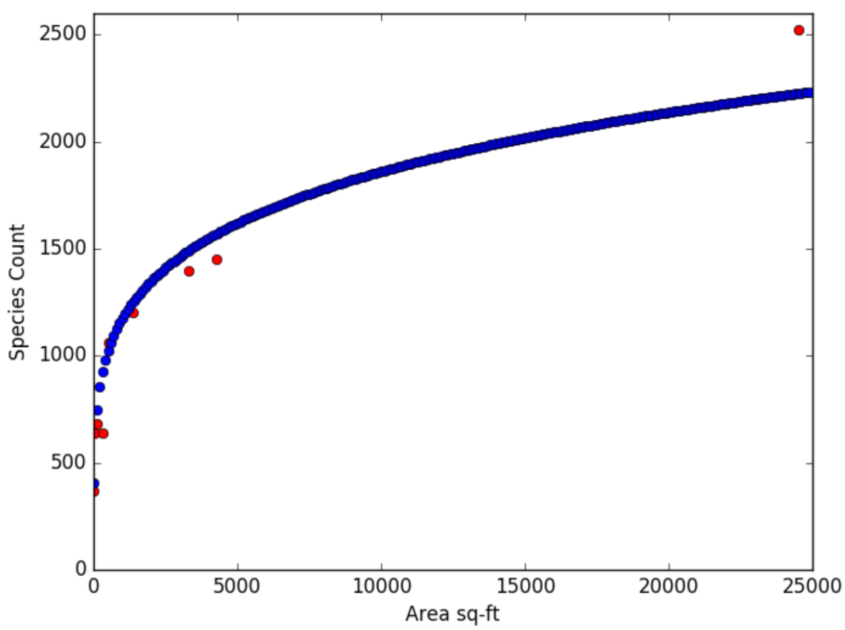
Draw a line that best fits the data and estimate its equation. We'll use the points (0, 2.47) and (3.2, 3.1) yields

$\mathbf{\log \mathit{species} = .2 \log \mathit{area} + 2.47 => \mathit{species} = 295\mathit{A}^{.2}}$

This is the so called **Power Function Model.**

Add this graph to the points with the following Python code:

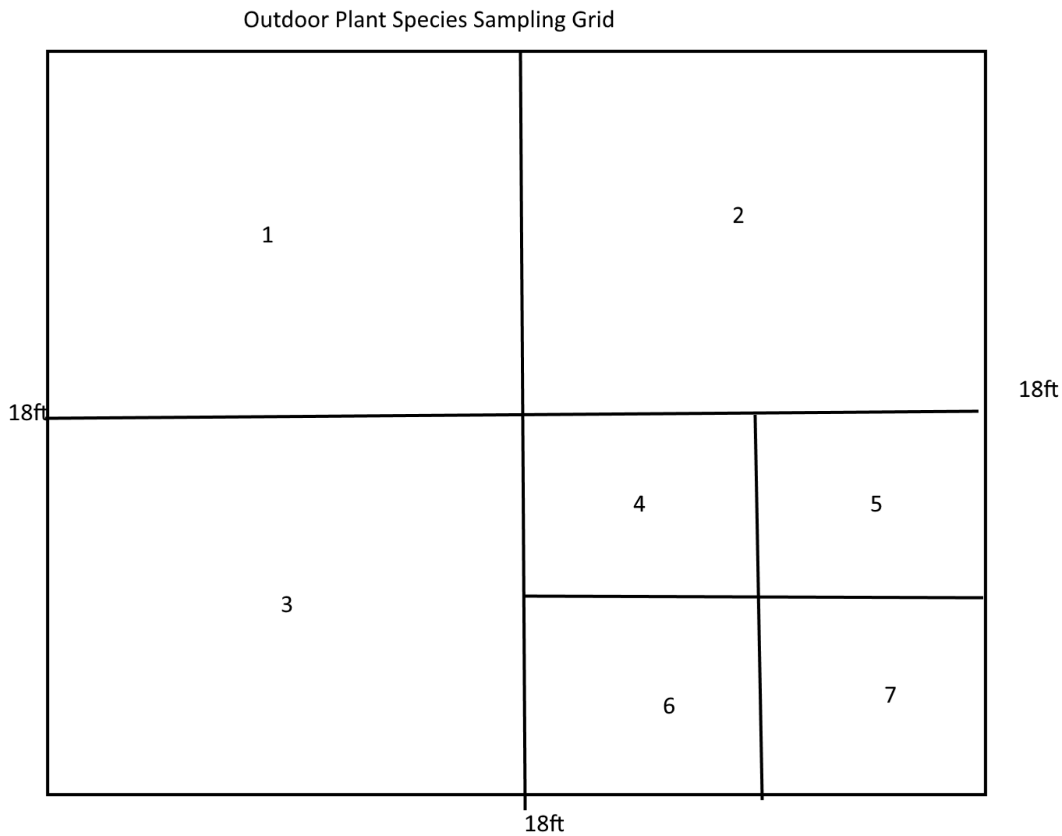
```
import matplotlib.pyplot as plt

import math

plt.xlabel('log Area sq-ft')

plt.ylabel('log Species Count')

xdata = [5.9,45,110,320,529,1386,3324,4260,24520]

ydata = [370,640,680,640,1060, 1200,1400,1450,2525]

x_data = []

y_data = []

for i in range(5,25000,100):

  x_data.append(i)

  y_data.append(295*math.pow(i,.2))

plt.plot(xdata,ydata ,'ro')

plt.plot(x_data,y_data,'bo')

plt.axis([0, 25000, 0, 2600])

plt.show()
```



The curve shows an increase of species with area, but at a decreasing rate. This

Now we will go outside…

Instructions Record the plant species and grid number on the plant library sheet.



- Based on the grid layout depicted above, students retrieve and record plant species from the Notre Dame Prep, Towson Maryland **NO MOW ZONE.**

- Students retrieving plants from the gridded region.

- Students recording plant species on the plant library board.

- Aside from the flora, fauna (a dead vole) was also discovered in the gridded region.

- Students and the completed plant library board.

Data and Analysis

Area sq feet	Number of plant species
20.25	5
81	15
324	24

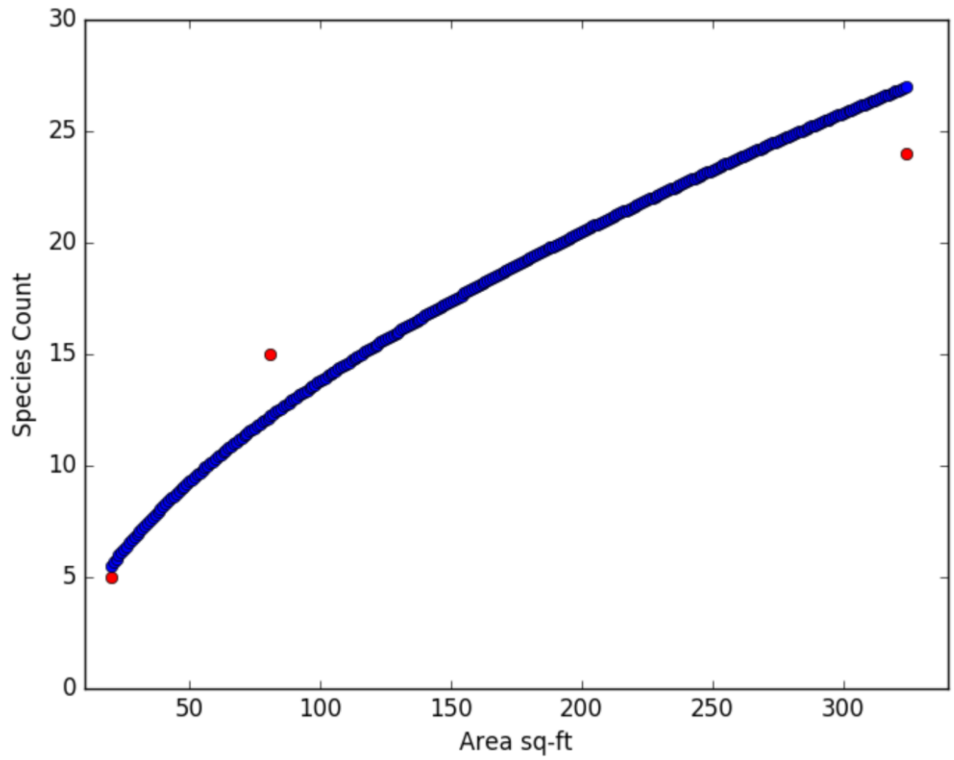
Fitting a curve to the data, as what was done for the data on coastal regions in California yields.

$$\boldsymbol{species = Area^{.57}}$$

This curve approximately a square root function, which is usually the model proposed when modeling species count versus areas. The coefficient of determination is .95. In other words, 95% of the variation in species count can be described by a power model in terms of area.

Python code depicting the plant data and fitted model

```
import matplotlib.pyplot as plt

import math

plt.xlabel('Area sq-ft')

plt.ylabel('Species Count')

xdata = [20.25,81,324]

ydata = [5,15,24]

x_data = []

y_data = []

for i in range(20,325,1):

  x_data.append(i)

  y_data.append(math.pow(i,.57))

plt.plot(xdata,ydata ,'ro')

plt.plot(x_data,y_data,'bo')

plt.axis([10, 340, 0, 30])

plt.show()
```



Summary

The **NO MOW ZONE** here at Notre Dame Prep can be thought of as an island of plant diversity, which makes it an ideal space to study the relationship between species count and area.

Many empirical studies have shown that that regardless of census design and habitat type, species–area relationships are often fit with a simple function. Here we used the commonly proposed power function model, which appears to fit our data well.
https://en.wikipedia.org/wiki/Species%E2%80%93area_relationship

Given the time, space, and person power, it would be interesting to conduct a larger area study. Other questions to pursue might involve the species counts versus latitude or rainfall totals.