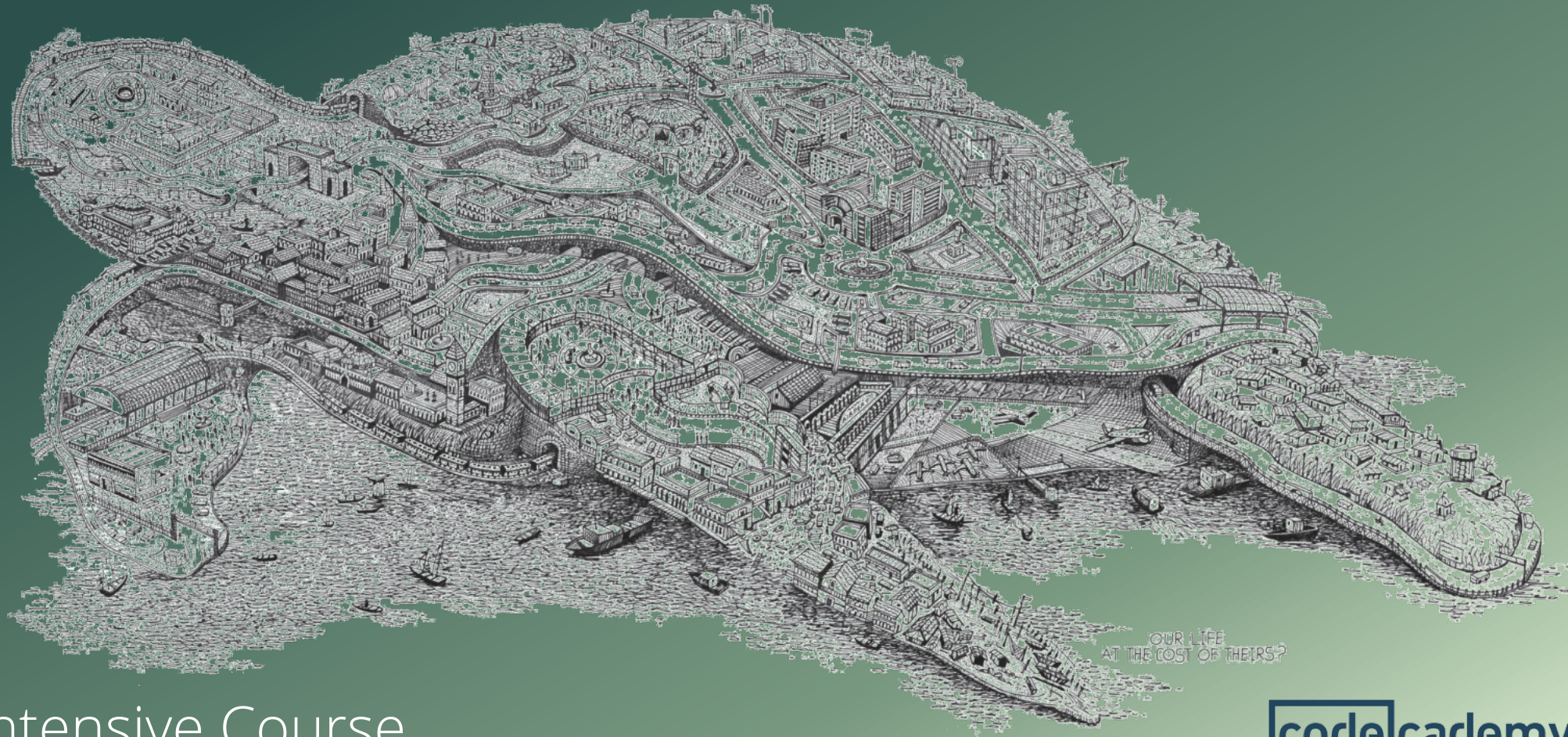


Capstone Option 2: Biodiversity for the National Parks



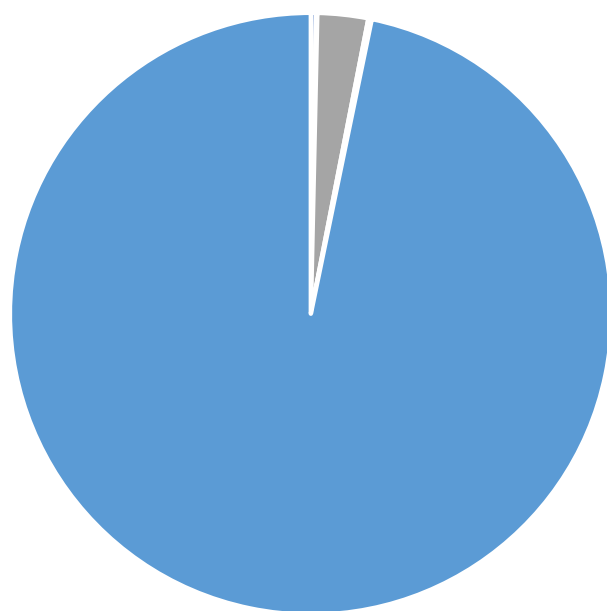
17 January 2018
Data Analysis Intensive Course

Data Descriptives – First impressions

Disclaimer: The attached .py file contains the copy pasted code which I produced online on codecademy, since I had trouble running pip on my machine.

A section describing the data in species_info.csv:

1. As the data shows, a large majority (5363) of the species are currently in no need for intervention



■ Endangered ■ In Recovery
■ Species of Concern ■ Threatened
■ No Intervention

2. There are 7 different types of species: 'Mammal' 'Bird' 'Reptile' 'Amphibian' 'Fish' 'Vascular Plant' 'Nonvascular Plant'

3. As can be seen below, Nonv. Plants (probably since they are plants) and Reptiles seem to be the least protected with some significance at first sight

Category	Not Protected	Protected	Percent Protected
Amphibian	72	7	88.6%
Bird	413	75	15.4%
Fish	115	11	8.7%
Mammal	146	30	17%
Nonvascular Plant	328	5	1.5%
Reptile	73	5	6.4%
Vascular Plant	4216	46	10.8%

More Data Descriptives and Recommendations

Park Name	Observations
Bryce National Park	250
Great Smoky Mountains National Park	149
Yellowstone National Park	507
Yosemite	282

4. Interestingly, sheep is also a common part of a plants name apparently. On the left you can see the number of sheep sightings per national park after excluding all plants.

5. While the *Ovis aries* is not endangered, apparently the *Ovis canadensis* (Bighorn Sheep) needs to be protected (*Species of Concern!*)

A recommendation for conservationists concerned about endangered species, based on your significance calculations

1 – All Species: I recommend that conservationists start paying stronger attention to reptiles, since these are significantly more endangered compared to other species, as proven by the significance tests.

2 – Sheep: To really see whether they were able to decrease the foot and mouth disease significantly, researchers should spend some time observing to for each resort, namely 1,75 weeks in Yellowstone and 3.56 weeks in Bryce.

Significance Calculations and Sample Sizing

A section describing the significance calculations that you did for endangered status between different categories of species:

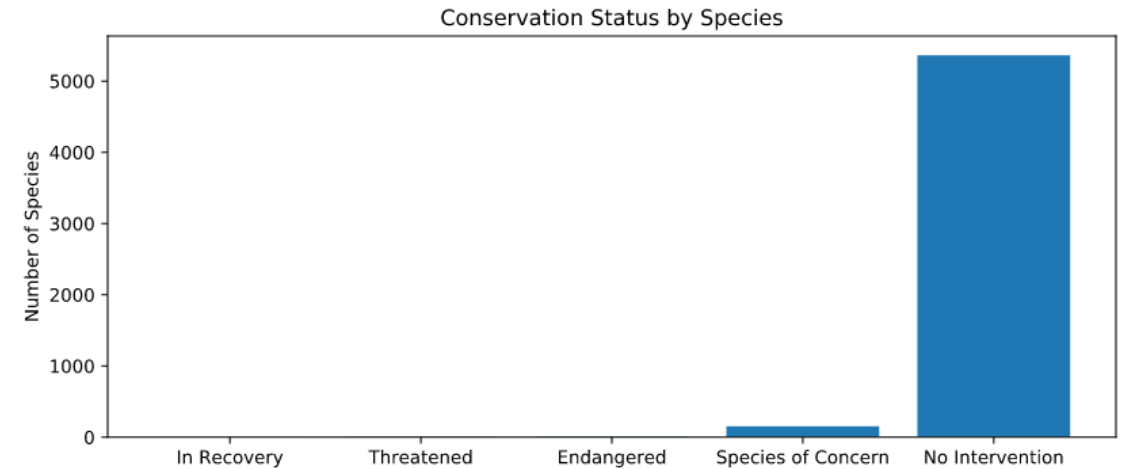
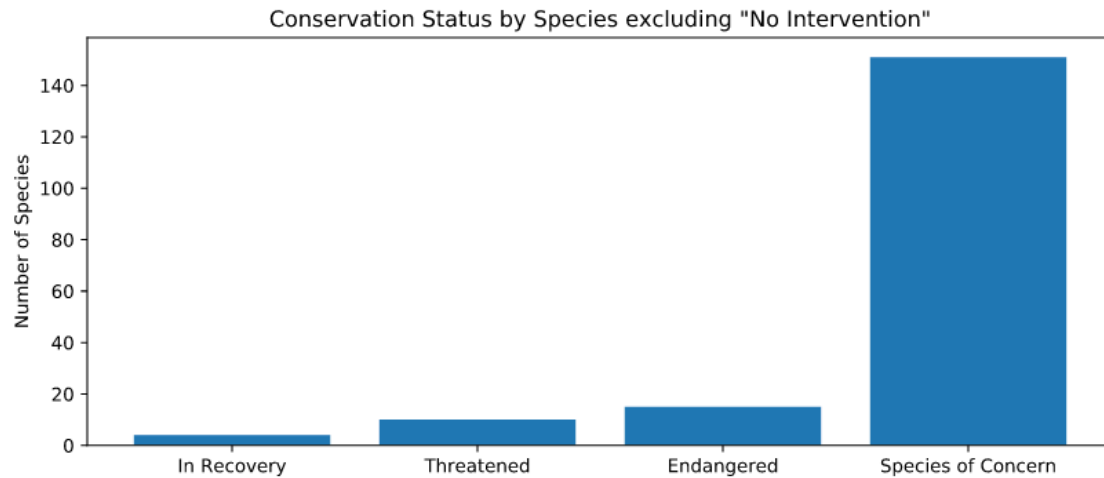
1. When comparing Mammals to Birds, the H_0 -Hypothesis **cannot even be rejected at a 10% significance level** (pvalue = 68.8%), meaning that there is no significant indication that the difference in protection percentage is not just pure chance at all.
2. When comparing Mammals to Reptiles however, the **H_0 -Hypothesis can be rejected at a reasonable 5% significance level** (pvalue = 3,8%), meaning that apparently Reptiles are indeed significantly less protected than Mammals.

A section describing the sample size determination that you did for the foot and mouth disease study:

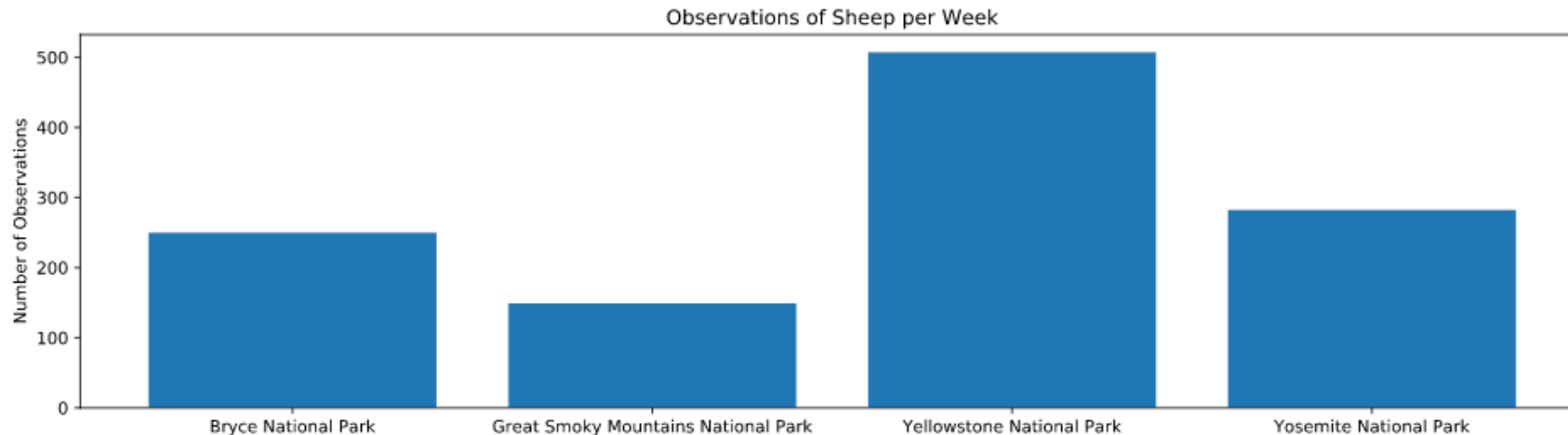
When determining the sample size, the **baseline** should be set to the current amount of infections (**15%**), leading to a **minimum detectable effect of 33%** given a **5% change**. According to the calculator I used, this leads to a suggested **sample size of 890** at a **90% confidence level**. To reach this, the scientists have to observe all sheep for **1,75 weeks in Yellowstone** and **3.56 weeks in Bryce**.

Disclaimer: I believe that the sample size page had some issues on codecademy.com: 1. The wrong Calculator was shown, imo this is the right one: [Link](#) 2. The Get code option doesn't work 3. The results seem to be unachievable using the calculator on the page or even the one I suggest

Overview of all Graphs



1, 2: As can be seen above, species of Concern make by far the largest chunk of somewhat endangered species, while if we look at the overall conversation status that still seems extremely little compared to the species with no intervention needed / happening.



3: As can be seen above, Bryce and Yosemite National Parks have similar amounts of sightings while Yellowstone knocks them out of the park ;)

Thank you for your attention!

Questions?

