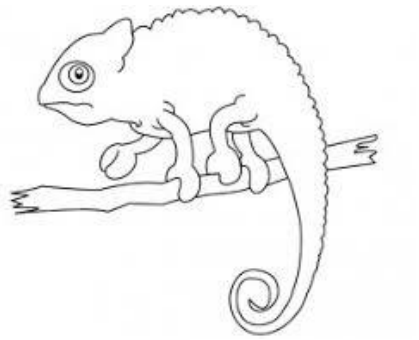
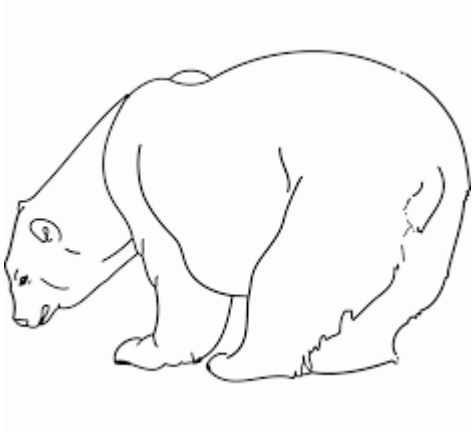


Biodiversity Project

National Park Service



By
Flavio Pasquali

Data Analysis Project – Part I

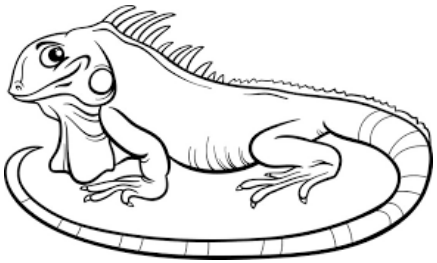
Purpose: To perform data analysis on the conservation statuses of different species and to find out if there are any patterns to the types of species that become endangered.

Data Source : **species_info.csv**

Total Rows (records) : **5543**

Species number: **7**

Conservation Status: **4**

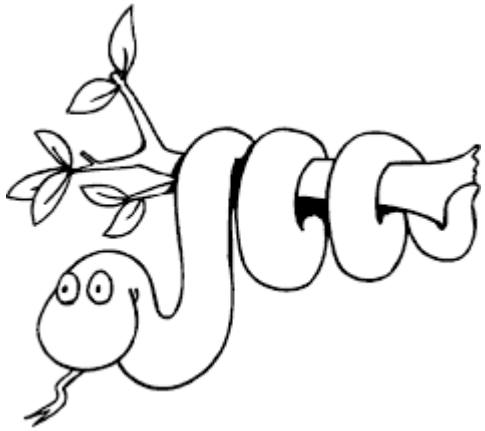


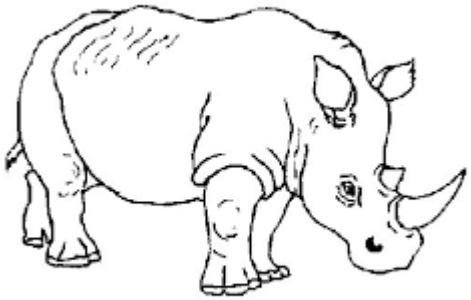
Dataframe:

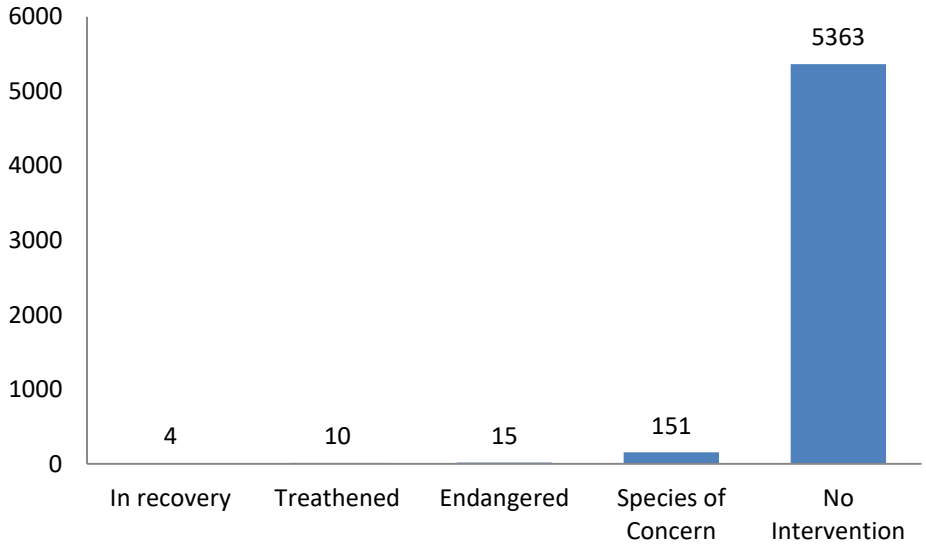
[‘**category**’,’**scientific_name**’,**common_names**’,’**conservation_status**’]

Conservation Status by Species

Conservation status	Scientific name
Endangered	15
In Recovery	4
No Intervention	5363
Species of Concern	151
Threatened	10







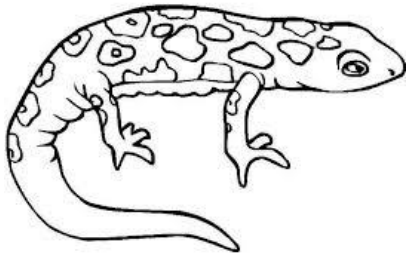
Chi-Squared Test for Significance

It looks like Mammals are more likely to be endangered than Birds, but is it a significant difference?

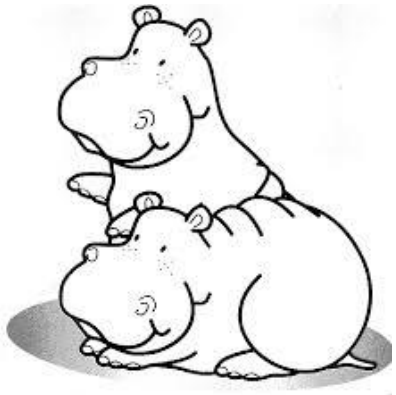
This difference isn't significant! (pval: 0,68759)

Is the difference between Reptile and Mammal significant?

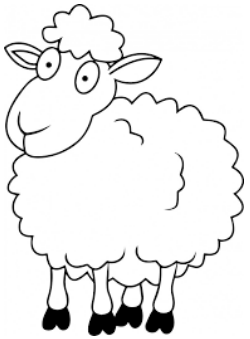
This difference is significant! (pval: 0,38355)



Category	Not Protected	Protected	% Protected
Amphibian	72	7	0.088608
Bird	413	75	0.153689
Fish	115	11	0.087302
Mammal	146	30	0.170455
Nonvascular Plant	328	5	0.015015
Reptile	73	5	0.064103
Vascular Plant	4216	46	0.010793



Data Analysis Project – Part II



Purpose: To analyze the observation and species DataFrames to help track sheep locations.

Data Source : **ovservations.csv**
species_info.csv

Dataframe: **ovservations**
['scientific_name', park_name',observations']

Dataframe: **species_info**
['category','scientific_name',common_names','conservation_status']

Merging Sheep and Observation DataFrames

Species Info

category	Scientific name	Common names	Conservation status

Observations

Scientific name	Park name	observations

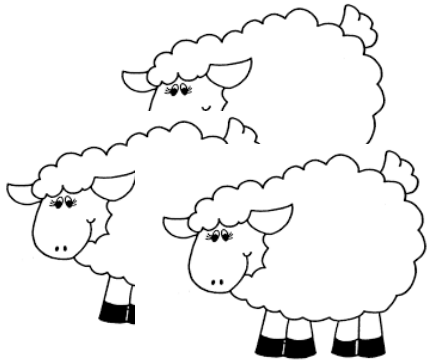
Data Cleaning and Merging DataFrames

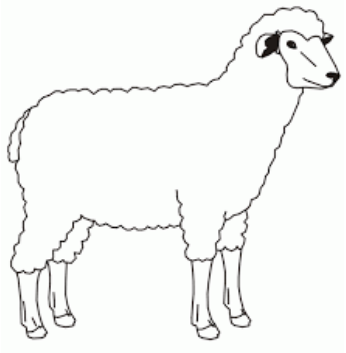
category	Scientific name	Common names	Conservation status	Is protected	Is Sheep	Park name	observations

Grouping by Park name and adding up the observations

Total sheep sightings (across all three species) were made at each national park?

	Park Name	Observations
0	Bryce National Park	250
1	Great Smoky Mountains National Park	149
2	Yellowstone National Park	507
3	Yosemite National Park	282

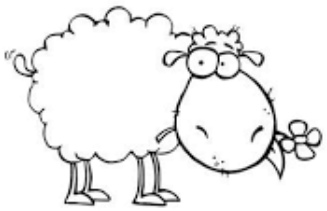




Foot and Mouth Reduction Effort - Sample Size Determination

Premises

- Last year it was recorded that 15% of sheep at Bryce National Park have foot and mouth disease.
- Detect at least 5% of reduction
- Level of significance (90%)



Results:

1. Baseline conversion rate: 15%
2. Minimum Detectable Effect: 33,33%
3. Sample Size: 870
4. Yellow weeks observing: 1,7
5. Bryce weeks observing: 3,48