Capstone Option 2:
Biodiversity for the National Parks

Data Analysis for better understanding biodiversity and conservation status of animals in national parks of USA

#Data resource collected from national parks of USA

#5824 species from 7 categories will be analysed

Nan Sheng 1/19/2018

Biodiversity Analyst of National Parks Service

CATALOGUE

Glimpse at data collected

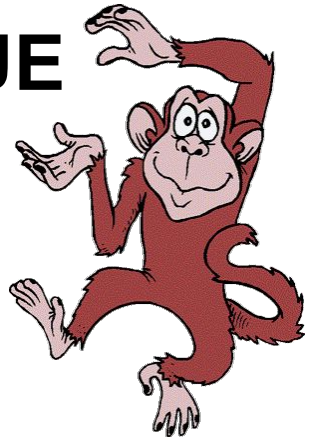
Conservation Status

- 4 categories of ‘conservation_status’
- Reorganised data
- Graph of Conservation Status by Species
- Conservation Status by category

Question: Are certain types of species more likely to be endangered?

In Search of Sheep

Foot and Mouth Reduction Effort



- Glimpse at data collected

- Animals are segmented according to their category, scientific name, common name and conservation status
- Data of 5824 species from 7 categories is presented

	category	scientific_name	common_names	conservation_status
0	Mammal	Clethrionomys gapperi gapperi	Gapper's Red-Backed Vole	nan
1	Mammal	Bos bison	American Bison, Bison	nan
2	Mammal	Bos taurus	Aurochs, Aurochs, Domestic Cattle (Feral), Domesticated Cattle	nan
3	Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	nan
4	Mammal	Cervus elaphus	Wapiti Or Elk	nan

● Conservation Status

➢ 4 categories of ‘conservation_status’

- ‘Endangered’ - 16 species
- ‘In recovery’ - 4 species
- ‘Species of Concern’ - 161 species
- ‘Threatened’ - 10 species

Where are the other species ?

	conservation_status	scientific_name
0	Endangered	16
1	In Recovery	4
2	Species of Concern	161
3	Threatened	10

‘Nan’ #5633 species → ‘no intervention’ is not included

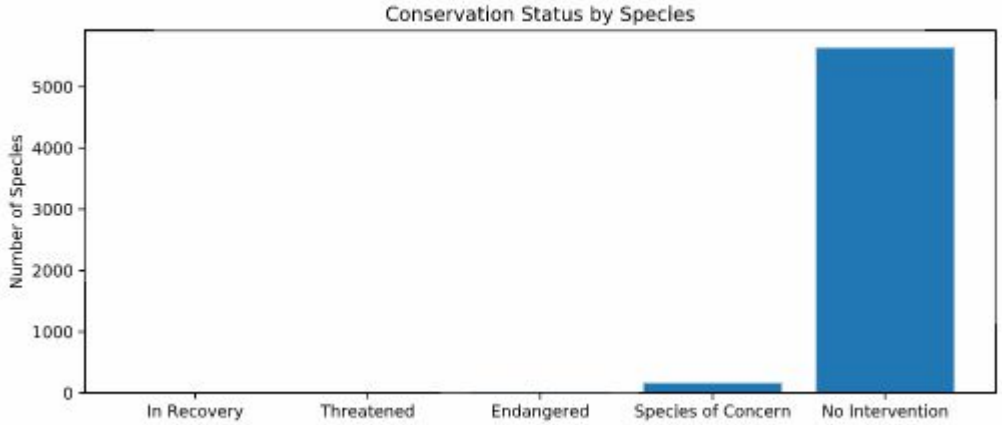
➢ Reorganised data

- 5633 ‘no intervention’ species
  - 96% among total

	conservation_status	scientific_name
0	Endangered	16
1	In Recovery	4
2	No Intervention	5633
3	Species of Concern	161
4	Threatened	10

➢ Graph of Conservation Status by Species

- ➔ Investigations on status of the 5633 species are needed.
- ➔ Possibly species without intervention are calling for protection
- ➔ As ‘No Intervention’ accounts 96% share, data of other species in the rest conservations is negligible in the graph

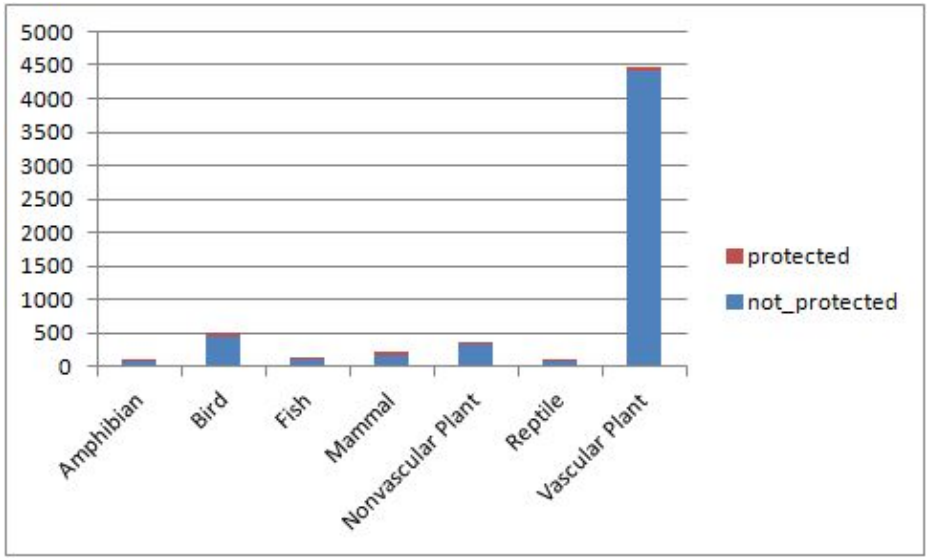


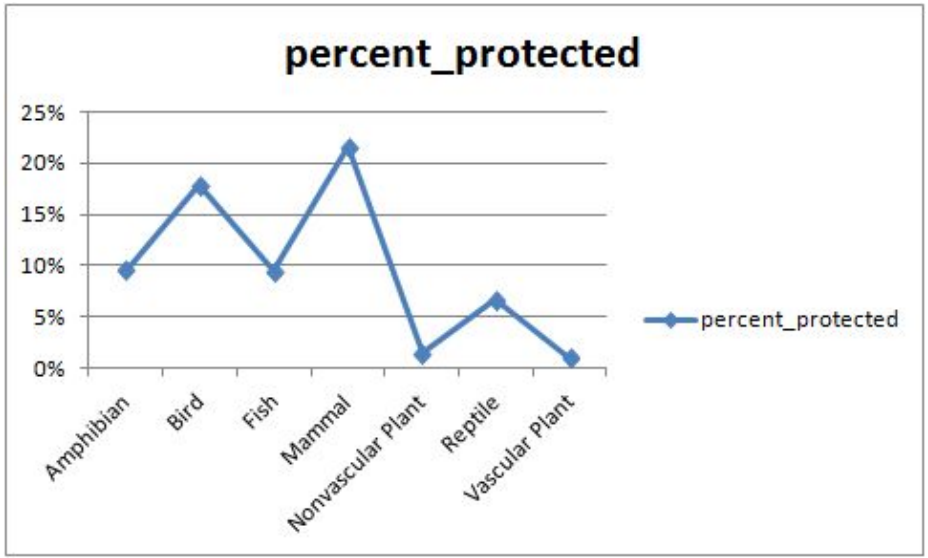
I can see nothing here. Let’s review the data in other way !

➢ Conservation Status by category

We provide O2 and absorb CO2. We deserve more care, hooman!

- ❑ More than 90% species are not under protection
- ❑ Mammal and bird are better protected than other species
- ❑ Vascular Plant are least protected

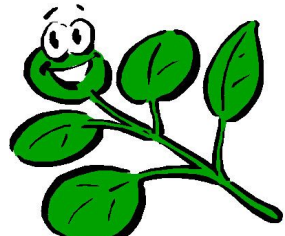




- **Question: Are certain types of species more likely to be endangered?**
  - ❏ Chi-Squared Test was applied as data is categorical
  - ❏ Pval = 5.51082804731e-89

Does this mean **“There is certain types of species are more likely to be endangered”** ?

- ➔ My answer is ‘depending’
- ➔ Since the conservation status of most species is ‘nan’, which means whether they are in danger or not is not clear, more specific investigations and data are needed.



In Search of Sheep

Baa Baa Baa

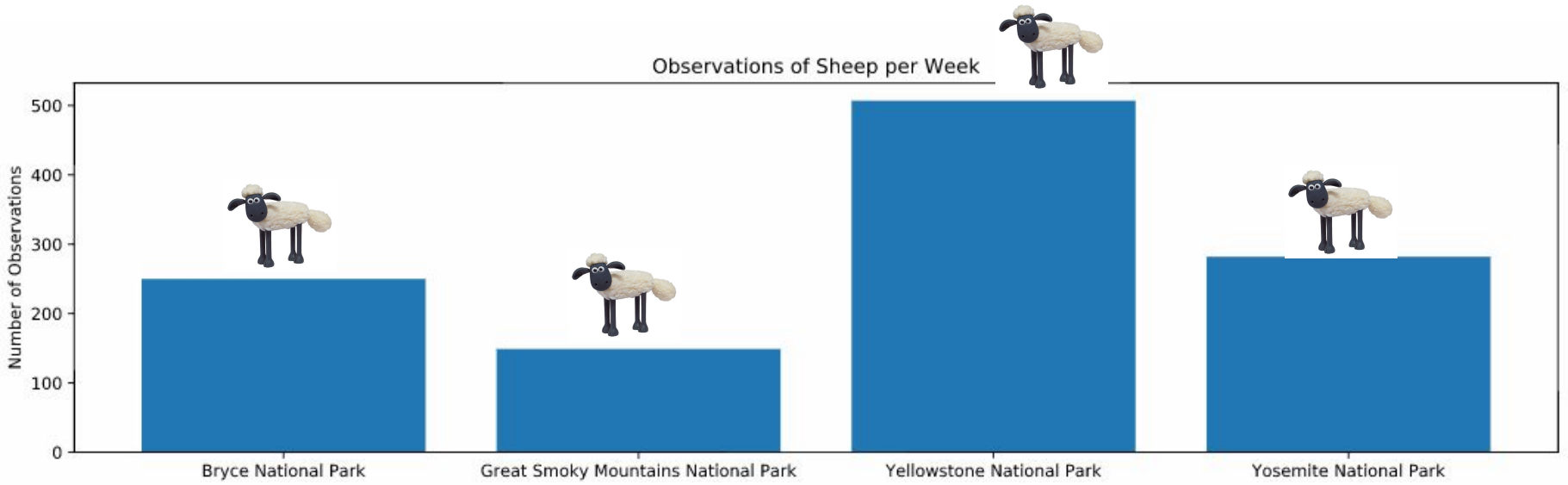
➢ A team of ruminant-enthused scientist has been tracking the movements of various species of sheep across different national parks and data is presented in table below.

	scientific_name	park_name	observations	category	common_names	conservation_status	is_protected	is_sheep
0	Ovis canadensis	Yellowstone National Park	219	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
1	Ovis canadensis	Bryce National Park	109	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
2	Ovis canadensis	Yosemite National Park	117	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
3	Ovis canadensis	Great Smoky Mountains National Park	48	Mammal	Bighorn Sheep, Bighorn Sheep	Species of Concern	True	True
4	Ovis canadensis sierrae	Yellowstone National Park	67	Mammal	Sierra Nevada Bighorn Sheep	Endangered	True	True

➢ Let’s zoom in and see how many sheep got observed in each national park

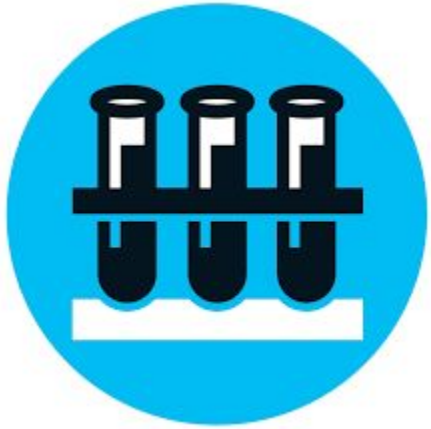
	park_name	observations
0	Bryce National Park	250
1	Great Smoky Mountains National Park	149
2	Yellowstone National Park	507
3	Yosemite National Park	282

- Numbers can be understand better in graph below
- More sheep have been seen in Yellowstone National Park



Foot and Mouth Reduction Effort

- Park Rangers at Yellowstone National Park have been running a program to reduce the rate of foot and mouth disease at that park.
- The scientists want to test whether or not this program is working.
- They want to be able to detect reductions of at least **5%** .



How large the sample should be for testing whether 5% is achieved or not ?

➔ sample size per variant is **510**

How long time it will take for getting this large simple in each national park?

Park Name	**Week(s) Needed for Observation**
Bryce National Park	2
Yellowstone National Park	1