Biodiversity in National Parks

Created By: Alec F. for Codecademy Introduction to Data Analysis Course

Data Analyzed

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

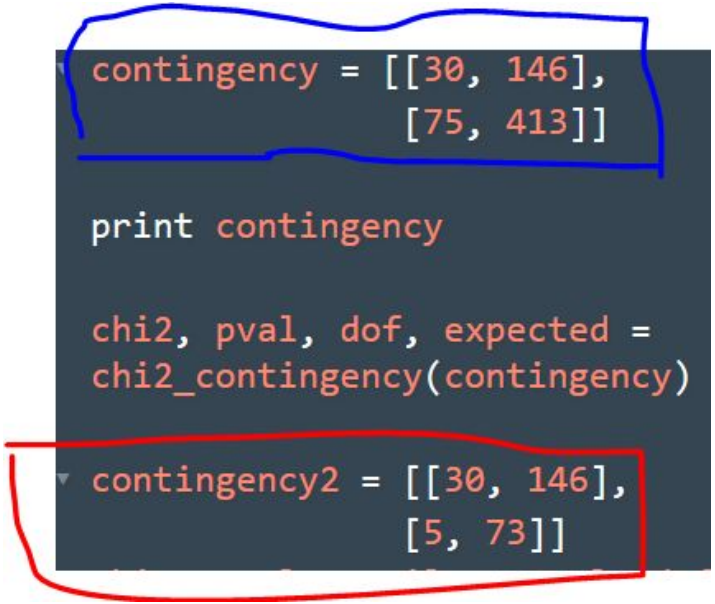
- Contained information for all species in the Parks
- Includes identifying factors of species such as Scientific and common name
- Key data: indicates if species has conservation status

Key Takeaways

- 5541 Unique species
- Mammals and Birds have highest percentage of conservation status with 17% and 15%
- Mammals are significantly more likely to be endangered than reptiles, but not significantly more likely to be endangered than birds

Significance Calculation

- To determine significance of conservation status:
  - Chi-Squared test conducted
  - Contingency test contained number of protected and non-protected species (left and right columns respectively)
  - First tested Mammals and Birds (Blue)
  - Next tested Mammals and Reptiles (Red)
  - P-val of 0.68 and 0.03 respectively
    - Since 0.03 is less than 0.05, it is significant



Foot and Mouth Disease

- Using a sample size calculator and data of the number of sheep with foot and mouth disease in parks
- Determined 870 is the sample size per variant
- Using this and the data on observations by park, determined number of weeks to determine significant data:
  - Yellowstone: 1.7 weeks
  - Bryce: 3.5 weeks