

# Capstone Project

Biodiversity for the National Parks

By: Jaime M Davilis

# Notes on the species dataframe

- ❖ 5824 rows in the dataframe, the majority of the species required no intervention; only 180 required some intervention:
  - 15 endangered
  - 4 in recovery
  - 151 species in concern
  - 10 threatened
- ❖ Most (4216) species were vascular plants
- ❖ Mammals, followed by birds, were the species under the most protection, whereas plants were under the least protection
- ❖ There was no significant difference between mammals and birds likelihood of being endangered, but there was between reptiles and mammals

# Recommendations for Conservationists

- ❖ Using the chi square tests, it does appear that certain species are more likely to become endangered:
  - Reptiles - the pvalues for reptiles vs mammals and reptiles vs vascular plants were both  $< 0.05$ , although the reptiles vs birds was slightly larger than 0.05; they seem most likely to become endangered
  - Birds and mammals seems least likely, as the pvalue was more than 0.68 for birds vs mammals
  - Interestingly, the chi square test for vascular plants vs mammals, and vascular plants vs non-vascular plants were both zero, indicating there is a very significant difference between these two groups, and that mammals could require more protection along with reptiles than both vascular and non-vascular plants

# Foot and Mouth Disease - Sample determination

- ❖ I first determined the baseline conversion rate: this was 15% based on data collected at the national park in the prior year
- ❖ A statistical significance of 90% was chosen by default; a higher significance would have increase the sample size, and a lower rate would lower the sample size
- ❖ The minimum detectable effect was 33.3%; this was calculated by determining the smallest percentage the park rangers wanted to detect (5%), multiplying that by 100 (to change the result to a percentage), and dividing by the baseline percentage:  $(100 * 0.05) / .15 = 33.33\%$
- ❖ The proper sample size for this determination was 870.