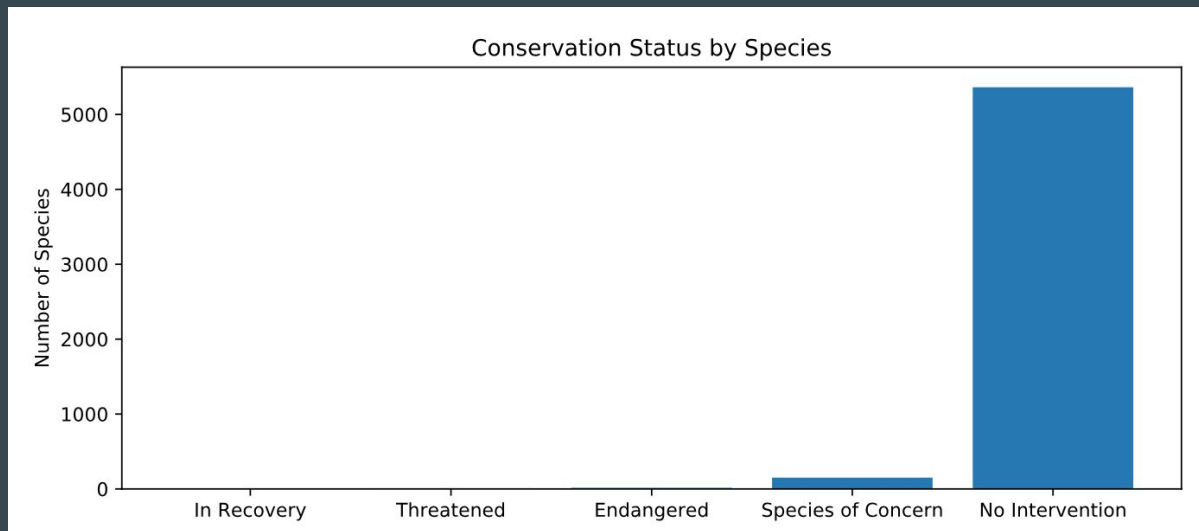# Biodiversity

● ● ●

Rebecca Sleath

Species Data

The csv file 'species.info' gives the scientific name, category, common name and in some cases, conservation status for 5,543 species found in the National Parks..

Most species have no conservation status assigned, but for those that do, Species of Concern' is most common with 151, and 'In Recovery' the rarest (4).

Endangered Species

Some categories of species appeared more likely to be listed as endangered than others. 17% of mammals and 15% of birds had protected statuses, compared to only 1-2% for the two plant groups.

```
   category          not_protected   protected   percent_protected
0  Amphibian                    72           7            0.088608
1  Bird                        413          75            0.153689
2  Fish                        115          11            0.087302
3  Mammal                      146          30            0.170455
4  Nonvascular Plant           328           5            0.015015
5  Reptile                      73           5            0.064103
6  Vascular Plant             4216          46            0.010793
```

Significance Test (Chi-Squared)

Using a chi-squared test, the difference between the percent of endangered birds was not found to be significantly different to the percentage of endangered mammals at a 95% confidence level. However, the difference between the percent of endangered reptiles and mammals was significant, with a P-value of 0.038.

Based on these findings, I would recommend that:

- Protection measures be focussed on mammals over reptiles
- Researchers ensure that the relative lack of conservation statuses for non-mammals is not down to a bias, such as in ease of collecting data.

Foot and Mouth

Scientists wanted to test if a treatment was working.

The rate of foot and mouth disease was 15% at one of the parks last year and others are not known, so the baseline rate of disease was 15%.

Scientists are looking for at least a 5 percentage point drop in rates of disease, so using 5%/15% to find the minimum detectable difference gives 33%.

These figures along with a 90% confidence gives a sample size per variant of 890. That means about 1 week in Yellowstone National Park or 2 weeks in Bryce would be enough to be reasonably sure whether the treatment was working.