CODECADEMY
Biodiversity in US National Parks

Daniel Rodriguez Ponce

Dataframes

Two dataframes were used in this project: species and conservation. They were pretty large dataframes.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5824 entries, 0 to 5823
Data columns (total 4 columns):
 #   Column               Non-Null Count   Dtype
---  ------               --------------   -----
 0   category             5824 non-null    object
 1   scientific_name      5824 non-null    object
 2   common_names         5824 non-null    object
 3   conservation_status  191 non-null     object
dtypes: object(4)
memory usage: 182.1+ KB
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 23296 entries, 0 to 23295
Data columns (total 3 columns):
 #   Column           Non-Null Count   Dtype
---  ------           --------------   -----
 0   scientific_name  23296 non-null   object
 1   park_name        23296 non-null   object
 2   observations     23296 non-null   int64
dtypes: int64(1), object(2)
memory usage: 546.1+ KB
```

Information contained in the species and observations dataframes respectively:

Out[9]:

	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), Dom...	NaN
3	Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	NaN
4	Mammal	Cervus elaphus	Wapiti Or Elk	NaN

In [10]:
```
#Exploration of observations dataframe with head
observations.head()
```

Out[10]:

	scientific_name	park_name	observations
0	Vicia benghalensis	Great Smoky Mountains National Park	68
1	Neovison vison	Great Smoky Mountains National Park	77
2	Prunus subcordata	Yosemite National Park	138
3	Abutilon theophrasti	Bryce National Park	84
4	Githopsis specularioides	Great Smoky Mountains National Park	85

This is the table and plot of the different conservation statuses among protected species

```
Out[14]: Species of Concern    161
         Endangered             16
         Threatened             10
         In Recovery             4
         Name: conservation_status, dtype: int64
```


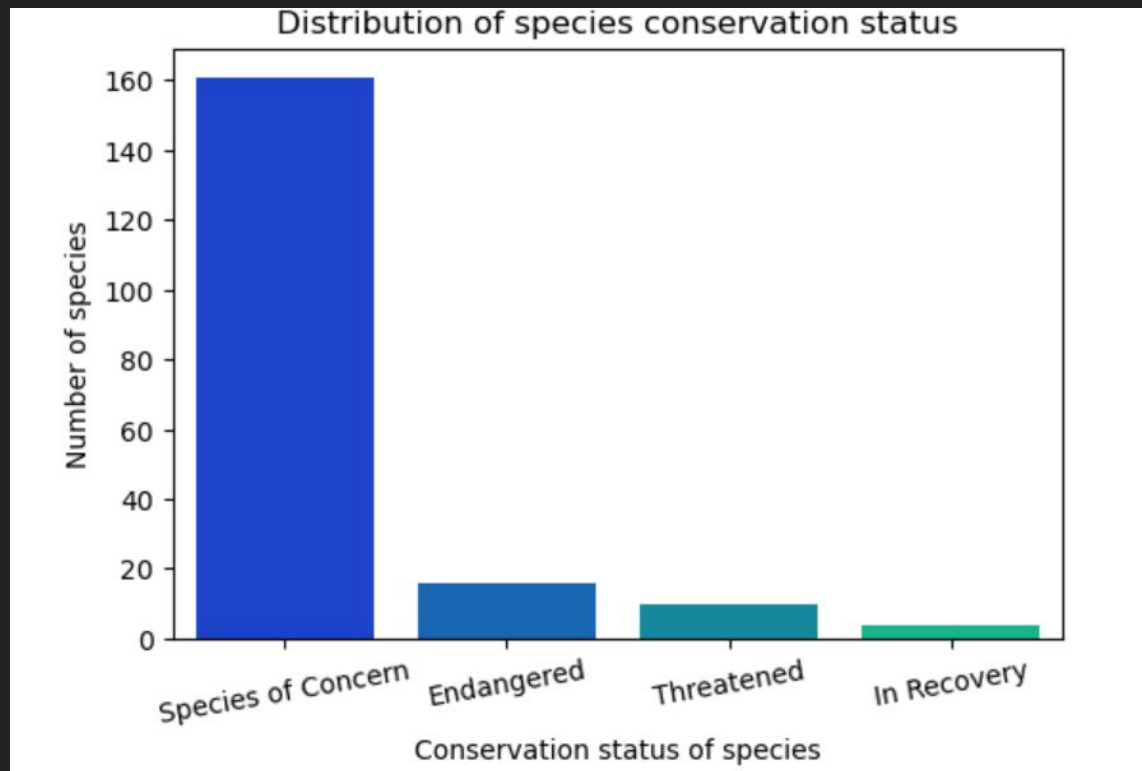Distribution of species conservation status

Table with the number of species among different categories found at US National Parks and table of different conservation statuses for different categories of

```
Vascular Plant        4470
Bird                   521
Nonvascular Plant      333
Mammal                 214
Fish                   127
Amphibian               80
Reptile                 79
Name: category, dtype: int64
```

category	Amphibian	Bird	Fish	Mammal	Nonvascular Plant \
conservation_status					
Endangered	1.0	4.0	3.0	7.0	NaN
In Recovery	NaN	3.0	NaN	1.0	NaN
Species of Concern	4.0	72.0	4.0	28.0	5.0
Threatened	2.0	NaN	4.0	2.0	NaN

category	Reptile	Vascular Plant
conservation_status		
Endangered	NaN	1.0
In Recovery	NaN	NaN
Species of Concern	5.0	43.0
Threatened	NaN	2.0

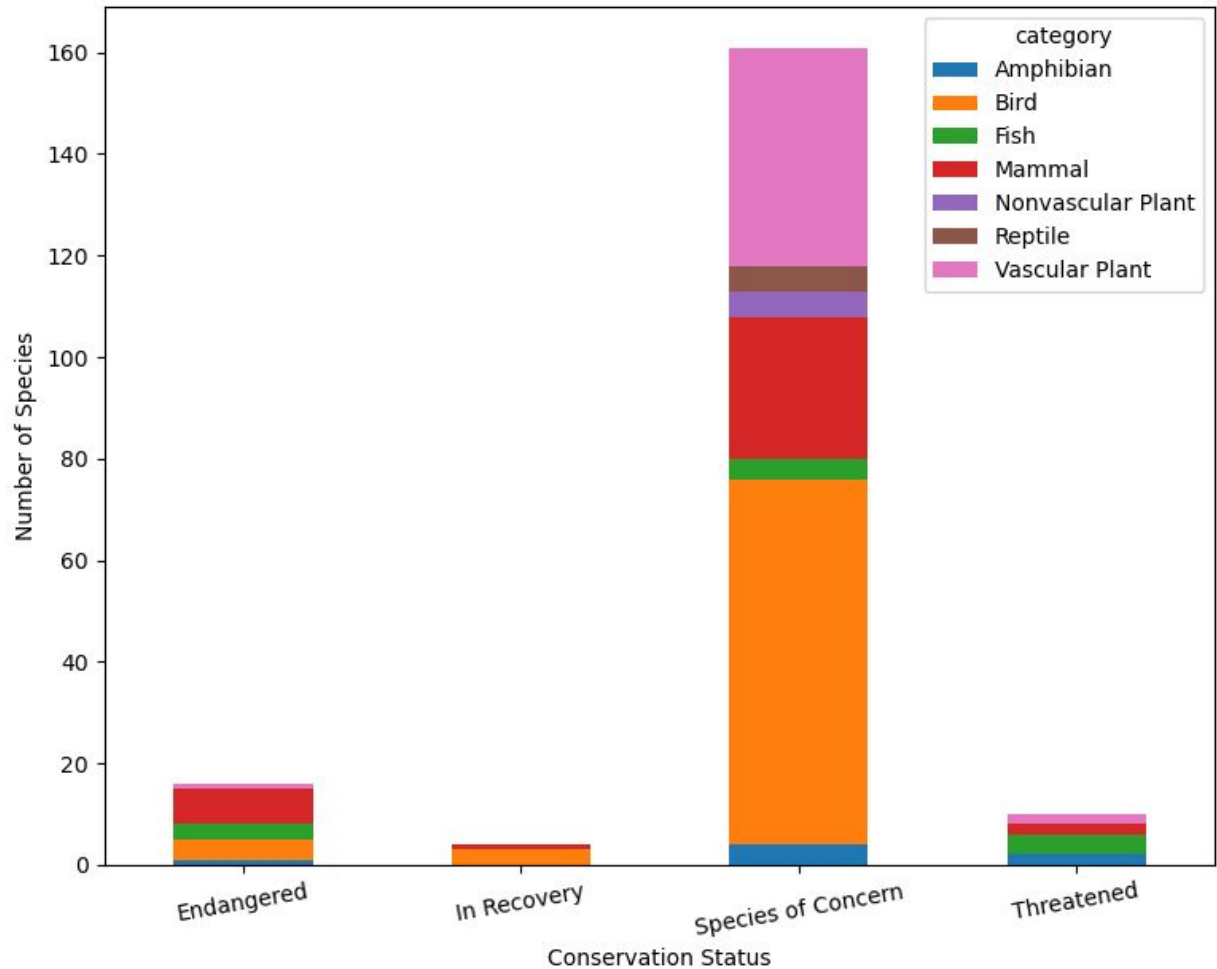Table of conservation statuses per category

Table that shows percentage of protected species vs total number of species per category

need_protection category	Not Protected	Protected	% Protected
Amphibian	72	7	8.860759
Bird	413	75	15.368852
Fish	115	11	8.730159
Mammal	146	30	17.045455
Nonvascular Plant	328	5	1.501502
Reptile	73	5	6.410256
Vascular Plant	4216	46	1.079305

Species most found at Smoky Mountains National Park

```
The most prevalent species at Great Smoky Mountains National Park are:
            scientific_name  observations        category  \
0   Sonchus asper ssp. asper          147  Vascular Plant
1         Lespedeza violacea          146  Vascular Plant
2       Botrychium dissectum          144  Vascular Plant
3           Festuca subulata          141  Vascular Plant
4           Branta hutchinsii          139            Bird

                                 common_names
0                            Spiny Sowthistle
1                           Violet Bush-Clover
2   Cutleaf Grape-Fern, Cut-Leaved Grape Fern
3            Bearded Fescue, Nodding Fescue
4                               Cackling Goose
```

The most prevalent species found at Yosemite National Park are:

```
The most prevalent species at Yosemite National Park are:
                  scientific_name  observations          category  \
0       Ivesia shockleyi var. shockleyi           223   Vascular Plant
1               Lipocarpha micrantha           222   Vascular Plant
2                   Trillium rugelii           219   Vascular Plant
3   Artemisia arbuscula ssp. arbuscula           219   Vascular Plant
4               Castilleja pallescens           214   Vascular Plant

                                    common_names
0                 Shockley's Ivesia, Sky Mousetail
1   Dwarf Bulrush, Smallflower Halfchaff Sedge, Sm...
2                       Southern Nodding Trillium
3           Gray Low Sagebrush, Little Sagebrush
4       Pale Indian Paintbrush, Pale Paintbrush
```

The most prevalent species found at Bryce National Park are:

```
The most prevalent species at Bryce National Park are:
         scientific_name   observations          category  \
0   Valerianella radiata             176   Vascular Plant
1      Melilotus indica             173   Vascular Plant
2    Ericameria bloomeri             172   Vascular Plant
3     Draba oligosperma             167   Vascular Plant
4       Martes americana             167           Mammal

                           common_names
0       Beaked Corn-Salad, Corn Salad
1                          Sourclover
2    Bloomer's Goldenbush, Rabbitbush
3         Fewseed Draba, Glacier Draba
4                         Pine Marten
```

The most prevalent species found at Yellowstone National Park are:

```
The most prevalent species at Yellowstone National Park are:
          scientific_name  observations            category  \
0    Lycopodium tristachyum           321     Vascular Plant
1        Castilleja miniata           317     Vascular Plant
2  Dracocephalum parviflorum          316     Vascular Plant
3        Cryptantha fendleri          316     Vascular Plant
4         Bidens tripartita          314     Vascular Plant

                                             common_names
0                    Deep-Root Clubmoss, Ground Cedar
1    Giant Red Indian Paintbrush, Great Red Indian ...
2                                  American Dragonhead
3    Fendler's Cryptantha, Sanddune Catseye, Sanddu...
4          Threelobe Beggarticks, Three-Lobe Beggarticks
```

CONCLUSIONS

The US National Parks System has done a great job. Of a number of 5541 species spotted in its parks, only 191 need some form of conservation measures.

Out of those 191 species that need protection, 161 are only of concern, without being threatened or endangered. 4 are in recovery.

The most endangered species at all parks is mammals, followed by birds.