Capstone Option 2: Biodiversity for the National Parks

Elizabeth Gray

	conservation_status	scientific_name
0	Endangered	15
1	In Recovery	4
2	Species of Concern	151
3	Threatened	10

	conservation_status	scientific_name
0	Endangered	15
1	In Recovery	4
2	No Intervention	5363
3	Species of Concern	151
4	Threatened	10

	category	scientific_name	common_names	conservation_status
0	Mammal	Clethrionomys gapperi gapperi	Gapper's Red-Backed Vole	No Intervention
1	Mammal	Bos bison	American Bison, Bison	No Intervention
2	Mammal	Bos taurus	Aurochs, Aurochs, Domestic Cattle (Feral), Domesticated Cattle	No Intervention
3	Mammal	Ovis aries	Domestic Sheep, Mouflon, Red Sheep, Sheep (Feral)	No Intervention
4	Mammal	Cervus elaphus	Wapiti Or Elk	No Intervention

```
species_count = species.nunique()
print species_count.head()

category               7
scientific_name     5541
common_names        5504
conservation_status    4
dtype: int64

print species.category.unique()

['Mammal' 'Bird' 'Reptile' 'Amphibian' 'Fish' 'Vascular Plant'
 'Nonvascular Plant']

print species.conservation_status.unique()

  [nan 'Species of Concern' 'Endangered' 'Threatened' 'In Recovery']
```

```
conservation_counts =
species.groupby('conservation_status').scientific_name.nunique().reset_index()
print conservation_counts
```

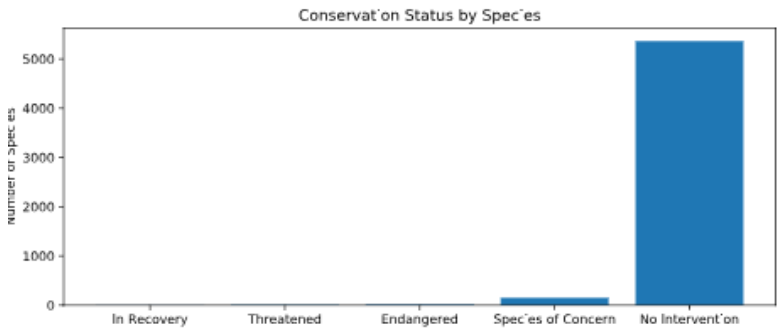
```
 conservation_status  scientific_name
 0         Endangered              15
 1        In Recovery              4
 2  Species of Concern             151
 3         Threatened              10
```

```
species.fillna('No Intervention', inplace = True)
conservation_counts_fixed =
species.groupby('conservation_status').scientific_name.nunique().reset_index()
print conservation_counts_fixed
```

```
conservation_status  scientific_name
0         Endangered              15
1        In Recovery              4
2     No Intervention             5363
3  Species of Concern             151
4         Threatened              10
```

```
protection_counts = species.groupby('conservation_status')\
   .scientific_name.nunique().reset_index()\
   .sort_values(by='scientific_name')

plt.figure(figsize=(10, 4))
ax = plt.subplot()
plt.bar(range(len(protection_counts)),protection_counts.scienti
fic_name.values)
ax.set_xticks(range(len(protection_counts)))
ax.set_xticklabels(protection_counts.conservation_status.value
s)
plt.ylabel('Number of Species')
plt.title('Conservation Status by Species')
labels = [e.get_text() for e in ax.get_xticklabels()]
plt.show()
```


Conservat˙on Status by Spec˙es
Number of spec˙es
5000
4000
3000
2000
1000
0
In Recovery
Threatened
Endangered
Spec˙es of Concern
No Intervent˙on

```
species['is_protected'] = species.conservation_status != 'No
Intervention'

category_counts = species.groupby(['category',
'is_protected']).scientific_name.nunique().reset_index()

print category_counts.head()

category_pivot =
category_counts.pivot(columns='is_protected',
                      index='category',
                      values='scientific_name')\
                      .reset_index()

print category_pivot
category_pivot.columns = ['category', 'not_protected',
'protected']
print category_pivot

category_pivot['percent_protected'] =
category_pivot.protected/(category_pivot.protected +
category_pivot.not_protected)

print category_pivot
```

	category	not_protected	protected
0	Amphibian	72	7
1	Bird	413	75
2	Fish	115	11
3	Mammal	146	30
4	Nonvascular Plant	328	5
5	Reptile	73	5
6	Vascular Plant	4216	46

	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

```
from scipy.stats import chi2_contingency

contingency = [[30, 146],
               [75, 413]]

pval = chi2_contingency(contingency)[1]
print(pval)
# No significant difference because pval > 0.05

contingency_reptile_mammal = [[30, 146],
                              [5, 73]]

pval_reptile_mammal =
chi2_contingency(contingency_reptile_mammal)[1]
print(pval_reptile_mammal)
# Significant difference! pval_reptile_mammal < 0.05
```

Recommendations

- Certain types of species are more likely to be endangered than others based on the data provided. In particular, it appears that Mammals are significantly more likely to be endangered than reptiles.
- Therefore, the Parks may consider doing more research into reptile species present in the parks to determine if there are reptiles that would fall into the category of endangered.
- Another options would be to research why more Mammals (and Birds) are endangered than the other species.