Examples_Chapter4

September 29, 2021

1 Example 4.3 of the textbook

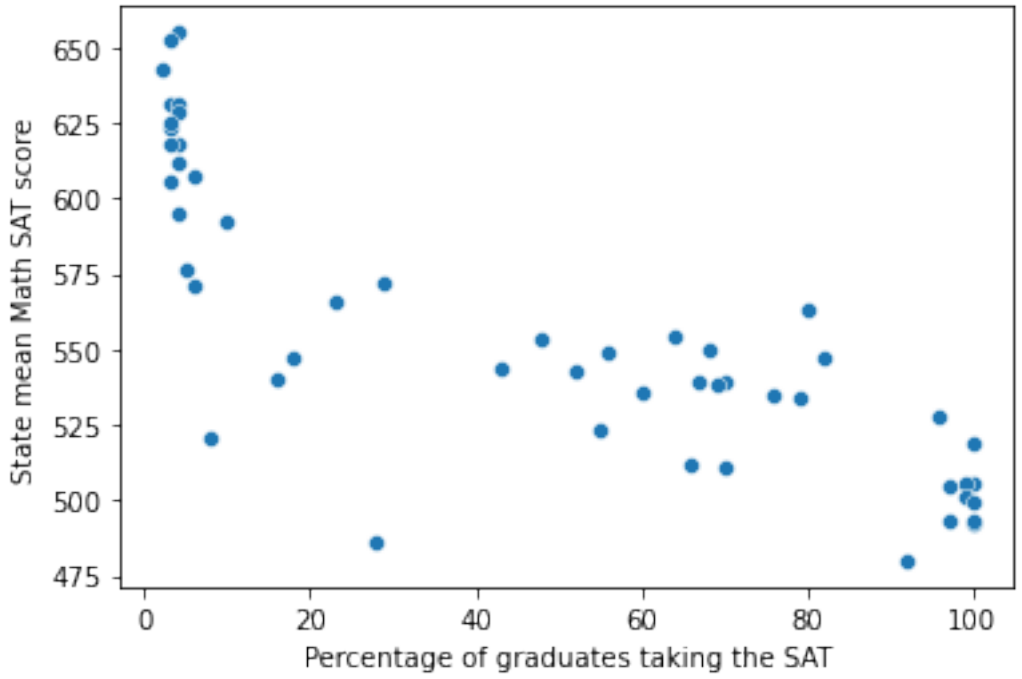
[1]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from scipy import stats
```

[2]:
```
mathsat = pd.read_csv("eg04-03mathsat.csv")
mathsat.head()
```

[2]:

	State	PctSAT	MathSAT2018
0	Alabama	6	571
1	Alaska	43	544
2	Arizona	29	572
3	Arkansas	5	576
4	California	60	536

[3]:
```
sns.scatterplot(x = "PctSAT", y = "MathSAT2018", data = mathsat)
plt.xlabel("Percentage of graduates taking the SAT")
plt.ylabel("State mean Math SAT score")
plt.show()
```



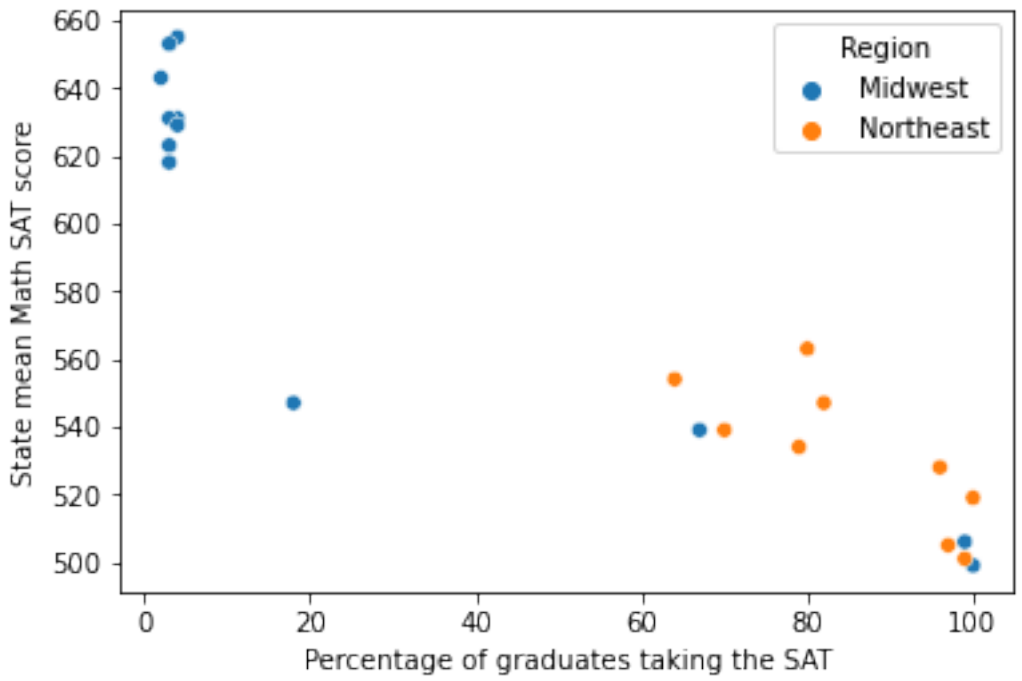
2 Example 4.4 of the textbook (correction: the midwest has 12 states not 13)

[4]:
```
sat = pd.read_csv("eg04-04Midwest_Northeast.csv")
sat.head()
```

[4]:

	State	PctSAT	MathSAT2018	Region
0	Minnesota	4	655	Midwest
1	Wisconsin	3	653	Midwest
2	NorthDakota	2	643	Midwest
3	Kansas	4	631	Midwest
4	Iowa	3	631	Midwest

[5]:
```
sns.scatterplot(x = "PctSAT", y = "MathSAT2018", hue="Region", data = sat)
plt.xlabel("Percentage of graduates taking the SAT")
plt.ylabel("State mean Math SAT score")
plt.show()
```

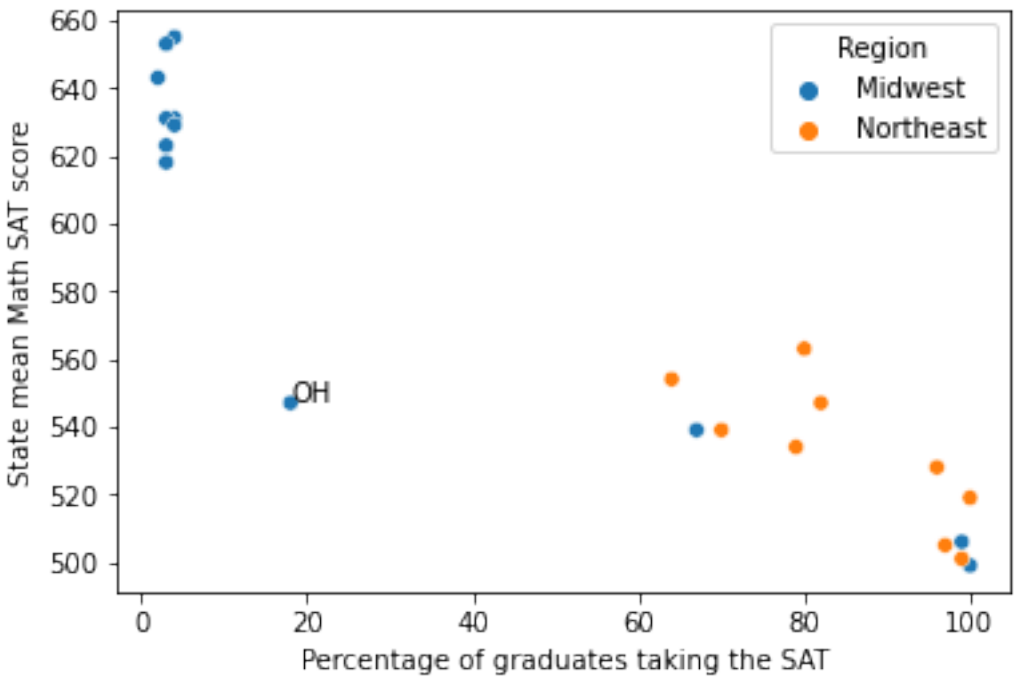


[6]:
```
sat[sat.State=="Ohio"]
```

[6]:

	State	PctSAT	MathSAT2018	Region
8	Ohio	18	547	Midwest

[7]:
```
sns.scatterplot(x = "PctSAT", y = "MathSAT2018", hue="Region", data = sat)
plt.xlabel("Percentage of graduates taking the SAT")
plt.ylabel("State mean Math SAT score")
## adding text into the plot to indicate the point corresponding to Ohio
plt.text(x=18, y=547, s='OH')
plt.show()
```



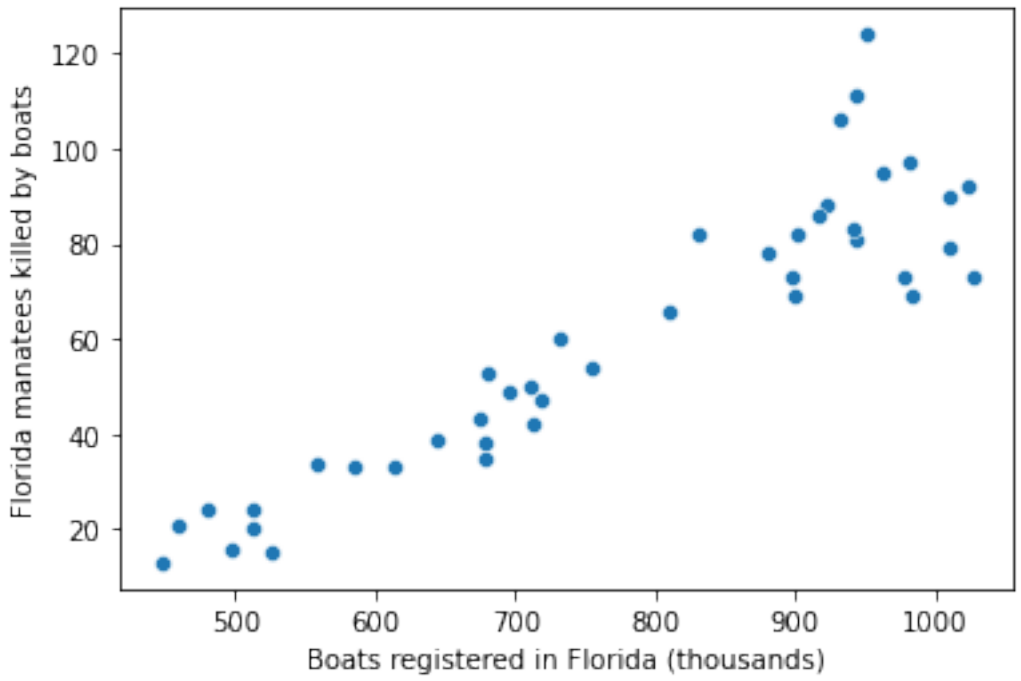
3 Figure 4.7 (a) of the textbook

[8]:
```
florida = pd.read_csv("eg04-05manatee.csv")
florida.head()
```

[8]:

	Year	Boats	Kills
0	1977	447	13
1	1978	460	21
2	1979	481	24
3	1980	498	16
4	1981	513	24

[9]:
```
sns.scatterplot(x = "Boats", y = "Kills", data = florida)
plt.xlabel("Boats registered in Florida (thousands)")
plt.ylabel("Florida manatees killed by boats")
plt.show()
```



[10]:
```
stats.pearsonr(x=florida['Boats'], y=florida['Kills'])[0]
```

[10]: 0.9189057628743729

[11]:
```
#help(stats.pearsonr)
```

[ ]:
```

```