Honey Production

February 7, 2022

0.0.1

As you may have already heard, the honeybees are in a precarious state right now. You may have seen articles about the decline of the honeybee population for various reasons. You want to investigate this decline and how the trends of the past predict the future for the honeybees.

[ ]:
```
from sklearn.linear_model import LinearRegression
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
```

[ ]:
```
honey=pd.read_csv(r'C:\Users\shaw\Downloads\honeyproduction.csv')
honey.head()
```

[ ]:

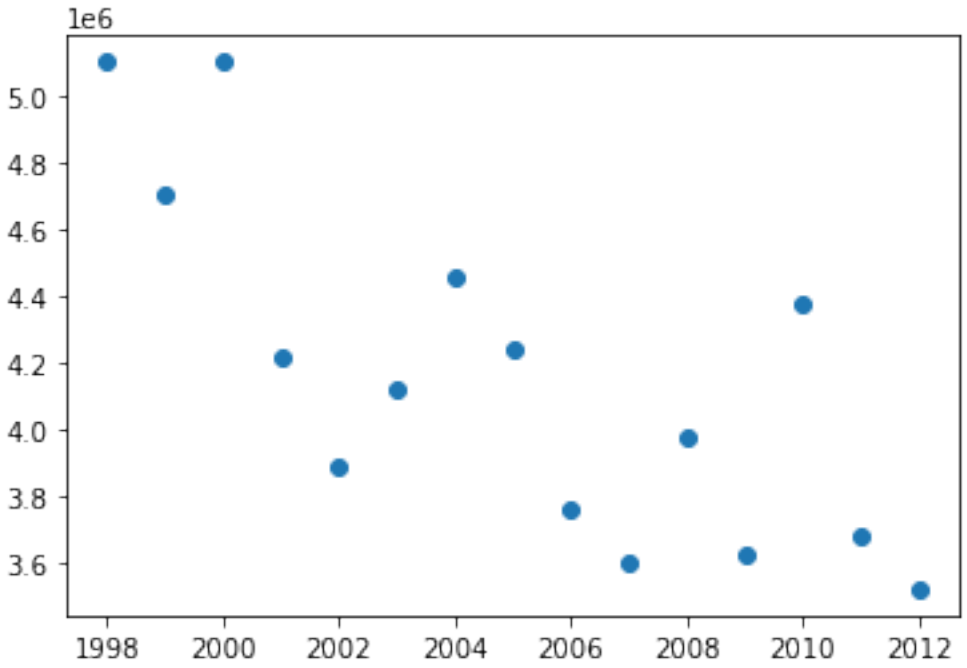
	state	numcol	yieldpercol	totalprod	stocks	priceperlb	prodvalue	year
0	AL	16000.0	71	1136000.0	159000.0	0.72	818000.0	1998
1	AZ	55000.0	60	3300000.0	1485000.0	0.64	2112000.0	1998
2	AR	53000.0	65	3445000.0	1688000.0	0.59	2033000.0	1998
3	CA	450000.0	83	37350000.0	12326000.0	0.62	23157000.0	1998
4	CO	27000.0	72	1944000.0	1594000.0	0.70	1361000.0	1998

[ ]:
```
prod_per_year = honey.groupby('year').totalprod.mean().reset_index()
prod_per_year.head()
```

[ ]:

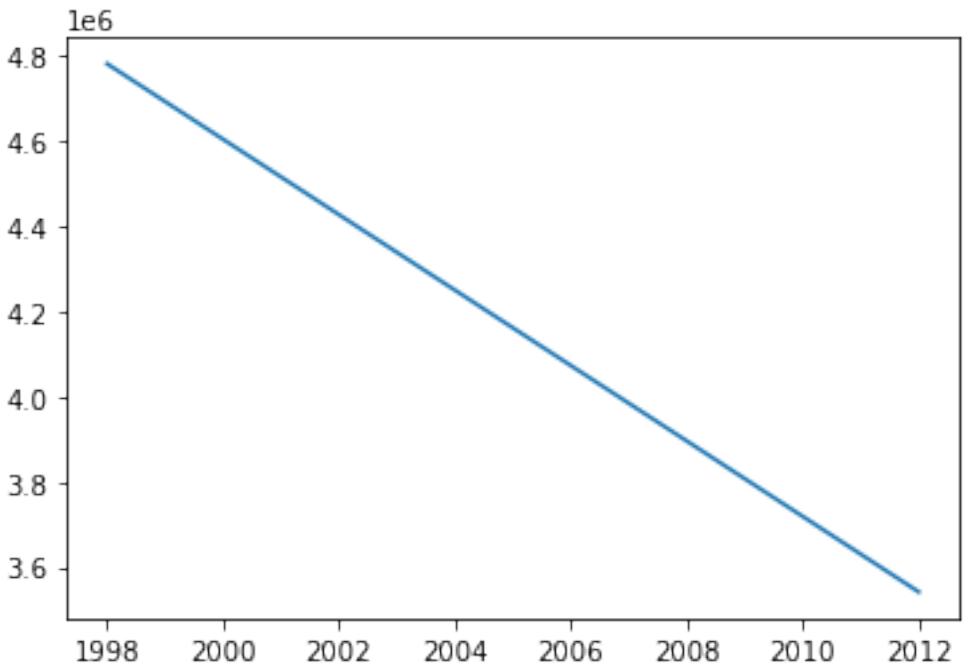
	year	totalprod
0	1998	5.105093e+06
1	1999	4.706674e+06
2	2000	5.106000e+06
3	2001	4.221545e+06
4	2002	3.892386e+06

[ ]:
```
x=prod_per_year['year']
x = x.values.reshape(-1, 1)
y=prod_per_year['totalprod']
plt.scatter(x,y);
```



[ ]:
```
regr=LinearRegression()
regr.fit(x,y)
print(regr.coef_,regr.intercept_)
y_predict=regr.predict(x)
plt.plot(x,y_predict)
plt.show()
```

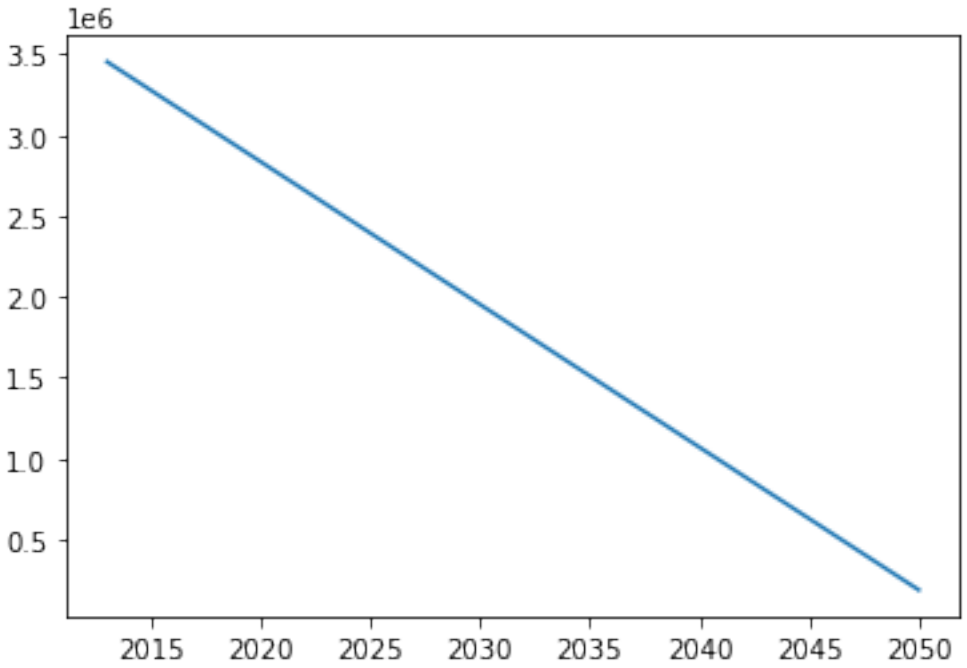
```
[-88303.18915238] 181208083.10732982
```



0.0.2 So, it looks like the production of honey has been in decline, according to this linear model. Let's predict what the year 2050 may look like in terms of honey production.

Our known dataset stops at the year 2013, so let's create a NumPy array called X_future that is the range from 2013 to 2050. The code below makes a NumPy array with the numbers 1 through 10

[ ]:
```
x_future=np.array(range(2013,2051))
x_future=x_future.reshape(-1,1)
future_predict=regr.predict(x_future)
plt.plot(x_future,future_predict)
plt.show()
```


1e6
3.5
3.0
2.5
2.0
1.5
1.0
0.5
2015
2020
2025
2030
2035
2040
2045
2050