NAME: DIVYA STUTI
E-MAIL: divyastuti.DS@gmail.com
COURSE NAME: Data Science with Python Career Program (ChatGPT Included)
ASSIGNMENT: ADVANCED PYTHON (MAJOR)

Topic: To visualize how honey production has changed over the years (1998-2021) in the United States.

Background: In 2006, global concern was raised over the rapid decline in the honey bee population, an integral component of American honey agriculture. Large numbers of hives were lost to Colony Collapse Disorder, a phenomenon of disappearing worker bees causing the remaining hive colony to collapse. Speculation as to the cause of this disorder points to hive diseases and pesticides harming the pollinators, though no overall consensus has been reached. The U.S. used to locally produce over half the honey it consumes per year. Now, honey mostly comes from overseas, with 350 of the 400 million pounds of honey consumed very year originating from imports. This dataset provides insight into honey production supply and demand in America from 1998 to 2021.

Objective: To visualize how honey production has changed over the years (1998-2021) in the United States.

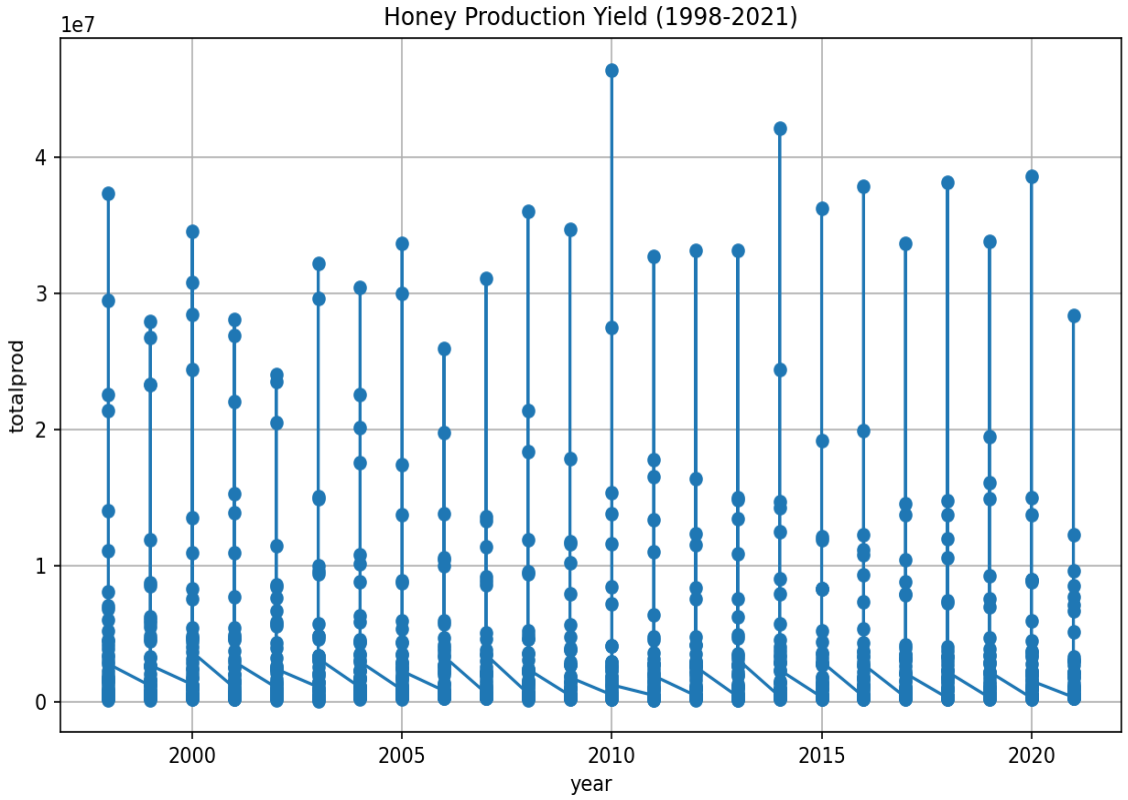
Q1) How has honey production yield changed from 1998 to 2021?

honeyproduction 1998-2021.csv | Python Major Assignment.py

Python Major Assignment.py > ...

```python

import pandas as pd
import matplotlib.pyplot as plt

file_path = 'honeyproduction 1998-2021.csv'

# Read the dataset into a Pandas DataFrame
df = pd.read_csv(file_path)

# Filter data for the years 1998 to 2021
filtered_df = df[(df['year'] >= 1998) & (df['year'] <= 2021)]

# Plotting the total production over the years
plt.figure(figsize=(10, 6))
plt.plot(filtered_df['year'], filtered_df['totalprod'], marker='o')
plt.title('Honey Production Yield (1998-2021)')
plt.xlabel('year')
plt.ylabel('totalprod')
plt.grid(True)
plt.show()

```

NAME: DIVYA STUTI
E-MAIL: divyastuti.DS@gmail.com
COURSE NAME: Data Science with Python Career Program (ChatGPT Included)
ASSIGNMENT: ADVANCED PYTHON (MAJOR)



Insight:

- Year 2010 marks the highest production yield.

- During 1998- 2006, huge dots suggest that, honey production was significantly high as compared to honey production during 2010 onwards.

- Honey production declined significantly in 2006 due to decline in honey bee population.

- After 2010, Production yield has decreased due to collapse of bee colonies and increased dependency on imports.

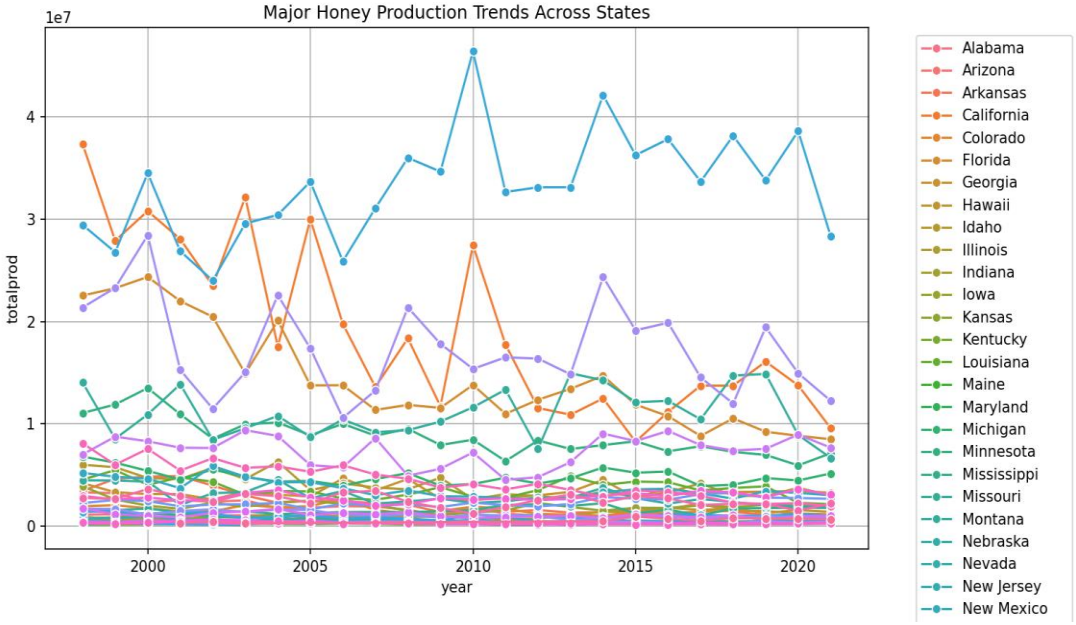
NAME: DIVYA STUTI
E-MAIL: divyastuti.DS@gmail.com
COURSE NAME: Data Science with Python Career Program (ChatGPT Included)
ASSIGNMENT: ADVANCED PYTHON (MAJOR)

Q2) Overtime, what are the major production trends across the states?

honeyproduction 1998-2021.csv | Python Major Assignment.py

Python Major Assignment.py > ...

```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

file_path = 'honeyproduction 1998-2021.csv'

# Read the dataset into a Pandas DataFrame
df = pd.read_csv(file_path)

# Group data by state and calculate mean honey production for each state over time
state_production = df.groupby('State')['totalprod'].mean().reset_index()

# Plotting the major production trends across states
plt.figure(figsize=(16, 12))
sns.lineplot(x='year', y='totalprod', hue='State', data=df, marker='o')
plt.title('Major Honey Production Trends Across States')
plt.xlabel('year')
plt.ylabel('totalprod')
plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left')  # Adjust legend position
plt.grid(True)
plt.show()

```



NAME: DIVYA STUTI
E-MAIL: divyastuti.DS@gmail.com
COURSE NAME: Data Science with Python Career Program (ChatGPT Included)
ASSIGNMENT: ADVANCED PYTHON (MAJOR)

Insights:

- New Mexico, California, and Florida are 3 major producers and contributed most in honey production in USA.

- Alabama and Arizona contribute least to the honey production in USA.

- Overall production in the state of New Mexico remained consistent over the period. On the other hand, in California honey production has declined steeply.

- Total honey production in Florida has also declined.

- In year 2010, New Mexico was the highest Honey Producing state.

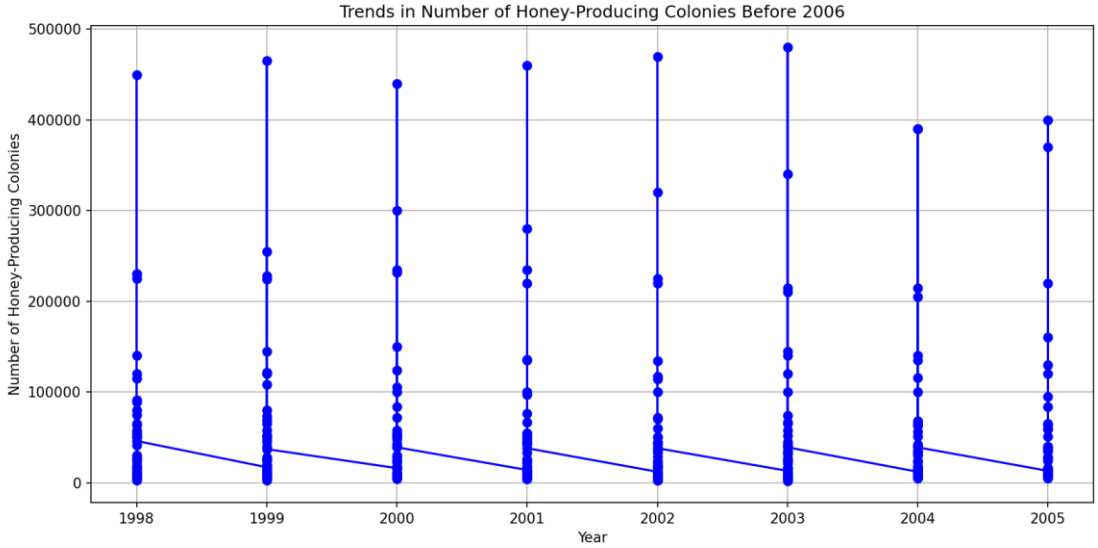
- Honey Production in New Mexico was lesser before 2010 but increased significantly after 2010.

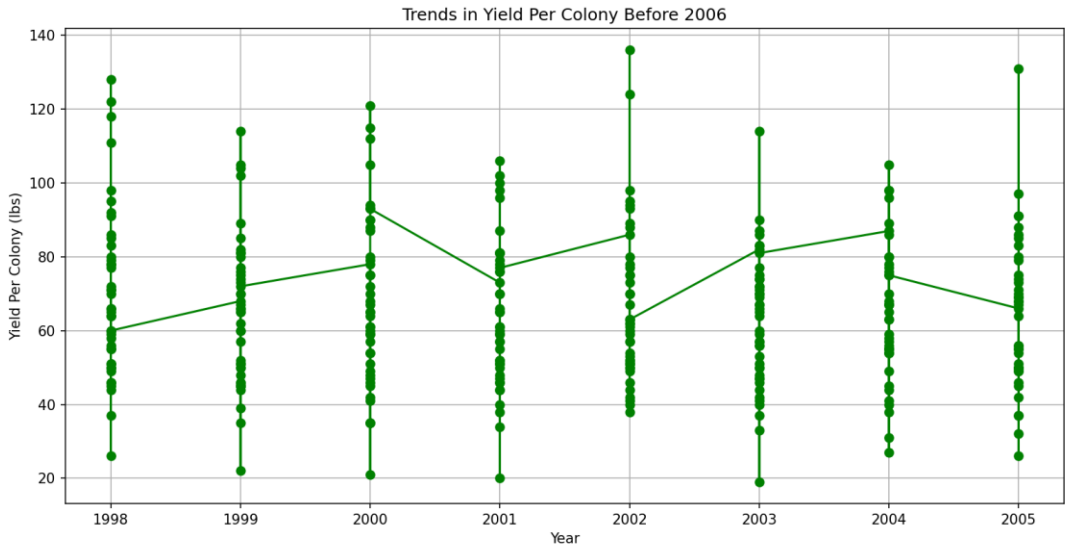
- Honey production yield has reduced for most of the states from 2004 until 2020, except in New Mexico, perhaps due to Colony Collapse Disorder.

- Most prominent honey producing states are- California, Florida, North Dakota and South Dakota and Montana.

Q3) Does the data show any trends in terms of the number of honey-producing colonies and yield per colony before 2006, which was when concern over Colony Collapse Disorder spread nationwide?

NAME: DIVYA STUTI
E-MAIL: divyastuti.DS@gmail.com
COURSE NAME: Data Science with Python Career Program (ChatGPT Included)
ASSIGNMENT: ADVANCED PYTHON (MAJOR)





Insight:

- Number of honey-producing colonies across the country shows a declining trend.

- The yield per colony shows a fluctuating trend. Yield increased during 1998-2000, 2001-2002, 2002-2004 while decreased sharply during 2000-2001 amd 2004-2005.

- Thus, we can infer that, decline in number of honey bee colony is not the only cause for decline in total honey production but the yield per colony is also responsible.

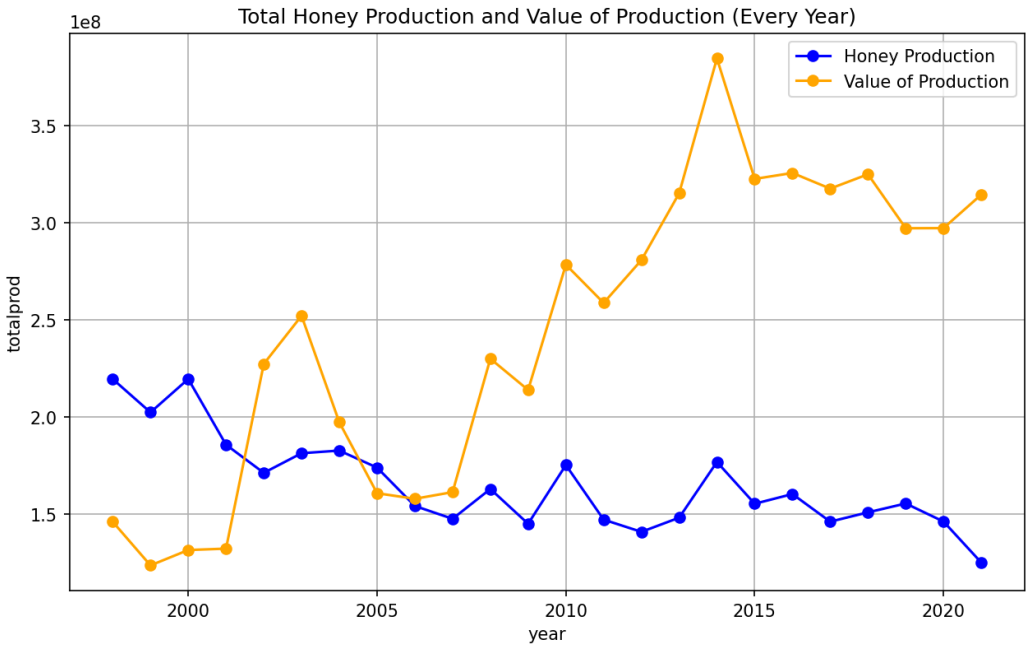
NAME: DIVYA STUTI
E-MAIL: divyastuti.DS@gmail.com
COURSE NAME: Data Science with Python Career Program (ChatGPT Included)
ASSIGNMENT: ADVANCED PYTHON (MAJOR)

Q4) Are there any patterns that can be observed between total honey production and the value of production every year?

honeyproduction 1998-2021.csv | Python Major Assignment.py

Python Major Assignment.py > ...

```python
import pandas as pd
import matplotlib.pyplot as plt

file_path = 'honeyproduction 1998-2021.csv'
df = pd.read_csv(file_path)

# Group the data by year and calculate the total honey production and value of production
yearly_totals = df.groupby('year').agg({'totalprod': 'sum', 'prodvalue': 'sum'}).reset_index()

# Line plot
plt.figure(figsize=(10, 6))
plt.plot(yearly_totals['year'], yearly_totals['totalprod'], marker='o', label='Honey Production', color='blue')
plt.plot(yearly_totals['year'], yearly_totals['prodvalue'], marker='o', label='Value of Production', color='orange')
plt.title('Total Honey Production and Value of Production (Every Year)')
plt.xlabel('year')
plt.ylabel('totalprod')
plt.legend()
plt.grid(True)
plt.show()
```



NAME: DIVYA STUTI
E-MAIL: divyastuti.DS@gmail.com
COURSE NAME: Data Science with Python Career Program (ChatGPT Included)
ASSIGNMENT: ADVANCED PYTHON (MAJOR)

Insight:

- Overall honey production has been declining over the years.

- During 1998-2001, value of production was less as compared to honey production. It increased sharply in 2002 and remained high during 2002-2003.

- 2006 onwards, value of production witnessed an increasing trend on the other hand, honey production witnessed a decreasing trend.

- After 2010, Value of production has sharply increased and remained high. On the other hand, Honey Production has been consistently decreasing.

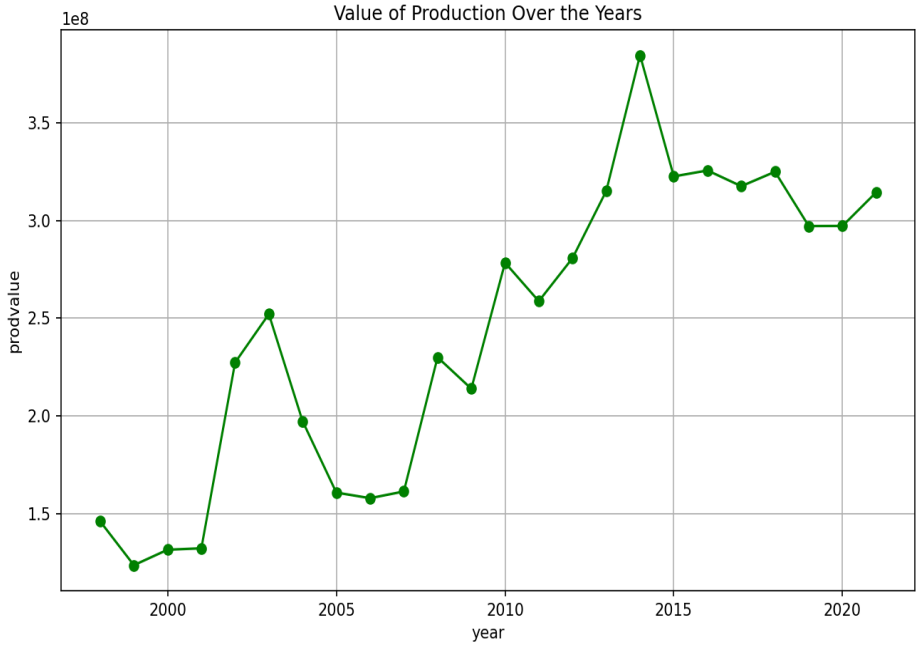
Q5) How has the value of production, which in some sense could be tied to demand, changed every year?

honeyproduction 1998-2021.csv | Python Major Assignment.py

Python Major Assignment.py > ...

```
1
2   import pandas as pd
3   import matplotlib.pyplot as plt
4
5   file_path = 'honeyproduction 1998-2021.csv'
6   df = pd.read_csv(file_path)
7
8   # Group the data by year and calculate the total value of production
9   yearly_value = df.groupby('year')['prodvalue'].sum().reset_index()
10
11  # Line plot for the value of production over the years
12  plt.figure(figsize=(10, 6))
13  plt.plot(yearly_value['year'], yearly_value['prodvalue'], marker='o', color='green')
14  plt.title('Value of Production Over the Years')
15  plt.xlabel('year')
16  plt.ylabel('prodvalue')
17  plt.grid(True)
18  plt.show()
19
```

NAME: DIVYA STUTI
E-MAIL: divyastuti.DS@gmail.com
COURSE NAME: Data Science with Python Career Program (ChatGPT Included)
ASSIGNMENT: ADVANCED PYTHON (MAJOR)



Insight:

- Value of production per pound has increased over time, even though honey production has declined.

- Due to reduced honey production, supply has declined. But value of production has increased due to demand of honey.