**Name: Shivaram Babar**
**Email: shivaramb1702@gmail.com**
**Course Name: Data Science with Python Career Program (ChatGPT Included)**
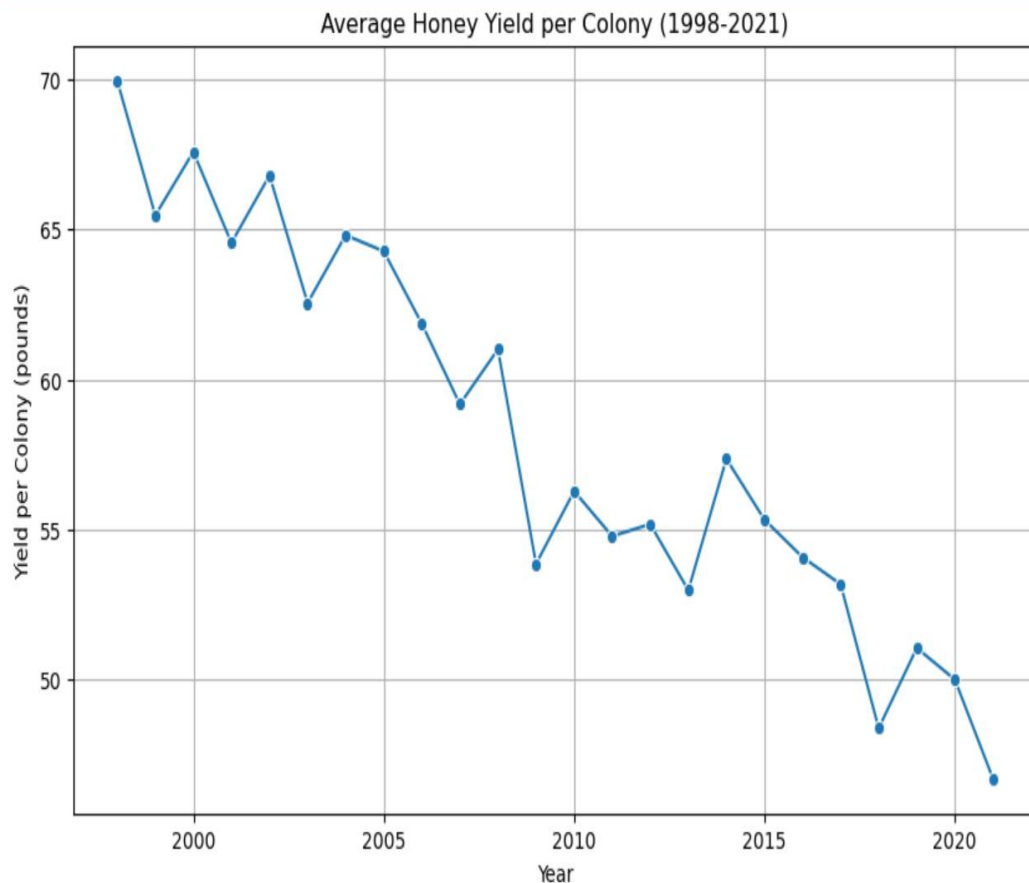**Assignment Tittle : Advanced Python [Major]**

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

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

yield_per_year = data.groupby('year')['yieldpercol'].mean().reset_index()

plt.figure(figsize=(10, 6))
sns.lineplot(x='year', y='yieldpercol', data=yield_per_year, marker='o')
plt.title('Average Honey Yield per Colony (1998-2021)')
plt.xlabel('Year')
plt.ylabel('Yield per Colony (pounds)')
plt.grid(True)
plt.show()
```


Average Honey Yield per Colony (1998-2021)

## Q2) Over time, what are the major production trends across the states?

```python
#Q2) Over time, what are the major production trends across the states?

if 'totalprod' in data.columns:
    # Sum total production by state and year
    total_production_by_state = data.groupby(['year', 'State'])['totalprod'].sum().reset_index()

    # Identify top 10 states with the highest total production overall
    top_states = total_production_by_state.groupby('State')['totalprod'].sum().nlargest(10).index

    # Filter data to include only the top states
    filtered_data = total_production_by_state[total_production_by_state['State'].isin(top_states)]

    plt.figure(figsize=(15, 10))
    sns.lineplot(data=filtered_data, x='year', y='totalprod', hue='State', marker='o')
    plt.title('Total Honey Production Across Top States (1998-2021)')
    plt.xlabel('Year')
    plt.ylabel('Total Production (pounds)')
    plt.legend(title='State', bbox_to_anchor=(1.05, 1), loc='upper left')
    plt.grid(True)
    plt.show()
else:
    print("Column 'totalprod' not found in the dataset.")
```
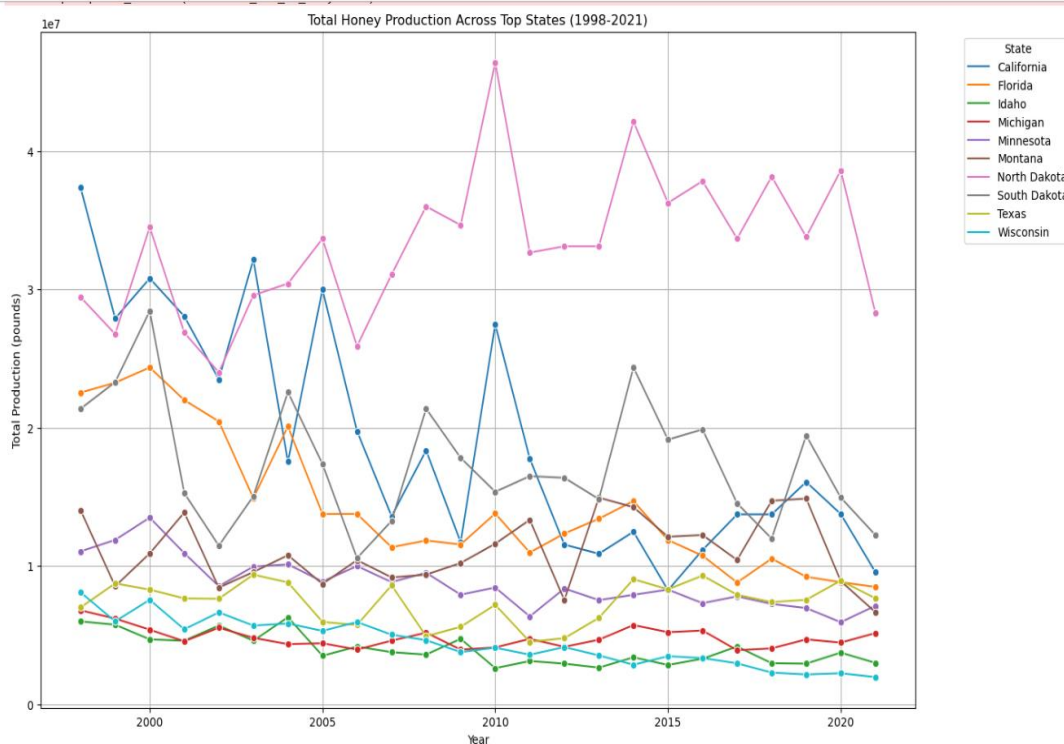
**Name: Shivaram Babar**
**Email: shivaramb1702@gmail.com**
**Course Name: Data Science with Python Career Program (ChatGPT Included)**
**Assignment Tittle : Advanced Python [Major]**

**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?**
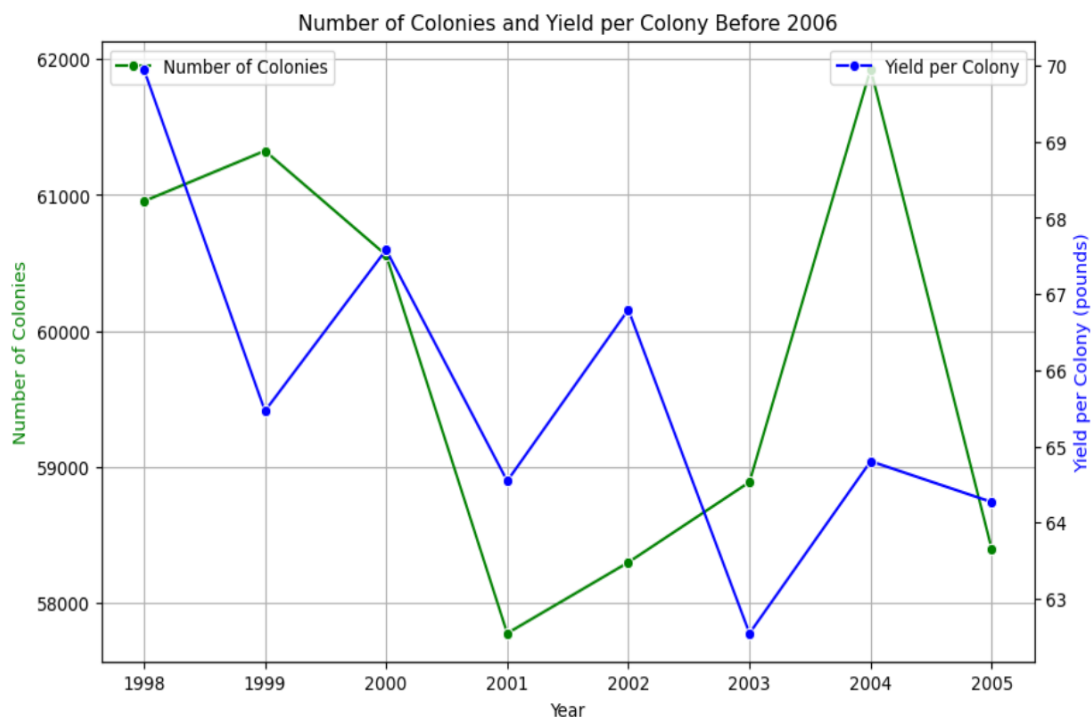
```python
#Q3) Does the data show any trends in terms of the number of honey-producing colonies and yield per colony before 2006,which was when concern over Colony

pre_2006_data = data[data['year'] < 2006]
pre_2006_summary = pre_2006_data.groupby('year').agg({'numcol': 'mean', 'yieldpercol': 'mean'}).reset_index()

fig, ax1 = plt.subplots(figsize=(10, 6))

ax2 = ax1.twinx()
sns.lineplot(x='year', y='numcol', data=pre_2006_summary, ax=ax1, color='g', marker='o', label='Number of Colonies')
sns.lineplot(x='year', y='yieldpercol', data=pre_2006_summary, ax=ax2, color='b', marker='o', label='Yield per Colony')

ax1.set_xlabel('Year')
ax1.set_ylabel('Number of Colonies', color='g')
ax2.set_ylabel('Yield per Colony (pounds)', color='b')
plt.title('Number of Colonies and Yield per Colony Before 2006')
ax1.grid(True)
plt.show()
```

**Name: Shivaram Babar**
**Email: shivaramb1702@gmail.com**
**Course Name: Data Science with Python Career Program (ChatGPT Included)**
**Assignment Tittle : Advanced Python [Major]**

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

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

production_value = data.groupby('year').agg({'totalprod': 'sum', 'prodvalue': 'sum'}).reset_index()

fig, ax1 = plt.subplots(figsize=(10, 6))

ax2 = ax1.twinx()
sns.lineplot(x='year', y='totalprod', data=production_value, ax=ax1, color='r', marker='o', label='Total Production')
sns.lineplot(x='year', y='prodvalue', data=production_value, ax=ax2, color='b', marker='o', label='Production Value')

ax1.set_xlabel('Year')
ax1.set_ylabel('Total Production (pounds)', color='r')
ax2.set_ylabel('Production Value (dollars)', color='b')
plt.title('Total Production and Production Value (1998-2021)')
ax1.grid(True)
plt.show()
```
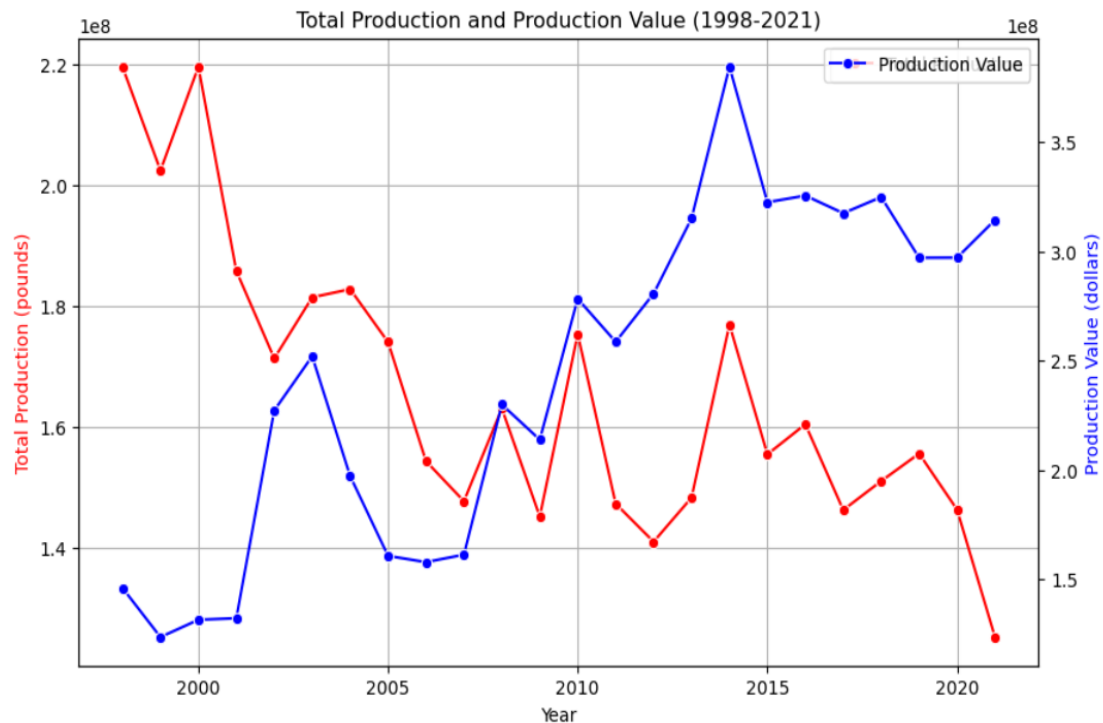


Total Production and Production Value (1998-2021)

**Name: Shivaram Babar**
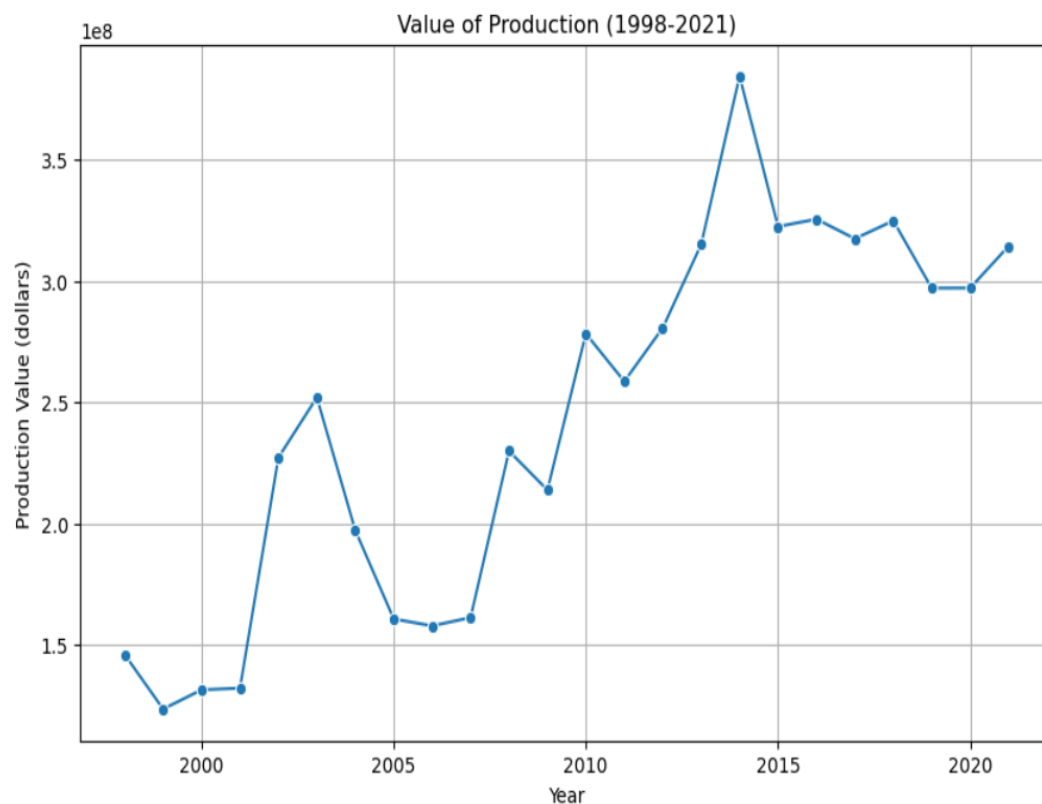**Email: shivaramb1702@gmail.com**
**Course Name: Data Science with Python Career Program (ChatGPT Included)**
**Assignment Tittle : Advanced Python [Major]**

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

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

plt.figure(figsize=(10, 6))
sns.lineplot(x='year', y='prodvalue', data=production_value, marker='o')
plt.title('Value of Production (1998-2021)')
plt.xlabel('Year')
plt.ylabel('Production Value (dollars)')
plt.grid(True)
plt.show()
```

## Q6) Construct there late dp lots using Seaborn and Matplot apply customization and derive insights from the visualization.

*#Q6)Construct there late dp lots using Seaborn and Matplot apply customization and derive insights from the visualization.*

```python
fig, ax1 = plt.subplots(figsize=(12, 8))

sns.lineplot(x='year', y='totalprod', data=production_value, ax=ax1, color='r', marker='o', label='Total Production')
sns.lineplot(x='year', y='numcol', data=data.groupby('year')['numcol'].mean().reset_index(), ax=ax1, color='g', marker='o', label='Number of Colonies')
sns.lineplot(x='year', y='yieldpercol', data=yield_per_year, ax=ax1, color='b', marker='o', label='Yield per Colony')

ax1.set_xlabel('Year')
ax1.set_ylabel('Values')
plt.title('Total Production, Number of Colonies, and Yield per Colony (1998-2021)')
plt.legend()
ax1.grid(True)
plt.show()
```