**Introduction & Purpose of the Project**

Honey bees play an essential role in pollinating plants, but they are facing serious threats from habitat loss, pesticides, and diseases. This dataset aims to track the health of bee populations across the United States over several decades, providing valuable insights into their decline and factors influencing it. The goal of this project is to create an interactive data visualization application that helps users explore key trends in honey bee population data, such as the number of colonies, losses due to various factors, and the factors correlated to colony loss.

**What Dataset? Motivation, and Data Cleaning**

The combination of data applications and interactive visualizations allows users to explore complex datasets, like the one concerning honey bee populations in the U.S., which is critical for understanding environmental issues. By examining the decline of bee populations, policies can be shaped to protect them. The process involved cleaning the bees dataset from Kaggle, merging it with the UsGeoCode data to include longitude and latitude for each state, and then filtering out states lacking this data. The cleaned data was stored in an SQL engine, and a visualization dataframe was created and connected to SQLite. The analysis answered key questions about the causes and trends of bee colony losses, the effectiveness of colony renovation, and the impact of pesticides, among other factors.

The analysis aimed to answer several key questions about honey bee colony losses, including identifying the leading causes of colony loss over time, determining which states have the worst losses, and exploring the trend of colony losses by quarter. Additionally, it sought to evaluate the effectiveness of renovating ("requeening") colonies in reducing losses, examine which factors are most correlated with colony loss, and understand the impact of pesticide use on bee populations. These questions guided the creation of visualizations to better understand the factors affecting bee populations and to inform potential solutions.

**Walkthrough of Your Data App (with Screenshots)**

* **Overview:**  
  The data app allows users to filter data based on year, state, and other factors. It displays a bar chart to visualize colony losses over time and presents detailed information per state and cause of colonies loss.
* **Screenshots:**
  + Screenshot 1: The main dashboard, showing the initial view of the app with interactive filters.

A screenshot of a web page

Description automatically generated

* + Screenshot 2: A filtered view of the data, after applying a filter for a specific year or region.

A screenshot of a computer

Description automatically generated

* + Screenshot 3: A detailed table displaying the data, with sortable columns.

A screenshot of a data table

Description automatically generated

* + Screenshot 4: The heatmap with the population data visualized for different states.

A map of the united states

Description automatically generated

* + Screenshot 5: Loss bees colonies per quarter

A screenshot of a computer

Description automatically generated

**Discuss Dashboard Design**

**Layout:** The layout is structured for simplicity and clarity, with filter options positioned at the top for easy access. The data visualizations, including a table, bar chart, donut bar, and line chart, are displayed prominently below for quick viewing.

**Colors and Design:** The design follows a bee-themed aesthetic to align with the dataset's focus. The bar chart, donut bar, and line chart are color-coded by state, making it easy to differentiate between the data points.

**Interactive Features:** Users can filter the data by year, state, and other variables. By hovering over the bar chart, donut bar, or line chart, additional information is displayed according to the selected year filter.

**Discuss Questions That Your Dashboards Help Answer**

**Question 1: What are the top causes of bee colony loss over time?**

**Question 2: Which states have the worst bee colony losses?**

**Question 3: What is the trend of bee colony losses per quarter?**

**Question 4: Is renovating ("requeening") colonies effective in reducing loss?**

**Question 5: Which factors are most correlated with colony loss?**

**Question 6: What impact does pesticide use have on bee colony loss?**

**Limits & Bias of Data**

* **Data Gaps:**  
  Some states or years may have incomplete data, which could introduce gaps in the analysis. Additionally, some years might not be representative due to missing data.
* **Bias in Data Collection:**  
  The data might be biased if certain states have more detailed reporting than others, or if certain factors (e.g., pesticide use or pest management) are underreported.
* **Sample Size:**  
  The data might only represent a subset of the entire U.S., and results may not be generalizable to other regions or countries.

**Conclusions & Future Work**

Summarize your findings and mention any plans for further improvements or exploration.

* **Conclusions:**  
  The interactive visualization clearly highlights the trends in honey bee colony losses and factors that contribute to their decline. Key insights include identifying regions with the most significant declines and understanding which issues (e.g., disease, pesticide use) affect colonies the most.
* **Future Work:**  
  Future improvements could involve incorporating real-time data, adding more advanced filtering options, and including more complex visualizations like trend lines to compare multiple factors simultaneously.