**Introduction & Purpose of the Project**

**Intro:**  
*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**

*Data applications combined with interactive visualizations allow users to explore and make sense of complex datasets. The dataset sheds light on the state of honey bee populations in the U.S., which is an urgent environmental issue. With the current decline in bee populations, understanding the factors contributing to their decline can inform agricultural and environmental policies aimed at preserving them. First, started cleaning the data by first saving the bees dataset from Kaggle and merged the data to UsGeoCode to get longitude and latitude for each state. Also, dropped the states with no longitude and latitude and then save it to SQL engine. Next, created the visualization data frame and connect it to sqlite.*

*Through the visualization dataframe, we answered our proposal questions and created graphs to show our results.*

***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?***

**4. 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.
  + Screenshot 2: A filtered view of the data, after applying a filter for a specific year or region.
  + Screenshot 3: A detailed table displaying the data, with sortable columns.
  + Screenshot 4: The bar chart with the population data visualized for different states.

**5. Discuss Dashboard Design**

* **Layout:**  
  *The layout is designed to prioritize ease of use and clarity. The filter options are located at the top, while the data visualizations (table and bar chart,map) are prominently displayed below for quick access.*
* **Colors and Design:**  
  *We used a bee-themed layout to match our dataframe. The bar chart is color-coded by state to easily distinguish the data.*
* **Interactive Features:**  
  *Users can filter the data based on year, state, and other factors. Hovering over the bar chart displays additional information, and users can sort the data in the table by clicking on the column headers.*

**6. 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?***

**7. 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.*

**8. 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.*