SeasonWatch Project

Authors: Cecily Wang-Munoz, Aditya Chopra, Ngo Ngoc Thuy An, Brenda Kim

Supported by BU SPARK!

Fall 2024

Table of Contents

Project Overview

The SeasonWatch project investigates phenological changes in tree species across India, to understand potential climate-driven impacts. The goal is to identify shifts in timing and growth patterns for key phenological stages, such as flowering and fruiting, and to track changes in response to seasonal and climate variables.

Abstract

To assess the potential impacts of climate change on tree phenological stages, the SeasonWatch project investigates flowering, fruiting, and other seasonal stages in tree species across India, focusing on Kerala. Phenological events are essential indicators of ecosystem health and biodiversity, and their shifts can profoundly impact ecological interactions, including pollination, fruiting cycles, and species distribution. This study combines citizen-submitted observations with historical datasets to assess temporal changes over the past decade, providing insights into how tropical tree species respond to changing environmental conditions.

The analysis reveals advanced statistical and computational methodologies, including regression analysis to quantify relationships between phenological shifts and climate variables. Survival analysis is used to model the timing and speed of transitions between phenological stages, offering insights into the responsiveness of species to seasonal cues. Markov models estimate the probabilities of transitions between discrete phenological states, while clustering techniques group species based on shared phenological patterns, identifying trends of resilience or vulnerability across different species.

Our results reveal significant shifts in onset times for key phenological stages, with earlier flowering and fruiting observed predominantly in the summer and monsoon seasons. Statistical analyses demonstrate strong correlations between these shifts and climate trends, such as rising temperatures and altered precipitation patterns. Survival analysis indicates an accelerated pace of phenological transitions, while Markov modeling highlights predictable changes in state transition probabilities, suggesting an adaptive response to climate variability. Clustering analysis further distinguishes species with differing responses, identifying those most susceptible to environmental stressors and those demonstrating resilience.

This comprehensive approach highlights the critical role of citizen science in tracking large-scale ecological changes, offering a scalable approach to collecting phenological data. By identifying discrepancies between citizen observations and reference datasets, the project improves the accuracy of phenological monitoring and refines baseline datasets for future research. By focusing on vulnerable species and regions showing pronounced phenological shifts, such as the monsoon-dominated areas, this work contributes to the broader understanding of climate-driven ecological dynamics and supports targeted adaptation measures to mitigate the impacts of climate change on biodiversity.

---

2. **Documentation - Github Notebook Workflow**

```
/DS-Seasonwatch-Trees/
├── code
    |── Fall 2024 Code
        ├── Clustering_analysis_Kerala.ipynb          // Species Clustering Analysis
        ├── NaN_values.py                             //Removing N/A values from data
        ├── data_analysis.ipynb                       //Core analyses of species data
        ├── data_cleaning_version2.ipynb              //Initial data cleaning and processing
        ├── googlemaps.ipynb                //Geocoding and handling missing location data
        ├── mock_analysis
        ├── powerbi_visual_csv_creation.ipynb         //csv files used in powerbi
├── data
    |── Fall 2024 data
        ├── INDIA_top_30_species.csv
        ├── POWERBI_india_all_states.csv
        ├── POWERBIkerala_weekly_shift_analysis.csv
        ├── Clustering_analysis_Kerala.ipynb
        ├── POWERBIkeralaseasonwatch_map_analysis_top_30.csv
        ├── POWERBIphenophase_data collection accuracy.csv
        ├── avg_phenological_timing_data.csv
        ├── cleaned_alldata_version2.csv
        ├── seasonwatch_phenophase_shifts_analysis.csv
        ├── seasonwatch_processed_data.csv
        ├── summary_statistics_top_30_species.csv
        ├── top_30_species_all_data.csv
|── VISUALIZATIONS- fall 2024
        ├──average_onset_shift_by_season.csv
        ├── mango_visualizations_testing.ipynb        //Test visuals for mango species
        ├── onset_shifts_top_30_species.csv
        ├── slope_comparisons_all_30_species.csv
        ├── slope_comparisons_top_30_species.csv
        ├── top30species_geo_analysis_.ipynb
        ├── clustering_analysis_Kerala-3.ipynb
        ├── top_30_species_visual.ipynb               //Visualizations for the top 30 species
└── README.MD
└── Seasonwatch Final Report
```

---

3. Methodology

Data Processing

3.1 Initial Steps

- **Location Data Filling**:
  - Used Google Maps API to fill missing state names for rows with latitude/longitude data.
  - Implemented:
    - Caching: Minimized redundant API calls (20-hour runtime).
    - Error Handling: Handled failed requests and applied rate limits.
  - Standardization: Unified state name formats.
- **Handling Missing Data**:
  - Prior team dropped 200,000 rows unnecessarily.
  - Revised approach: Only 20,000 rows (<4%) with missing lat/long data were removed.

3.2 Top 30 Observed Species

- Using visualizations.ipynb, the top 30 species with the highest citizen observation counts were identified for focused analysis:

Jackfruit-Artocarpus heterophyllus	Rain tree-Albizia saman	Wood Apple-Aegle marmelos	Purple Bauhinia-Bauhinia purpurea
Mango (all varieties)-Mangifera indica	Chiku Sapodilla-Manilkara zapota	Country Fig-Ficus racemosa	Maulsari-Mimusops elengi
Teak-Tectona grandis	Banyan-Ficus benghalensis	Drumstick tree-Moringa oleifera	Gulmohur-Delonix regia
Tamarind-Tamarindus indica	Peepal-Ficus religiosa	Red Silk Cotton-Bombax ceiba	Pongam Tree-Pongamia pinnata
Indian Laburnum-Cassia fistula	Guava tree-Psidium guajava	Indian Almond-Terminalia catappa	True Ashoka-Saraca asoca
Amla-Phyllanthus emblica	Devil's Tree-Alstonia scholaris	Custard apple-Annona squamosa	
Jamun-Syzygium cumini	Chandada-Macaranga peltata	Gamar-Gmelina arborea	
Coconut palm-Cocos nucifera		Copper-pod-Peltophorum pterocarpum	
Neem-Azadirachta indica		Pride of India-Lagerstroemia speciosa	

---

4. Base Questions and Analysis Methods

1. How are trees changing because of climate change?

- **Regression Analysis**:

- **Objective**: Model the relationship between tree phenology (leafing, flowering, fruiting) and climate variables, such as temperature and precipitation, to identify long-term trends and climate impacts.
- **Method**: Perform linear regression where the onset time of each phenological stage is the dependent variable and climate variables are independent variables.
- **Visualization**: Generate trend lines for each phenology stage over time, with overlays of climate variable trends. This will illustrate if there's a significant change in onset timing as climate conditions evolve.

2. How fast do trees change in response to changing seasons? What is the onset time for flowering and fruiting in tropical species?

- **Descriptive Statistics**:
  - **Objective**: Calculate the average onset time of each phenological stage across species, seasons, and states.
  - **Method**: Compute summary statistics (mean, median, variance, and standard deviation) for onset times of leafing, flowering, and fruiting stages by species and region.
  - **Output**: Summary tables showing: `Species, State, Avg Leafing Time, Avg Flowering Time, Avg Fruiting Time`.
- **Survival Analysis**:
  - **Objective**: Estimate how quickly trees transition through different phenology stages within a season.
  - **Method**: Use survival analysis to determine the probability that trees will reach specific phenology stages by a certain point in the season, giving insight into the responsiveness of tree phenology to seasonal cues.

3. What is the probability of transition from one seasonal state of a tree to another?

- **Markov Model**:
  - **Objective**: Model the probability of trees transitioning between seasonal states (e.g., leaf buds to mature leaves).
  - **Method**: Define states as discrete stages (`none`, `few`, `many`) and build a transition matrix from historical data to represent the probabilities of moving between these states.
  - **Extension**: Simulate future probabilities of state transitions for each species based on the current dataset, which may reveal if climate changes are altering these transition probabilities.

4. Comparing Citizen Observations with Reference Data

- **Correlation Analysis**:
  - **Objective**: Quantify how closely citizen-input data aligns with reference data for phenology stages.
  - **Method**: Calculate correlations between citizen observations and reference data for each phenological stage.
- **Mean Error Calculation**:

- **Objective**: Measure the degree of deviation in citizen observations from the reference data.
- **Method**: Calculate mean error and standard deviation to assess the accuracy of citizen-submitted data.

- **Clustering Analysis**:
  - **Objective**: Group species with similar patterns in citizen and reference datasets to identify alignment or discrepancies.
  - **Method**: Perform clustering to group species by phenological patterns, then visualize clusters through scatter plots or heatmaps to highlight alignment and discrepancies between datasets.

5. Additional Questions and Visualizations

- **Shifts in Onset Times Over Years**:
  - **Time Series Analysis**: Track the average onset of flowering and fruiting over multiple years to determine if these events are occurring earlier or later, potentially influenced by climate change.
  - **Visualization**: Use line plots with trend lines to show onset shifts over time, including climate variable overlays for direct comparison.
- **Distribution of Onset Times by Species and State**:
  - **Box Plots**: Show the distribution of onset times across species and states, illustrating variations in timing.
- **Transition Probabilities and Clustering**:
  - **Markov Model with Transition Matrix**: For each species, model the likelihood of transitioning between states (e.g., flowering onset to fruiting) within a season, providing insights into phenological patterns.
  - **Clustering Visualization**: Scatter or heatmaps can display clusters of species with similar patterns, helping identify species whose timing may be especially climate-sensitive.

---

4. PowerBI Interactive Map

Link to the interactive PowerBI visual: PowerBI

1. **Objective**
   - How can we efficiently determine the usability and validity of large-scale tree behavior data for analysis?
2. **Final PowerBI Result**
   - The cleaned dataset was added into a PowerBI interactive visual that allows the dataset to be dynamically modified. Users are able to filter data points by year, state or the type of species in order to see how much data is within a certain region.
     - When a user applies a filter, all related charts, tables, or graphs update immediately to reflect the filtered dataset.

- The primary objective of this visualization was to display the distribution of data points across specific regions or years, enabling users to assess whether there is sufficient data for meaningful analysis.

5. Mock Analysis Outline with Expected Deliverables

1. **Summary Statistics Table**:
   - Tables summarizing the average timing for each phenological stage (leaf, flower, fruit) for each species and state, providing baseline timing information.
2. **Regression Analysis Output**:
   - Regression results showing the relationship between climate variables and onset timing.
   - Visualization: Trend line plots showing phenology shifts over time with climate overlays.
3. **Survival Analysis Results**:
   - Survival curves indicating the probability of reaching each phenology stage by season.
4. **Markov Model Transition Matrix**:
   - Transition matrix with probabilities of state changes within seasons.
5. **Citizen vs. Reference Data Comparison**:
   - Correlation and mean error metrics between citizen observations and reference data.
   - Clustering visualizations showing alignment and discrepancies in phenology timing between datasets.

---

6. Analysis Code Examples from Github

6.1 Seasonal Onset Shift Analysis

- **Objective**: Identify shifts in onset timing for mature leaves between the reference period (2014-2020) and post-2020.
- **Method**: Calculated average onset week for each species and season and measured shifts in weeks.

Flowchart:

```
1.  Load Data
●   Load the CSV file onset_shifts_top_30_species.csv into a DataFrame
    (onset_shifts_df).
2.  Copy DataFrame
●   Create a copy of the original DataFrame and assign it to
    significant_onset_shifts.
3.  Calculate Absolute Onset Shifts
●   Add a new column Absolute Onset Shift to the significant_onset_shifts
    DataFrame, which contains the absolute values of the Onset Shift (Weeks)
    column.
4.  Sort Data
●   Sort the significant_onset_shifts DataFrame in descending order based on the
    Absolute Onset Shift column.
5.  Export Results
```

- `Save the sorted DataFrame to a new CSV file named onset_shifts_all_30_species.csv. Ensure the index is excluded during the export.`

- **Findings**: Species such as *Copper-pod* and *Coconut palm* demonstrated early onset shifts in the Summer and Monsoon, suggesting a potential climate adaptation response.

2. Trend Slope Comparison Analysis

- **Objective**: Compare trend slopes for growth patterns between reference and post-2020 periods.
- **Method**: Calculated slope of mature leaf observations over time for each season and species, and measured changes in trend slope.

Flowchart:

1. **Load Data**
   - Load the CSV file `slope_comparisons_top_30_species.csv` into a DataFrame (`slope_comparisons_df`).
2. **Copy DataFrame**
   - Create a copy of the original DataFrame and assign it to `significant_trend_changes`.
3. **Calculate Absolute Slope Differences**
   - Add a new column `Absolute Slope Difference` to the `significant_trend_changes` DataFrame, which contains the absolute values of the `Slope Difference` column.
4. **Sort Data**
   - Sort the `significant_trend_changes` DataFrame in descending order based on the `Absolute Slope Difference` column.
5. **Export Results**
   - Save the sorted DataFrame to a new CSV file named `slope_comparisons_all_30_species.csv`. Ensure the index is excluded during the export.

- **Findings**: Species like *Red Silk Cotton* and *Coconut Palm* exhibited significant decreases in trend slopes, especially in Winter and Monsoon, hinting at possible environmental stress.

3. Seasonal Average Shifts by Species and Season

- **Objective**: Determine which seasons show the most consistent shifts across species.
- **Method**: Aggregated onset shifts by season to find seasonal patterns.

Flowchart:

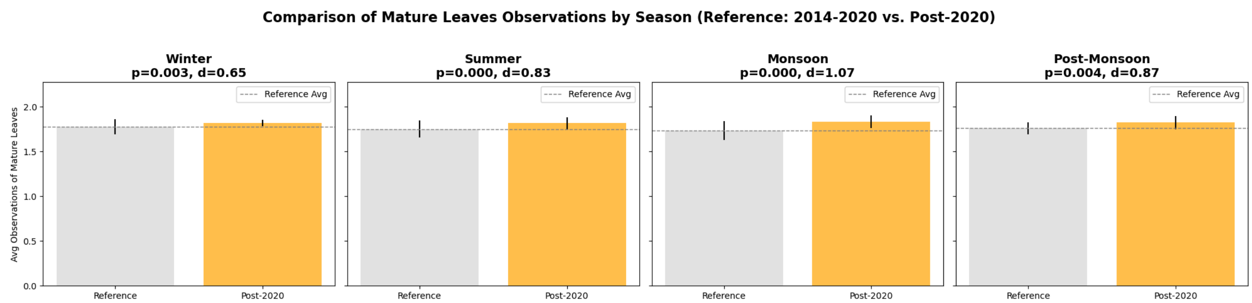
1. **Group Data by Season**
   - Group the `onset_shifts_df` DataFrame by the `Season` column.
2. **Calculate Mean Onset Shift**
   - Compute the average (`mean`) of the `Onset Shift (Weeks)` column for each season.
3. **Reset Index**

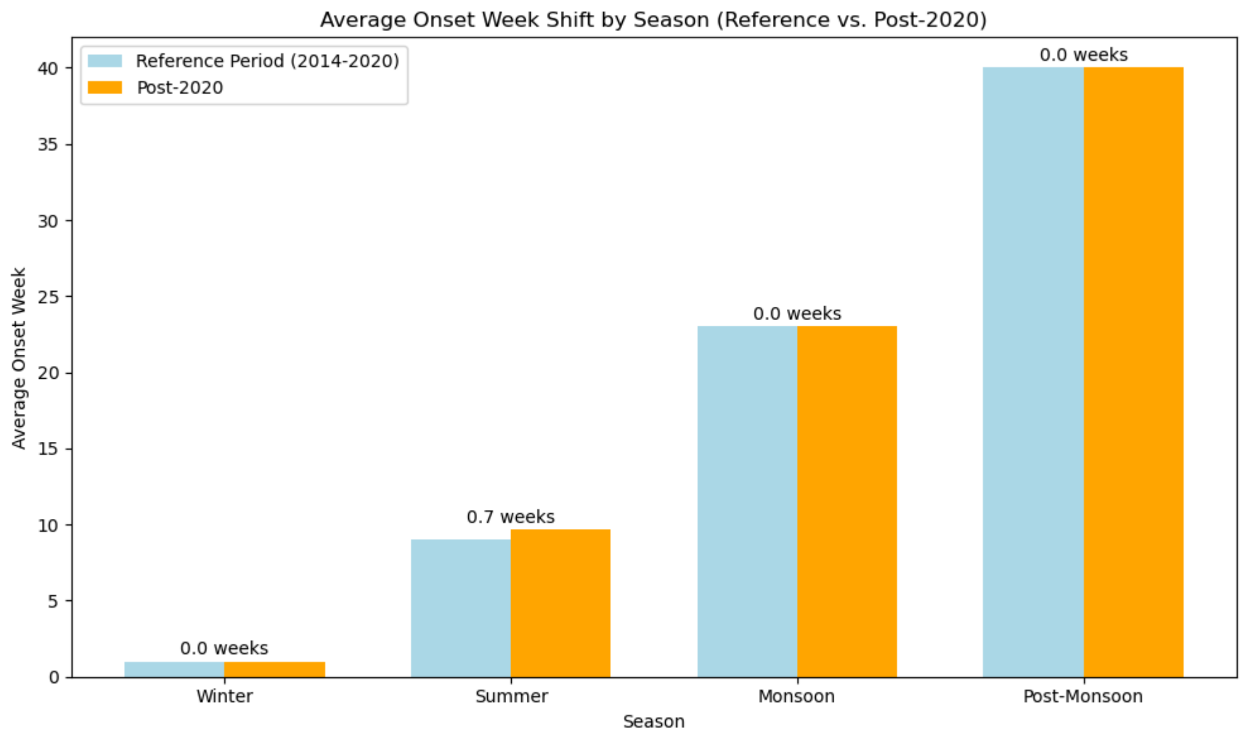
- Convert the grouped data into a standard DataFrame using `.reset_index()` to make it easier to handle.

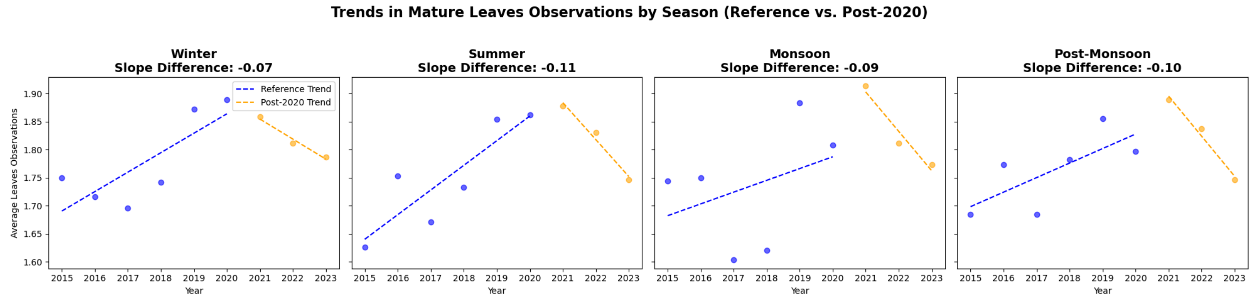
4. **Export Results**
   - Save the resulting DataFrame (`seasonal_onset_shift_avg`) to a CSV file named `average_onset_shift_by_season.csv`. Ensure the index is excluded during the export.

- Findings: Monsoon and Winter show slight negative onset shifts, indicating possible trends toward earlier growth, while Summer showed a slight positive shift.

below images are derived from 'mango_visualizations.ipynb' and only represent MANGO SPECIES ONLY







7. Detailed Analyses and Code Implemented

PRELIMINARY Base Question Findings

1. How are trees changing because of climate change?

By comparing the onset shifts and trend slopes between the reference period (2014-2020) and post-2020, we observe that some species, particularly **Copper-pod** and **Coconut palm**, have shown earlier onset weeks, especially in the Summer and Monsoon seasons. These earlier growth patterns could be an adaptation to changing climate conditions, like warmer temperatures or shifting rainfall patterns.

2. What is the onset time for flowering and fruiting in tropical species?

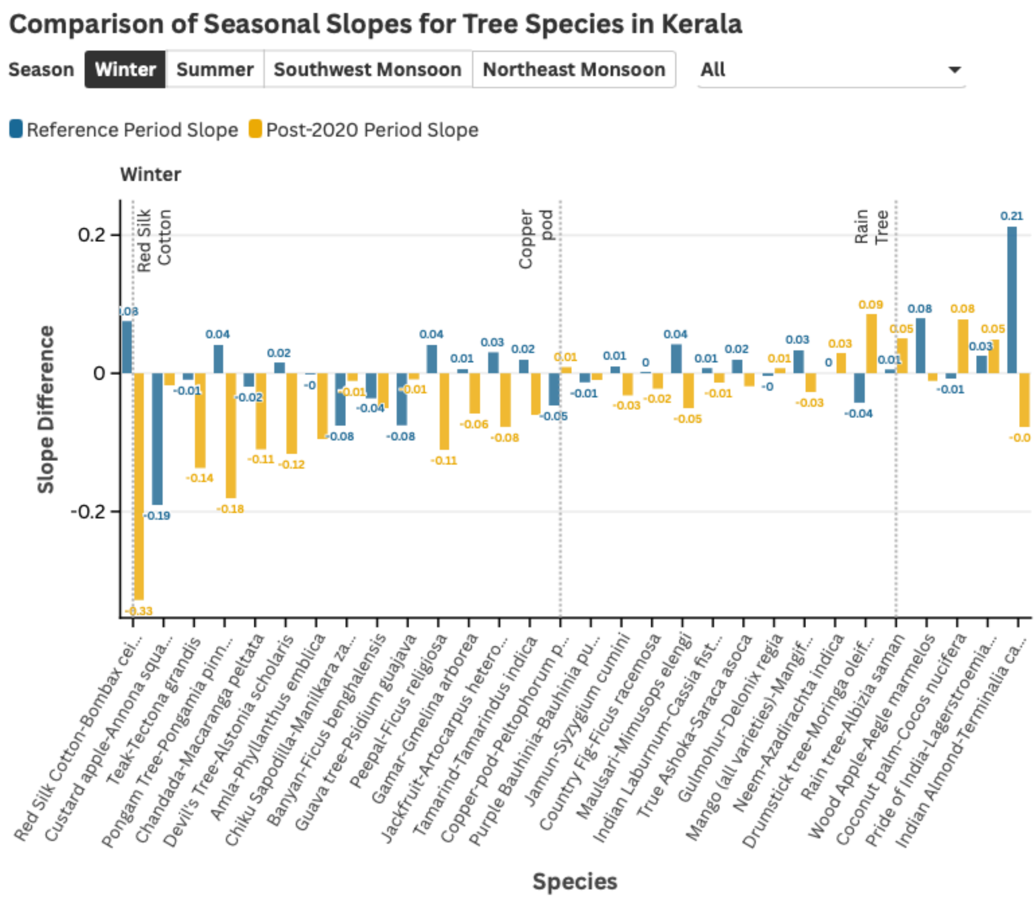
The onset shifts calculated for the top 30 observed species provide an average onset week for each season, allowing us to identify how early or late these phenological stages occur on average. For example, Summer onset shifts were generally earlier in post-2020 data, which could suggest a trend towards earlier development in response to climate signals.

3. How fast do trees change in response to changing seasons?

The slope comparisons reveal how quickly growth trends are changing over time, with several species showing decreased trend slopes in the post-2020 period. This suggests that certain species might be experiencing slowed or reduced growth rates, possibly due to altered seasonal conditions or increased environmental stress.

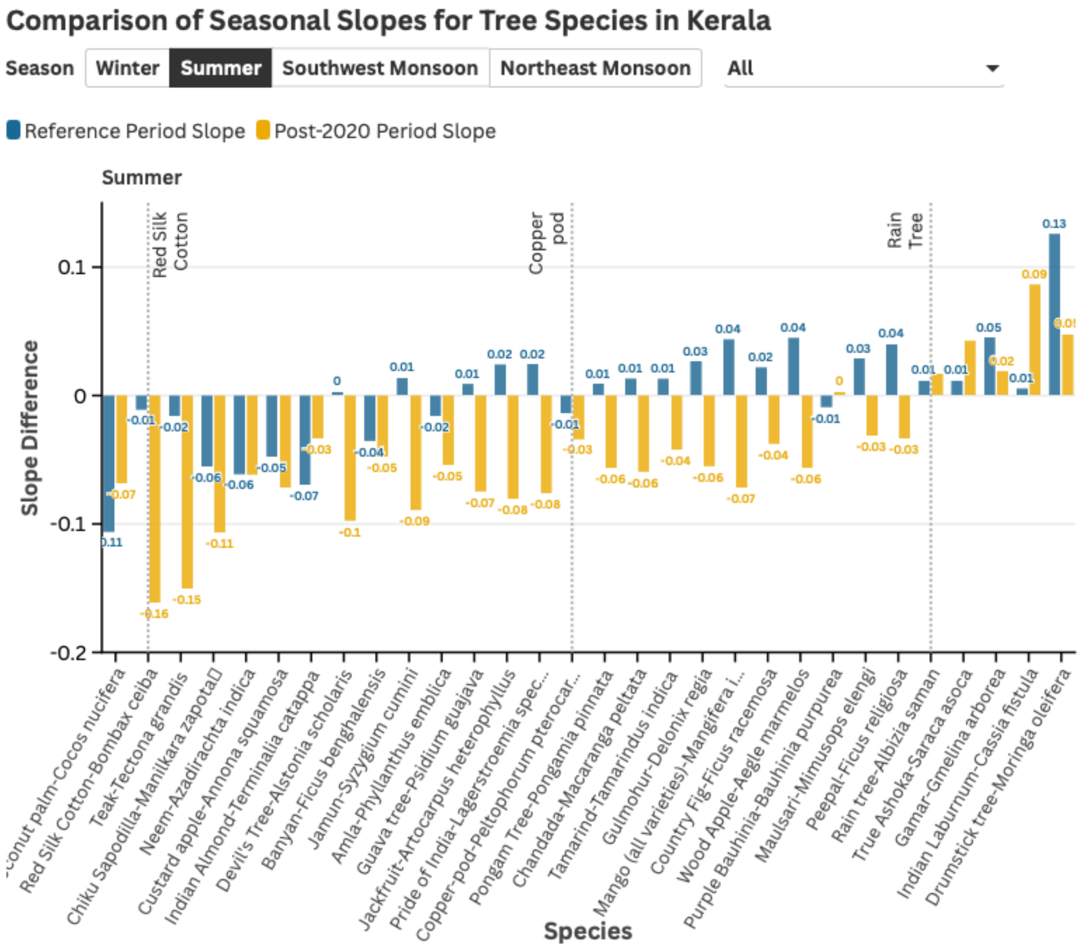
Winter Season

- ***Major Negative Shifts:*** Species like **Coconut Palm** and **Red Silk Cotton** show significant negative slope changes (possibility for delayed phenological events).
- ***Moderate Positive Shifts:*** Few species, such as **Rain Tree**, exhibit positive slope changes, indicating earlier phenological events.



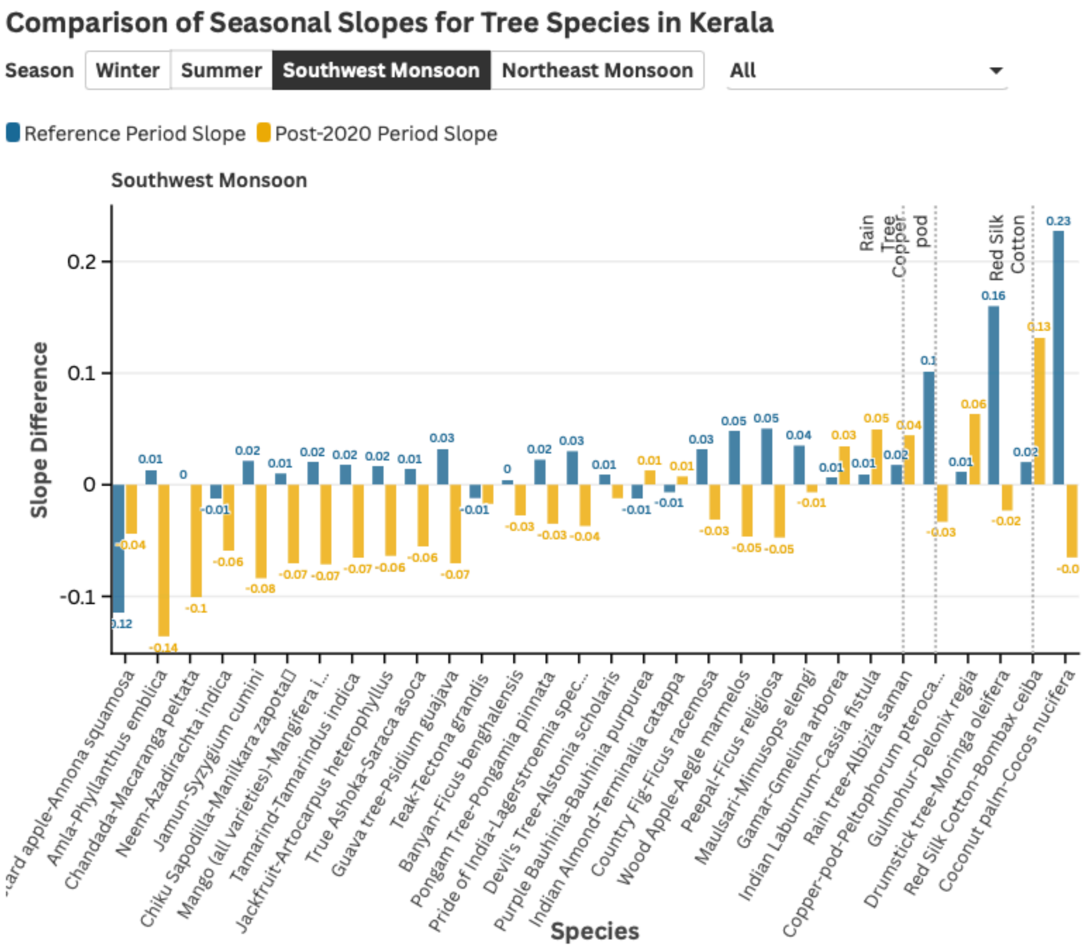
Summer Season

- ***Negative Trends:*** More species exhibit negative slope changes in the post-2020 period, such as **Custard Apple** and **Neem**.
- ***Positive Outliers:*** Species like **Rain Tree** and **Teak** show slight advancements (positive slopes).



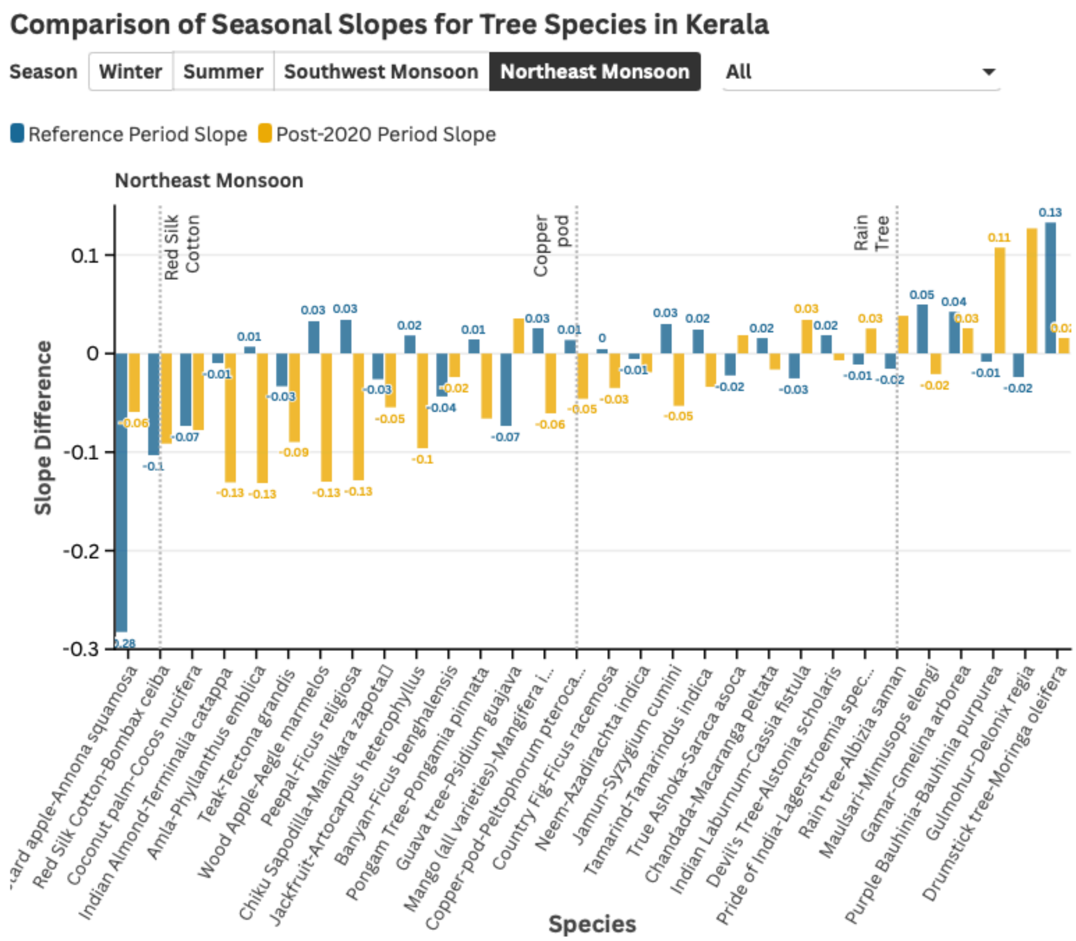
Southwest Monsoon

- ***Delays:*** Many species (e.g., **Coconut Palm**, **Neem**) show pronounced negative shifts during this season, indicating a widespread delay in monsoonal phenological events.



Northeast Monsoon

- ***Delayed Events:*** Negative slopes dominate, particularly for species like **Coconut Palm**.
- ***Positive Advancements:*** A few species, including **Rain Tree**, show early phenological trends.



4. How has the probability of flowering and fruiting in a given season changed since 2014?

While this question is designed to be answered probabilistically, our onset and trend analyses already show that flowering and fruiting stages are shifting earlier in some seasons. This indirectly suggests that the likelihood of these phenological events happening in a given season may be increasing or decreasing due to climate changes. Further probability-based modeling could refine this insight.

5. Do certain seasons show more pronounced shifts in timing or growth patterns?

According to the average onset shifts, Summer and Monsoon seasons display more pronounced shifts towards earlier onset timing. This pattern suggests that these seasons may be experiencing stronger climate-driven impacts, which is further supported by the general trend of reduced slopes in growth patterns across these seasons.

Detailed Insights Stemming from Base Questions

1. Significant Early Onset Shifts in Summer and Monsoon

- Species with Largest Shifts: **Copper-pod (Peltophorum pterocarpum)**, **Coconut palm (Cocos nucifera)**, and I**ndian Almond (Terminalia catappa)** show significant shifts in onset timing, particularly in the **Summer** and **Monsoon** seasons, with onset weeks moving earlier by up to 3 weeks.
- Possible Interpretation: This trend of earlier onset in Summer and Monsoon could suggest that these species are responding to changing seasonal cues, possibly due to warmer or more favorable conditions occurring earlier in these seasons.

2. Mixed Onset Shifts in Post-Monsoon and Winter

- In **Post-Monsoon** and **Winter**, the average onset shifts are generally smaller and less consistent across species, with both earlier and later onset observed.
- *Key Observation:* For instance, **Pride of India (Lagerstroemia speciosa)** showed a delayed onset of around 1.3 weeks in Post-Monsoon, while **Chiku Sapodilla (Manilkara zapota)** in Winter had an earlier onset by 1 week.
- Implication: The variability in Winter and Post-Monsoon onset shifts may indicate a more complex interaction with climate, where some species experience delays due to potentially cooler or altered post-monsoon conditions.

3. Notable Changes in Growth Trends in Winter and Monsoon

- Top Species with Trend Changes: **Red Silk Cotton (Bombax ceiba)** in Winter and **Coconut Palm (Cocos nucifera)** in Monsoon experienced the most substantial trend changes, with slopes decreasing by over 0.29, indicating a reduction in growth rates in these seasons post-2020.
- Broader Pattern: Many species, including **Pongam Tree (Pongamia pinnata)**, **Drumstick Tree (Moringa oleifera)**, and **Peepal (Ficus religiosa)**, showed decreasing trends across Winter and Monsoon.
- **Potential Climate Impact**: This decrease in trend slopes for growth in Winter and Monsoon suggests that these species may be experiencing stress or limitations in their typical growth patterns, possibly due to less favorable conditions (e.g., reduced rainfall or temperature changes) in these seasons post-2020.

4. Average Seasonal Shift Patterns

- Overall Seasonal Impact: On average, Summer shows a positive shift of 0.32 weeks, while Monsoon and Winter show slight negative shifts, with Monsoon having the most substantial change of -0.35 weeks.
- Implications for Seasonal Cycles: These shifts indicate that species in Kerala may be adapting to earlier summer conditions, while Monsoon and Winter may be shifting slightly later or becoming less predictable.

---

7. Recommendations for Client so far…

1. **Update Reference Timelines**:
    - Adjust reference onset times for species that show consistent timing changes, particularly in Summer and Monsoon, to ensure accuracy for citizen observation validation.
2. **Focus on Climate-Responsive Species**:
    - Monitoring species with substantial timing shifts and trend changes, such as *Copper-pod* and *Coconut palm*, may provide insights into climate adaptation processes in tropical species.
3. **Expand Citizen Data Validation Protocols**:
    - Modify validation approaches to account for observed shifts in growth patterns and earlier onsets, aligning citizen observation data with updated timelines.

4. **Extend Analysis to Additional Regions in India**:
   - Expanding this analysis to other states will help identify if these patterns are consistent across India or localized to Kerala, enabling a more comprehensive understanding of climate impacts on phenology.

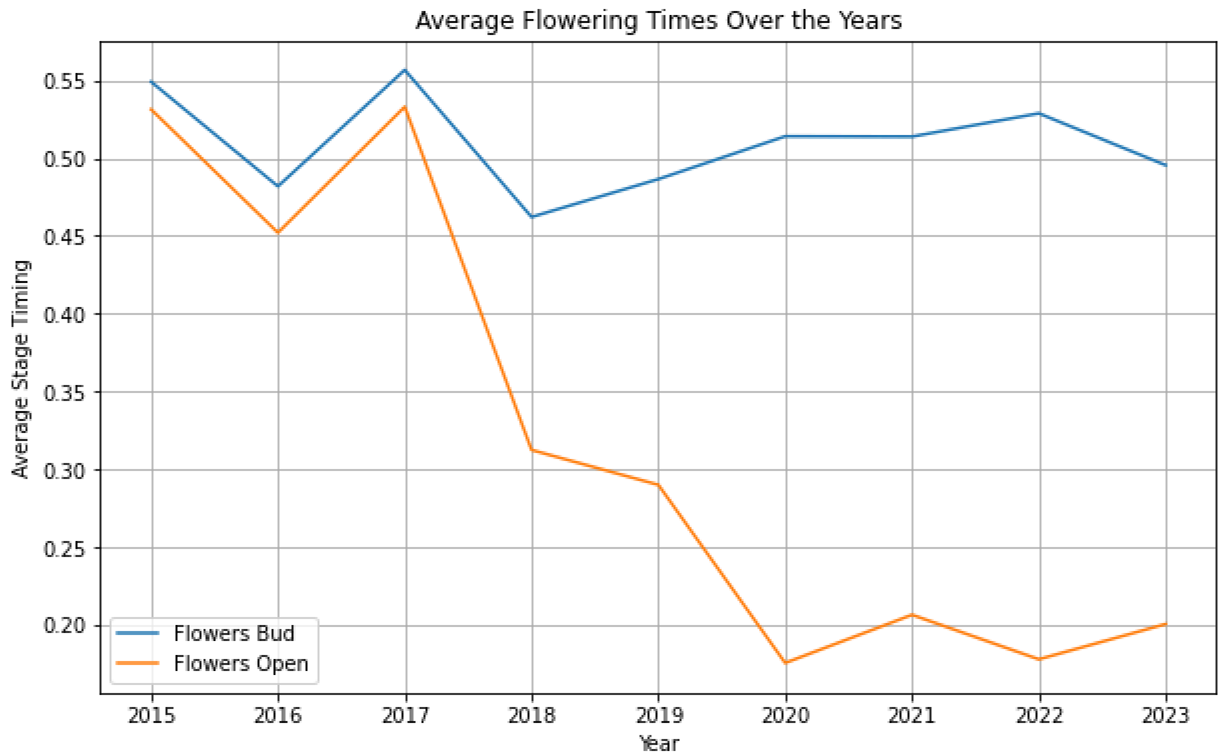
8. Documentation for 'data_analysis.ipynb'

Summary Statistics for Stages:

- The analysis filters invalid values (like negative stage counts) for accurate statistics.
- Grouped data by Species_name and State_name calculates the average timing for various stages—Leaves_fresh, Flowers_bud, and Fruits_ripe.
- The results are stored in avg_stages_top_30 and structured into a summary table summary_table_top_30, which is exported as "**summary_statistics_top_30_species.csv".**

Time Series Analysis:

- A time series analysis implemented to observe trends in phenology stages over time.
- The dataset is grouped by Year, and average timings for flowering, leafing, and fruiting are calculated for each year, creating the foundation for visualizations.
- Using matplotlib, a plot is generated to visualize average flowering times (Flowers_bud and Flowers_open) across years, aiding in the analysis of any shifts or trends over time.



Markov Modeling Introduction:

- **State Mapping**: A mapping is established for phenology states: 'None' (-2) is mapped to 0, 'Few' (0) to 1, and 'Many' (1) to 2. This prepares the data for transition analysis across these states.
- **Transition Matrix**: Probabilities for state transitions (e.g., 'None' to 'Few' or 'Many') are defined and normalized. The transition matrix captures the likelihood of a tree transitioning between different phenological states over time.

Markov Matrix Interpretation:

- The rows of the transition matrix represent initial states, while columns represent the potential states trees may transition into. For example, if a tree starts in the 'None' state, there is a 54.79% chance it remains in that state, a 10.10% chance of moving to 'Few,' and a 35.11% chance of moving to 'Many.'

```
Markov Transition Matrix:

[[0.54788012 0.10099692 0.35112296]

 [0.1534788  0.52888315 0.31763805]

 [0.18053093 0.10761835 0.71185072]]

From None:

  To None: 0.55
```

```
  To Few: 0.10
  To Many: 0.35
From Few:
  To None: 0.15
  To Few: 0.53
  To Many: 0.32
From Many:
  To None: 0.18
  To Few: 0.11
  To Many: 0.71
```

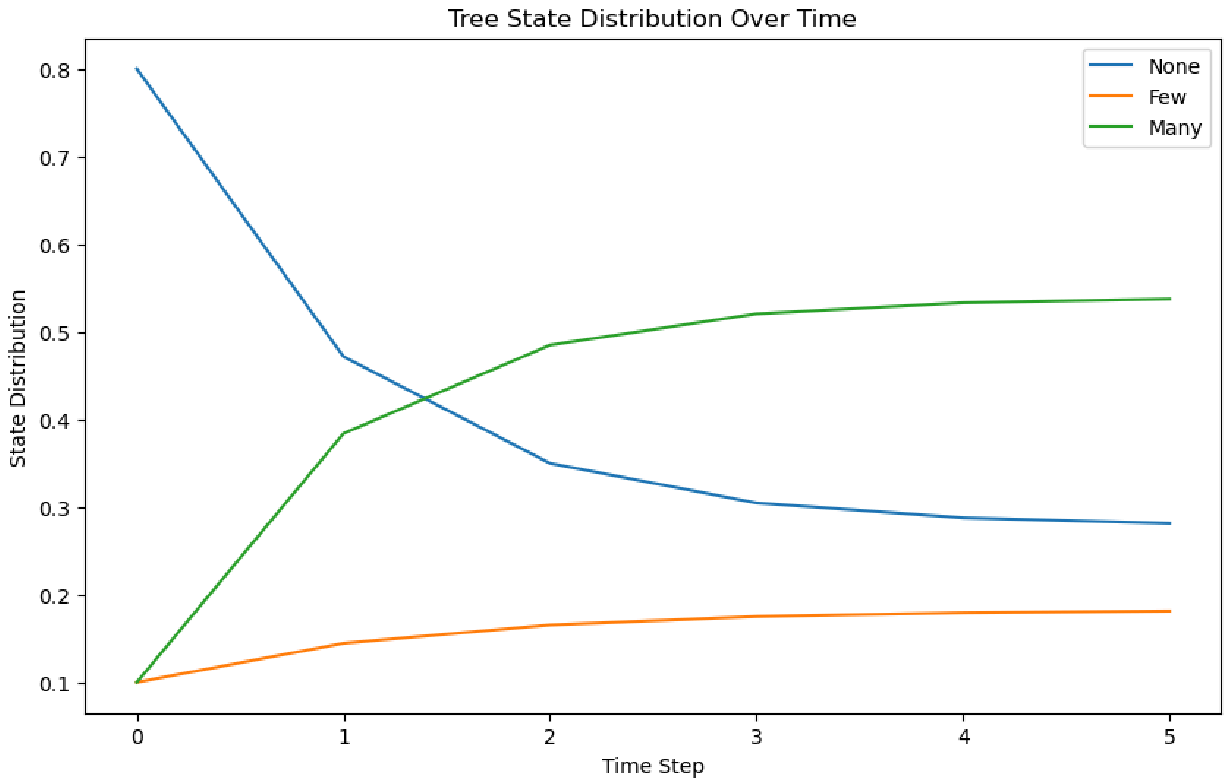
Simulation of State Transitions:

- An initial state distribution (e.g., 80% in 'None') is defined, and the model simulates changes over five time steps. The results show how tree states evolve with time, allowing observation of phenological shifts.

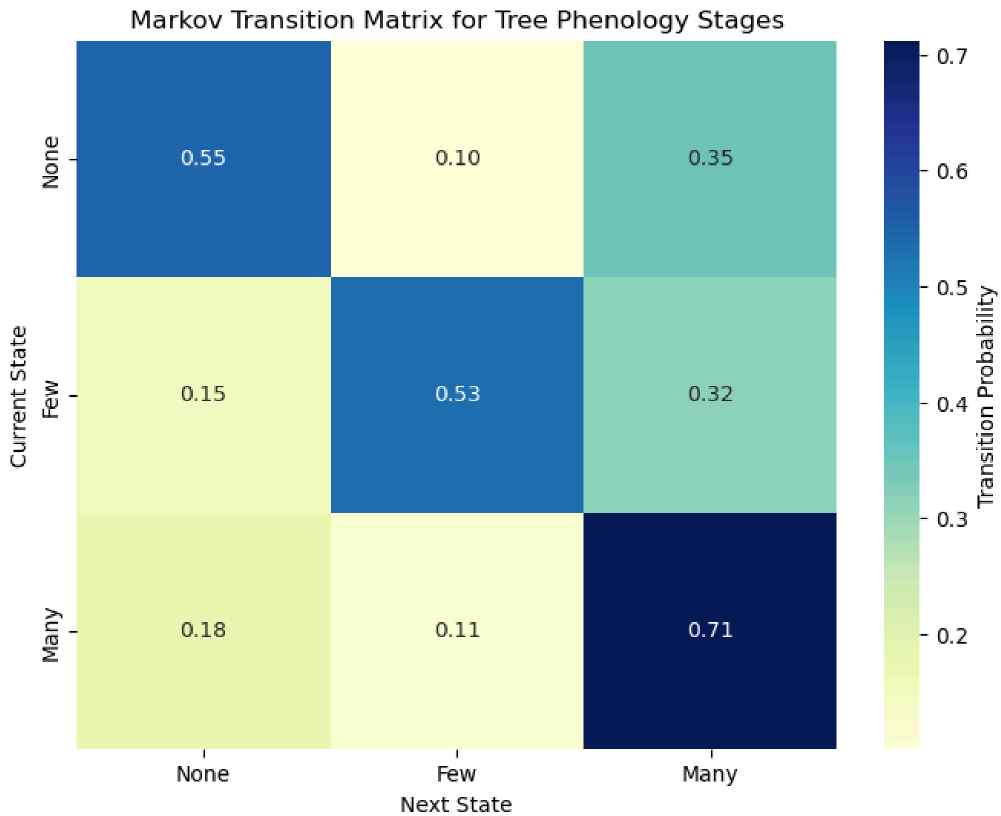
```
Step 0: [0.8 0.1 0.1]
Step 1: [0.47170507 0.14444769 0.38384725]
Step 2: [0.34990379 0.16534571 0.4847505 ]
Step 3: [0.30459485 0.17495581 0.52044934]
Step 4: [0.28769067 0.17930422 0.53300511]
Step 5: [0.2813633  0.18124798 0.53738871]
```

Visualization of State Changes:

- A time series displaying the distributions of states ('None', 'Few', 'Many') across the simulated time steps, visualizing how tree states transition over time.



Transition Matrix Heatmap: the probabilities of state transitions comparing the likelihood of each transition.



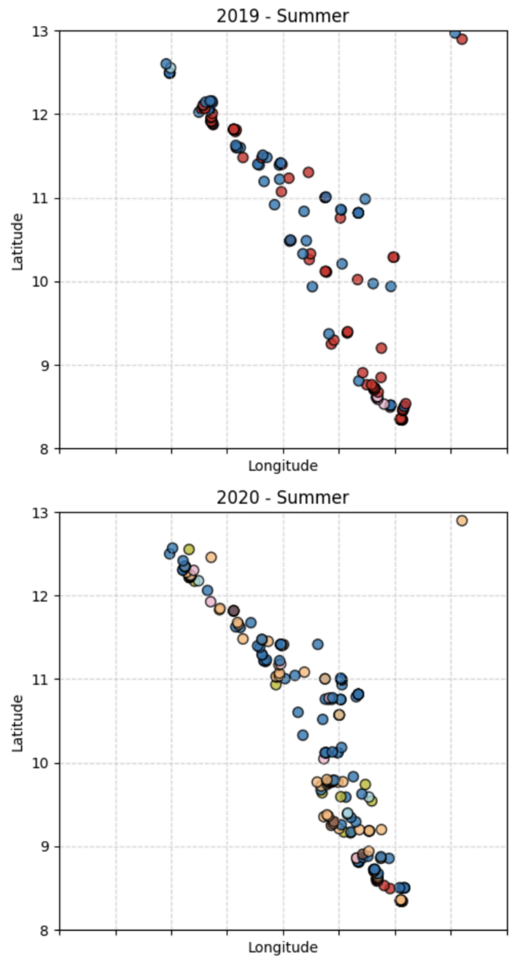
9. Geographical Analysis using DBSCAN

- Utilized DBSCAN to run clustering algorithms on the plants in order to show change in climate that could be affecting the plants in Kerala.
- The colors in the visualization illustrate the similarities in seasons throughout the years.
    - There is a similar climate from 2015-2017, but there is a change in climate starting from 2018.
- **Clustering timeline**
    - Visualizing each of the top 30 species to see if the clustering of such species differs throughout the years.
    - Adding seasons to see at which seasons plants are affected by the climate the most. Using such data, we can focus on a time period more closely or a species more closely

Season Watch would like us to assess and understand the shifts of seasonal behavior of common trees by comparing citizen submitted tree data against the reference database.

This is achieved by comparing citizen-submitted observations of tree phenology (such as leafing, flowering, and fruiting times) with a historical reference database. The goal is to **identify discrepancies, infer actual shifts over a decade, and update the reference database based on emerging patterns** to improve its accuracy and reliability.

Seasonal Clustering of Jackfruit (2018-2023)



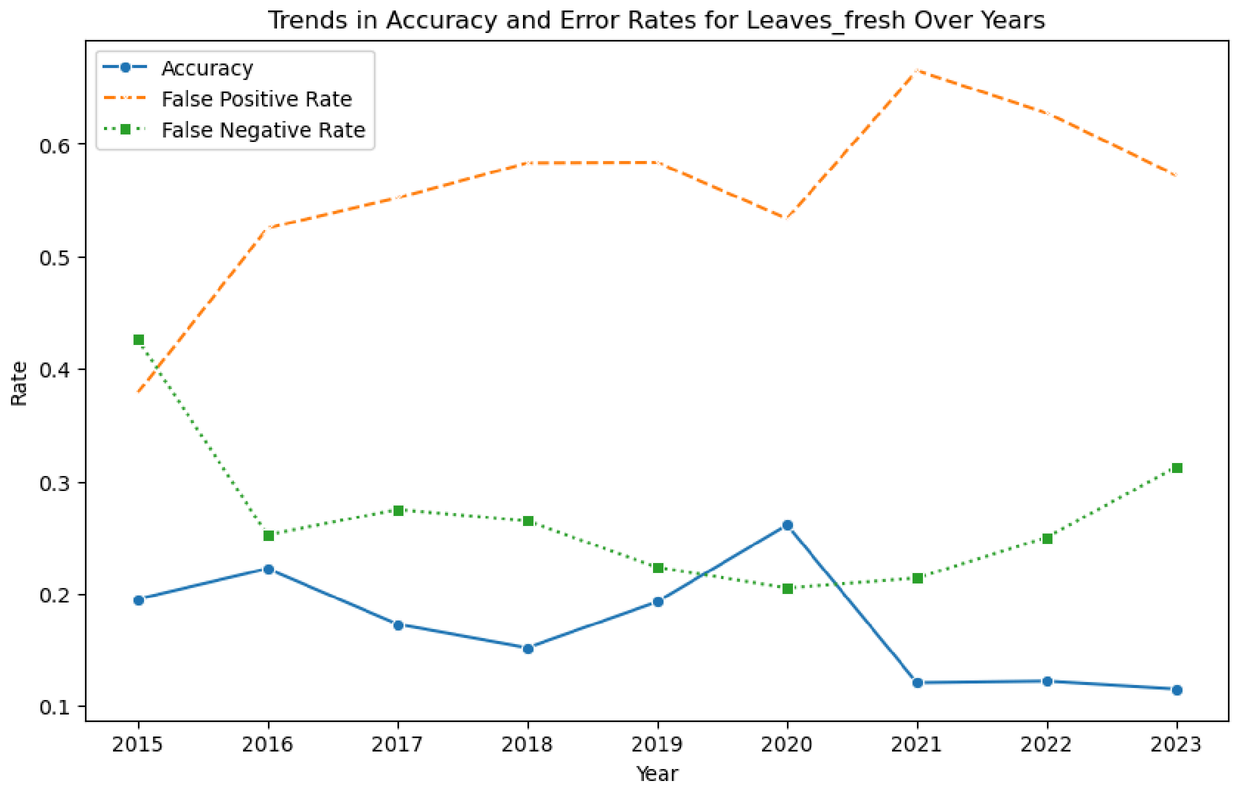
- **Each color is a cluster, and each cluster represents trees that behave similarly** (similar leaf mature times, flowering times, etc.)
- Looking at 2019, there are a lot more clusters of different colors which indicates a change in how trees are responding to the environment
- The blue dots (-1) values represents outliers which could indicate climate-driven disruptions

The code in the github is dynamic and present visualizations similar to the ones above when searched for specific species

NEXT STEPS FOR CLUSTERING ANALYSIS

We believe that the clustering analysis parameters (DBSCAN's Neighborhood Radius) can be altered to present a more informative visualization of Kerala. Adjusting the 'min_samples' can influence the size and number of clusters; test different values to balance noise points and meaningful clusters. The points within the visualizations could be improved to better show distinction. There could be additional features to extend clustering beyond latitude and longitude to include: altitude, seasonal temperature and precipitation patterns.

10. Discrepancies, Accuracy, in -2 vals of data (the rest resides in 'data_analysis.ipynb' code in GitHub')



---

11. Comparison of Seasonal Slopes for Tree Species in Kerala: Reference and Post-2020 Period Slopes Across Seasons:

Link to interactive visual: https://public.flourish.studio/visualisation/20379276/
Link to Github code for slope: [GitHub]()

Methodology

1. **Data Segmentation**:
   - The dataset is divided into two periods:
     - **Reference Period (2014-2020)**: Data between the years 2014 and 2020.
     - **Post-2020 Period**: Data after 2020.
   - Seasons are defined by week ranges, such as "Winter," "Summer," etc.
2. **Linear Regression for Slopes**:

    - The scipy.stats.linregress function is used to compute the slope and intercept for linear trends in the data.
    - For each species and season, data points within the relevant week range are filtered.
    - The linregress function is applied to determine the slope of the trend for each period.
    - The linregress function from scipy.stats is used to calculate the slope for each species and season.
    - The function takes the Year column (as the independent variable) and a phenological indicator (e.g., Leaves_mature) as the dependent variable.
3. **Comparison of Slopes**:
    - Slopes for the reference and post-2020 periods are compared to observe any significant differences or shifts in trends.
4. **Handling Missing Data**:
    - If there isn't enough data for regression within a specific season or species, the slope is set to NaN to avoid incorrect results.

Overall Patterns

1. **Significant Seasonal Variations**:
    - Slopes differ drastically between the reference period and the post-2020 period, indicating shifts in phenological patterns across species and seasons.
    - For example, *Red Silk Cotton* (*Bombax ceiba*) in **Winter** shows a sharp decline in slope (-0.4036), suggesting major changes in behavior during this season.
2. **Species-Specific Insights**:
    - Certain species, such as *Copper-pod-Peltophorum pterocarpum* and *Mango (Mangifera indica)*, show mixed slope differences across seasons, reflecting variability in their seasonal responses.
    - Species like *Custard Apple (Annona squamosa)* in **Winter** experienced a steep positive slope difference (+0.1725), indicating increased activity during this season.
3. **Seasonal Stability vs. Shifts**:
    - **Winter**: Significant slope changes for many species, with both positive (e.g., *Custard Apple*, *Copper-pod*) and negative shifts (e.g., *Red Silk Cotton*). This suggests Winter is a season of marked phenological transitions post-2020.
    - **Monsoon**: Fewer extreme changes, though some species (e.g., *Copper-pod-Peltophorum pterocarpum*) exhibit large negative differences, indicating a decline in activity.
    - **Post-Monsoon**: A mix of increases and decreases in slope, with *Purple Bauhinia* (+0.1159) showing significant increases in activity.

Implications

1. **Climate Change Impacts**:
    - The post-2020 period shows marked changes in phenological activity, likely due to climate change effects such as altered rainfall patterns, temperature fluctuations, and ecological disturbances.
2. **Species Sensitivity**:

    - Some species (*Red Silk Cotton, Mango*) show more drastic slope changes across seasons, making them potential indicators of climate sensitivity.
    - Conversely, species with smaller slope differences (e.g., *Neem, Banyan*) suggest resilience or slower adaptation.
3. **Seasonal Adaptation**:
    - Large differences in seasonal slopes indicate that species may be adjusting their phenophases to align with changing environmental cues, such as shifting flowering or fruiting timings.

Key Examples

- **Indian Almond (Terminalia catappa)**:
    - Large slope decrease in **Winter** (-0.2911) and **Post-Monsoon** (-0.1208), indicating reduced phenological activity.
    - Slight positive changes in **Monsoon** (+0.0141), suggesting potential adaptation to wetter conditions.
- **Copper-pod (Peltophorum pterocarpum)**:
    - Notable positive slope difference in **Winter** (+0.0559), indicating increased activity.
    - Significant declines in **Monsoon** (-0.1349) and **Post-Monsoon** (-0.0596), reflecting a drop in activity during these seasons.
- **Rain Tree (Albizia saman)**:
    - Consistent positive slope differences across all seasons, especially **Post-Monsoon** (+0.0537), indicating an overall increase in activity.

Conclusion

- The data highlights clear phenological shifts in many species and seasons, with potential links to climate change.
- These insights are critical for understanding tree behavior, ecosystem responses, and the need for further research into adaptive strategies for vulnerable species.

---

12. Weekly Seasonal Transitions: Probability Modeling for the Top 30 Tree Species

Link to interactive visual: https://public.flourish.studio/visualisation/20360646/

1. Significant Shifts in Onset Weeks

The visual captures percentage changes in average phenological onset weeks between the pre-2020 Reference Period and the Post-2020 Period. Several species exhibit notable shifts, including:

- **Chiku Sapodilla (Manilkara zapota):**

- Northeast Monsoon onset advanced by 50%, the largest shift observed.
- **Rain Tree (Albizia saman):**
    - Northeast Monsoon onset advanced by 33.33%.
- **Coconut Palm (Cocos nucifera):**
    - Monsoon onset advanced by 17.73%, indicating a strong response to rainfall variability.

2. Seasonal Patterns

- **Winter:**
  Minimal changes across most species suggest stability during this season, with exceptions like Copper-pod and Indian Almond advancing significantly.
- **Summer and Monsoon:**
  Earlier onset timings are observed for species such as Coconut Palm (-17.73%) and Chiku Sapodilla (-1.43%), likely reflecting temperature and precipitation shifts.
- **Post-Monsoon:**
  Most species exhibit minimal shifts, indicating reduced climate impact during this season.

3. Species-Specific Insights

- **Advancers:**
    - *Chiku Sapodilla* and *Rain Tree* show pronounced advances, especially during Monsoon and Post-Monsoon.
    - *Coconut Palm* and *Copper-pod* consistently advance across multiple seasons.
- **Resilient Species:**
  Species like *Mango (Mangifera indica)* and *Jackfruit (Artocarpus heterophyllus)* exhibit negligible shifts, suggesting a slower response to environmental changes.

4. Implications for Climate Change

Earlier onset timings highlight the sensitivity of key species to changing climate cues, such as:

- Altered rainfall patterns (Monsoon).
- Shifts in seasonal temperature thresholds (Summer).
  These phenological shifts may disrupt ecosystem dynamics, including pollination, fruiting, and seed dispersal.

5. Recommendations

- **Monitoring:** Focus on species with significant advances (*Chiku Sapodilla*, *Rain Tree*, *Coconut Palm*) to assess their adaptation strategies.
- **Conservation:** Target regions and seasons with the largest shifts (e.g., Monsoon) for adaptive management and conservation planning.