

**BHARATIYA VIDYA BHAVAN'S**  
**SARDAR PATEL INSTITUTE OF TECHNOLOGY**  
(Empowered Autonomous Institute Affiliated to University of Mumbai)  
[Knowledge is Nectar]

**Advance Data Visualization**

<b>UID</b>	2021700012
<b>Name</b>	Rahul Chalwadi
<b>Aim</b>	Design interactive dashboards and create visual storytelling using D3.js on a dataset related to Environment/Forest cover, covering basic and advanced charts

**Objectives:**

- To understand how to use D3.js for data visualization.
- To implement basic charts like Bar chart, Pie chart, Histogram, Timeline chart, Scatter plot, and Bubble plot.
- To implement advanced charts like Word chart, Box and whisker plot, Violin plot, Regression plot (linear and nonlinear), 3D chart, and Jitter.
- To draw observations and insights from each chart.
- To create an interactive storytelling dashboard using the above visualizations.

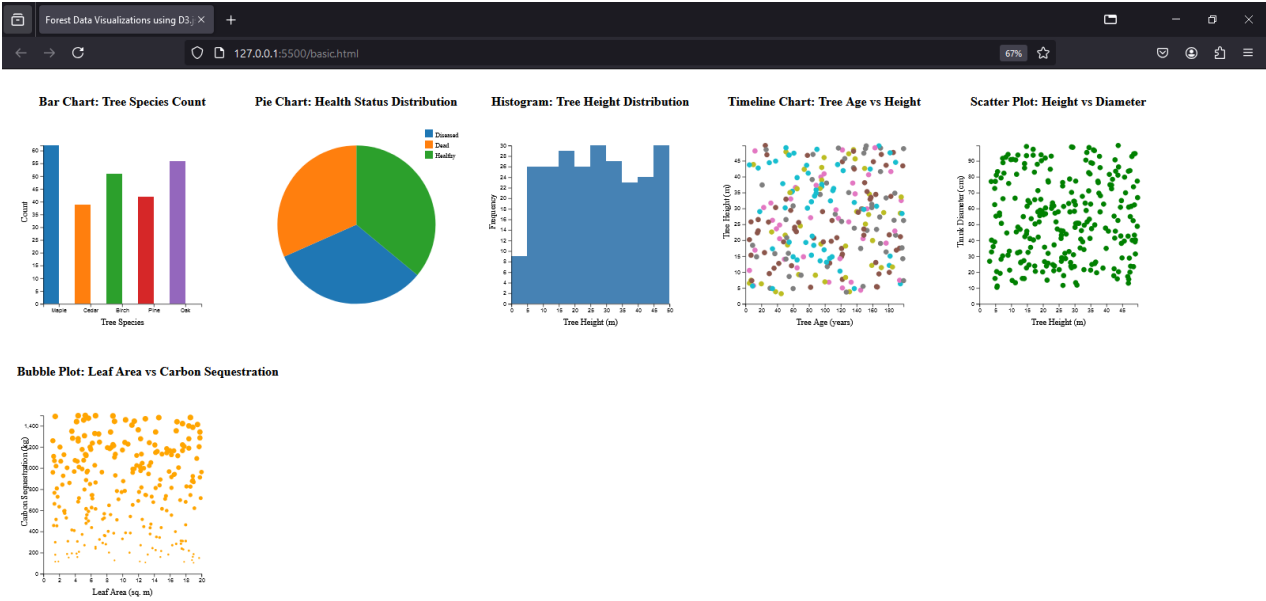
**Dataset:** Forest Dataset

**Column Attributes:**

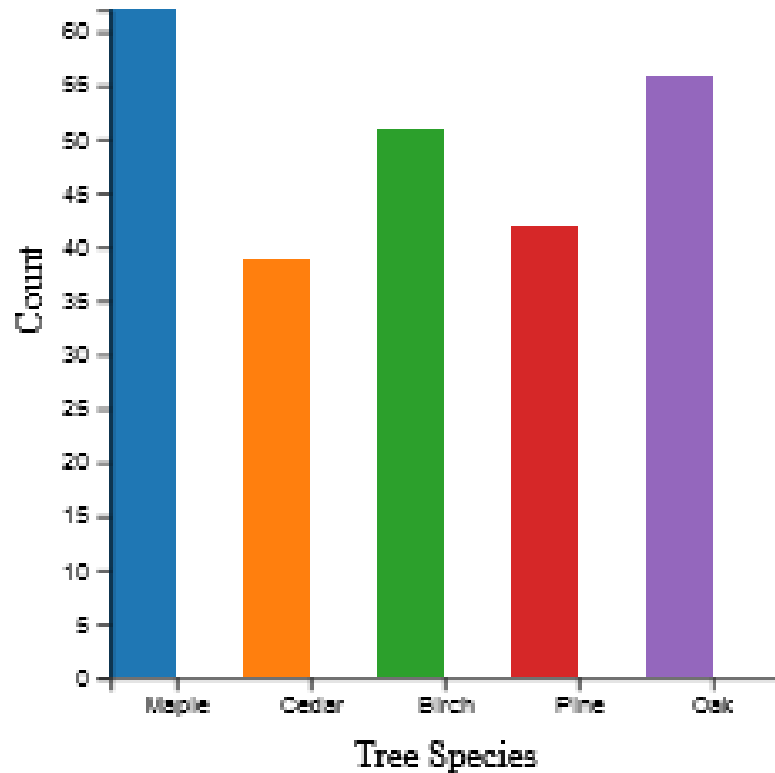
- **Tree ID:** Unique identifier for each tree (Categorical, e.g., T001, T002).
- **Tree Species:** Different species of trees (Categorical, e.g., Oak, Pine, Birch).
- **Tree Height (m):** Height of the trees in meters (Numerical, e.g., 5.3, 12.7).
- **Tree Age (years):** Age of the tree (Numerical, e.g., 50, 120).
- **Trunk Diameter (cm):** Diameter of the tree's trunk at breast height (Numerical, e.g., 30.5, 45.2).
- **Leaf Area (sq. m):** Total leaf surface area of the tree (Numerical, e.g., 12.4, 8.7).
- **Tree Health Status:** Health status of the tree (Categorical, e.g., Healthy, Diseased, Dead).
- **Annual Rainfall (mm):** Rainfall in the area where the tree is located (Numerical, e.g., 1200, 1500).
- **Soil Type:** Type of soil in which the tree is growing (Categorical, e.g., Clay, Sandy, Loamy).

- **Carbon Sequestration (kg):** The amount of carbon the tree sequesters annually (Numerical, e.g., 500, 900).

Visualizations:



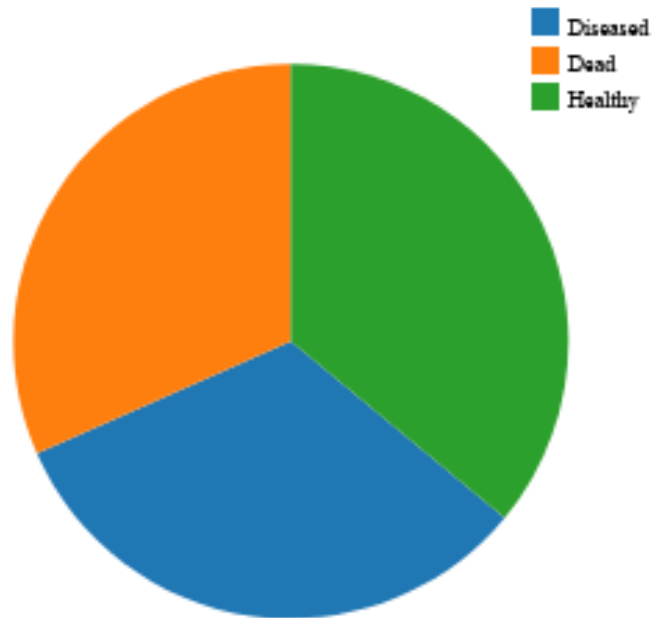
## Bar Chart: Tree Species Count



### Observation:

- Maple trees have the highest count, with around 60 trees, indicating they are the most common species in this dataset.
- Cedar trees have the lowest count among the listed species, with fewer than 40 trees.
- Oak and Birch trees have similar counts, both ranging between 45 and 50, suggesting a relatively balanced presence of these species.

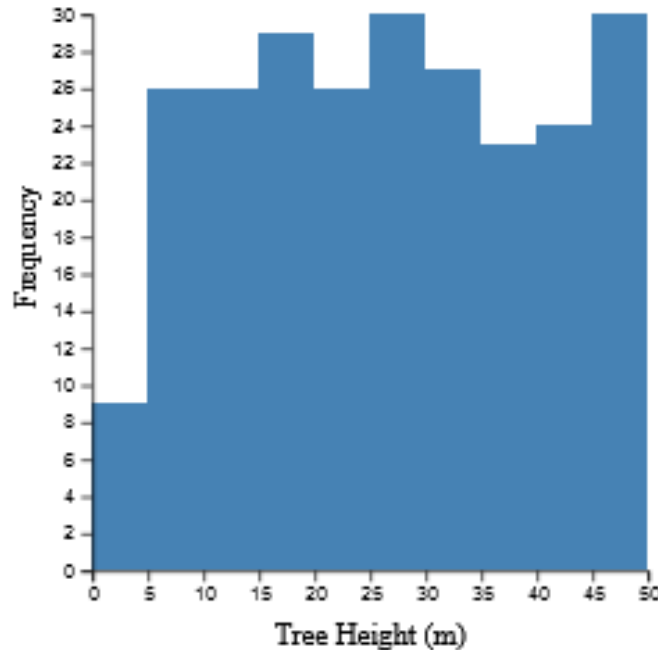
## Pie Chart: Health Status Distribution



### Observation:

- Healthy trees make up the largest portion of the pie chart, indicating that most trees in the dataset are in good health.
- Diseased trees form a slightly smaller proportion compared to healthy trees, suggesting a significant presence of trees that are not in optimal health.
- Dead trees occupy the smallest segment of the chart, indicating that there are fewer trees that have completely perished compared to those that are diseased or healthy.

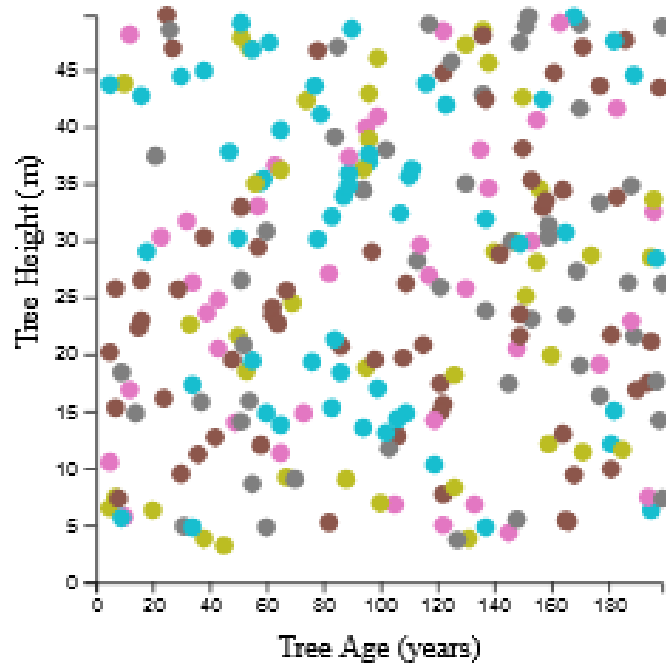
## Histogram: Tree Height Distribution



### Observation:

- Tree heights are fairly evenly distributed across different ranges, with each height interval having a relatively similar frequency.
- The height range of 10 to 15 meters and 45 to 50 meters have slightly higher frequencies, suggesting more trees fall within these height ranges compared to others.
- There are fewer trees in the 5 to 10 meter range, indicating a lesser presence of shorter trees in the dataset.

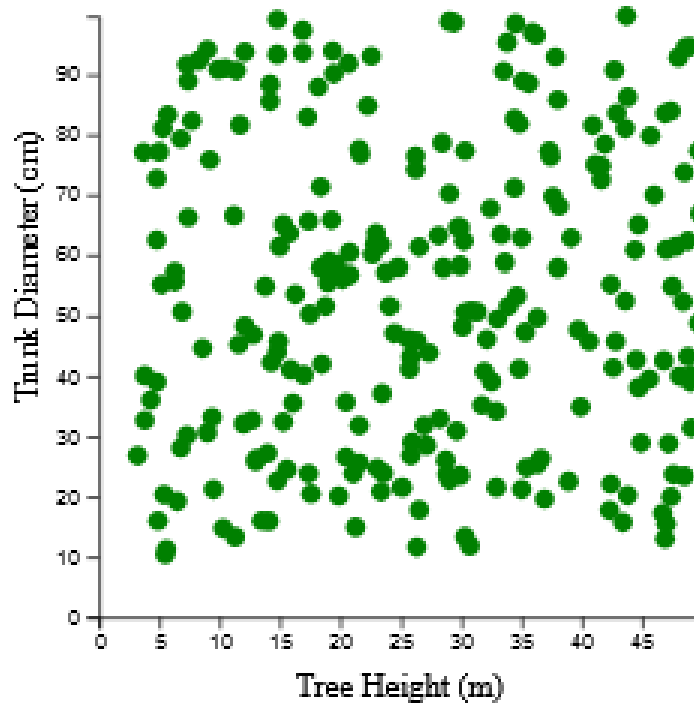
## Timeline Chart: Tree Age vs Height



### Observation:

- General trend: A positive correlation exists between tree age and height, indicating older trees tend to be taller.
- Variation: However, there is also a significant amount of variation, suggesting other factors besides age influence height.
- Clustering: The data points are somewhat clustered, possibly indicating different growth patterns among groups of trees.

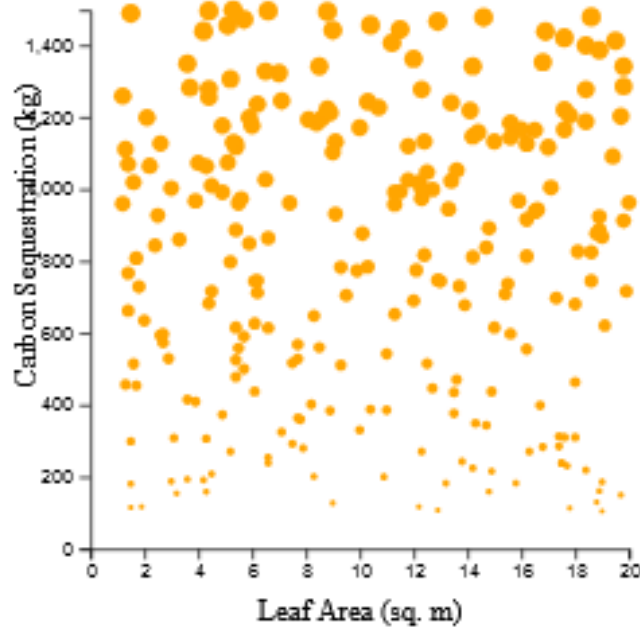
## Scatter Plot: Height vs Diameter



### Observation:

- General trend: A positive correlation exists between tree height and diameter, indicating taller trees tend to have larger diameters.
- Variation: However, there is also a significant amount of variation, suggesting other factors besides height influence diameter.
- Clustering: The data points are somewhat clustered, possibly indicating different growth patterns among groups of trees.

## Bubble Plot: Leaf Area vs Carbon Sequestration



### Observation:

- General trend: There appears to be a positive correlation between leaf area and carbon sequestration, indicating that trees with larger leaf areas tend to sequester more carbon.
- Scatter: However, there is also a significant amount of scatter in the data points, suggesting that other factors besides leaf area influence carbon sequestration.
- Size variation: The size of the bubbles, representing the amount of carbon sequestered, varies, further emphasizing the influence of other factors on carbon sequestration.



## Word Chart: Tree Species



- Larger fonts like "Maple" and "Oak" may indicate higher frequency or importance

### CONCLUSION:

In conclusion, this dataset provides valuable insights into the relationships between tree species, physical characteristics, and environmental factors such as soil type and rainfall. By analyzing attributes like tree height, age, trunk diameter, and health status, we can better understand the factors that contribute to carbon sequestration and overall tree health.