Advance Data Visualization

UID	2022701010
Name	Hawaiza Siddiqui
Batch	Batch K
Aim	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.

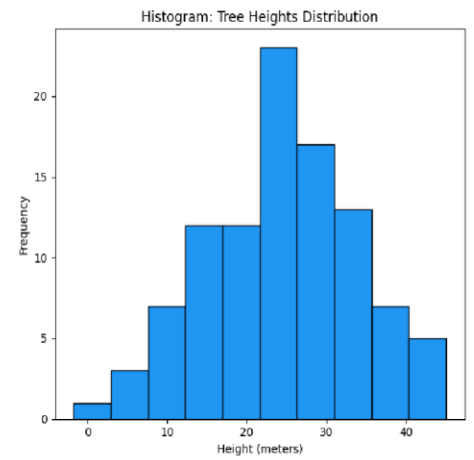- To implement advanced charts like Word chart, Box and whisker plot, Violin plot, Regression plot (linear and nonlinear), 3D chart, and Jitter.

- To draw observations and insights from each chart.

- To create an interactive storytelling dashboard using the above visualizations.

Dataset: Forest Dataset
Visualizations -

Bar Chart: Tree Species Count

Pie Chart: Tree Health Status

Healthy 33.0%
Dead 43.4%
Diseased 23.6%

Histogram: Tree Heights Distribution

Scatter Plot: Tree Age vs Height

Scatter Plot: Tree Height vs Trunk Diameter

Bubble Plot: Leaf Area vs Carbon Sequestration

Tree Species Count

| Tree Species	Count
Maple	59
Oak	80
Pine	93
Cedar	94
Birch	35
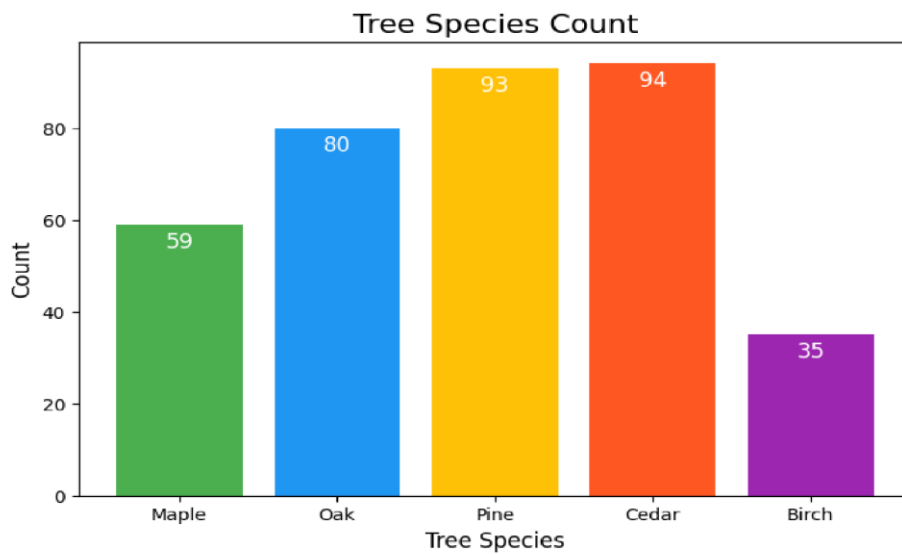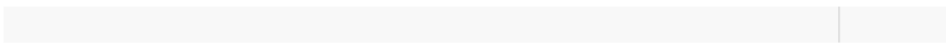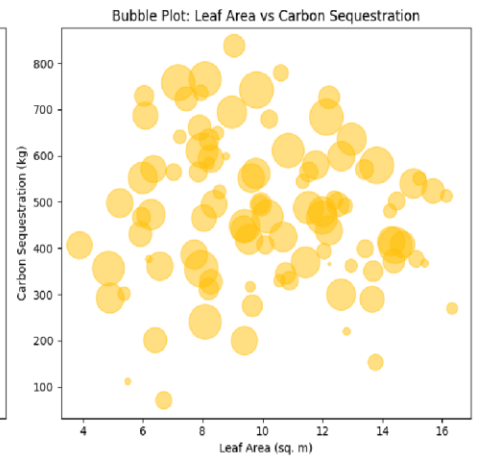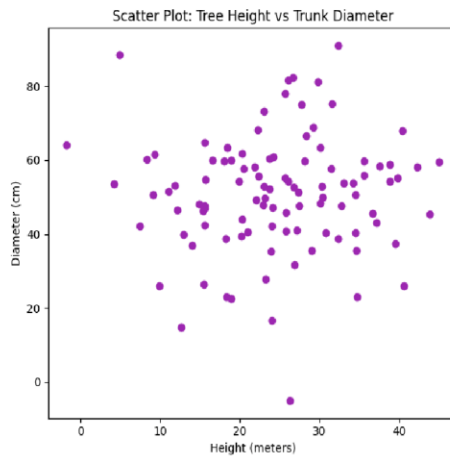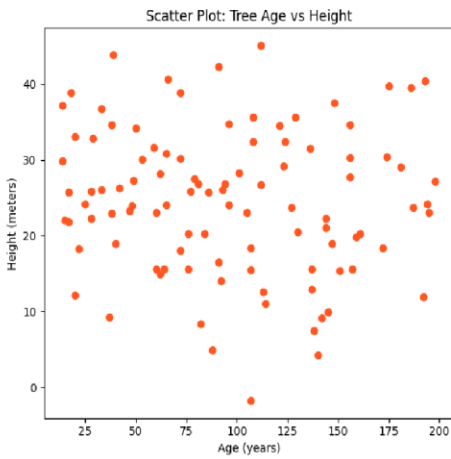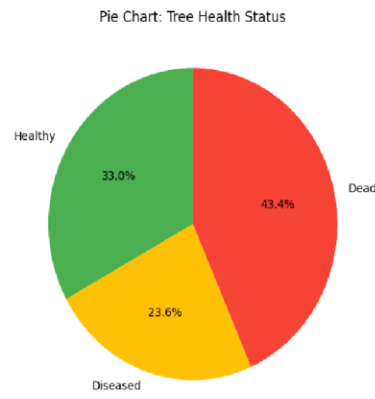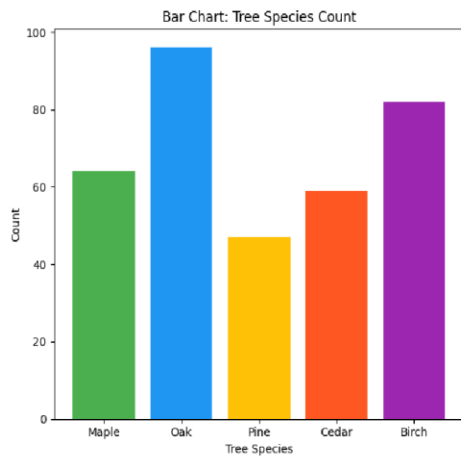
Observation:

1. The chart shows how many trees there are of different kinds.
2. Cedar trees are the most common.
3. Birch trees are the least common.
4. Pine and Oak trees are about the same.
5. Maple trees are a little less common than Pine and Oak.
6. The chart helps us see how many of each kind of tree there are.

Tree Health Status
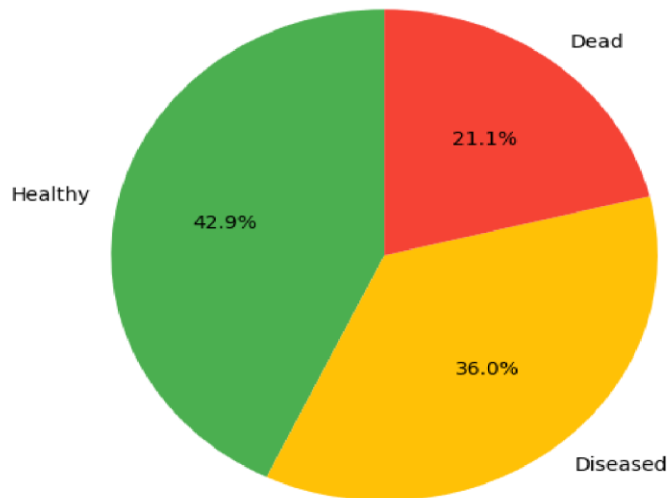
Observation:

- **Healthy:** 42.9% of the trees are in good health.
- **Diseased:** 36% of the trees are affected by diseases or conditions.
- **Dead:** 21.1% of the trees have perished.

Overall Assessment:

The pie chart reveals that while a majority of the trees are healthy, a significant portion are either diseased or dead. This suggests a need for further investigation into the factors contributing to tree mortality and the implementation of appropriate conservation measures.

Distribution of Tree Heights
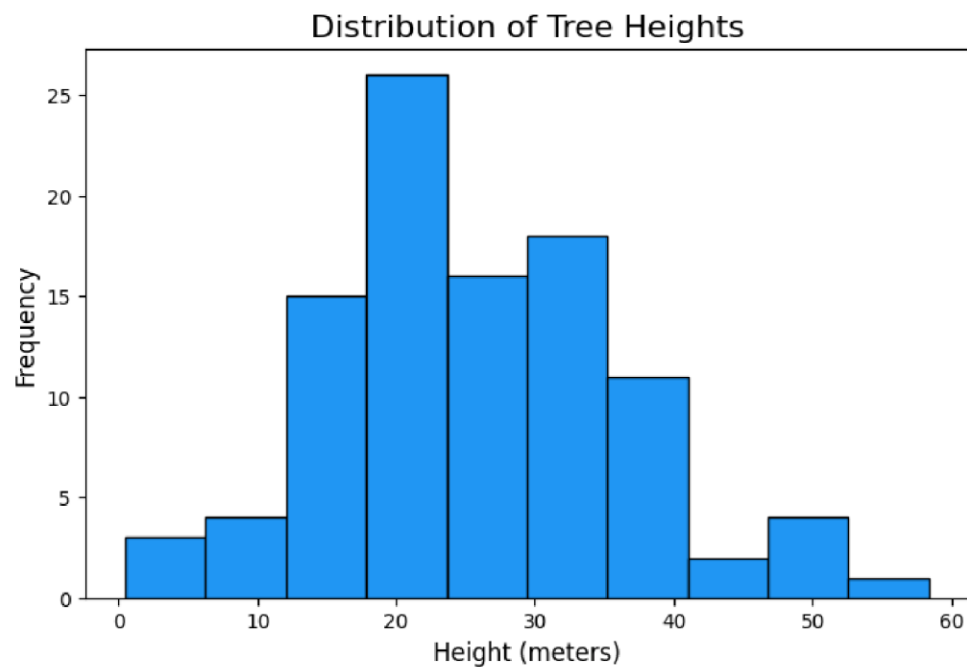
Observation:

- **Most trees are between 20 and 30 meters tall.** This suggests that this height range is the most common for the trees in the sample.
- **There are a few very tall trees.** This indicates that the distribution is skewed to the right, meaning there are some outliers that are pulling the average height up.
- **The trees have a variety of heights.** This suggests that there is a moderate amount of variability in the tree heights, as evidenced by the spread of the bars in the histogram.

Overall Interpretation:

The histogram provides a clear visualization of the distribution of tree heights. The right-skewness suggests that the forest may have some older, taller trees, while the central tendency and variability give insights into the overall growth patterns of the trees in the population.
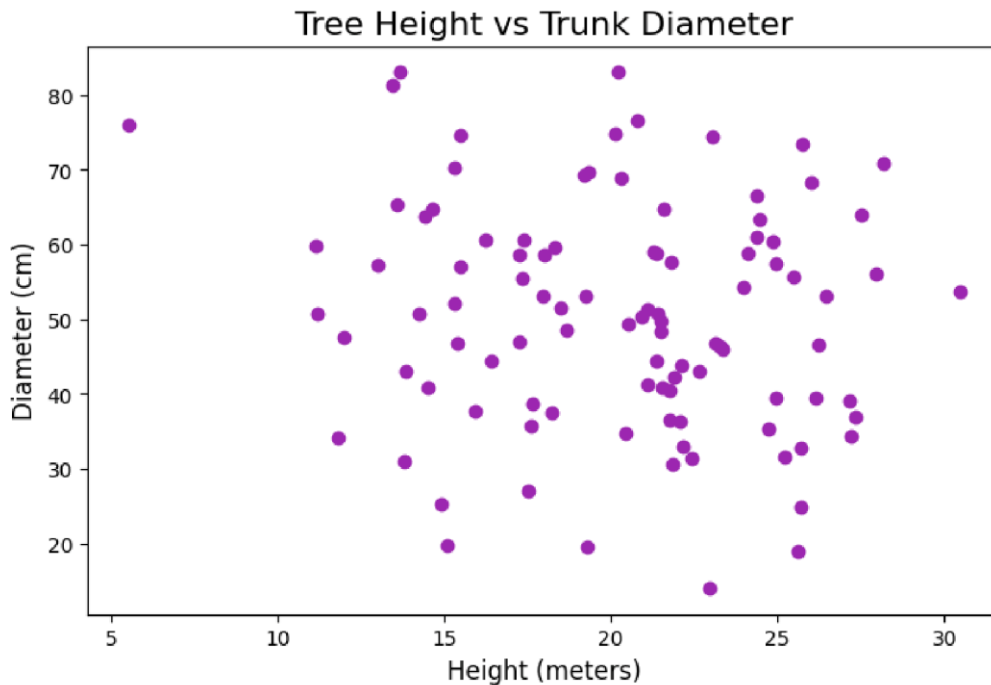

Tree Age vs Height
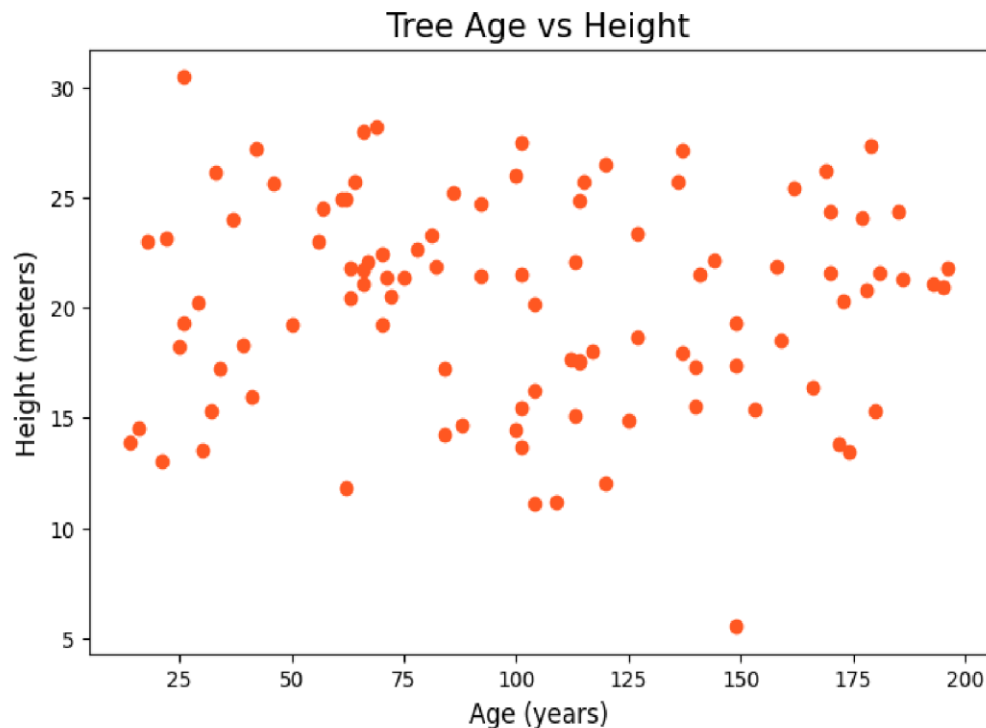
Observation:

- **General Trend:** There appears to be a positive correlation between tree age and height. As trees get older, they tend to grow taller.
- **Scatter:** The points are somewhat scattered, indicating that while there is a general trend, there are individual variations in growth rates.
- **Outliers:** A few points deviate significantly from the general trend, suggesting that factors other than age might influence tree height.

Overall Interpretation:

The scatter plot suggests a relationship between tree age and height, with older trees generally being taller. However, individual variations and potential outliers indicate that other factors, such as genetics, environmental conditions, or competition, may also play a role in tree growth.

Observation:

- **Positive Correlation:** There appears to be a positive correlation between tree height and trunk diameter. As trees get taller, they tend to have thicker trunks.
- **Scatter:** The points are somewhat scattered, indicating that while there is a general trend, there are individual variations in growth patterns.
- **Outliers:** A few points deviate significantly from the general trend, suggesting that factors other than height might influence trunk diameter.

  Overall Interpretation:

The scatter plot suggests a relationship between tree height and trunk diameter, with taller trees generally having thicker trunks. However, individual variations and potential outliers indicate that other factors, such as genetics, environmental conditions, or competition, may also play a role in tree growth.
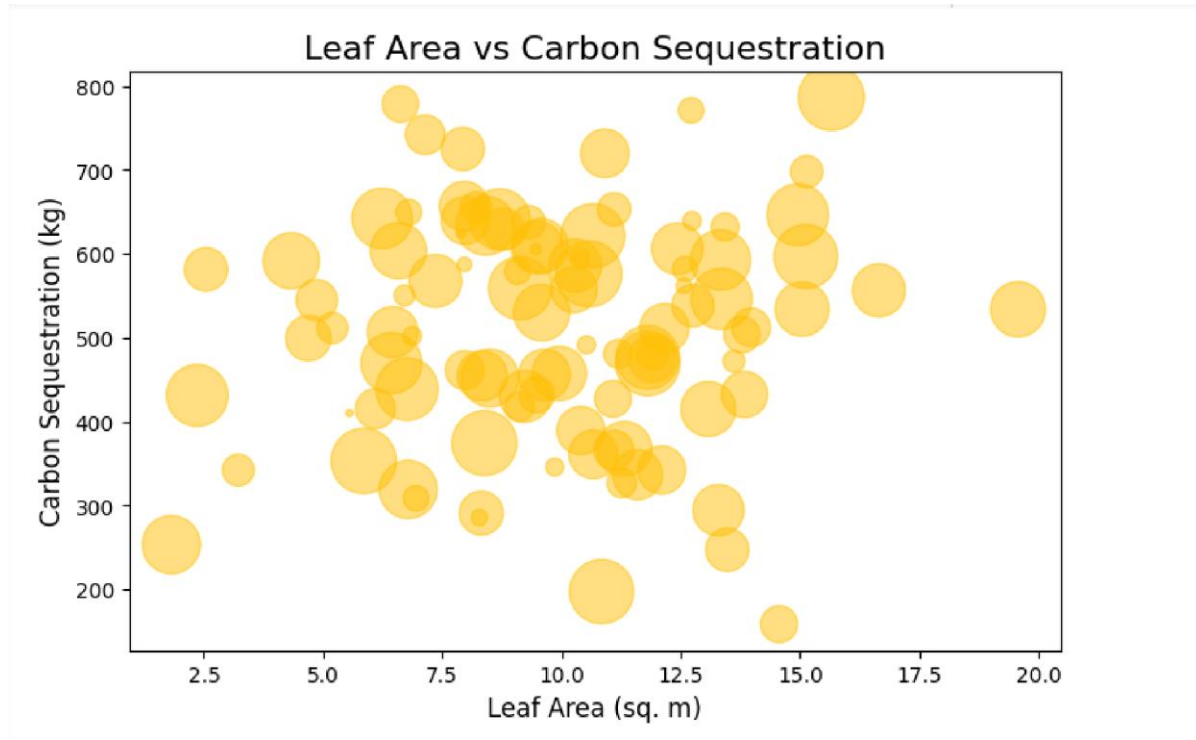

Observation:

- **Positive Correlation:** There appears to be a positive correlation between leaf area and carbon sequestration. Trees with larger leaf areas tend to sequester more carbon.

- **Scatter:** The points are somewhat scattered, indicating that while there is a general trend, there are individual variations in carbon sequestration rates.

- **Size of Bubbles:** The size of the bubbles represents an additional variable (likely the number of trees or another factor). Larger bubbles indicate a higher concentration or quantity of trees with the corresponding leaf area and carbon sequestration values.

Oak Oak
Cedar Cedar
Maple Maple
Birch Birch
Pine Pine

The words "Oak", "Maple", and "Birch" are the biggest words in the word cloud.

CONCLUSION:

In summary, this dataset offers important insights into the connections between tree species, physical attributes, and environmental factors like soil type and rainfall. By examining characteristics such as tree height, age, trunk diameter, and health status, we can gain a deeper understanding of the elements that influence carbon sequestration and the overall health of trees.