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

Department of Computer Engineering

Advanced Data Visualization

UID	2021300114
Name	Harditya Shah
Class and Batch	BE Comps B – Batch L

Experiment 8:

Aim: To design interactive dashboards using Power BI for visualizing and analyzing an Animal/Wildlife/Marine dataset, employing both basic and advanced charts to uncover insights and trends.

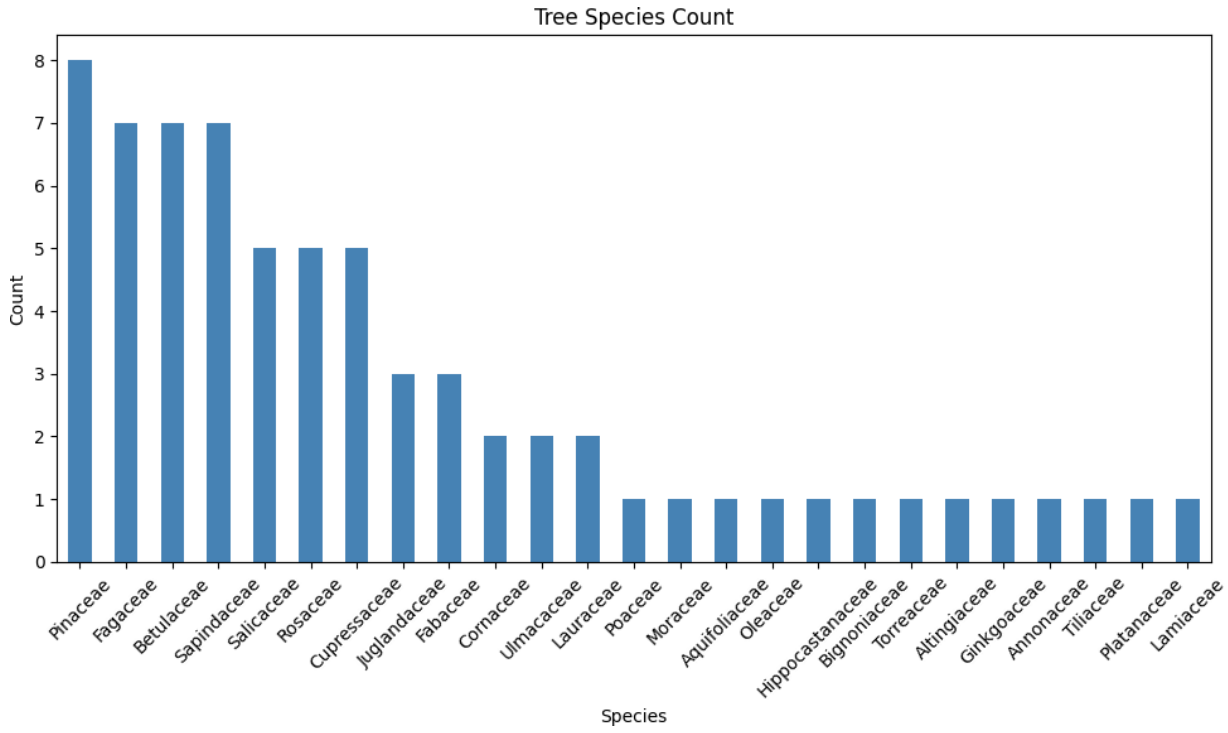
DATASET LINK: **Trees in the City of Melbourne**

Dataset Columns :

1. com_id: A unique identifier for each tree in the dataset.
2. common_name: The common name of the tree species (e.g., Oak, Maple).
3. scientific_name: The scientific name of the tree species.
4. genus: The genus to which the tree species belongs.
5. family: The family classification of the tree species.
6. diameter_breast_height: The diameter of the tree measured at breast height (typically 1.3 meters above ground). This is an important metric for assessing tree health and size.
7. year_planted: The year the tree was planted. This helps in determining the age of the tree.
8. date_planted: The exact date the tree was planted (not always used in analysis).
9. age_description: A categorical description of the tree's age (e.g., Young, Mature, Old).
10. useful_life_expectency: A description of the expected useful life of the tree (e.g., short, long).

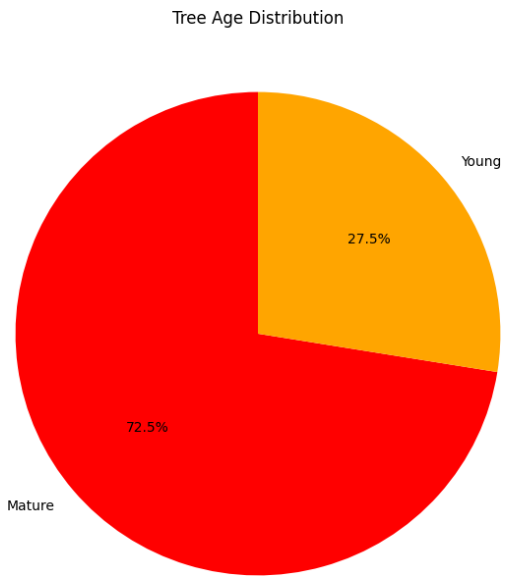
11. useful_life_expectency_value: A numerical representation of the expected useful life (in years).
12. precinct: The precinct or area where the tree is located.
13. located_in: More specific location details, potentially including park or street names.
14. uploaddate: The date the data was uploaded to the dataset.
15. Coordinate location: The coordinates for mapping the tree's location (not always used in visualizations).
16. latitude: The latitude coordinate of the tree's location.
17. longitude: The longitude coordinate of the tree's location.
18. easting: The easting coordinate for mapping.
19. northing: The northing coordinate for mapping.
20. geolocation: A combination of latitude and longitude for precise location identification.

BASIC VISUALIZATIONS:



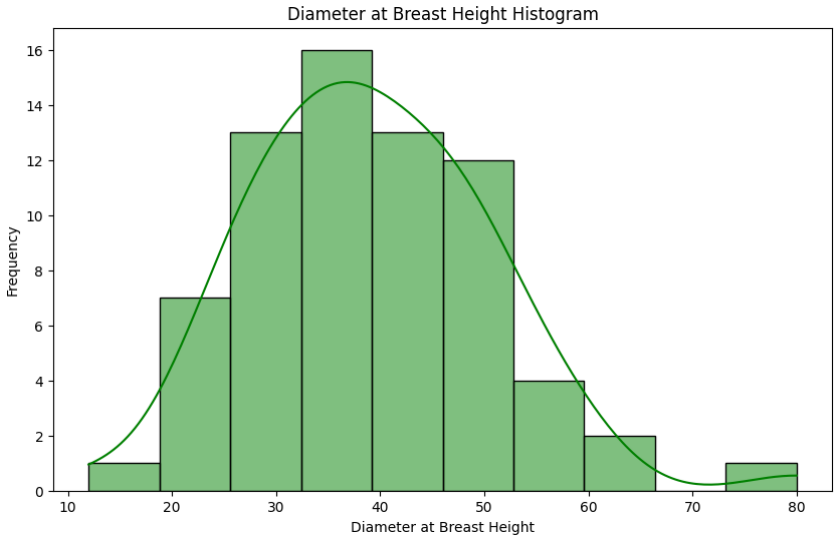
Bar Chart: Tree Species Count

Insight: This chart shows the number of trees for each species. A high count of pinaceae indicates its popularity in landscaping. If certain species are thriving, consider promoting their planting in other areas.



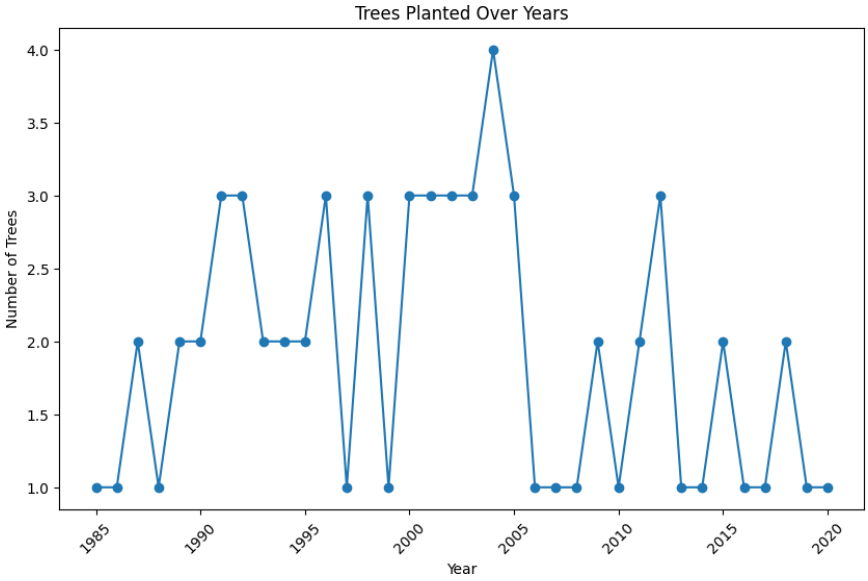
Pie Chart: Tree Age Distribution

Insight: This chart represents the distribution of trees by age category. A balanced distribution suggests a healthy mix of all which is beneficial for biodiversity. But this chart shows more mature trees and less young.



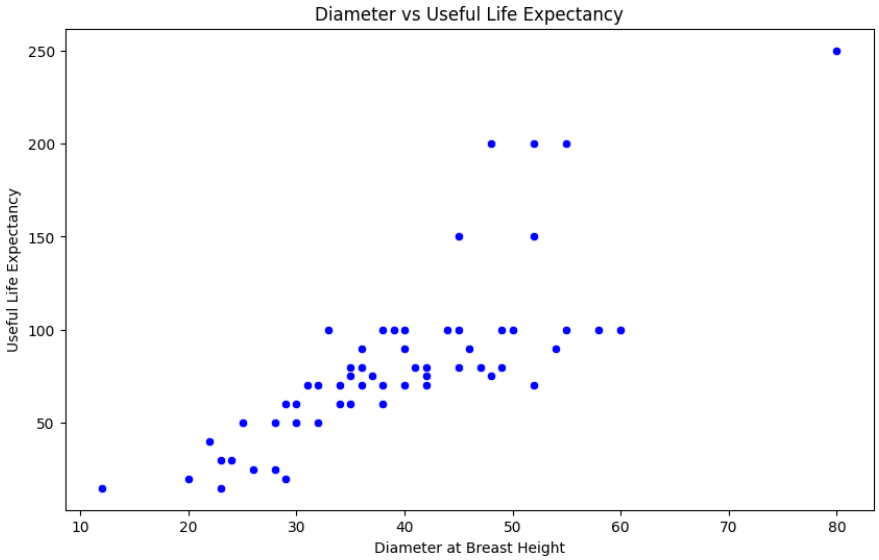
Histogram: Diameter at Breast Height

Insight: The histogram displays the frequency distribution of tree diameters. A concentration of trees at 30- 40 diameter ranges indicates successful growth conditions.



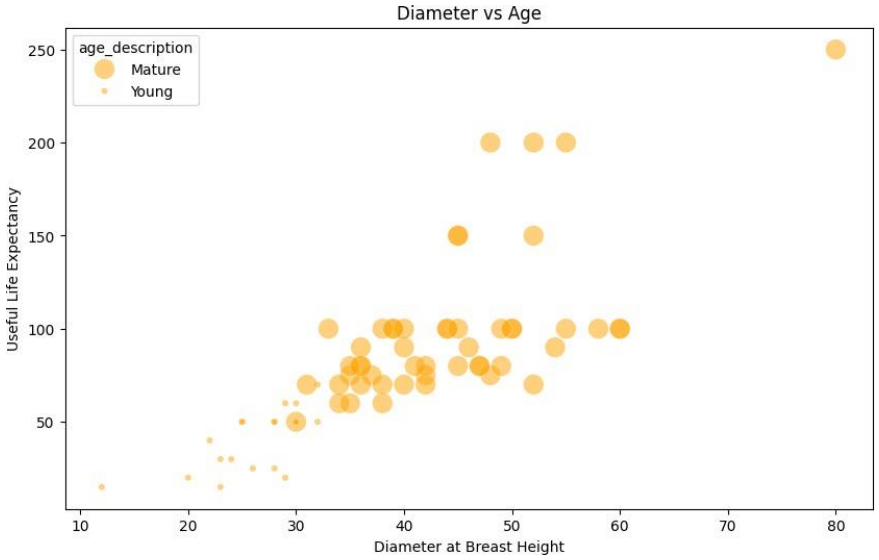
Timeline Chart: Trees Planted Over Years

Insight: This line chart shows the number of trees planted each year. A growing trend in 2000 indicates increasing efforts in tree planting, while a decline after 2005 suggests a need for renewed initiatives.



Scatter Plot: Diameter vs. Useful Life Expectancy

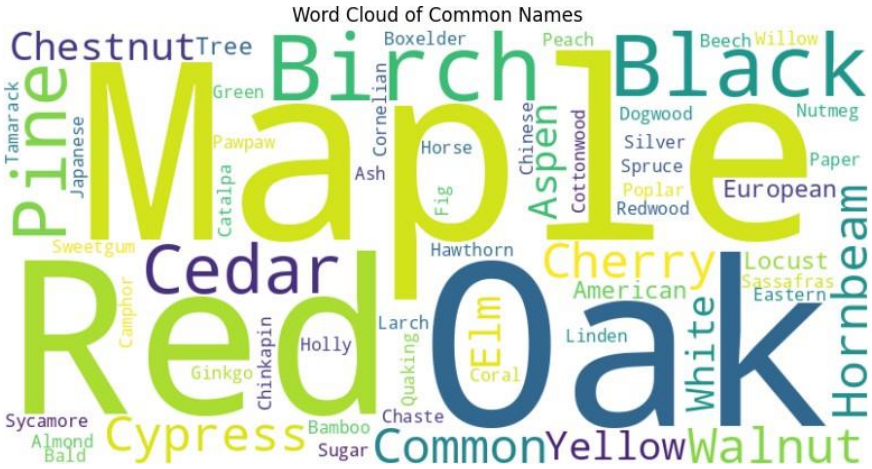
Insight: This scatter plot illustrates the relationship between tree diameter and life expectancy. A positive correlation suggests that larger trees tend to have longer life expectancies.



Bubble Plot: Diameter vs. Age

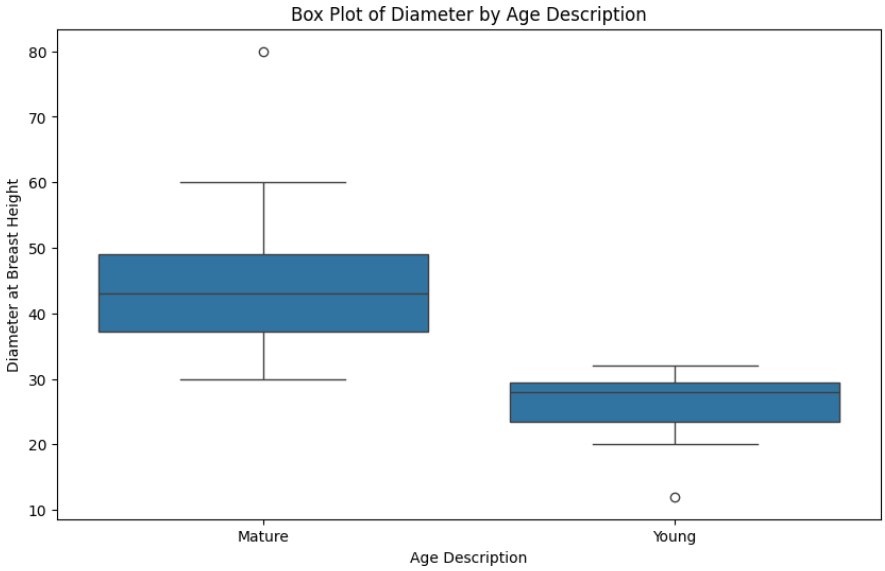
Insight: The bubble plot adds another dimension by representing age categories with varying bubble sizes. This highlights the relationship between size, age, and species.

Advanced visualizations



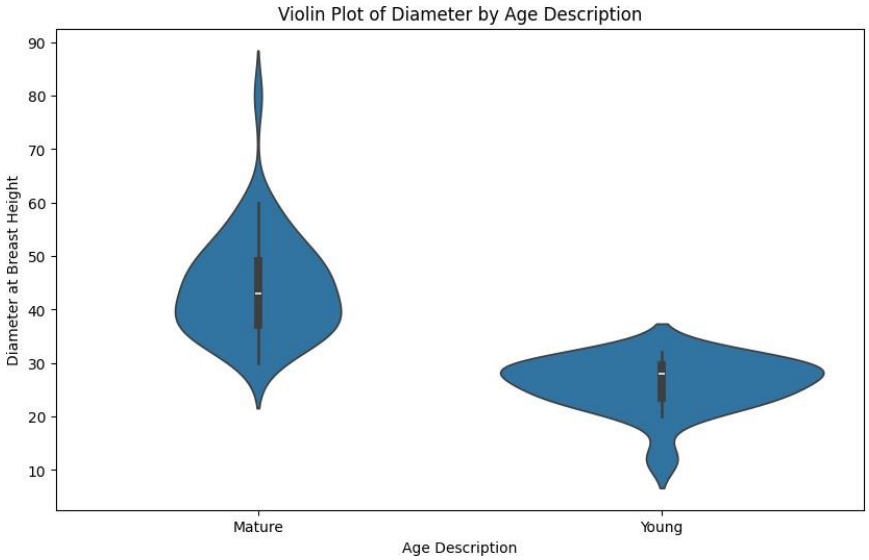
Word Cloud

Insight: This word cloud visualizes the most common species names, helping to identify the dominant species in the dataset.



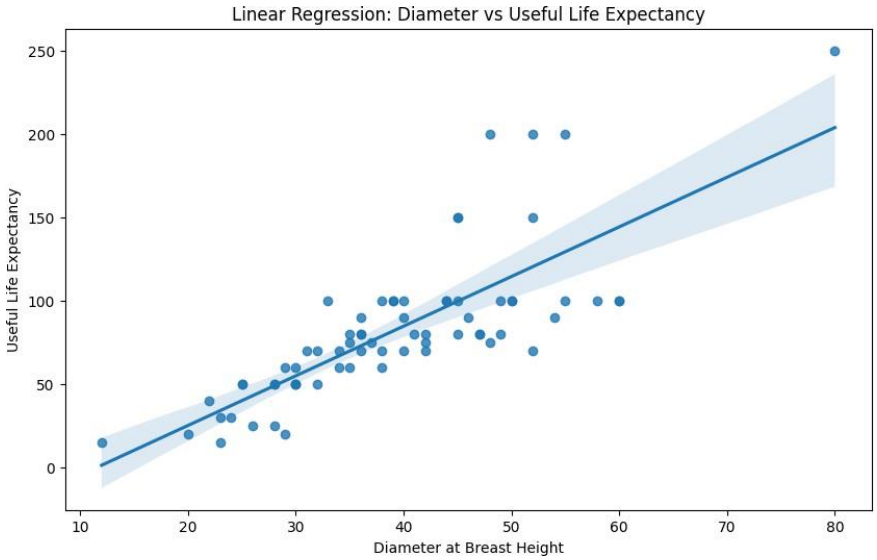
Box Plot: Diameter

Insight: This box plot summarizes the distribution of diameters across age categories, showing the median, quartiles, and potential outliers.



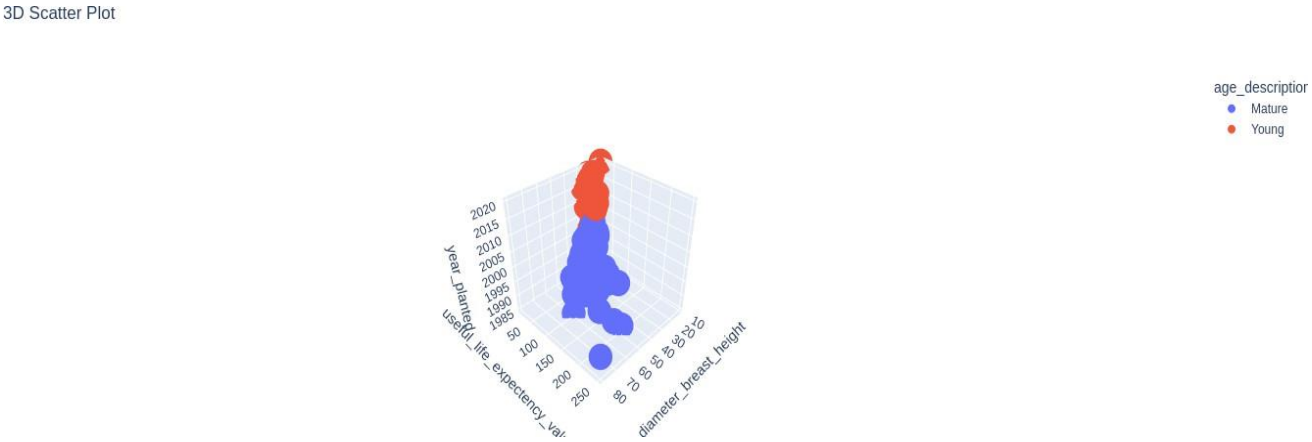
Violin Plot

Insight: This violin plot shows the distribution of diameters, giving insight into both the density of data points and summary statistics. This shows a higher existence of the mature plants according to Diameter.



Linear Regression Plot

Insight: This plot indicates that there's a statistically significantly positive relationship between diameter and life expectancy, supporting predictive modeling.



3D Scatter Plot

Insight: This provides a more comprehensive view of relationships among diameter, life expectancy, and year planted according to the age description.

Jitter Plot

Insight: This jitter plot displays the distribution of useful life expectancy across age categories, showing potential trends of mature plants having more life expectancy than young.

Conclusion:

These visualizations provide a comprehensive overview of the dataset, enabling insights into tree species, growth patterns, and management strategies. They can be instrumental in guiding conservation efforts, urban forestry practices, and environmental planning.