Nagar Yuwak Shikshan Sanstha's
Yeshwantrao Chavan College of Engineering
(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)
Hingna Road, Wanadongri, Nagpur - 441 110
NAAC A++
Ph.: 07104-237919, 234623, 329249, 329250 Fax: 07104-232376, Website: www.ycce.edu

Department of Artificial Intelligence & Data Science

Vision of the Department

To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration.

Mission of the Department

To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problem-solving skills through emerging technologies.

Session 2025-2026

Vision: Dream of where you want.	**Mission:** Means to achieve Vision

Program Educational Objectives of the program (PEO): (broad statements that describe the professional and career accomplishments)

PEO1	**Preparation**	**P: Preparation**	**Pep-CL abbreviation pronounce as Pep-si-lL easy to recall**
PEO2	**Core Competence**	**E: Environment (Learning Environment)**	
PEO3	**Breadth**	**P: Professionalism**	
PEO4	**Professionalism**	**C: Core Competence**	
PEO5	**Learning Environment**	**L: Breadth (Learning in diverse areas)**	

Program Outcomes (PO): (statements that describe what a student should be able to do and know by the end of a program)

Keywords of POs:
Engineering knowledge, Problem analysis, Design/development of solutions, Conduct Investigations of Complex Problems, Engineering Tool Usage, The Engineer and The World, Ethics, Individual and Collaborative Team work, Communication, Project Management and Finance, Life-Long Learning

PSO Keywords: Cutting edge technologies, Research

"I am an engineer, and I know how to apply engineering knowledge to investigate, analyse and design solutions to complex problems using tools for entire world following all ethics in a collaborative way with proper management skills throughout my life." *to contribute to the development of cutting-edge technologies and Research*.

Integrity: I will adhere to the Laboratory Code of Conduct and ethics in its entirety.

Name and Signature of Student and Date
(Signature and Date in Handwritten)

Nagar Yuwak Shikshan Sanstha's
Yeshwantrao Chavan College of Engineering
(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)
Hingna Road, Wanadongri, Nagpur - 441 110
NAAC A++
Ph.: 07104-237919, 234623, 329249, 329250 Fax: 07104-232376, Website: www.ycce.edu

Department of Artificial Intelligence & Data Science

Vision of the Department

To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration.

Mission of the Department

To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problem-solving skills through emerging technologies.

Session	2025-26 (ODD)	**Course Name**	BIG DATA AND HADOOP-LAB
Semester	7 AIDS	**Course Code**	22ADS704
Roll No	03	**Name of Student**	**Debasrita Chattopadhyay**

Practical Number	09
Course Outcome	1. Understand big data analytics and its business applications. 2. Analyse the HADOOP and Map Reduce technologies associated with big data analytics. 3. Apply Big Data analytics Using Pig and Hive.
Aim	Analysing Various Data Visualization Methods Using R
Problem Definition	Analysing Various Data Visualization Methods Using R
Theory (100 words)	Data visualization is the visual representation of data that allows users to identify, for example, patterns, trends, and relationships in a data set. R is a statistical programming language that provides advanced functionality for data visualization with libraries that include ggplot2, lattice, and the base R plotting functions. Dataviz can be, for example, a single plot of each of simple histograms, bar charts, scatter plots, or advanced dataviz such as boxplots, heatmaps or interactive dashboards. Choosing the correct visual representation of data is important because it depends on the type of data (categorical or numerical), what the distribution is and the purpose of the analysis. Data visualization makes the data more interpretable, permits communicating the information and valuable insights, and can assist in future decision making.
Procedure and Execution (100 Words)	Steps of Implementation: - • Install and load required libraries: ggplot2, dplyr, lattice. • Load the dataset using read.csv() or other appropriate function. • Check data structure and summary with str() and summary(). • Handle missing values or clean data if necessary. • Create histograms with geom_histogram() for numerical distributions. • Create boxplots with geom_boxplot() to detect outliers. • Create density plots with geom_density() for smooth distribution visualization.

	• Create scatter plots with geom_point() and optional geom_smooth() for correlations. • Compute correlation matrix and visualize with corrplot(). • Create bar charts with geom_bar() for categorical data counts. • Create pie charts using pie() for categorical proportions. • Use lattice plots like histogram() or xyplot() for multivariate analysis. • Use faceted plots in ggplot2 with facet_wrap() or facet_grid() for category-wise comparisons.

	Code:
	# 1. Load built-in dataset
	data <- iris
	head(data)
	
	# 2. Install and load necessary libraries
	install.packages(c("ggplot2", "corrplot", "lattice"))
	library(ggplot2)
	library(corrplot)
	library(lattice)
	
	# 3. Histogram - Petal Length Distribution
	ggplot(data, aes(x = Petal.Length, fill = Species)) +
	  geom_histogram(binwidth = 0.3, color = "black", alpha = 0.7) +
	  ggtitle("Histogram of Petal Length by Species")
	
	# 4. Boxplot - Compare Sepal Width across Species
	ggplot(data, aes(x = Species, y = Sepal.Width, fill = Species)) +
	  geom_boxplot() +
	  ggtitle("Boxplot of Sepal Width by Species")
	
	# 5. Scatter Plot - Sepal Length vs Petal Length
	ggplot(data, aes(x = Sepal.Length, y = Petal.Length, color = Species)) +
	  geom_point(size = 3) +
	  ggtitle("Sepal Length vs Petal Length")
	
	# 6. Correlation Heatmap for numeric columns
	numeric_data <- data[, 1:4]
	corr_matrix <- cor(numeric_data)
	corrplot(corr_matrix, method = "color", type = "upper", tl.col = "black",

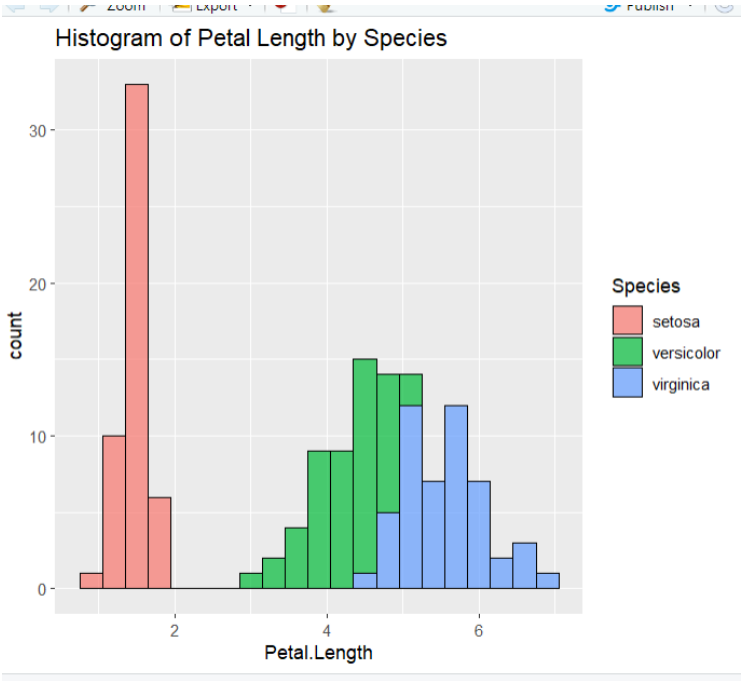
```
        addCoef.col = "black", tl.srt = 45)

# 7. Density Plot - Petal Width by Species
ggplot(data, aes(x = Petal.Width, fill = Species)) +
  geom_density(alpha = 0.6) +
  ggtitle("Density Plot of Petal Width by Species")

# 8. Lattice Plot - Sepal Length vs Sepal Width by Species
xyplot(Sepal.Length ~ Sepal.Width | Species, data = data,
       layout = c(3,1), main = "Sepal Length vs Sepal Width by
Species")
```

Output:

```
> head(data)
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1          5.1         3.5          1.4         0.2  setosa
2          4.9         3.0          1.4         0.2  setosa
3          4.7         3.2          1.3         0.2  setosa
4          4.6         3.1          1.5         0.2  setosa
5          5.0         3.6          1.4         0.2  setosa
6          5.4         3.9          1.7         0.4  setosa
> -|
```

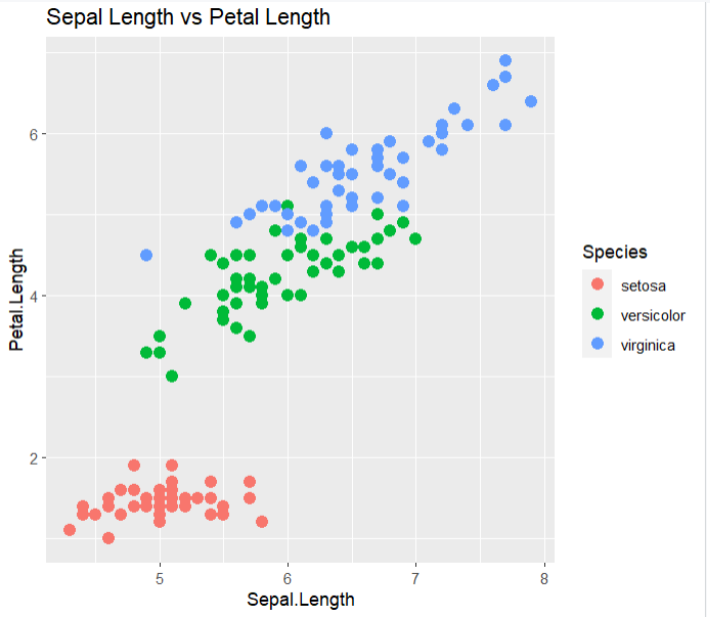


Vision of the Department

To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration.

Mission of the Department

To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problem-solving skills through emerging technologies.

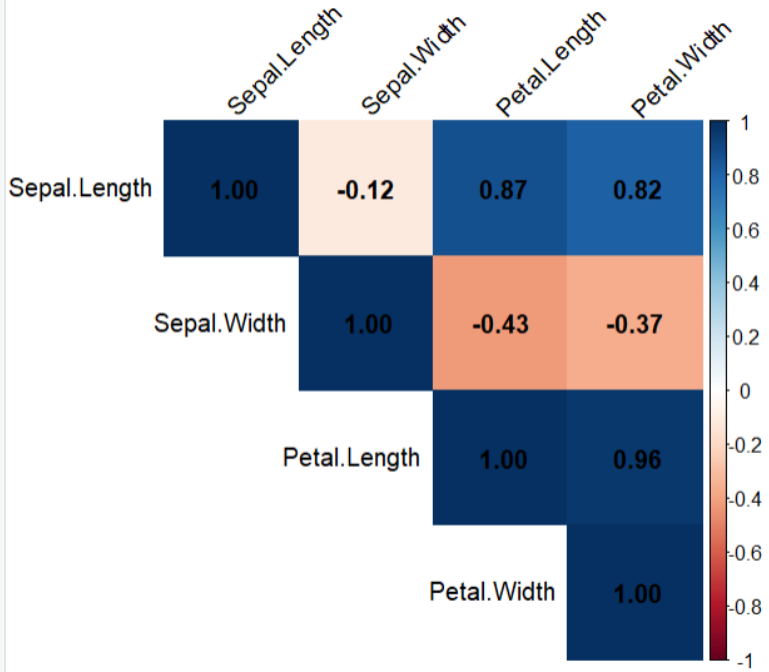


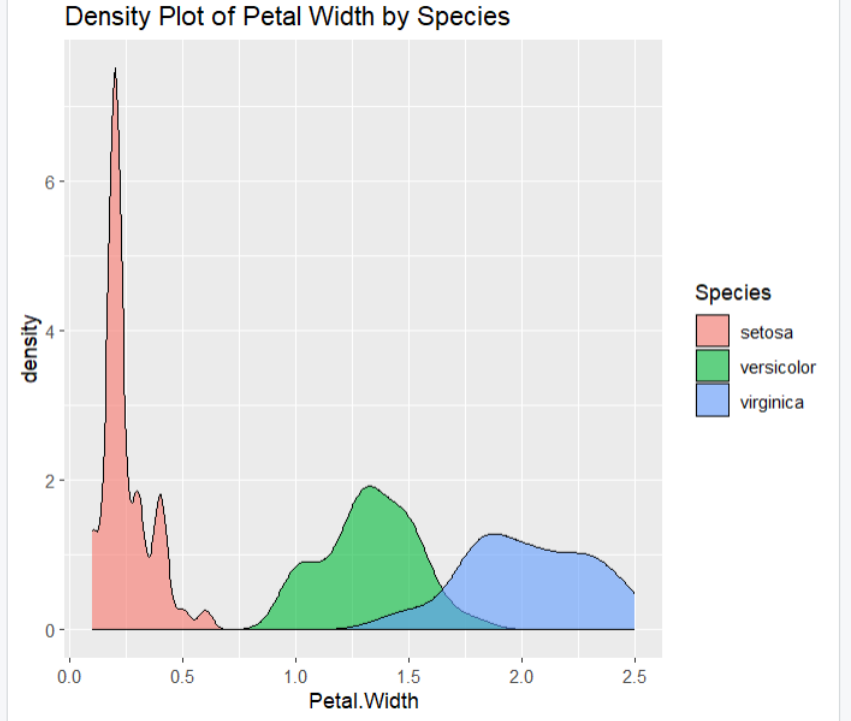
Vision of the Department

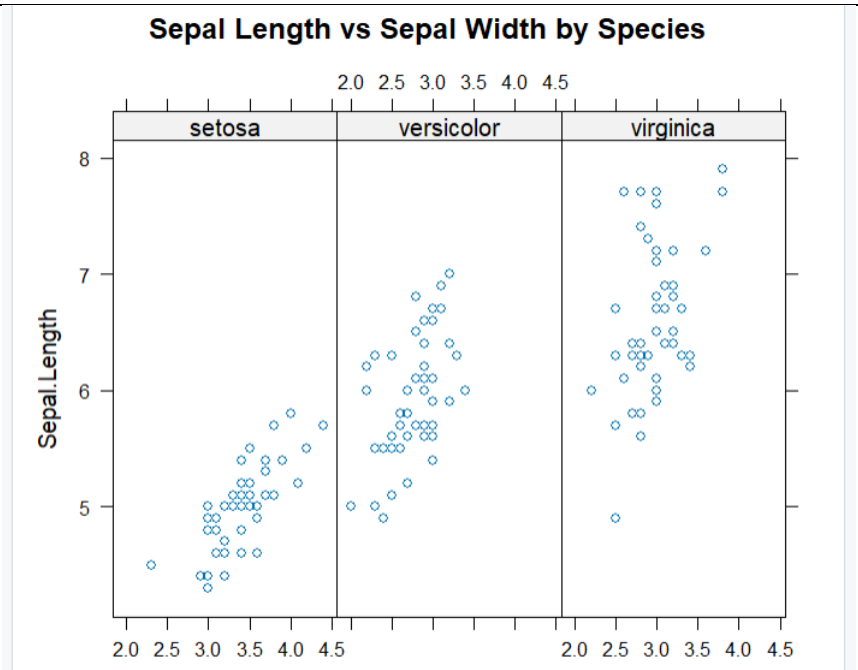
To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration.

Mission of the Department

To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problem-solving skills through emerging technologies.





	
Output Analysis	**Histogram**: Displays the frequency distribution of a numerical variable, helping detect skewness, normality, or clusters. **Boxplot**: Highlights median, quartiles, and outliers, allowing comparison between categories. **Density plot**: Smooth representation of distribution, useful for comparing multiple groups. **Scatter plot**: Shows correlation or trends between two numerical variables. Regression line can indicate direction and strength of relationship. **Heatmap (Correlation matrix)**: Visualizes correlations among numerical variables. Strong positive or negative correlations are easily spotted. **Bar/Pie charts**: Summarize categorical data frequencies or proportions. Easy to interpret at a glance. **Lattice and Faceted plots**: Useful for multivariate analysis, showing variable relationships across multiple categories.
Link of student Github profile where lab assignment has been uploaded	
Conclusion	Analyzing Various Data Visualization Methods Using R implemented successfully.

Nagar Yuwak Shikshan Sanstha's
Yeshwantrao Chavan College of Engineering
(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)
Hingna Road, Wanadongri, Nagpur - 441 110
NAAC A++
Ph.: 07104-237919, 234623, 329249, 329250 Fax: 07104-232376, Website: www.ycce.edu

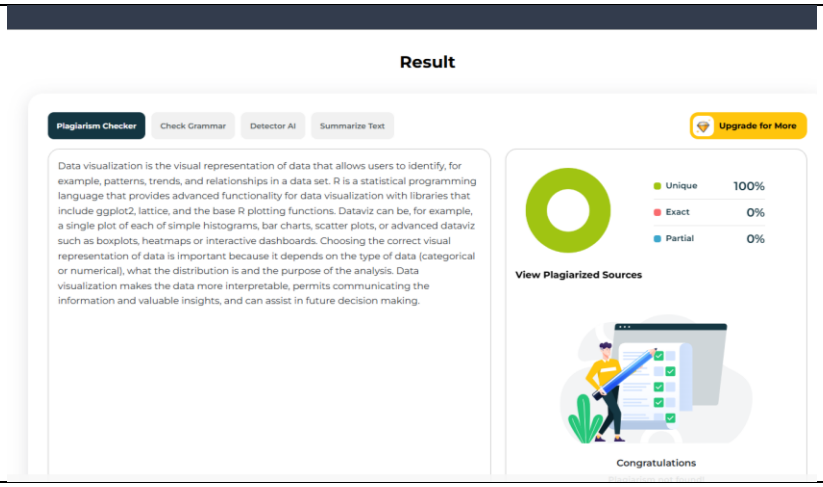
Department of Artificial Intelligence & Data Science

Vision of the Department

To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration.

Mission of the Department

To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problem-solving skills through emerging technologies.

Plag Report (Similarity index < 12%)	
Date	**16 /10 / 25**