OPEN ACCESS

Eco-Shift 2026: International Conference on Global Climate Action
(From Root Causes to Revolutionary Remedies for Sustainable Development)

Int. Jr. of Contemp. Res. in Multi. *PEER-REVIEWED JOURNAL* *Volume 5 [Special Issue 2] Year 2026*



Research Article

Compact Data Visualization for Feature Relationships and Class Separability

Shridhar B. Dandin [1*], Santosh Kumar Das [2], Megha Sinha [3]

[1] Department of Computer Science and Engineering, Sarala Birla University, Ranchi, Birla Knowledge City, Mahilong, Purulia Road, Jharkhand, India

[2] Department of Computer Science and Engineering, Sarala Birla University, Ranchi, Birla Knowledge City, Mahilong, Purulia Road, Jharkhand, India

[3] Department of Computer Science and IT, Arka Jain University, Jamshedpur, Jharkhand, India

Corresponding Author: * Shridhar B. Dandin

DOI: https://doi.org/10.5281/zenodo.23641572

Abstract

Data visualization is mandatory for different applications, such data interpretation, exploratory analysis, pattern identification, etc. It is rapidly used in the domain of Artificial Intelligence & Machine Learning (AI/ML), healthcare and medical analytics, cybersecurity, finance, IoT, etc. Although it is effectively used in several areas, but it has some issues such as large volume of data, choosing the right visualization technique, data quality, user interpretation, etc. In this paper, five visualization methods are used to illustrate a compact multi-view visualization in the form of an analytical. Finally, the results show that each chart illustrates a specific insight. Among all visualizations, scatter visualization provides richer information with the context of feature relationship and class separability.

Manuscript Information

- **ISSN No:** 2583-7397
- **Received:** 10-09-2026
- **Accepted:** 23-10-2026
- **Published:** 25-10-2026
- **IJCRM:**5(SP2); 2026: 212-219
- **©2026, All Rights Reserved**
- **Plagiarism Checked:** Yes
- **Peer Review Process:** Yes

How to Cite this Article

Dandin B S, Das K S, Sinha M. Compact Data Visualization for Feature Relationships and Class Separability. Int J Contemp Res Multidiscip. 2026;5(SP2): 212-219.

Access this Article Online

www.multiarticlesjournal.com

KEYWORDS: Data Visualization, Visual Analytics, Exploratory Data Analysis, Feature Relationships, Class Separability.

© 2026 Shridhar B. Dandin, Santosh Kumar Das, Megha Sinha. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND). https://creativecommons.org/licenses/by-nc-nd/4.0/

1. INTRODUCTION

In the last few decades, the application of data visualization increases rapidly in different areas [1]. The main reason behind this, it easily explores the data analysis and transform the categorical and numerical information in the form of graphical representation. It identifies observation, distribution, pattern, relationship, etc. [2-4]. Despite its benefits and widespread application, it has several challenges like user interpretation, real-time data handling, data quality, choosing the right chart, large volume of data. Data interpretation indicates different users interpret differently with the context of same visualization [5]. It raises due to poorly or complex chart design. Data handling in real-time become more difficult, due to continuously changing the data and its related parameters. Data quality becomes poor when the users deal with incorrect, missing and inconsistent data that creates misleading visualization [6]. Choosing right visualization at right time also become difficult due to variation of data. Sometime, huge dataset becomes the graph as crowded which difficult to understand [7].

Motivation

The motivation of this paper, to use different visualization technique in Iris dataset. It helps to represent the different characteristics and relationship within the dataset is more effectively. Finally, it aims that which visualization provides the better insight with the context of feature variation and comparison, relationship and distribution.

Contribution

The main contribution of this paper is as follows:

a) The six visualization techniques are used for comprehensive data analysis
b) The comparison is based on different abilities of representation as variation, frequency, distribution, etc.
c) The study also investigates the relationship among different parameters.
d) The study is also identifying the more effective visualization.

Organization of the Paper

Section 2 discusses several existing works as a literature review. Section 3 describes the main methodology of the proposed method and in the Section 4 discusses performance evaluation of the proposed method. Finally, in the last section i.e. Section 5 discusses conclusions of the paper.

2. LITERATURE REVIEW

In the last few years, several works have been proposed in the area of data visualization. Some of the works are discussed in this section. T. Xiao et al. [8] proposed a method for a social network service to support a document embedding system and topic transition visualization. It analyses the abstract of the document based on two-dimensional space. The visualization of the work uses time series analysis for word analysis. Finally, it analyses different data such as user ratings, COVID-19 data, and nuclear disaster data. V. Molchanov and L. Linsen [9] proposed a method for preserving shape in star coordinates. In this method, all dimensions are explored within the whole set of attributes. Analysis tasks detect clusters trends and outliers. The method focuses on minimizing loss of visualization related information. In this paper, the authors used 2D visual space technique along with an encoding method for mapping more than one-dimensional data. It improves the tool for exploration by linear mapping. M. Fan and L. Zhou [10] proposed a method for medical visualization aimed design purposes. The proposed method consists of cognitive habits and analytics logic for distinguishing among the design, target user analysis, and classifying the task based on descriptive analysis. Finally, it analyses three models of medical visualization. F. Hong et al. [19] proposed a method for ensemble data aimed visual analytics. The main purpose of this work is to realize set analysis with large-scale data. It is also used to find relation among data. It helps to design two-dimension matrix for ensemble data. The visualization contains similar feature in subspace for each intersection. G. Wu et al. [12] proposed a method for visual analysis in the context of damage data from simulations. It uses an integrated visual method for defect analysis as clustering and identification and graph tracking. It overcomes the issue of blur modes and defect dynamics in the visualization. J. Li et al. [13] proposed a method for geoscience data analysis which is based on interactive visualization. The work is based on more than one parameter with a time series analysis dataset. It contains observation-based subset selection by a domain expert. It illustrates high-level analysis in compact view mode. The spatial information consists of attribute and temporal aspect for center visualization. K. B. Schloss et al. [14] proposed a method for colormaps to support data visualization. In this method, the colormaps were influenced by relations between colors within colormaps and colors of the background. The proposed method analysed color quantity mapping to map colors in data visualization and helps to resolve several conflicts of data visualization. R. A. Rensink and G. Baldridge [15] proposed a perception method for scatterplots. The method analysed the visual perception correlation based on classical psychophysical techniques. The accuracy is measured by parameters of scatterplots for lower and upper values. Finally, it outperforms the result based on visualization. W. Wu et al. [16] proposed a visualization method for family classification and imbalanced analysis. The method is based on malware visualization that uses matrix suitability and transition probability analysis. The CNN model is used for input image analysis and prediction. Finally, it outperforms the results in the context of classification and compatibility.

Proposed Method

In this section, the proposed method is described based on Iris dataset. There are several visualization techniques are used in model to predict and analyzed several features. Initially, dataset is loaded into the Python module then apply for data cleaning and parameterized analyse. Each visualization method provides

© 2026 Shridhar B. Dandin, Santosh Kumar Das, Megha Sinha. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND). https://creativecommons.org/licenses/by-nc-nd/4.0/

a specific analytical purpose and helps to provides different insights based on parameters. Some statistical analysis is used to generate insight according to parameters of the dataset. The qualify assessment is also highlight for enabling the user examine for variable aspects. It provides informed decision making that helps interpret data effectively. It helps to present complex information in interpretable form. The overall analysis workflow shown in the Fig. 1. The algorithm of this model is shown in Algorithm 1.

Algorithm 1: Methodology of the Comparative Visualization Technique

Input: Iris dataset Output: Multiple visualizations and its related observations
Step 1: Start **Step 2:** Load the Iris dataset by pd.read_csv("iris.csv"). **Step 3:** Select the required attributes as Species, Petal Length and Petal Width. **Step 4:** Check the dataset for missing values. **Step 5:** Separate the analyzed observations based on species. **Step 6:** Find the average petal length for each species. **Step 7:** Illustrates a line plot to observe petal-length variation across samples. **Step 8:** Illustrates a bar chart to compare the average petal length of each species. **Step 9:** Illustrates a histogram to analyses the distribution of petal length. **Step 10:** Illustrates the range-frequency chart for dividing petal-length values. **Step 11:** Illustrates a pie chart to show the proportion of each Iris species. **Step 12:** Illustrates a scatter plot using petal length and petal width with different color-coded species. **Step 13:** Compare the visualizations based on features selection as distribution, comparison. Variation, relationship and species separation. **Step 14:** Identify the most insightful visualization for analytical purpose. **Step 15:** Stop

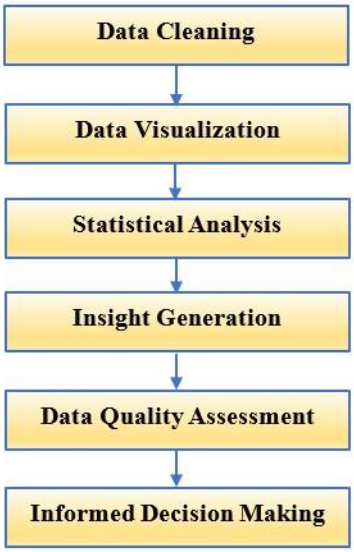


Figure 1. Data Analysis Workflow for Informed Decision-Making.

Performance Evaluation

In this section performance evaluation is discuss, there are several parameters are used in this simulation analysis. The Table 1 shows dataset and experimental parameters. The visualization environment and its related software is shown in Table 2. The simulation analysis is done in the laboratory system which configuration shown in Table 3.

Table 1: Dataset and Experimental Parameters.

Parameter	Specification
Dataset	Iris Dataset
Total samples	150
Number of classes	3

© 2026 Shridhar B. Dandin, Santosh Kumar Das, Megha Sinha. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND). https://creativecommons.org/licenses/by-nc-nd/4.0/

Samples per class	50
Numerical features	4
Features used for analysis	Petal Length, Petal Width
Data analysis type	Exploratory Data Analysis
Visualization methods	Line Chart, Bar Chart, Histogram, Pie Chart, Scatter Plot

Table 2: Software and Visualization Environment.

Parameter	Specification
Programming Language	Python 3.x
Data Processing	Pandas
Visualization Library	Matplotlib
Development Environment	Jupyter Notebook / Python IDE
Operating System	Windows 10/11
Data Format	CSV
Visualization Output	2D graphical plots

Table 3: Hardware Configuration.

Parameter	Specification
Processor	Intel Core i5 / equivalent
RAM	8 GB
Storage	256 GB SSD
System Architecture	64-bit
Display	Full HD / equivalent
Graphics Processing	Integrated Graphics
Internet Requirement	Not required for visualization execution

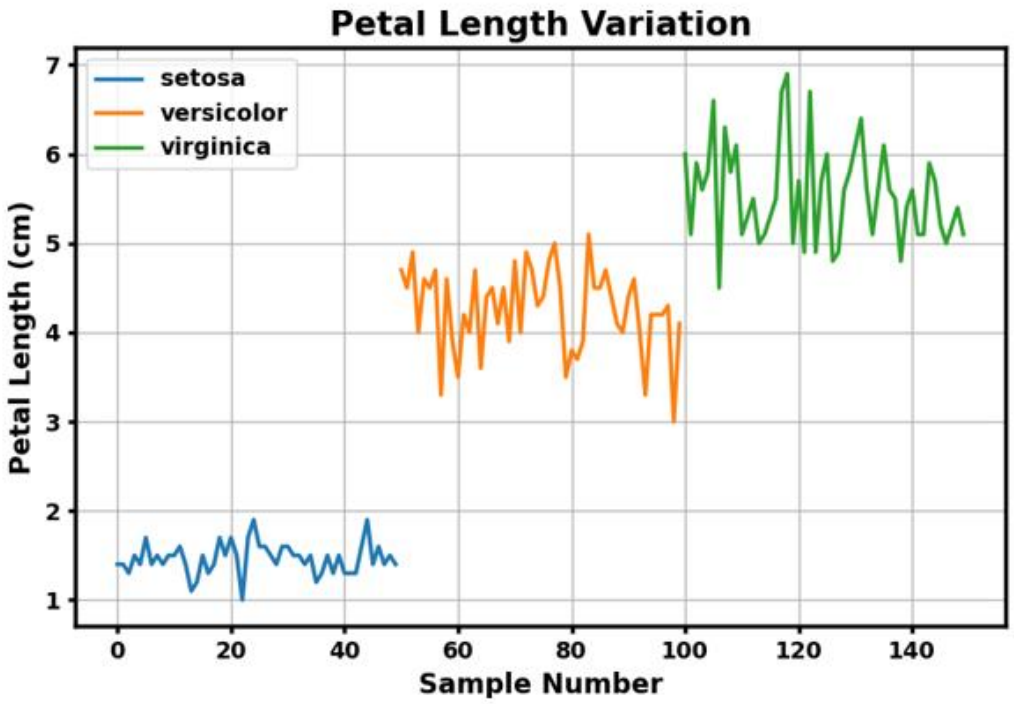


Figure 2. Comparative Analysis of Petal Length for Iris Species**.**

Fig. 2 shows comparative analysis of Petal length for Iris species where X-axis show sample number or index of the dataframe. And Y-axis shows Petal length in cm. The range of the X is between 0 to 149 where distribution is based on three species as 0 to 49 Setosa, 50 to 99 Versicolor and 100 to 149 Virginica. The approximate range of the Petal length is 1.0 cm to 7.0 cm. Three color lines indicates as Setosa in blue color, Versicolor in orange color, and Virginica in green color. The

© 2026 Shridhar B. Dandin, Santosh Kumar Das, Megha Sinha. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND). https://creativecommons.org/licenses/by-nc-nd/4.0/

Petal length of Setosa distributed between 1.0 to 1.9 cm, Versicolor distributed between 3.0 to 5.1 cm, and Virginica distributed beween 4.5 to 6.9 cm. The Setosa remains at the lowest value due to shortest Petals. Versicolor lies middle and indicate medium variation. Virginica remains highest value due to longest Petals.

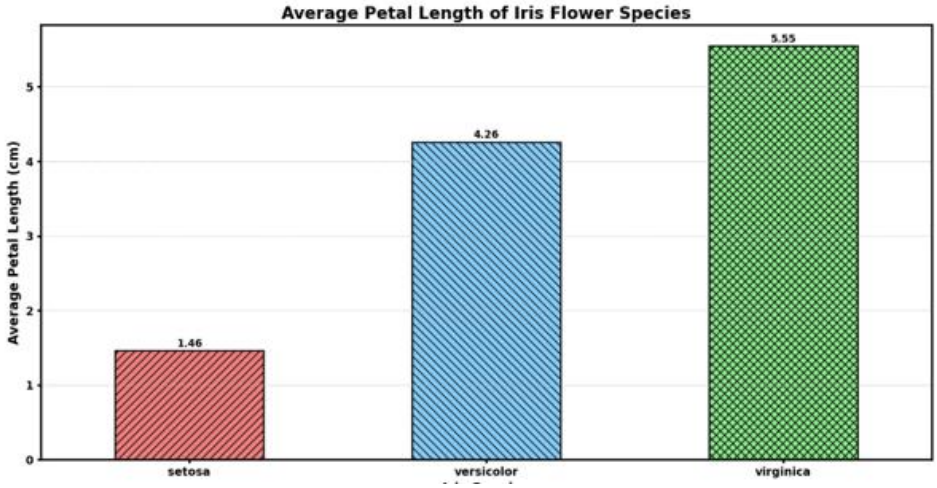


Figure 3. Comparative Analysis of Average Petal Length among Iris Species.

Fig. 3 shows comparative analysis of average Petal length of three Iris species. The bar chart clearly indicates that Petal length is vary among three different species. Virginica is 1.29 cm longer than Versicolor and 4.09 cm longer than Setosa. And same ways Versicolor's Petal is longer 2.80 cm than Setosa. This variation illustrates that Petal length is very important for morphological features analysis in different species. Hence, this discrimination is a strong attribute to classify or identify Iris species in machine leaning areas through data visualization.

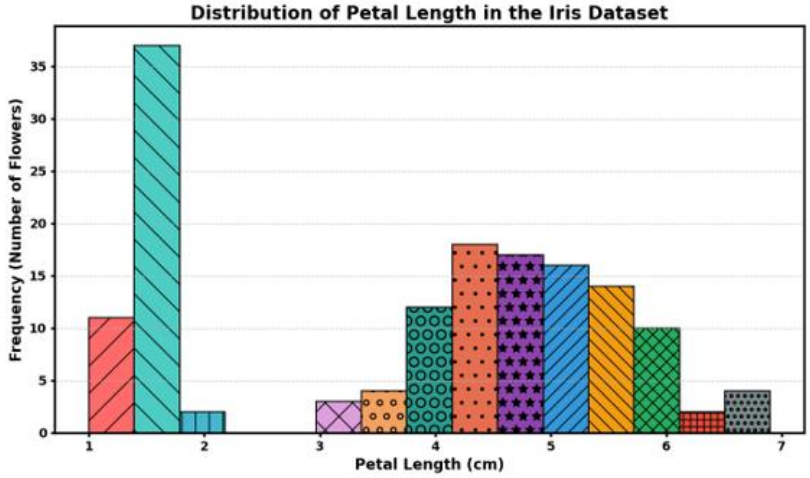
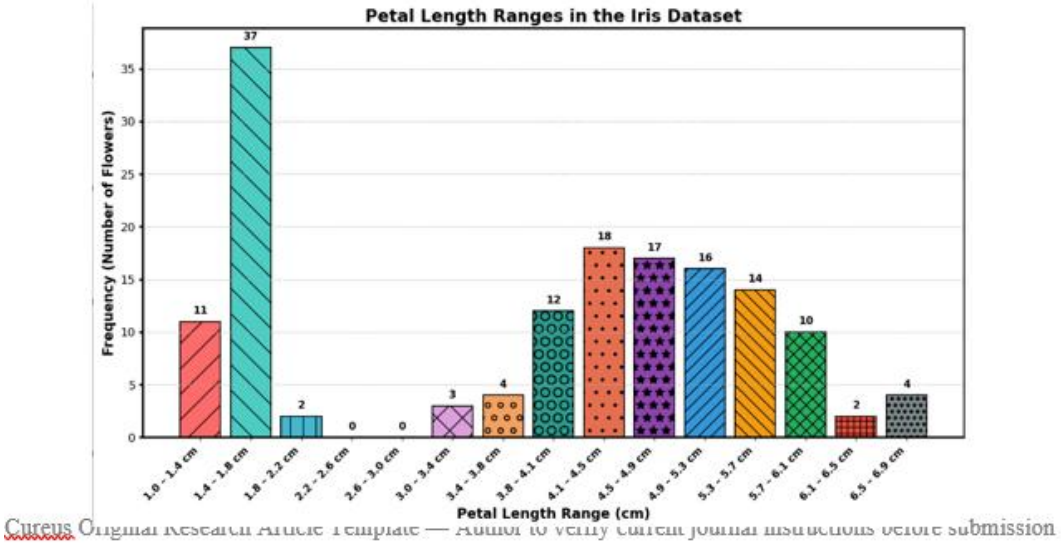


Figure 4. Distribution of Petal Lengths of Iris Flowers Using a Histogram

Fig. 4 shows histogram graph where X-axis indicates Petal length in cm and Y-axis indicates frequence as number of flowers. The value range of Petal length is given from 1 to 7 cm. In the graph, the tallest bar has the frequency approximately 37. It indicates that around 37 flowers having the Petal length in this interval. The second tallest bars are between 4 to 6 cm that indicates several flowers having the Petal length in this range. The range between 2 to 3 cm, very few observations are existed. The histogram in this case useful, because by looking Iris dataset, it is difficult to observe overall pattern and measurement. The histogram graph converts several patterns and measurements in the form of compact visualization.

© 2026 Shridhar B. Dandin, Santosh Kumar Das, Megha Sinha. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND). https://creativecommons.org/licenses/by-nc-nd/4.0/



Cureus Original Research Article Template — Author to verify current journal instructions before submission

Figure 5. Frequency Distribution of Petal Length Ranges in the Iris Dataset.

Fig. 5 shows frequency distribution of Petal length based on different range with frequency. It illustrated that highest frequency between range 1.4 to 1.8 cm as 37 flowers, between 4.1 to 4.5, 18 flowers, between 4.5 to 4.9, 17 flowers, between 4.9 to 5.3, 16 flowers, and few ranges, there is no flowers. The overall graph indicates the Iris flower have different Petal length.

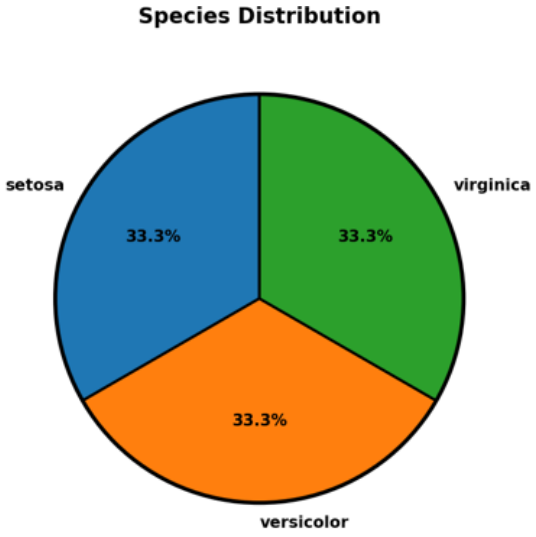


Figure 6. Species Distribution in the Iris Dataset.

Fig. 6 shows percentage distribution of each species in the Iris dataset. Each of the species represents 33.3% in overall dataset. It is easy to understand the proportion in the dataset. It also shows, the contribution of each species in dataset and it balanced across the three species.

© 2026 Shridhar B. Dandin, Santosh Kumar Das, Megha Sinha. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND). https://creativecommons.org/licenses/by-nc-nd/4.0/

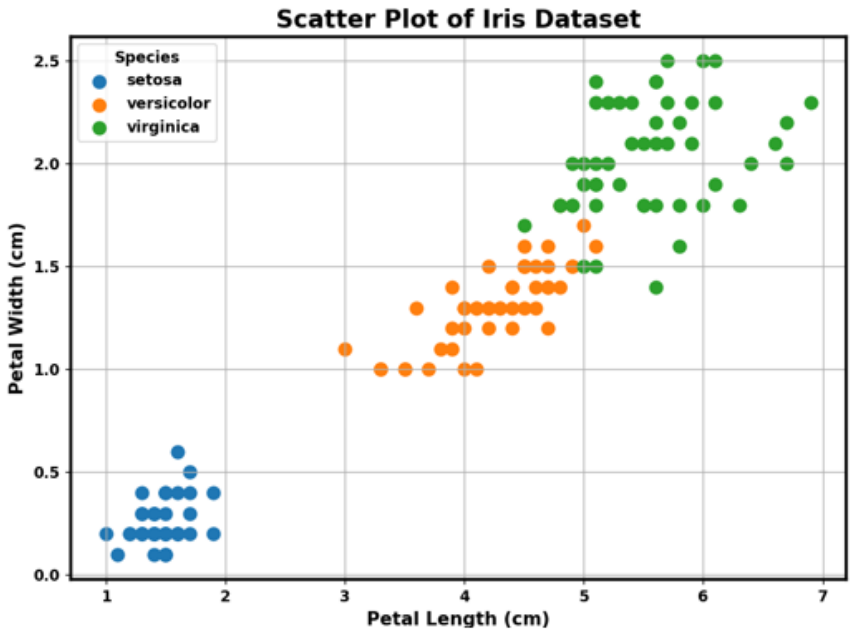


Figure 7. Relationship Between Petal Length and Petal Width Across Iris Species.

Fig. 7 shows relationship between Petal Length and Petal Width in cm of different flowers. In this graph, three colors dots are showing as three different species such as Setosa, Versicolor, and Virginica. Total 150 dots for each species carry 50 dots. Blue Setosa dots in X-axis approximately 1.0 to 1.9 cm and Y axis approximately 0.1 to 0.6 cm, total blue dots are concatenate in lower-left corner. Orange Versicolor dots in X-axis approximately 3.0 to 5.1 cm and Y axis approximately 1.0 to 1.8 cm, total orange color dots are concatenate in the middle area. Green Virginica dots in X axis approximately 4.5 to 6.9 cm and in Y axis 1.4 to 2.5 cm, total green dots are concatenate in the upper right area. Hence, Setosa is clearly illustrated in different, while Versicolor and Virginica have few overlapping, but distinguishable in reasonable. The comparative analysis of different visualization techniques is shown in the Table 4.

Table 4: Comparative Analysis of Visualization Techniques

Visualization	Primary Purpose	Variables Represented	Feature Relationship	Species Separation	Distribution Insight	Individual Sample Visibility
Line Chart	Sample-wise variation	1 numerical + sample order	Limited	Low	Moderate	High
Bar Chart	Average comparison	1 numerical + species	Not represented	High	Low	Low
Histogram	Frequency distribution	1 numerical	Not represented	Low	High	Low
Pie Chart	Species composition	1 categorical	Not represented	Moderate	Low	None
Scatter Plot	Relationship and separation	2 numerical + 1 categorical	Directly visible	Directly visible	Visible	High

The scatter plot provides the most comprehensive representation for examining the relationship between petal length and petal width while simultaneously preserving species-level information. Unlike the other visualizations, it represents individual observations and allows patterns, clusters, overlap, and species separation to be examined within the same visualization.

3. CONCLUSIONS

The comparative analysis demonstrates that each visualization technique provides a distinct perspective on the Iris dataset. The line chart effectively represents sample-wise variation, the bar chart facilitates comparison of average petal lengths, the histogram reveals the frequency distribution, and the pie chart illustrates species composition. Among the evaluated techniques, the scatter plot provides the richest analytical representation by simultaneously displaying petal length, petal

© 2026 Shridhar B. Dandin, Santosh Kumar Das, Megha Sinha. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND). https://creativecommons.org/licenses/by-nc-nd/4.0/

width, and species information. It enables direct observation of feature relationships, clustering patterns, and species separability. Therefore, for exploratory analysis involving multiple numerical features and categorical classes, the scatter plot offers greater interpretive value than the individual univariate visualization techniques considered in this study.

REFERENCES

1. Han L, Song J, Chen S, Zhong Z. Visualization and interaction techniques for single-text digital reading: a survey. Comput Graph. 2025;104481.
2. Wang J, Deng Z, Deng D, Wang X, Sheng R, Cai Y, et al. Empowering multimodal analysis with visualization: a survey. Comput Sci Rev. 2025; 57:100748.
3. Shahzad M, Tah JH, Younas M, Almukhtar A. Technologies and techniques in digital twins for real-time data visualisation in building maintenance: a state-of-the-art review. J Infrastruct Intell Resil. 2025;100185.
4. Das SK, Dey N, Crespo RG, Herrera-Viedma E. A non-linear multi-objective technique for hybrid peer-to-peer communication. Inf Sci. 2023; 629:413-439.
5. Lo LYH, Gupta A, Shigyo K, Wu A, Bertini E, Qu H. Misinformed by visualization: what do we learn from misinformative visualizations? Comput Graph Forum. 2022;41(3):515-525.
6. Lo LYH, Qu H. How good (or bad) are LLMs at detecting misleading visualizations? IEEE Trans Vis Comput Graph. 2025;31(1):1116-1125.
7. Hamed MS, Kwan P, Klich M, Aurisano J, Rajabiyazdi F. Identifying challenges in designing, developing and evaluating data visualizations for large displays. arXiv [Preprint]. 2024. arXiv:2403.14802.
8. Xiao T, Oda N, Onoue Y. Visualization of topic transitions in SNSs through document embedding and dimensionality reduction. J Vis. 2023;26(6):1405-1419.
9. Molchanov V, Linsen L. Shape-preserving star coordinates. IEEE Trans Vis Comput Graph. 2019;25(1):449-458.
10. Fan M, Zhou L. A design study process model for medical visualization. J Vis. 2025;1-21.
11. Hong F, Shao H, Mei X, Pang A, Yuan X. Visual analytics of ensemble data using coupled subspaces. J Vis. 2024;27(5):867-884.
12. Wu G, Lin D, Wang H, Liu L. Visual analysis of defect clustering in 3D irradiation damage simulation data. J Vis. 2022;25(1):31-45.
13. Li J, Meng ZP, Huang ML, Zhang K. An interactive visualization approach to the overview of geoscience data. J Vis. 2017;20(3):433-451.
14. Schloss KB, Gramazio CC, Silverman AT, Parker ML, Wang AS. Mapping color to meaning in colormap data visualizations. IEEE Trans Vis Comput Graph. 2019;25(1):810-819.
15. Rensink RA, Baldridge G. The perception of correlation in scatterplots. Comput Graph Forum. 2010;29(3):1203-1210.
16. Wu W, Peng H, Xu C, Liu Y, Li L. A malware visualization method based on transition probability matrix suitable for imbalanced family classification. Appl Intell. 2025;55(3):236.

Creative Commons (CC) License

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution–Non-Commercial–No Derivatives 4.0 International (CC BY-NC-ND 4.0) license. This license permits sharing and redistribution of the article in any medium or format for non-commercial purposes only, provided that appropriate credit is given to the original author(s) and source. No modifications, adaptations, or derivative works are permitted under this license.

© 2026 Shridhar B. Dandin, Santosh Kumar Das, Megha Sinha. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY NC ND). https://creativecommons.org/licenses/by-nc-nd/4.0/