Computer-Based Data Visualization Analysis for Simplifying Complex Information

Salsabila Nasution[1*], Fatwa Aulia[2], Saprina Putri Utama Ritonga[3], Anggi Jelita Sitepu[4], Supiyandi[5,6]

[1,2,3,4,6]Faculty of Science and Technology, Computer Science Study Program, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia
[5]Faculty of Computational Science and Digital Intelligence, Information Technology, Universitas Pembangunan Panca Budi, Medan, Indonesia
E-mail: [1]*salsabilanasution862@gmail.com, [2]fatwaauliaa08@gmail.com, [3]saprinaputriutamaritonga@gmail.com, [4]anggisitepu006@gmail.com, [5,6]supiyandi.mkom@gmail.com
*E-mail Corresponding Author: salsabilanasution862@gmail.com

Abstract

This study aims to analyze global temperature data by employing computer visualization as a tool to simplify complex information. The dataset was obtained from Kaggle, specifically the Global Land Temperatures by City dataset, which contains monthly average temperature data from various cities worldwide. The methods applied include data preprocessing, descriptive statistical analysis, and data visualization using the Python programming language with the Pandas, Matplotlib, and Seaborn libraries. The visualization results reveal an upward trend in the global average temperature from 1900 to 2020, with an increase of approximately 1°C, indicating the occurrence of global warming. Computer visualization has proven to be effective in helping researchers and policymakers better understand temperature change patterns compared to numerical table-based analysis. Therefore, this study emphasizes that the application of computer visualization is an efficient solution for presenting and analyzing large-scale data, making it more interpretable.

Keywords: *Computer visualization; global temperature data; Python; climate change; big data analysis.*

I. INTRODUCTION

With the rapid advancement of information and communication technology, the amount of data generated in today's digital era has grown exponentially. Characterized by large volumes, high velocity, and diverse formats, this explosion of data represents the phenomenon of big data. Such conditions make the processes of data processing and analysis increasingly challenging, as traditional methods can no longer extract meaningful insights. Computer visualization—an approach that presents data in visual forms such as graphs, maps, diagrams, or interactive dashboards—serves as an effective way to address these challenges (Elvania, 2025).

Computer visualization plays a crucial role in big data analysis because it transforms complex numerical data into visual forms that are easier to comprehend. Through visualization, users can recognize relationships between variables, detect temporal patterns, and understand the dynamics of phenomena more efficiently than by merely examining raw data tables. This method simplifies large datasets by converting them into targeted and communicative visual information, thereby facilitating both analysis and decision-making processes.

In the modern era of data-driven research, visualization has become an essential part of scientific analysis across multiple disciplines, including climate science, epidemiology, and environmental monitoring. The ability to translate large and complex datasets into visual insights allows researchers to uncover hidden correlations that might otherwise remain unnoticed in traditional statistical tables. Furthermore, visual representation supports transparency and reproducibility in scientific communication, as visual patterns are easier to verify and interpret.

Global temperature data is one example of big data that requires visualization. The *Global Land Temperatures by City* dataset, available on the Kaggle platform, contains monthly average temperature records from thousands of cities worldwide, dating back to the 18th century. Due to its vast temporal and spatial dimensions, this dataset is difficult to analyze when presented solely in tabular form. Therefore, a visual approach is necessary to interpret the data more effectively (Setyawan, 2025).

The topic of global temperature analysis is highly significant, as it is directly related to the issues of global warming and climate change. According to the Intergovernmental Panel on Climate Change (IPCC), the Earth's surface temperature has increased by approximately 1.09°C compared to the pre-industrial era, significantly affecting ecosystems and the balance of the environment. These environmental changes also have socioeconomic implications, influencing agriculture, urban development, and energy demand. In this context, computer visualization plays an essential role in making scientific information more accessible and comprehensible. Techniques such as line charts, heatmaps, and scatter plots are among the common forms of computer-based visualization (Nugroho, 2024).

Moreover, integrating visualization with programming languages such as Python provides flexibility and scalability for handling large datasets. Libraries like Pandas, Matplotlib, and Seaborn enable researchers to preprocess, analyze, and visualize climate data efficiently. Compared to commercial software, open-source visualization tools also enhance accessibility and collaboration, allowing researchers from different fields to reproduce and extend existing analyses.

Beyond its analytical function, computer visualization also serves as a scientific communication tool. It helps convey research findings effectively through visual designs that account for human perception factors such as color, contrast, and layout. Visual-based analysis enables policymakers and the public to grasp complex climate data intuitively, supporting evidence-based decision-making and awareness of global warming trends.

Based on this background, this study focuses on two main research questions: (1) how computer visualization can simplify global temperature data to make it more understandable, and (2) to what extent visualization techniques can reveal patterns and trends in global temperature change over time. This research is expected to provide a visual representation of global temperature changes over time, highlight the role of computer visualization in big data analysis, and serve as a reference for future studies in climate and environmental science (Kurniawan, 2023).

II. RESEARCH METHODOLOGY

The data used in this study were obtained from the *Global Land Temperatures by City* dataset developed by Berkeley Earth. This dataset comprises over one million rows of data, recording the monthly average temperatures of various cities worldwide from 1743 to 2013. Each entry includes the date (dt), average temperature (*AverageTemperature*), measurement uncertainty (*Average Temperature Uncertainty*), city name, country, and geographical coordinates. The dataset was selected due to its extensive temporal coverage and broad spatial dimensions, making it well-suited for analysis using a computer visualization approach (Ardian & Ainy, 2025).

Before conducting the analysis, the dataset underwent several preprocessing stages to ensure its quality and consistency. The data cleaning process was crucial because the dataset's large size and long historical span often resulted in missing values and irregularities in measurement intervals. Additionally, adjustments were necessary to standardize temperature units, address incomplete records, and limit the time frame to a relevant modern period for accurate climate comparison.

This research was conducted using the Python programming language, executed in the Google Colab environment. Python was chosen because it provides a comprehensive, efficient, and open-source ecosystem for data analysis and visualization. Several libraries were utilized, including Pandas for data manipulation, Matplotlib for graphical representation, and Seaborn for advanced statistical visualization (Darma, 2024).

To represent both tropical and temperate climate regions, four major cities—Jakarta, London, New York, and Tokyo—were selected as case studies. These cities were chosen because they reflect distinct climatic characteristics and socioeconomic contexts, allowing for a comparative analysis of temperature variations across different hemispheres and climate zones.

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df=
pd.read_csv("GlobalLandTemperaturesByCity.cs
v")
df = df.dropna(subset=['AverageTemperature'])
df['dt'] = pd.to_datetime(df['dt'])
df['year'] = df['dt'].dt.year
```

DOI: https://doi.org/10.64803/ikosstemi.v1i1.43


```
df = df[(df['year'] >= 1900) & (df['year'] <=
2020)]
cities = ['Jakarta', 'London', 'New York', 'Tokyo']
data_cities = df[df['City'].isin(cities)]
print(data_cities.describe())
plt.figure(figsize=(10,5))
sns.lineplot(data=data_cities,x='year',
y='AverageTemperature', hue='City')
plt.title("Tren Suhu Rata-Rata Tahunan (1900–
2020)")
plt.xlabel("Tahun")
plt.ylabel("Suhu Rata-Rata (°C)")
plt.grid(True)
plt.show()
```

The code above demonstrates how the *Global Land Temperatures by City* dataset was read and cleaned. This process includes removing missing values, converting the date column into the *datetime* format, filtering the data from 1900 to 2020, and generating visualizations that show the annual average temperature trends in four major cities: Jakarta, London, New York, and Tokyo.

After data preprocessing using the Pandas library, missing values in the *AverageTemperature* column were removed, and the date format was converted to the *datetime* type. The dataset was then filtered to include only the years 1900–2020. These steps ensured that the resulting analysis was both accurate and relevant in representing contemporary temperature changes (Ghivary et al., 2023).

III. RESULTS AND DISCUSSION

Based on the analysis using the *Global Land Temperatures by City* dataset, four main types of visualizations were produced: line chart, bar chart, heatmap, and boxplot.

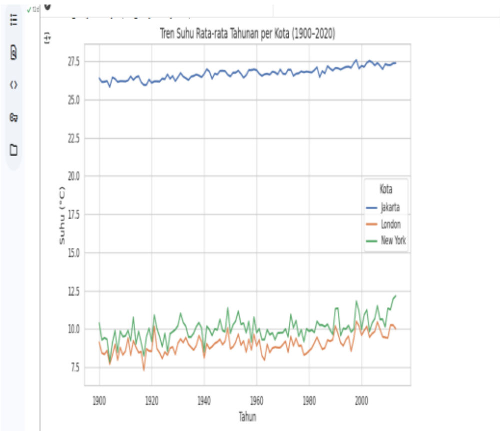

Figure 1. Annual Average Temperature Trends by City (1900-2020)

The four major cities—Jakarta, London, New York, and Tokyo—exhibited annual average temperature trends from 1900 to 2020. The graph shows a consistent upward trend in temperature over time. This pattern indicates that between the 20th and early 21st centuries, a steady global warming phenomenon occurred (Adam et al., 2025).

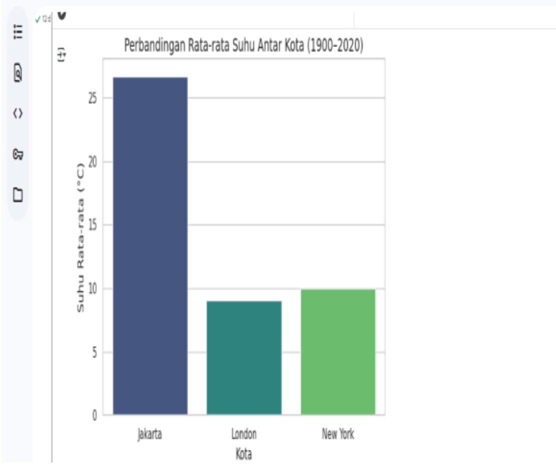

Figure 2. Comparison of Inter-City Average Temperatures (1900–2020)

The average temperatures during the observation period were compared across cities using a bar chart. The results show that tropical cities such as Jakarta have higher average temperatures compared to temperate cities like London and New York. Meanwhile, Tokyo exhibits greater temperature variation due to its more extreme seasonal changes (Nurwati et al., 2025).

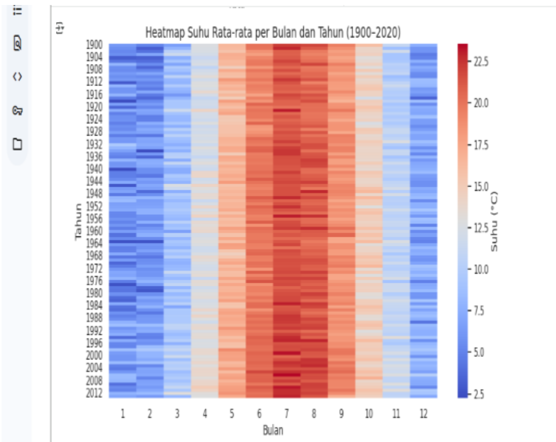

Figure 3. Monthly and yearly average temperature heatmap

The average temperatures were distributed monthly and annually through a heatmap visualization. The heatmap reveals consistent seasonal patterns, with the highest temperatures occurring in mid-year (July–August) and the lowest in early-year (January–February), particularly for northern cities such as New York and London. The blue color represents lower temperature values (Santhi, 2023).

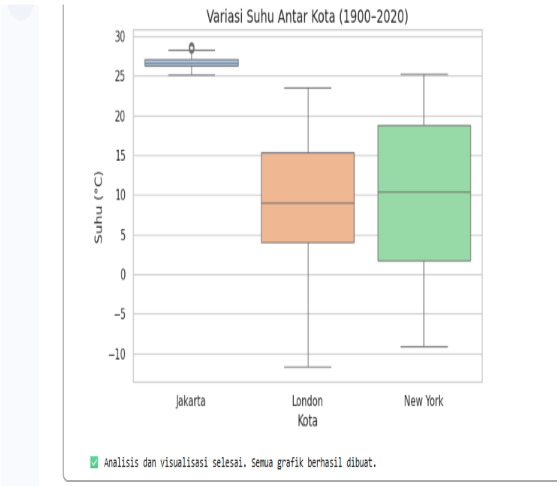

Figure 4. Boxplot Showing Inter-City Temperature Variations

The boxplot illustrates temperature variations among cities, showing that tropical cities have narrower temperature ranges compared to four-season cities, which display greater variability.

The visual analysis results indicate a consistent upward trend in global average temperatures from 1900 to 2020. Based on the calculation of annual average temperature values, an overall increase of approximately 1.02°C was observed during this period, with an average rate of 0.009°C per year. The most significant temperature rise occurred after the 1980s, corresponding with the increase in industrial activities and global greenhouse gas emissions. This pattern is consistent with the reports from the Intergovernmental Panel on Climate Change (IPCC, 2021) and the National Oceanic and Atmospheric Administration (NOAA, 2023), which state that the past two decades represent the warmest period in modern recorded history (Fasadena et al., 2025).

In addition, the results also reveal the influence of the urban heat island effect in major metropolitan areas such as Tokyo and New York, where urbanization and population density contribute to local temperature intensification.

The computer visualizations generated using Python and the Seaborn library have made it easier to observe global temperature change patterns, as complex tabular data can be transformed into more interpretable graphical representations. This technique enables researchers to clearly identify long-term patterns, regional temperature differences, and seasonal variations. Therefore, the implementation of computer visualization serves not only as a presentation tool but also as an exploratory analytical instrument that strengthens the understanding of climate change phenomena (Saputra et al., 2025).

IV. CONCLUSION

By applying computer visualization to global temperature data, complex information can be presented in a more simplified and informative manner. The use of computer visualization has proven effective in analyzing large-scale data and presenting patterns and trends in an easily interpretable format. The visualization results—consisting of line charts, bar charts, heatmaps, and boxplots—collectively demonstrate a global temperature increase of approximately 1°C since 1900, confirming the ongoing phenomenon of global warming as reported by IPCC and NASA.

Furthermore, the findings highlight the significant role of computer visualization in transforming complex numerical data into meaningful visual insights that facilitate scientific understanding and policy decision-making. This research reinforces that computer visualization is not only a presentation tool but also a powerful analytical instrument for big data interpretation.

For deeper analysis, future studies are encouraged to incorporate additional environmental variables such as carbon emissions, humidity, and precipitation. Moreover, interactive visualization tools such as Tableau or Plotly can be employed to enhance analytical flexibility and user engagement in exploring dynamic climate data.

V. RECOMMENDATIONS

For a more in-depth analysis, future research should incorporate additional environmental variables such as carbon emissions, humidity, and precipitation. Moreover, more dynamic datasets can be explored using interactive visualization tools such as Tableau or Plotly.

VI. REFERENCES

Adam, M. F. S., Putra, B., & Puteri, S. I. (2025). *Eksplorasi dan Analisis Data Mining untuk Prediksi Pola Konsumen*. https://prosiding.sentimeter.nusaputra.ac.id/index.php/prosiding/article/view/66

Ardian, Z., & Ainy, W. N. (2025). *ANALISIS DAN VISUALISASI DATA MONITORING DAFTAR TUNGGU PELANGGAN PT PLN (PERSERO) UP3 LHOKSEUMAWE MENGGUNAKAN GOOGLE LOOKER*

STUDIO Analysis and Visualization of Customer Waiting List Monitoring Data at PT PLN (Persero) UP3 Lhokseumawe Using Google Loo. *11*(1), 148–154.

Darma, S. (2024). *Analisis Visualisasi Arsitektur dalam Perencanaan Kota*. 1–10.

Dr. Adi Nugroho, S. T. M. M. S. I. (2024). *Data Science Menggunakan Bahasa R: Deep Learning*. Penerbit Andi. https:/ /books. google.co.id/ books?id =EkgsEQAAQBAJ

Elvania, N. C. (2025). *BUKU AJAR PENGOLAHAN AIR BERSIH*. Penerbit Widina. https://books.google.co.id/books?id=sghMEQAAQBAJ

Fasadena, N. S., Riono, S. H., Hadiatullah, D. R., Jumasa, H. M., Ganiem, L. M., Nurul Fadhillah S, Anwar, K., Wali, M., Rohmah, N. L. N., & Ali, M. M. (2025). *Cyber Media*.

Ghivary, R. Al, Wulandari, N., Srikandi, N., Publik, D. A., & Jakarta, U. M. (2023). Peran Visualisasi Data Untuk Menunjang Analisa Data the Role of Data Visualisation To Support Population. *Jurnal Administrasi Publik*, *1*(1), 57–62.

Johannes Kurniawan. (2023). Analisis Dan Visualisasi Data. In *Angewandte Chemie International Edition, 6(11), 951–952.* (Vol. 3, Issue 1). https://medium.com/@arifwicaksanaa/pengertian-use-case-a7e576e1b6bf

Nurwati, N., Dalimunthe, R. A., Apriandani, B., & Pebriani, E. (2025). Pengenalan Konsep Big Data Dan Analisis Data Untuk Siswa Sma. *Jurnal Pemberdayaan Sosial Dan Teknologi Masyarakat*, *5*(1), 254. https://doi.org/10.54314/jpstm.v5i1.3826

Santhi, T. (2023). https://ejurnal.umri.ac.id/index.php/SEIS/index e-ISSN: 2809-0950. *Jurnal Software Engineering and Information System*, *3*(2), 66–73.

Saputra, P. S., Putra, Y. D., Bali, P. N., Udayana, U., & Korespondensi, P. (2025). *EKSPLORASI DATA WINE PUTIH DAN KARAKTERISTIK KIMIAWI EXPLORING WHITE WINE DATA AND CHEMICAL CHARACTERISTICS USING*. *4*(2), 1–6.

Setyawan, F. E. B. (2025). *Metode Penelitian Konsep dan Analisis*. UMMPress. https://books.google.co.id/books?id=SZJdEQAAQBAJ