MORTALITA

Volume 1 Issue 1 (2024), Pages 23-29
Mortalita: Journal of Mathematics and Its Applications
p-ISSN xxxx-xxxx, *e*-ISSN xxxx-xxxx
https://ejournal.darunnajah.ac.id/index.php/mortalita

Data Visualization of PT Aneka Tambang Tbk Stock Price Report Through Data Science Approach Using Google Colab

[1]Alrafiful Rahman* and [2]Abdul Hariest Putra

[1,2]*Universitas Darunnajah, Jakarta, Indonesia*

*Email Corresponding author: alrafifulrahman@darunnajah.ac.id

Info Artikel

Diterima : 00/00/2024 Direvisi :00-00-2024 Disetujui : 00-00-2024 Diterbitkan : 00-00-2024

DOI: https://doi.org/ 10.61159/mortalita

ABSTRACT

This research discusses data visualization using Google Colab with a data science approach on the PT Aneka Tambang Tbk stock price report. The purpose of this research is to show that good data visualization can provide deeper insights into sales patterns and trends in the share price of PT Aneka Tambang Tbk and help in making better decisions in the future. This research uses an online store sales dataset of PT Aneka Tambang Tbk stock prices manipulated using Python and Google Colab. Several types of graphs and data visualizations were created, including bar charts and line charts. Through the data visualization created, the information obtained from the dataset can be more easily understood and used to make decisions. The results showed that Google Colab is a very effective and useful tool in creating data visualization on the PT Aneka Tambang Tbk stock price report. In this research, data science techniques are used to gain deeper insights into the stock price data of PT Aneka Tambang Tbk. The results of data visualization can help in improving understanding of PT Aneka Tambang Tbk stock price data and provide deeper insights for better decision making in the future.

Keywords: Data Visualization, Data Science, PT Aneka Tambang Tbk Stock Price, Google Colab

1. INTRODUCTION

In today's digital era, data has become one of the most valuable assets for organizations and individuals [1]. Especially in the financial sector, data plays a crucial role in strategic decision-making [2]. Companies such as PT Aneka Tambang Tbk (ANTAM) rely on stock price data to analyze market trends, identify investment opportunities, and manage risks. Data visualization is an effective tool to facilitate understanding and interpretation of complex information from stock price data. Therefore, data visualization is becoming increasingly important in communicating the information contained in the data [3].

Data visualization allows stakeholders to quickly recognize patterns, anomalies, and relationships in data that may not be visible in tabular format or raw numbers. Data visualization with a data science approach is the process of using techniques and methods from data science to help produce better and more informative data visualizations [4]. In this context, the use of a data science approach is very relevant. Data science combines statistical, mathematical, and computational techniques to extract meaningful insights from data. Tools such as Google Colab provide an efficient and collaborative platform for analyzing and visualizing data at scale. This approach involves data collection, data preprocessing, data analysis, and data visualization in a holistic manner.

Volume 1 Issue 1 (2024), Pages 23-29
Mortalita: Journal of Mathematics and Its Applications
p-ISSN xxxx-xxxx, *e*-ISSN xxxx-xxxx
https://ejournal.darunnajah.ac.id/index.php/mortalita

Common methods used in data science approaches include clustering, regression, classification, factor analysis, and others [5]. In addition, visualization techniques commonly used in this approach are heat maps, histograms, scatter plots, and others [6]. In this case, Python libraries such as Pandas, Matplotlib, Seaborn, and Plotly can help in data processing and visualization creation [7].

There are several related studies related to data visualization in the context of online stores. This research [8] [9] [10] [11] discusses data visualization in the context of customer behavior in online stores. The methodology used is data processing using SQL and visualization using Python libraries such as Matplotlib and Seaborn. This research [12] [13] [14] also discusses data visualization in the context of customer transactions in online stores. The methodology used is data processing using Python libraries such as Pandas and visualization using libraries such as Matplotlib and Plotly. Research [15] [16] [17] Data Visualization of Online Store Sales Reports Through a Data Science Approach Using Google Colab.

Overall, these studies show that data visualization can help online businesses optimize their sales performance by analyzing customer behavior, transactions, and sales trends. In the context of online stores, the use of Python libraries such as Pandas, Matplotlib, and Seaborn is very helpful in data processing and visualization. This research aims to demonstrate how ANTAM stock price data visualization can be done using a data science approach. Using Google Colab, this research will show the steps in processing stock price data, conducting explorative analysis, and producing informative visualizations.

2. METHODS

The first step in performing data visualization is to collect the data to be used. In this case, we will use online store sales data in CSV form. This data contains information about product sales in a certain period of time, including product name, price, quantity sold, and date of sale.

After collecting the data, the next step is to process the data to prepare it for visualization. In this case, we will use Pandas, a Python library for data manipulation. The next step is to visualize the data. In this case, we will use Matplotlib, a Python library for data visualization. Matplotlib makes it possible to create different types of graphs, such as bar charts, line graphs, and scatter plots. The following are the design stages of this data visualization research.

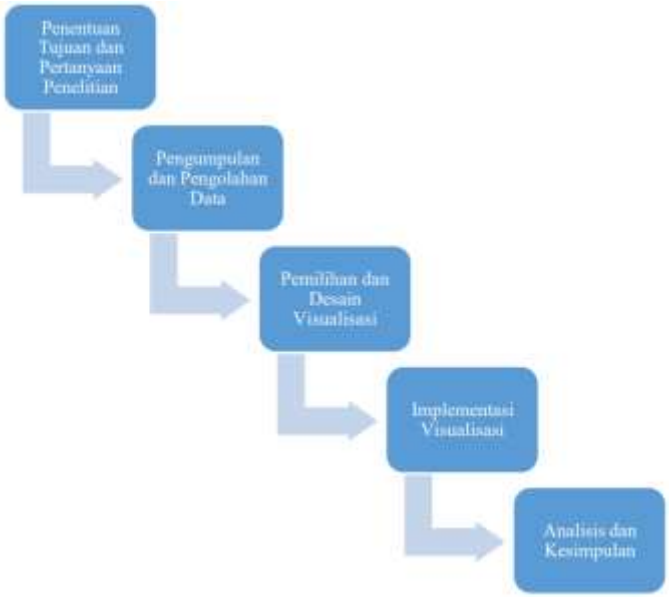

Figure 2.1. Stages of Research Design and Methods

Volume 1 Issue 1 (2024), Pages 23-29
Mortalita: Journal of Mathematics and Its Applications
p-ISSN xxxx-xxxx, *e*-ISSN xxxx-xxxx
https://ejournal.darunnajah.ac.id/index.php/mortalita

1. Determination of Research Objectives and Questions
   This is an important early stage in data visualization research design. Researchers must define clear objectives and research questions to be able to produce data visualizations that are relevant and useful for business. Questions that can be asked in the context of stock prices include: "Has there been an increase or decrease in stock prices in a certain period?" or "What is the trend of stock price sales over the past few months?"
2. Data Collection and Processing
   This stage involves collecting data from data sources relevant to the research objectives. The data is then processed using software such as Microsoft Excel or Python libraries such as Pandas to prepare the data before visualization. At this stage, it is important to pay attention to data integrity and quality so that the resulting data visualization can be trusted.
3. Visualization Selection and Design
   This stage involves selecting the right type of visualization for the research question posed. The type of visualization used can vary from bar charts, line charts, to pie charts. In addition, researchers also need to consider the aesthetic factor of the visualization so that it is easily understood by the target audience.
4. Visualization Implementation
   This stage involves the implementation of visualizations using Python libraries such as Matplotlib or Plotly. At this stage, researchers must ensure that the visualizations produced are in accordance with the objectives and research questions posed.
5. Analysis and Conclusion
   This stage involves analyzing the visualization results to draw conclusions that are useful to the business. In the context of an online store, conclusions can be drawn such as the best-selling products and sales trends that occur over a period of time.

3. RESULTS AND DISCUSSION

In this research, the dataset used is the market data of PT Aneka Tambang Tbk on a website https://finance.yahoo.com/quote/ANTM.JK/financials/ which is stored in CSV (Comma-Separated Values) format. This dataset contains information about sales transactions such as transaction date, Open (share price at the opening of trading), High (highest share price achieved in one day), Low (lowest share price achieved in one day), Close (share price at the close of trading), Adj Close (closing price of stock market trading that has been adjusted when corporate action occurs), and Volume (number of shares traded in a certain period). This dataset consists of 235 rows of data and 7 columns of data. The data columns contained in the dataset are as follows:

Tabel 3.1. Dataset used

No	Column	Description
1	Open	share price at the opening of trading
2	Close	share price at close of trading
3	High	highest share price reached in a single day
4	Adj. Close	the closing price of stock market trading that has been adjusted when the company's corporate action occurs
5	Low	lowest share price reached in one day
6	Volume	number of shares traded in a certain period

In the context of data visualization, there are some common mathematical formulas used to calculate various metrics. Here are some mathematical formulas that are often used in data visualization

Volume 1 Issue 1 (2024), Pages 23-29
Mortalita: Journal of Mathematics and Its Applications
p-ISSN xxxx-xxxx, *e*-ISSN xxxx-xxxx
https://ejournal.darunnajah.ac.id/index.php/mortalita

:

1. Mean
   The average can be calculated by summing up all data values and dividing by the number of data. data.
$$Mean = \frac{(x_1 + x_2 + \cdots + x_n)}{n}$$
2. Median
   Median is the middle value of the data sorted from the smallest value to the largest or vice versa. If the number of data is odd, the median is the middle value. If the number of data is even, the median is the average of the two middle values.
$$Median = \frac{n+1}{2}$$
3. Variance
   Variance measures how far data values are spread out from the mean.
$$Variance = \frac{\sum(x_i - \mu)^2}{n}$$
4. Standard Deviation
   Standard deviation is the square root of variance and is used to measure how far the data values are spread out from the mean.
$$Standard\ Deviation = \sqrt{Variance}$$
5. Correlation
   Correlation measures how close the relationship is between two variables. Correlation values range from -1 to 1, with a value of 1 indicating a strong positive relationship, a value of 0 indicating no relationship, and a value of -1 indicating a strong negative relationship.
$$Correlation = \frac{\sum\left((x_i - \mu_x)(y_i - \mu_y)\right)}{n(\sigma_x)(\sigma_y)}$$

This research will produce data visualizations such as bar charts for the number of products by category, line charts for total sales per month. This data visualization will help in understanding the stock price data and help in making better decisions. The following is an example of data visualization that can be generated using Google Colab for the PT Aneka Tambang Tbk stock price report:
1. Line chart visualization of PT Aneka Tambang Tbk Share Price Per Month

MORTALITA
Volume 1 Issue 1 (2024), Pages 23-29
Mortalita: Journal of Mathematics and Its Applications
p-ISSN xxxx-xxxx, *e*-ISSN xxxx-xxxx
https://ejournal.darunnajah.ac.id/index.php/mortalita

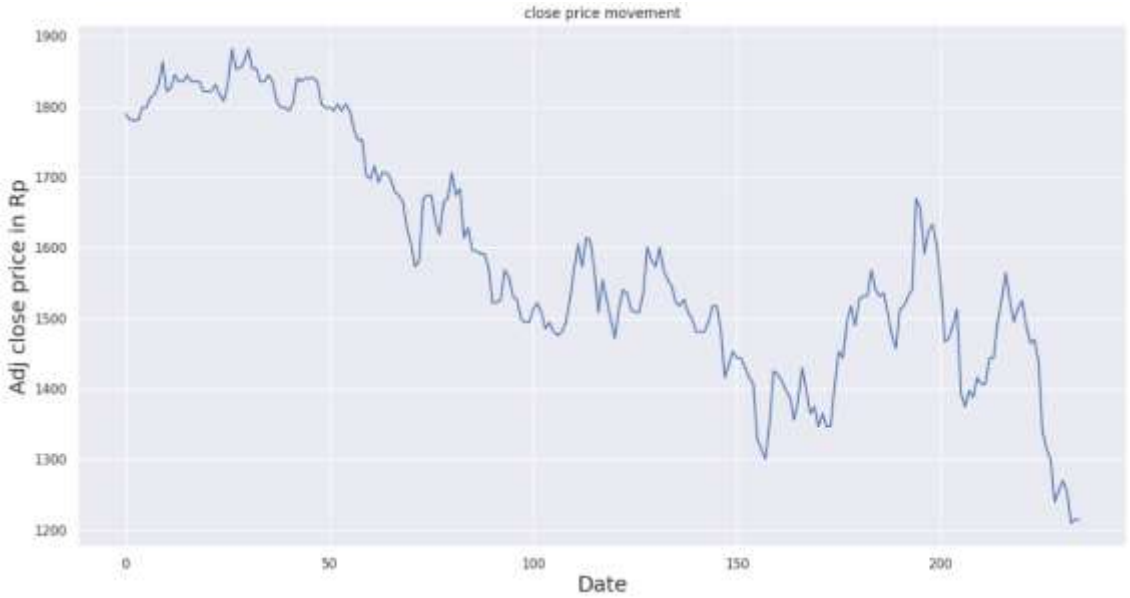

Figure 3.1. PT Aneka Tambang Tbk Stock Price Line Chart

The visualization results above show the share price of PT Aneka Tambang Tbk over the past few months. In this visualization, it can be seen that date 0-50 has the highest price of 1900, while date 200 has the lowest share price. This visualization can help to know the share price of PT Aneka Tambang Tbk in planning open and close strategies in certain months.

Volume 1 Issue 1 (2024), Pages 23-29
Mortalita: Journal of Mathematics and Its Applications
p-ISSN xxxx-xxxx, *e*-ISSN xxxx-xxxx
https://ejournal.darunnajah.ac.id/index.php/mortalita

2. Bar chart visualization of PT Aneka Tambang Tbk Share Price Per Month

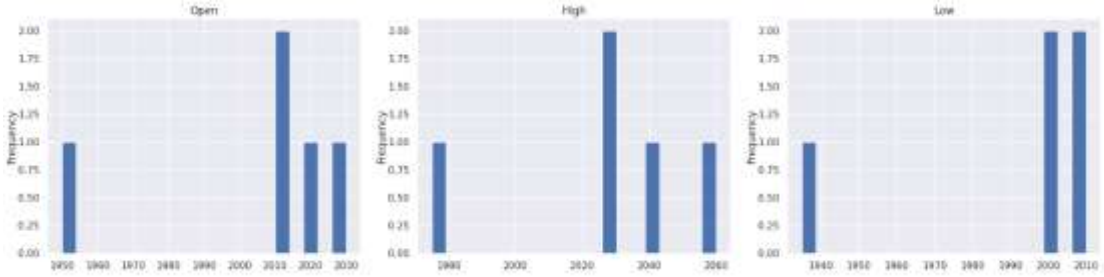

Table 3.2. ***Bar Chart of PT Aneka Tambang Tbk Share Price***

The visualization results above show that Open (stock price at the opening of trading) has the highest stock price in 2010, High (highest stock price achieved in one day) has the highest stock price between 2020 and 2040, Low (lowest stock price achieved in one day) has the highest stock price in 2010 and 2000.

In using Google Colab for data visualization, there are various Python libraries that can be used, such as Pandas, Matplotlib, and Plotly, to help process data and produce informative data visualizations. By using Google Colab, data visualization can be generated easily and quickly without requiring software installation on the local computer.

CONCLUSION

Based on the results of the research that has been done, it can be concluded that Google Colab is a very effective and useful tool in making data visualization on the PT Aneka Tambang Tbk Stock Price report. In this research, the PT Aneka Tambang Tbk Stock Price dataset has been used and manipulated using Python and Google Colab. Some types of graphs and data visualizations that have been made include bar charts and line charts.

Through the data visualization created, the information obtained from the dataset can be more easily understood and used to make decisions. In addition, using data science techniques in data visualization can provide more information about sales patterns and trends in the Share Price of PT Aneka Tambang Tbk.

Thus, it can be concluded that data visualization using Google Colab with a data science approach can help in improving understanding of PT Aneka Tambang Tbk Stock Price online store sales data and provide deeper insights for better decision making in the future.

REFERENCES

[1] Astuti, D. (2019). Penentuan Strategi Promosi Usaha Mikro Kecil Dan Menengah (UMKM) Menggunakan Metode CRISP-DM dengan Algoritma K-Means Clustering. Journal of Informatics, Information System, Software Engineering and Applications (INISTA), 1(2), 60–72. https://doi.org/10.20895/inista.v1i2.71

[2] Diartono, D. A., Zuliarso, E., Sulastri, S., & Anis, Y. (2022). PELATIHAN PEMROGRAMAN BERBASIS KOMPUTASI AWAN UNTUK MENDUKUNG PEMBELAJARAN JARAK JAUH BAGI SISWA SMK Negeri 4 KENDAL. Jurnal Pengabdian Masyarakat Intimas (Jurnal INTIMAS): Inovasi Teknologi Informasi Dan Komputer Untuk Masyarakat, 2(1), 1–6. https://doi.org/10.35315/intimas.v2i1.8841

[3] Lizana, H. I. N., & Ridho, F. (2021). Implementasi dan Evaluasi Visualisasi Data Interaktif pada Publikasi Laporan Bulanan Data Sosial Ekonomi Indonesia. Seminar Nasional Official Statistics, 2021(1), 947–957. https://doi.org/10.34123/semnasoffstat.v2021i1.795

[4] Kurniawan, C. (2019). A Survey on Big Data Analytics Model. ITEJ (Information Technology

Engineering Journals), 4(1), 1–13. https://doi.org/10.24235/itej.v4i1.46

[5] Pakan, P. D. (2020). PENERAPAN PCA PADA KLASIFIKASI DATA MIKRO ARAY GEN MENGGUNAKAN JARINGAN SYARAF TIRUAN. Jurnal Ilmiah Flash, 5(2), 24. https://doi.org/10.32511/jiflash.v4i1.204

[6] Harahap, M., Rozi, F., Yennimar, Y., & Siregar, S. D. (2021). Analisis Wawasan Penjualan Supermarket dengan Data Science. Data Sciences Indonesia (DSI), 1(1), 1–7. https://doi.org/10.47709/dsi.v1i1.1173

[7] Nugraha, W., & Syarif, M. (2023). Teknik Weighting untuk Mengatasi Ketidakseimbangan Kelas Pada Prediksi Churn Menggunakan XGBoost, LightGBM, dan CatBoost. Techno.Com, 22(1), 97–108. https://doi.org/10.33633/tc.v22i1.7191

[8] Fedushko, S., & Ustyianovych, T. (2022). E-Commerce Customers Behavior Research Using Cohort Analysis: A Case Study of COVID-19. Journal of Open Innovation: Technology, Market, and Complexity, 8(1), 12. https://doi.org/10.3390/joitmc8010012

[9] Maulani, J., & Amin, M. (2019). RANCANG BANGUN SISTEM INFORMASI JASA PEMBUATAN PAKAIAN DENGAN ALGORITMA PEMPROGRAMAN TERSTRUKTUR. Technologia: Jurnal Ilmiah, 10(2), 85. https://doi.org/10.31602/tji.v10i2.1811

[10] Qisthiano, M. R., Kurniawan, T. B., Negara, E. S., & Akbar, M. (2021). Pengembangan Model Untuk Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu dengan Metode Naïve Bayes. JURNAL MEDIA INFORMATIKA BUDIDARMA, 5(3), 987. https://doi.org/10.30865/mib.v5i3.3030

[11] Santoso, E. B., & Nugroho, A. (2019). Analisis Sentimen Calon Presiden Indonesia 2019 Berdasarkan Komentar Publik Di Facebook. Eksplora Informatika, 9(1), 60–69. https://doi.org/10.30864/eksplora.v9i1.254

[12] Kimm, H., Paik, I., & Kimm, H. (2021). Performance Comparision of TPU, GPU, CPU on Google Colaboratory Over Distributed Deep Learning. 2021 IEEE 14th International Symposium on Embedded Multicore/Many-Core Systems-on-Chip (MCSoC), 312–319. https://doi.org/10.1109/MCSoC51149.2021.00053

[13] Stancin, I., & Jovic, A. (2019). An overview and comparison of free Python libraries for data mining and big data analysis. 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 977–982. https://doi.org/10.23919/MIPRO.2019.8757088

[14] Subroto, V. K., & Endaryati, E. (2021). BUSINESS INTELLIGENCE DAN KESUKSESAN BISNIS di ERA DIGITAL. Dinamika: Jurnal Manajemen Sosial Ekonomi, 1(2), 41–47. https://doi.org/10.51903/dinamika.v1i2.45

[15] Gelar Guntara, R. (2023). Pemanfaatan Google Colab Untuk Aplikasi Pendeteksian Masker Wajah Menggunakan Algoritma Deep Learning YOLOv7. Jurnal Teknologi Dan Sistem Informasi Bisnis, 5(1), 55–60. https://doi.org/10.47233/jteksis.v5i1.750

[16] Viégas, F. B., & Wattenberg, M. (2007). Artistic Data Visualization: Beyond Visual Analytics (pp. 182–191). https://doi.org/10.1007/978-3-540-73257-0_21

[17] Wahjoerini, W., Pamurti, A. A., & Prabowo, D. (2022). PELATIHAN PEMBUATAN VISUALISASI DATA SPASIAL BAGI SISWA SMA WALISONGO SEMARANG. SELAPARANG: Jurnal Pengabdian Masyarakat Berkemajuan, 6(3), 1126. https://doi.org/10.31764/jpmb.v6i3.933