Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th June 2025. Vol. 4. No. 3; e-ISSN: 2808-4519

# Prediction of Electricity kWh Sales in Pontianak City Using Linear Regression Method

**Rido Safaryansyah[1*], Alda Cendekia Siregar[2], Istikoma[3]**

[1,2,3] *Universitas Muhammadiyah Pontianak*

*ridosafaryansyah@gmail.com[1*], alda.siregari@unmuhpnk.ac.id[2], istikoma@unmuhpnk.ac.id[3]*

**Abstract**

This study presents the development of a web-based system to predict monthly electricity sales (in kWh) in the city of Pontianak using the Simple Linear Regression method. The main objective is to build a system capable of estimating electricity demand for the latest period and projecting the required kWh for the following month. The system uses 24 months of historical electricity sales data as the basis for prediction, allowing it to identify trends and patterns over time. After applying the regression calculation, the system predicted the next month's electricity sales to be 93,394,700 kWh. This result indicates that the system's prediction aligns with historical trends, demonstrating the model's reliability and potential accuracy. The relationship between the independent and dependent variables used in the model is linear and causal, making this method suitable for forecasting electricity consumption. Additionally, the system includes data visualization features on the website to enhance user understanding and simplify analysis. These visual tools help stakeholders to interpret predictions more effectively. Overall, the system serves as a practical and efficient solution to support electricity demand planning, resource management, and decision-making processes for local authorities and energy providers in Pontianak.

***Keywords***: *Prediction of kWh electricity sales, simple linear regression method, website based.*

## 1. Introduction

Electricity consumption in Pontianak City is divided into several sectors, including residential, industrial, office, and public facilities. Many government institutions rely on electricity services such as offices, housing, and educational institutions [1]. Currently, PLN ULP Pontianak, as the electricity distribution center, still uses manual methods like Microsoft Excel to determine electricity sales targets. This manual process has several drawbacks, such as limited scalability, risk of human error, limited analysis capabilities, vulnerability to data loss, lack of visualization, and difficulty in formatting and data migration.

Therefore, a web-based system is needed to help PLN ULP Pontianak manage operations and efficiently predict monthly electricity sales targets. This research aims to assist PLN ULP Pontianak in predicting monthly electricity sales using a structured web-based system. The study is titled "Electricity kWh Sales Prediction in Pontianak City Using Simple Linear Regression." This method is appropriate due to its simplicity, which aligns with the characteristics of the data and the purpose of the analysis. The prediction results will provide estimates of electricity usage at specific time points based on historical data. The difference between simple and multiple linear regression lies in the number of independent variables involved. Simple linear regression includes one independent variable (X) and one dependent variable (Y), aiming to find the best-fit line that describes their relationship, while multiple linear regression involves more than one independent variable (X1, X2, X3...) and one dependent variable (Y), aiming to build the best model to represent the relationships among variables.

The simple linear regression method has been proven effective through previous research, such as the study titled “Implementation of Linear Regression Method for Electricity Demand Prediction at PLN Rayon Sintang Web-Based” [2]. Based on this background, one solution is to use the data mining process. Data mining is used to predict the next month’s electricity sales, allowing the relevant parties to analyze electricity sales trends and adjust operational strategies accordingly. It applies data analysis techniques and algorithms to identify patterns, relationships, and trends in historical data for future prediction. Having monthly electricity consumption forecasts will help PT PLN (Persero) ULP Pontianak make better operational decisions and determine electricity sales targets for the following month.

## 2. Theoritical Basis

### 2.1. Literature review

Each research has a specific focus that differentiates it from previous studies, even if the themes are similar. In this study, the author conducted a literature review and found several related works. However, this study presents a different focus in terms of location, research object, theoretical framework, and method used. The first related study is titled "Prediction of Unilever Product Sales Using Linear Regression Method", which aimed to develop a web-based system to predict sales using simple linear regression. The system used MEA and

MAPE to evaluate prediction errors, utilizing 15 months of Unilever sales data from January 2021 to March 2022. The dataset was split into 12 months for training and 3 months for testing. The results of the prediction for the three future periods matched the manual linear regression calculation, with the lowest MAPE value of 1% on the Sunsilk Conditioner product, indicating high accuracy of the method [3]. The second study, "Application of Linear Regression for Sales Forecasting", analyzed sales at Vansquare Fish Store, which faced challenges in stock management due to space limitations and increasing sales trends. Linear regression was used to forecast sales using data from November 2020 to October 2021. Error calculations were done using MAD, MSE, and MAPE. The system was built using PHP and MySQL as the database [4]. The third study, titled "Prediction of International Travelers During the COVID-19 Pandemic Using Simple Linear Regression Method", focused on forecasting international arrivals as they declined during the pandemic. The independent variable (X) was the number of travelers, and the dependent variable (Y) was the number of COVID-19 cases in Indonesia. The prediction results using linear regression produced values of 471296.2073 and 3.010223266 [5]. The fourth study, "Prediction of Mobile Phone Sales in Store X Using Linear Regression Algorithm", aimed to forecast sales of mobile phones for the next three months. The method used the number of phones sold as the dependent variable and sales period as the independent variable. RMSE and Relative Error were used to assess prediction performance. For the Entry category, sales were predicted at 84, 86, and 88 units over three months, with an RMSE of 10.36 and Relative Error of 19.11%. For the Mid category, RMSE was 7.50 and Relative Error was 32.97%, showing the method's applicability with 24 months of data [6]. The fifth study, "Application of Data Mining to Predict Best-Selling Shoe Products Using Simple Linear Regression Method", was conducted on sales data from PT Sepatu Bata from January to December 2019. The study used simple linear regression to identify the direction and strength of the relationship between independent and dependent variables and to predict the dependent variable based on changes in the independent variable. A data mining-based system was designed using PHP and MySQL to optimize product sales forecasting [7].

### 2.2. Data Mining

Data mining, also known as Knowledge Discovery in Database (KDD), involves extracting information from large datasets. This process of searching for information identifies patterns and trends in the data, and the results of the mining process provide easily understandable information. The algorithms used in data mining are highly varied, and the choice of method or algorithm depends on the goals and the Knowledge Discovery in Database (KDD) process [9]. The technique used in the data mining process in this study is simple linear regression.

### 2.3. Prediction

Prediction is a process of estimating or forecasting systematically about what is most likely to happen in the future based on past and present information, with the aim of minimizing errors. A prediction does not have to be identical to the actual event, but rather strives to find the closest possible answer to what will occur [10].

### 2.4. Electricity

Electricity is a form of energy produced by the flow of electrons through conductors like copper or aluminum wires. It is used for various purposes, such as powering machines, lighting rooms, generating heat, and operating electronic devices. Electricity is generated from sources like generators, batteries, and solar panels, with most electricity today being produced by generators. It is divided into two types: DC (direct current) and AC (alternating current), with PLN supplying electricity to customers via AC

### 2.5. Website

A website is a service used on computers connected to the internet, utilizing hypertext facilities to display data such as sound, multimedia, text, animation, and other data [11]. The web uses the HTTP (Hypertext Transfer Protocol) protocol as the foundation for communication. Users can access websites through web browser applications, such as Google Chrome, Mozilla Firefox, or Safari, by entering a URL (Uniform Resource Locator) or using a search engine to find the desired webpage [12].

### 2.6. Python

Python is a high-level programming language that is interpreted, interactive, object-oriented, and can operate on almost all platforms: Mac, Linux, and Windows. Python is considered easy to learn due to its clear syntax, the availability of ready-to-use modules, and efficient high-level data structures. The Python distribution includes facilities like the shell in Linux. The typical installation location of Python is "/usr/bin/python," although it may vary [13].

### 2.7. Visual Studio Code

Visual Studio Code is a lightweight yet powerful source code editor that runs on desktops. It offers built-in support for JavaScript, scripting, and Node.js, with a wide range of extensions available for other languages, including C++, C#, Python, and PHP. A cross-platform version of Atom's code-editing components, it is based on JavaScript and HTML5. This fully-featured integrated development environment (IDE) is designed for developers working with Microsoft's open cloud technologies. Visual Studio Code uses open-source .NET tools to support ASP.NET C# code, Omnisharp .NET developer tools, and Roslyn compilers. Its user-friendly interface follows a common explorer style, with a left panel showing files and folders, and a right panel displaying the contents of the open file.

## 3. Result and Discussion

In this study on the application of simple linear regression for predicting electricity energy demand, there are several steps that need to be carried out, namely problem identification, data collection, needs analysis, system implementation design, and finally, testing.

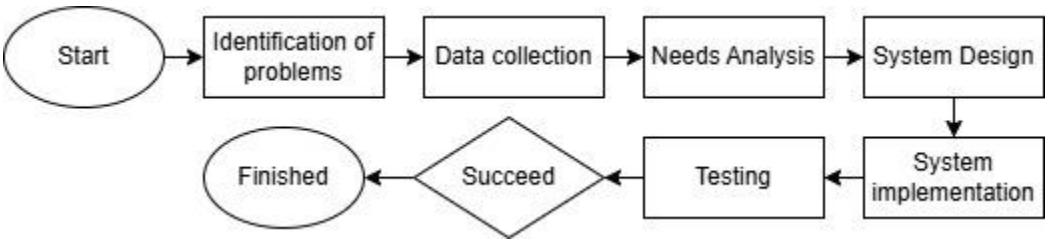


**Fig. 1:** Research Flow Chart

### 3.1. Problem Identity

The researcher identified several issues at PT. PLN (Persero) ULP Pontianak, particularly regarding the importance of accurately calculating electricity sales targets and available supply. As the central electricity distributor in Pontianak City, the company requires a reliable system to ensure that both operational and sales management run smoothly. This is essential for maintaining service quality and meeting the electricity demands of the community effectively.

### 3.2. Data collection

The data collection method for this research involves gathering historical electricity usage data to support demand forecasting. Data was obtained through observation and surveys from the Energy Transaction Division of PT. PLN (Persero) Pontianak, covering monthly electricity sales and customer numbers from January 2022 to June 2024. The next step is to describe the data through analysis using descriptive statistics such as mean, median, and distribution. The quality of the data—its completeness, accuracy, and consistency—will also be evaluated to ensure its reliability for future demand prediction.

### 3.3. Needs Analysis

This stage involves identifying the necessary requirements for the research. The study uses the Y variable representing electricity kWh sales in Pontianak over a 24-month period from June 2022 to June 2024, and the X variable representing the upcoming month periods. The system is expected to perform simple linear regression analysis to generate output in the form of predicted kWh sales.

### 3.4. System Design

The system implementation in this research involves modeling the linear regression method using Visual Studio Code, as well as managing the data within the same environment to generate electricity kWh sales predictions in Pontianak. This stage also includes designing the output to be displayed on a website, making it easier for users to view. Additionally, the system allows users to input datasets to be processed for research purposes.

### 3.5. System implementation

At this stage, the components that have been previously designed are built based on the design process to produce outputs that align with the initial plans, including the database structure. This includes the design of Data Flow Diagrams (DFDs) and the application's user interface. The implementation phase begins with the development of the user interface, the creation of the dataset, and the implementation of the selected method into the application. After the system has been fully implemented, it proceeds to the testing phase to ensure that all features and functions work as intended.

### 3.6. System Testing

After creating the testing design in the form of functional system scenarios, testing is carried out according to the sequence of the test code. If the system successfully produces the expected output, it is concluded that the system functions as intended. However, if it fails to meet expectations, the system is considered not in accordance with the desired results or deemed invalid. The results of the testing for the 12 scenarios are presented in the table.

**Table 1:** Test Results on Scenarios

| ID | Hasil Pengujian | Kesimpulan |
|---|---|---|
| **T01** | *Website* dapat ditampilkan pada *web browser* | ***valid*** |
| **T02** | Hasil prediksi dapat dilihat pada *website* | ***valid*** |
| **T03** | Hasil Dataset dapat dilihat pada *website* | ***valid*** |
| **T04** | ***Website* dapat melakukan prediksi** | ***valid*** |

This conclusion is drawn based on the results of the testing analysis that has been conducted, as well as from the analysis during the application development and modeling process. Therefore, the results of the system testing provide a reference for drawing a valid conclusion in accordance with the outcomes of the previous modeling, allowing for the formulation of a final conclusion.

## 4. Conclusion

The conclusion drawn from this research aligns with the problem formulation: Based on the results of this study, it can be concluded that data mining techniques can be used to design a prediction system for electricity kWh sales in Pontianak for the upcoming month using the simple linear regression method. Additionally, it can be concluded that the final results meet the objectives, as the prediction output on the website is able to predict and display the sales figures for the next month, as well as visualize the linear line connecting the historical kWh sales data and the number of customers.

## 5. Advice

The continuous improvement and regular updating of data should be given serious consideration as essential steps to enhance the overall accuracy and reliability of electricity consumption forecasts. At the same time, developing a general yet comprehensive understanding of how the prediction of electricity sales in kilowatt-hours (kWh) is carried out using this particular method can play a crucial role in enabling relevant stakeholders—such as policymakers, utility providers, and energy planners—to fully appreciate the significance of data-driven insights in supporting long-term, sustainable energy planning and decision-making processes.

## References

[1] L. Di and K. Lubuklinggau, "Implementasi Data Mining Prediksi Kebutuhan Tenaga," vol. 19, no. 2, pp. 197–206, 2020.

[2] N. Nazeriandy, Y. Syahra, and M. Syaifudin, "Penerapan Data Mining Untuk Memprediksi Penggunaan Daya Listrik Pada PT.PLN (Persero) Rayon Medan Selatan Dengan Menggunakan Metode Regresi Linier Berganda," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 20, no. 1, p. 20, 2021, doi: 10.53513/jis.v20i1.2431.

[3] A. Anggrawan, N. Azmi, U. Bumigora, and I. Anthonyangrawan, "Prediksi Penjualan Produk Unilever Menggunakan Metode Regresi Linear," *J. Bumigora Inf. Technol.*, vol. 4, no. 2, pp. 123–132, 2022, doi: 10.30812/bite.v4i2.2416.

[4] S. Monica and A. Hajjah, "Penerapan regresi linier untuk peramalan penjualan," *J. Tek. Inform. Kaputama*, vol. 6, no. 2, pp. 777–779, 2022, [Online]. Available: Binjai

[5] D. Rahmawati, T. Kristanto, B. F. Setya Pratama, and D. B. Abiansa, "Prediksi Pelaku Perjalanan Luar Negeri Di Masa Pandemi COVID-19 Menggunakan Metode Regresi Linier Sederhana," *J. Inf. Syst. Res.*, vol. 3, no. 3, pp. 338–343, 2022, doi: 10.47065/josh.v3i3.1507.

[6] Y. Aqsho Ramadhan, A. Faqih, and G. Dwilestari, "Prediksi Penjualan Handphone di Toko X menggunakan Algoritma Regresi Linear," *J. Inform. Terpadu*, vol. 9, no. 1, pp. 40–44, 2023, doi: 10.54914/jit.v9i1.692.

[7] D. A. Pohan, M. H. Dar, and Irmayanti, "Penerapan Data Mining untuk Prediksi Penjualan Produk Sepatu Terlaris Menggunakan Metode Regresi Linier Sederhana," *J. Nas. Inform. dan Teknol. Jar.*, vol. 2, pp. 2–6, 2022, [Online]. Available: https://doi.org/10.30743/infotekjar.v6i2.4795

[8] A. Prakarsya and A. Prambayaun, "Implementasi Data Mining Untuk Prediksi Penyebaran Virus Hiv / Aids Di Bandar Lampung," *J. Sist. ...*, vol. 3, no. 2, pp. 18–26, 2020, [Online]. Available: https://www.ejournal.lembahdempo.ac.id/index.php/SISKOMTI/article/view/120

[9] P. Purwadi, P. S. Ramadhan, and N. Safitri, "Penerapan Data Mining Untuk Mengestimasi Laju Pertumbuhan Penduduk Menggunakan Metode Regresi Linier Berganda Pada BPS Deli Serdang," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 18, no. 1, p. 55, 2019, doi: 10.53513/jis.v18i1.104.

[10] G. N. Ayuni and D. Fitrianah, "Penerapan metode Regresi Linear untuk prediksi penjualan properti pada PT XYZ," *J. Telemat.*, vol. 14, no. 2, pp. 79–86, 2019, [Online]. Available: https://journal.ithb.ac.id/telematika/article/view/321

[11] S. Arfida and B. Waseso, "Pengujian Sistem Prediksi Penjualan Makanan dengan Menggunakan Metode Black-Box," *J. Tek.*, vol. 16, no. 2, pp. 289–2935, 2022.

[12] A. Zahir, "PENGEMBANGAN MEDIA PEMBELAJARAN LIVE STREAMING PENGETAHUAN KOMPUTER BERBASIS WEBSITE Kata kunci : Media Pembelajaran , Website , Live Streaming," vol. 9, 2019.

[13] Widi Hastomo, Nur Aini, Adhitio Satyo Bayangkari Karno, and L.M. Rasdi Rere, "Metode Pembelajaran Mesin untuk Memprediksi Emisi Manure Management," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 11, no. 2, pp. 131–139, 2022, doi: 10.22146/jnteti.v11i2.2586.