International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

HOUSEHOLD ELECTRICITY CONSUMPTION PATTERN ANALYSIS

[1] N Bhargavi, [2] B Manjunadh, [3] M Sai Akhil, [4] P Kalyan, [5] B Sudheer
[1] AssistantProfessor, [2345] Students
Department of CSE(Data Science)
Siddhartha Institute of Technology & Sciences, Narapally

bhargavi.cse@siddhartha.co.in , 24tq1a6748@siddhartha.co.in, 24tq1a6731@siddhartha.co.in, 24tq1a6736@siddhartha.co.in, 25TQ5A6706@siddhartha.co.in,

Abstract

The Household Electricity Consumption Pattern Analysis project focuses on understanding and analyzing the growing demand for electricity in modern households using a data-driven approach. With the increasing use of electrical appliances, air conditioning, and electronic devices, electricity consumption has risen significantly, leading to higher energy costs, frequent peak load conditions, and increased stress on power grids.

This project utilizes the publicly available Individual Household Electric Power Consumption dataset and employs Python programming language for data processing and analysis. The dataset includes minute-level measurements of global active power, voltage, current, and appliance-level consumption over several years. Data preprocessing steps such as cleaning, handling missing values, and converting timestamps into a unified datetime format are performed. The data is then resampled into hourly and daily intervals for better analysis.

Exploratory Data Analysis (EDA) and visualization techniques are used to identify consumption patterns across different times of the day, days of the week, and seasons. Peak load periods are detected using statistical threshold-based methods applied to global active power. These peaks are further analyzed to determine critical hours of high energy demand.

I. Introduction

Electricity is one of the most essential resources in modern society, and its consumption has increased significantly due to rapid urbanization, improved living standards, and the widespread use of electrical and electronic appliances in households. Residential consumers contribute a major share of total electricity demand, which often leads to peak load conditions and increased stress on power grids. These peak demand periods can result in power outages, voltage instability, and higher operational costs for utility providers, as they must maintain additional backup capacity and operate less efficient power generation systems.

Understanding household electricity consumption patterns is therefore crucial for efficient energy management and planning. Traditional billing systems only provide aggregated monthly or bi-monthly consumption data, which is insufficient for analyzing detailed usage behavior. However, with the advancement of smart metering technologies and the availability of high-resolution time-series data, it is now possible

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

to examine electricity usage at finer levels, such as hourly or daily intervals. This enables the identification of consumption trends across different times of the day, days of the week, and seasonal variations.

The Household Electricity Consumption Pattern Analysis and Peak Load Period Detection project focuses on analyzing detailed electricity usage data using the Python programming language. The study utilizes the "Individual Household Electric Power Consumption" dataset, which contains minute-level measurements of global active power, voltage, current, and appliance-level consumption over several years. Data preprocessing techniques such as cleaning, handling missing values, and converting timestamps into a unified datetime format are applied. The data is then resampled into hourly and daily intervals for better interpretation.

II. Literature Survey

A number of studies have examined residential electricity consumption patterns and their implications for energy efficiency and demand-side management. Research by Guo et al. analyzed smart meter data along with household characteristics to predict electricity usage patterns. Their findings highlighted that socio-demographic factors, appliance ownership, and lifestyle significantly influence load profiles and peak demand. Other studies have focused on descriptive analysis of household consumption, showing variations across urban and rural areas, income groups, and types of dwellings. These works emphasize that residential electricity demand is strongly affected by climatic conditions, behavioral habits, and appliance usage, underlining the importance of detailed consumption analysis for energy planning.

With the increasing availability of high-resolution smart meter data, recent research has adopted data mining and machine learning techniques to uncover hidden consumption patterns. For example, Kang and Reiner conducted a large-scale study on household electricity usage in Chengdu, China. Using unsupervised learning techniques, they identified distinct consumption clusters and analyzed behavioral responses to seasonal variations and extreme weather conditions. Their research demonstrated that time-series data can reveal multiple load profiles and classify households into different consumption groups, which is valuable for implementing targeted demand response strategies.

Similarly, Yu et al. proposed a residential load forecasting model that combines clustering techniques with deep neural networks. Their study showed that grouping households based on similar consumption patterns before applying predictive models significantly improves forecasting accuracy. This approach highlights the importance of understanding consumption behavior before performing prediction tasks.

In addition to analytical approaches, several researchers have explored visualization techniques to make electricity consumption data more accessible and understandable. Studies on visualization-based analysis demonstrate how spatial and temporal graphs can effectively represent electricity usage patterns across regions and time periods. These visual tools help non-experts interpret complex datasets and support better decision-making.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

In the context of developing countries such as India, various reports and academic works have analyzed household electricity consumption patterns, emphasizing the influence of income levels, appliance ownership, and housing conditions. These studies provide important insights into regional differences and highlight the need for customized energy management strategies.

Despite extensive research in this field, most existing studies focus on either socio-economic analysis, large-scale clustering, or advanced forecasting models. There is still a need for simple, reproducible approaches that demonstrate how raw electricity consumption data can be processed, visualized, and analyzed for practical applications such as peak load detection. Some open-source implementations using Python programming language and the "Individual Household Electric Power Consumption" dataset mainly focus on forecasting rather than explaining consumption patterns.

III. System Analysis

Electricity consumption in households is increasing rapidly due to modern lifestyle changes and increased use of appliances. Managing and analyzing this consumption is important for efficient energy utilization. Traditional systems provide only aggregated consumption data, which is insufficient for detailed analysis. There is a need to understand usage patterns at different time intervals such as hourly and daily. Identifying peak load periods helps in reducing energy costs and improving grid stability. A data-driven system can help analyze large volumes of electricity data effectively. Time-series analysis plays a key role in understanding consumption behavior. Visualization techniques make complex data easy to interpret. The system should be able to detect patterns, trends, and anomalies. It should also support decision-making for both consumers and utility providers. Efficient data processing and analysis tools are required. This project addresses these needs using modern analytics techniques.

Existing System

In the existing system, electricity consumption is recorded through traditional meters or basic digital meters. These systems provide only monthly or periodic readings of total energy usage. There is no detailed breakdown of consumption over time. Users are unaware of their peak usage periods and consumption patterns. Utilities rely on aggregated data for planning and billing purposes. There is limited use of advanced analytics or visualization tools. Data is often stored in raw form without proper processing. Peak load detection is not performed at the household level. There is no real-time monitoring of electricity usage. Decision-making is based on historical data without deep insights. Consumers cannot optimize their energy usage effectively. Overall, the system lacks intelligence and detailed analysis capabilities.

Disadvantages of Existing System

- Provides only aggregate consumption data
- No detailed time-based analysis (hourly/daily)
- Lack of peak load detection

- No real-time monitoring
- Poor visualization of electricity usage
- Limited awareness for consumers
- Inefficient energy utilization
- No support for demand-side management
- Difficult to identify consumption patterns
- No predictive or analytical capabilities
- Higher energy costs due to unmanaged usage
- Limited decision-making support

Proposed System

The proposed system focuses on analyzing household electricity consumption patterns using Python programming language. It uses the "Individual Household Electric Power Consumption" dataset for analysis. The system collects and preprocesses raw data by cleaning and handling missing values. It converts timestamps into a unified datetime format. The data is resampled into hourly and daily intervals for better analysis. Exploratory Data Analysis (EDA) is performed to identify patterns and trends. Visualization techniques such as line graphs and charts are used to represent consumption behavior. The system detects peak load periods using statistical threshold methods. It identifies critical hours of high electricity demand. The system provides meaningful insights for energy optimization. It is user-friendly and easy to interpret. Overall, it improves energy awareness and supports efficient usage.

Advantages of Proposed System

- Detailed time-based consumption analysis
- Identification of peak load periods
- Improved energy usage awareness
- Easy-to-understand visualizations
- Supports demand-side management
- Helps reduce electricity costs
- Efficient handling of large datasets
- Better decision-making support
- Detects trends and patterns effectively
- Scalable and flexible system

IV. Methodology

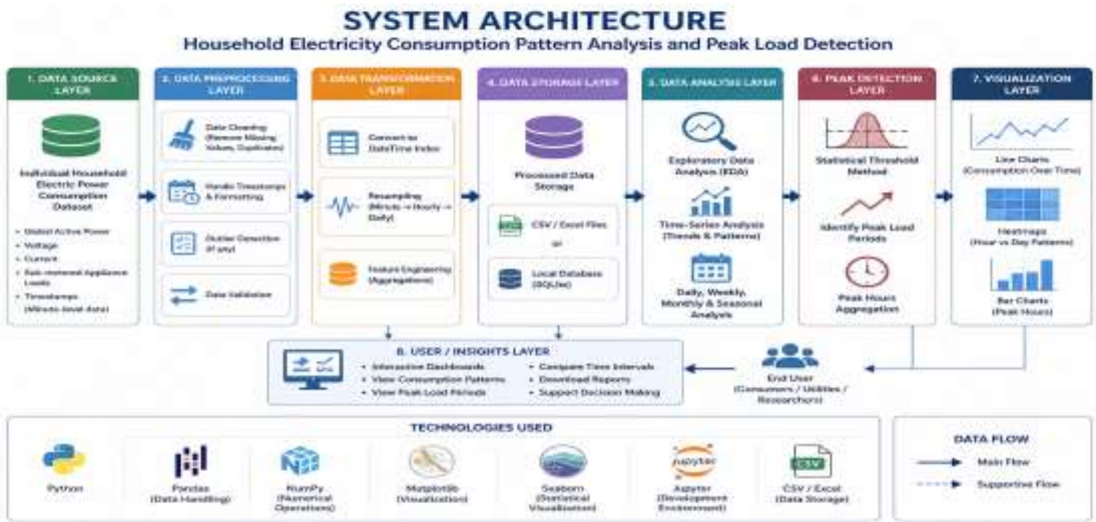
The project begins with collecting electricity consumption data from the dataset. The raw data is cleaned to remove missing or incorrect values. Data preprocessing is performed using libraries like Pandas. Timestamps are converted into a standard datetime format. The data is resampled into hourly and daily intervals. Exploratory Data Analysis is conducted to understand consumption behavior. Visualization techniques are used to represent trends and patterns. Statistical methods are applied to detect peak load periods. Threshold values are defined to identify high consumption levels. The results are analyzed to determine peak usage times. Insights are generated

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

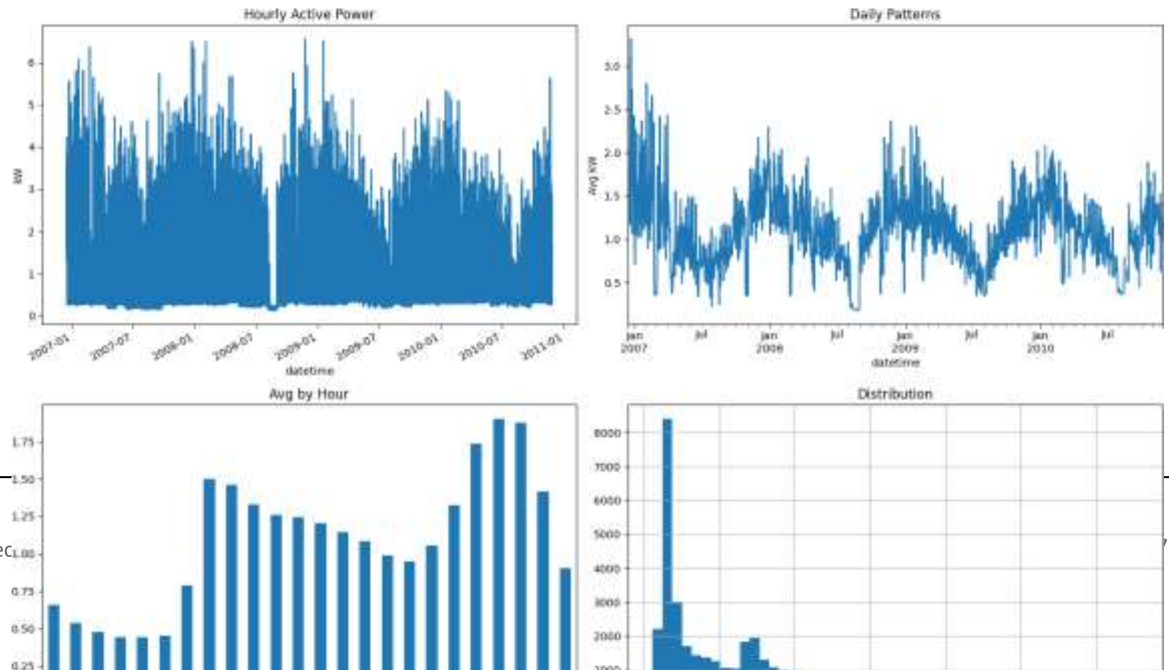
for energy optimization. Finally, the results are interpreted and presented in a clear format.

System Architecture

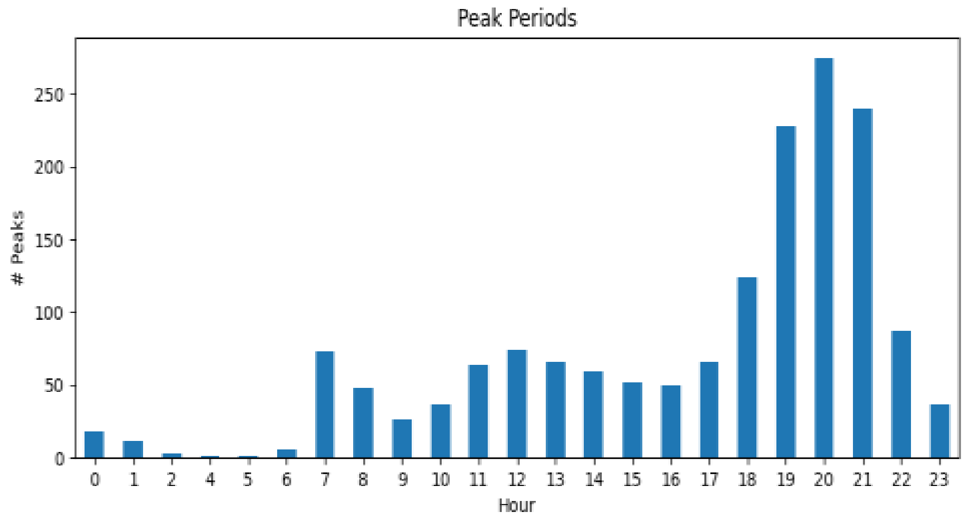
The system architecture consists of multiple layers for efficient data processing and analysis. The first layer is the Data Source Layer, which includes electricity consumption datasets. The second layer is the Data Preprocessing Layer, where cleaning and transformation are performed. The third layer is the Data Storage Layer, where processed data is stored. The fourth layer is the Data Analysis Layer, where time-series analysis is conducted. The fifth layer is the Peak Detection Layer, where high consumption periods are identified. The sixth layer is the Visualization Layer, where graphs and charts are generated. The system uses Python programming language for implementation. Users interact with the system through visual outputs. The architecture supports scalability and flexibility. It ensures efficient data flow between layers. Overall, it provides a complete solution for electricity consumption analysis.



V. Result and Output



International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper




```
  df_hourly = df.resample('H').agg({
Hourly shape: (34168, 7)
Top Peak Hours:
 hour
20    275
21    240
19    228
18    124
22     87
dtype: int64
```

VI. Conclusion

The Household Electricity Consumption Pattern Analysis and Peak Load Detection project successfully demonstrates how a data-driven approach can be used to understand and optimize residential electricity usage. By leveraging Python programming language, the system processes raw minute-level data from the "Individual Household Electric Power Consumption" dataset and transforms it into meaningful insights through data preprocessing, resampling, and time-series analysis.

The analysis revealed clear daily and weekly consumption patterns, with lower electricity usage during late-night and early-morning hours and higher demand during evening periods. These patterns reflect typical household behavior and validate findings from previous research. The implementation of a statistical threshold-based method enabled accurate detection of peak load periods, showing that electricity demand is concentrated during specific hours rather than being evenly distributed.

The results highlight the importance of demand-side management strategies, such as shifting energy usage away from peak periods and implementing time-of-use pricing. These strategies can help reduce stress on power grids, improve energy efficiency, and lower electricity costs for consumers.

References

1. Kumar, R. D., Prudhviraj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 through intensive investigation with supervised machine learning algorithm. In Handbook of Artificial Intelligence and Wearables (pp. 145-158). CRC Press.
2. Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2024, November). Machine Learning Techniques in Cloud Based Intrusion Detection. In The International Conference on Artificial Intelligence and Smart Environment (pp. 557-564). Cham: Springer Nature Switzerland.
3. Sv satyakrishna, shirisha rangu ,bhargavi nalacheruve.(2024) Prospective investigation on colorectal cancer with SMOTE on machine learning Algorithm
4. Dr.G.Vishnu Murthy, BhargaviNalacheruve 1Professor, Department of computer Science & engineering, Anurag University, TS, India. 2Student, Department of computer Science & engineering, Anurag University, TS, India.
5. V. N. S. Manaswini, K. K, C. Nigam, S. S. Ali, R. Niranjana, and Suman, "Real-Time Object Detection in Drone Surveillance Using YOLOv5," in Proc. 2025 3rd Int. Conf. IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 2025, pp. 1–6, doi: 10.1109/ICICAT68430.2025.11414670.
6. B. Soundarya, V. N. S. Manaswini, M. Ayyakrishnan, R. D. Kumar, "Contextual Analysis of Big Data Analytics in Intelligent Transportation Frameworks," in Intersection of Artificial Intelligence, Data Science, and Cutting-Edge Technologies: From Concepts to Applications in Smart Environment, Lecture Notes in Networks and Systems, vol. 1353, Cham: Springer, 2025, doi: 10.1007/978-3-031-88304-0_79.
7. R. D. Kumar, V. N. S.Manaswini, "Applications of blockchain in smart cities: detecting fake documents from land records using blockchain technology," in Blockchain for Smart Cities, Elsevier, 2021, pp. 105–117, doi: 10.1016/B978-0-12-824446-3.00017-X.

8. Tejavath Veeramma, Badarla Anil, Guguloth Ravinder, "An advanced movie recommender using collaborative filtering and sentiment analysis," International Research Journal of Modernization in Engineering Technology and Science, vol. 7, no. 7, July 2025, doi: 10.56726/IRJMETS81618.
9. Ravi Kumar Banoth, Ramana Murthy B V, "Automatic crop recommendation system using LightGBM and decision tree machine learning models," Journal of Machine and Computing, vol. 5, no. 1, pp. 343, Jan. 2025, doi: 10.53759/7669/jmc202505026.
10. Ravi Kumar Banoth, Dr. B.V. Ramana Murthy, "Smart agriculture through IoT and machine learning for analyzing carbon footprints," in Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE), Apr. 2025.
11. Ravi Kumar Banoth, B. V. Ramana Murthy, "Soil image classification using transfer learning approach: MobileNetV2 with CNN," SN Computer Science, vol. 5, art. no. 199, 2024, doi: 10.1007/s42979-023-02500-x.
12. Gaddam, S. Integrating Analytics into the Development Process: Bridging the Gap between Data Insights and Design Execution.
13. Reddy, S. K. R. Developing a Modular AI Framework to Enhance Scalability and Personalization in Next-Generation Reward Platforms.
14. Poojari, R. Frameworks for Data Management and Lineage in Large-Scale Healthcare Data Systems.
15. Purmani, S. S. R. (2025). Streamlining IT operations and service management with agile frameworks. European Journal of Advances in Engineering and Technology, 12(4), 76–81.
16. Viswanathan, V. (2024). Embedding Ethical Principles into Generative AI Workflows for Project Teams.
17. Mudusu, S. K. (2026, April 15). The secure intelligence framework: Architecting AI systems for a data-driven world. CIO (Foundry Expert Contributor Network).
18. Viswanathan, V. (2024). Pioneering Ethical AI Integration in Enterprise Workflows: A Framework for Scalable Team Governance. Available at SSRN 5375619.
19. Mudusu, S. K. (2025, June 3). Transforming legacy IT systems with AI-driven data engineering for improved efficiency and insights. Hampton Global Business Review (HGBR).
20. Gajula, S. (2026, March). Two Pillars of Banking Intelligence: A Comparative Analysis of AI Techniques for Fraud Prevention and Churn Mitigation. In 2026 14th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-6). IEEE.
21. Maturi, S. Y. (2023). Crowdsourced frontier: Unveiling autonomous adversarial cybercapabilities via open AI competition. International Journal of Intelligent Systems and Applications in Engineering, 11(1s), 275–284.
22. Chowdhury, A. K., Muhit, M. M. I., & Islam, M. M. (2023). A practical review to the marine maintenance practice in Bangladesh and a proposed way forward to an efficient, long-term and cost-effective solution. In Proceedings of the 13th International Conference on Marine Technology (MARTEC 2022). https://doi.org/10.2139/ssrn.4445071
23. Manoharan, D. (2025). Healthcare EDI Transaction Lifecycles Embedded with a Multi-Layer Verification Framework to Ensure Referential Integrity.

24. Ravishankara, M. (2026, February). CircuChain: Disentangling Competence and Compliance in LLM Circuit Analysis. In SoutheastCon 2026 (pp. 1-7). IEEE.
25. Doragacharla, V. R. (2023). Comprehensive Benchmarking Analysis of Auto Scaling Approaches in Cloud Native Streaming Pipelines During Flash Sales and Holiday Traffic Peaks. Available at SSRN 6566479.
26. P. Venkata Ramana. (2024). AI-driven predictive analytics in ERP systems for proactive supply chain optimization. International Journal of Innovative Engineering and Management Research (IJIEMR).
27. Kumar Adabala, P. (2021). Optimizing ERP Modernization: A Smart Data Migration Framework Approach. International Journal of Enhanced Research in Science, Technology &Amp; Engineering, 10(07), 61–72. https://doi.org/10.55948/ijerste.2021.0708
28. Kavuri, S. (2025). Critical Review of Software Testing Problems in the Current Decade. International Journal on Science and Technology, 16(2). https://doi.org/10.71097/ijsat.v16.i2.9469
29. Srikanth Kavuri. (2024). Probabilistic Generative Modeling for Synthesizing High-Coverage Test Data in Safety-Critical Software Applications. Computer Fraud and Security, 633–642. https://doi.org/10.52710/cfs.838
30. Venkata Pavan Kumar Gummadi. (2024). API Design and Implementation: RAML and OpenAPI Specification. Journal of Electrical Systems, 16(4), 76–85. https://doi.org/10.52783/jes.9329
31. Venkata Pavan Kumar Gummadi. (2025). MuleSoft's Role in Advancing Sustainable Digital Infrastructure: An Enterprise Integration Perspective. Journal of Information Systems Engineering and Management, 10(62s), 1313–1321. https://doi.org/10.52783/jisem.v10i62s.13783