

HOTEL BOOKING INSIGHTS AND ANALYTICS

¹Ch Roja, ²P Sagar, ³G Abhinay Kumar, ⁴K Yugendhar Chari

¹Assistant Professor, ^{2,3,4}Students

Department of Data Science

Siddhartha Institute of Technology & Sciences, Narapally

sits.rojachilukuri@siddhartha.co.in, 23TQ1A6719@siddhartha.co.in, 23TQ1A6755@siddhartha.co.in,
23TQ1A6726@siddhartha.co.in,

ABSTRACT

The Hotel Booking Insights and Analytics system is a data-driven application designed to analyze hotel reservation data and generate meaningful insights about booking patterns, customer behavior, cancellations, and hotel performance. Hotels generate large amounts of data through reservations, customer information, room types, booking channels, stay duration, and cancellation records. Analyzing this information can help understand business trends and improve operational planning.

The proposed system collects and processes hotel booking data containing attributes such as hotel type, booking date, arrival date, number of guests, room type, lead time, stay duration, customer type, booking channel, deposit type, and cancellation status. Data preprocessing techniques are applied to clean the dataset, handle missing values, remove duplicate records, and standardize inconsistent information. The processed data is then prepared for analytical operations.

The system performs exploratory data analysis to identify important patterns in hotel reservations. It can analyze booking volumes, cancellation rates, average stay duration, lead time, customer segments, seasonal demand, and booking channels. Interactive dashboards, charts, graphs, and KPI cards are used to present the results in an easy-to-understand format.

The analytics module can also compare hotel performance across different periods and customer categories. It can identify months with high or low booking demand, frequently selected room types, common booking channels, and patterns associated with cancellations. These insights can help hotel management understand customer behavior and plan resources more effectively.

Overall, the Hotel Booking Insights and Analytics system provides a centralized platform for converting raw reservation data into useful business insights. It reduces manual data analysis and supports data-driven decision-making for hotel operations. Future enhancements can include demand forecasting, cancellation prediction, personalized recommendations, dynamic pricing analysis, customer segmentation, and real-time booking analytics.

1. INTRODUCTION

The hotel industry handles a large number of reservations and customer interactions every day. Booking information contains valuable details about customer preferences, arrival dates, room types, length of stay, booking channels, and cancellation behavior. Analyzing these records can help hotels understand demand patterns and improve their operational efficiency.

Traditional hotel booking management systems primarily focus on creating reservations, maintaining customer details, and managing room availability. Although these functions are necessary for daily operations, they may not provide sufficient analytical information about historical booking trends. Hotel managers may need additional tools to analyze large datasets and understand business performance.

Booking patterns can vary depending on season, holidays, customer type, location, room availability, and booking channel. Cancellation behavior is also an important

factor because cancelled reservations can affect room occupancy and expected revenue. Understanding these patterns can help hotel management plan resources and identify areas requiring attention.

Data analytics provides methods for examining hotel booking records and discovering relationships between different variables. Descriptive analytics can calculate booking counts, cancellation rates, average stay duration, and other KPIs. Data visualization can represent these results using charts and dashboards, making complex booking information easier to interpret.

The proposed Hotel Booking Insights and Analytics system combines data preprocessing, exploratory analysis, visualization, and reporting in a centralized platform. It provides users with insights into booking trends, customer behavior, cancellations, room preferences, and seasonal demand. The system can serve as a foundation for future predictive analytics and intelligent hotel-management applications.

II. LITERATURE SURVEY

Hotel booking analytics has become an important application of data analytics because hospitality businesses generate large amounts of reservation and customer data. Historical booking datasets can contain information about arrival dates, departure dates, hotel types, room types, customer categories, booking channels, and cancellation status. Researchers and businesses use this information to understand hotel demand and operational performance.

Descriptive analytics is commonly used in hospitality to summarize booking activity. Metrics such as total bookings, cancellation percentage, average length of stay, lead time, and occupancy-related measures can provide an overview of hotel operations. These metrics can be grouped by month, customer type, hotel category, or booking channel to identify differences between segments.

Seasonality is an important factor in hotel demand analysis. Booking volumes may change depending on holidays, vacation periods, weekends, and other seasonal factors. Historical data can be analyzed using monthly and yearly trends to identify periods of high and low demand. Such analysis can support capacity planning and resource allocation.

Customer behavior analysis is another important area of hotel analytics. Information about customer type, number of guests, preferred room types, booking channels, and previous booking behavior can provide insights into different customer segments. Understanding these patterns can help hotels design appropriate services and improve customer experience.

Cancellation analysis is also widely studied because cancelled bookings can affect hotel occupancy and revenue planning. Factors such as lead time, deposit type, booking channel, customer type, and arrival period may be examined to understand cancellation patterns. Machine Learning techniques can later be used to build cancellation-prediction models based on historical booking data.

Modern hotel analytics systems increasingly combine data processing, visualization, forecasting, and predictive modeling. Interactive dashboards allow managers to examine booking trends and filter information according to different categories. The proposed Hotel Booking Insights and Analytics system follows this approach by integrating booking-data analysis, cancellation analysis, customer insights, visualization, and a foundation for future predictive capabilities.

III. SYSTEM ANALYSIS

The Hotel Booking Insights and Analytics system is designed to process historical hotel reservation data and generate useful analytical insights. The system accepts booking information such as hotel type, arrival date, departure date, booking date, room type, customer type, booking channel, lead time, and cancellation status. The input dataset is first validated to ensure that required fields and supported formats are available. A preprocessing module removes duplicate records and handles missing, invalid, or inconsistent values. Date-related attributes are standardized and transformed into useful features such as month, day, and season. The analytics engine calculates KPIs including total bookings, cancelled bookings, cancellation rate, average stay duration, and booking volume. Comparative analysis is performed across hotel types, room categories, customer segments, and booking channels. Trend analysis identifies high-demand and low-demand periods using historical booking data. Visualization components present analytical results through charts, graphs, tables, and KPI cards. Users can apply filters to analyze specific periods, hotel categories, customer types, or booking statuses. The system can generate analytical reports for hotel-management purposes. The collected data can also serve as a foundation for future Machine Learning models such as demand forecasting and cancellation prediction.

Existing System

Existing hotel booking systems mainly focus on reservation management, room availability, customer registration, check-in, check-out, and payment processing. These systems are useful for handling individual bookings and maintaining operational records. However, many basic systems provide limited analytical capabilities for understanding historical booking patterns. Hotel staff may need spreadsheets or separate reporting tools to analyze reservation data. Calculating cancellation rates, average stay duration, booking trends, and customer patterns manually can be time-consuming. Large datasets can make manual filtering and comparison difficult. Traditional systems may provide standard reports but may have limited interactive dashboards. Seasonal booking patterns may not be easily visible without additional analysis. Comparing different booking channels or customer segments may require multiple reports. Cancellation patterns can also be difficult to identify when data is distributed across operational records. Some systems focus primarily on transaction processing rather than business intelligence. Therefore, an integrated analytics platform can provide additional value by combining hotel booking information with automated analysis and visualization.

Disadvantages of Existing System

- Limited historical booking analysis.
- Manual preparation of analytical reports.
- Difficult to analyze large booking datasets.
- Limited interactive dashboards.
- Seasonal demand patterns may be difficult to identify.
- Cancellation trends may require manual processing.
- Difficult to compare customer segments.
- Limited booking-channel analysis.
- Repetitive spreadsheet calculations may be required.
- Data may be distributed across multiple systems.
- Limited predictive analytics capabilities.
- Difficult to obtain comprehensive business insights quickly.

Proposed System

The proposed Hotel Booking Insights and Analytics system provides a centralized platform for analyzing hotel reservation data and generating meaningful business insights. Users can upload historical booking datasets or connect the application to an authorized data source. The system validates the incoming data and performs preprocessing to handle missing values, duplicates, invalid entries, and inconsistent formats. Date fields are transformed to generate useful attributes such as month, season, weekday, and booking period. The analytics module calculates important metrics such as total bookings, cancellation rate, average stay duration, lead time, and booking volume. Users can compare booking activity across hotel types, room categories, customer segments, and booking channels. Trend analysis identifies seasonal variations and periods with higher or lower booking activity. Cancellation analysis highlights patterns associated with cancelled and completed reservations. Interactive dashboards display KPIs, charts, graphs, tables, and filtered results in a clear format. The reporting module can generate summaries for management and operational analysis. The system can later be extended with Machine Learning models for demand forecasting and cancellation prediction. Overall, the proposed system combines hotel booking analysis and business intelligence in a unified platform.

Advantages of Proposed System

- Provides centralized hotel booking analytics.
- Automates data cleaning and analysis.
- Identifies booking trends over time.
- Calculates cancellation rates automatically.
- Analyzes customer behavior patterns.
- Supports hotel and room-type comparisons.
- Analyzes different booking channels.
- Identifies seasonal demand patterns.
- Provides interactive dashboards.
- Reduces manual reporting effort.
- Supports customized filtering and analysis.
- Provides a foundation for predictive hotel analytics.

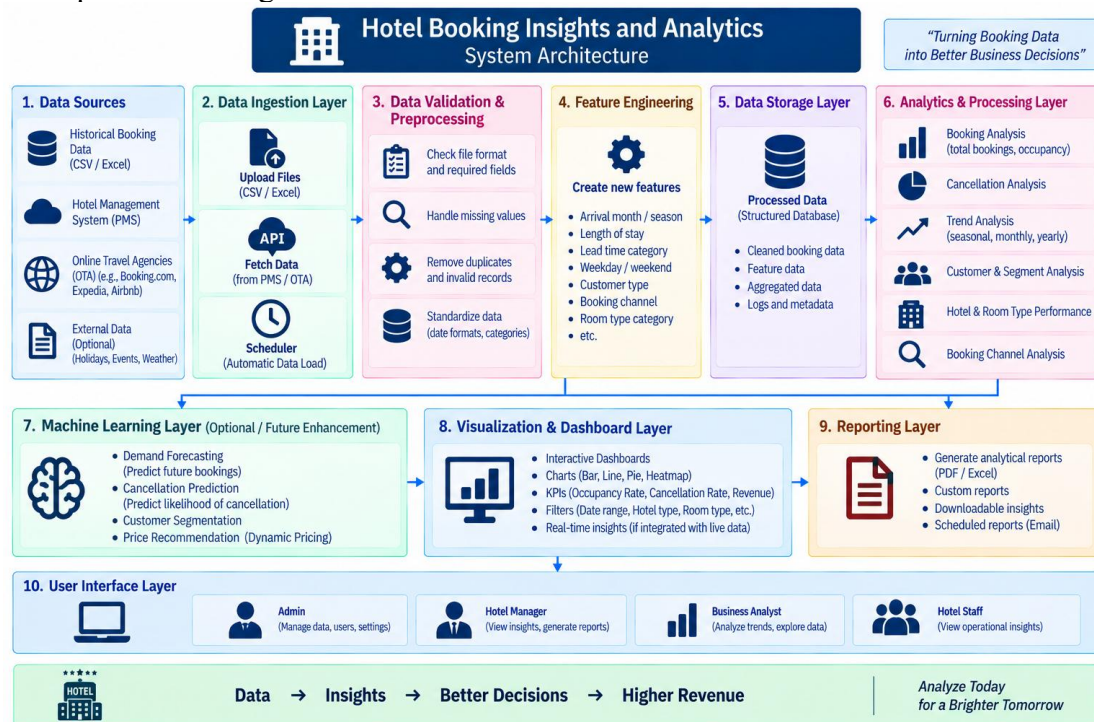
IV. METHODOLOGY

The methodology begins with collecting historical hotel booking records from an authorized dataset or database. The collected data is validated to check file structure, required attributes, data types, and completeness. Data preprocessing is performed to remove duplicate records and handle missing, invalid, or inconsistent values. Date fields are standardized and transformed into useful features such as arrival month, arrival weekday, season, and booking period. Exploratory Data Analysis is performed to understand booking distributions and identify important patterns. Booking analysis is used to examine reservation volumes across different hotels, room types, customer categories, and time periods. Cancellation analysis is performed to calculate cancellation rates and examine patterns associated with cancelled bookings. Customer and booking-channel analysis provides additional information about reservation behavior. Visualization techniques are used to represent the results through bar charts, line graphs, pie charts, tables, and interactive dashboards. Statistical metrics and KPIs are calculated to summarize important booking information. The system provides filtering and reporting capabilities so users can examine specific segments of the dataset. The complete methodology combines data collection, validation,

preprocessing, exploratory analysis, booking analysis, cancellation analysis, visualization, filtering, and reporting.

SYSTEM ARCHITECTURE

The system architecture consists of several interconnected layers that process hotel booking information and deliver analytical insights. The Data Source Layer provides historical reservation records from authorized datasets, databases, or APIs. The Data Input Layer accepts booking information through CSV, Excel, database, or API sources. The Data Validation Layer checks file format, required fields, data types, and basic data quality. The Data Preprocessing Layer handles missing values, duplicate records, inconsistent dates, and invalid entries. The Feature Engineering Layer creates useful attributes such as arrival month, season, weekday, lead-time category, and stay duration. The Data Storage Layer securely stores cleaned booking records and historical analytical information. The Analytics Processing Layer calculates booking counts, cancellation rates, average stay duration, lead time, and other KPIs. The Trend and Comparative Analysis Layer examines booking patterns across hotels, room types, customer segments, booking channels, and time periods. The Visualization Layer presents results through dashboards, charts, graphs, tables, and KPI cards. The Reporting Layer generates analytical summaries and reports for authorized users. The Security Layer manages authentication, authorization, and protection of sensitive booking and customer information. An optional Machine Learning Layer can later provide demand forecasting, cancellation prediction, customer segmentation, and other predictive insights.



V. RESULT AND OUTPUT



VI. CONCLUSION

The Hotel Booking Insights and Analytics system provides a data-driven solution for analyzing hotel reservation data and generating meaningful business insights. The system combines data preprocessing, exploratory analysis, visualization, and reporting to reduce the effort required for manual analysis.

The system analyzes important booking factors such as hotel type, arrival date, room type, customer type, booking channel, lead time, length of stay, and cancellation status. This helps identify booking patterns and understand how different factors are associated with hotel reservations and cancellations.

The analytical dashboard presents important information such as total bookings, cancelled bookings, cancellation rate, average stay duration, booking trends, and customer distribution. Charts, graphs, KPI cards, and filters make it easier for hotel management to understand large amounts of booking data.

The proposed system can help hotel managers and analysts monitor booking performance, identify seasonal demand patterns, compare customer segments, and examine cancellation behavior. The insights generated from historical data can support operational planning and business analysis.

In the future, the system can be enhanced with Machine Learning-based demand forecasting, cancellation prediction, dynamic pricing analysis, customer segmentation, personalized recommendations, real-time booking integration, and automated alerts. These enhancements can make the system more intelligent, scalable, and useful for modern hotel-management operations.

REFERENCES

- [1] Kumar, R. D., Prudhviraaj, 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 rang, 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.