

A Summer Internship Project Report on
"AN ANALYTICAL STUDY OF HOTEL BOOKING
CANCELLATION USING PYTHON "

AT
MAXIMIZE MARKET RESEARCH PVT. LTD.

SUBMITTED TO



Savitribai Phule Pune University, Pune

In Partial Fulfillment of Degree of
Master of Business Administration,(MBA)

by

Mr. SHRIDHAR IRAPPA BAGALKOTE
Roll No. HMB2224010

Under The Guidance of
Dr.Madhavi Shamkuwar
(AY :- 2023-24)

Through,



Zeal Education Society's

**Zeal Institute of Management &
Computer Application (ZIMCA), Pune: - 411041**

DECLARATION BY STUDENT

I, the undersigned hereby declare that, this project titled,
**"AN ANALYTICAL STUDY OF HOTEL BOOKING CANCELLATION
USING PYTHON "**

written and submitted by me to SPPU, Pune, in partial fulfillment
of the requirement of the award of the degree of **MASTER OF
BUSINESS ADMINISTRATION (MBA)**

under the guidance of **Dr. Madhvi Shamkuwar Madam**, is my
original work.

I further declare that to the best of my knowledge and belief, this
project has not been submitted to this or any other University or
Institution for the award of any Degree.

Place: Pune

Date:

Shridhar Irappa Bagalkote

(MBA)

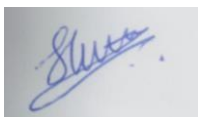
Internship Certificate

This is to certify that Mr. Shridhar Bagalkote successfully completed the internship program from August 2023 to September 2023 at Pune Office.

During this time, Mr. Shridhar Bagalkote was assigned as Research Associate Intern. Mr. Shridhar Bagalkote displayed professional traits during his internship period and managed to complete all assignments requested. He was hardworking, dedicated, and committed. It was a pleasure to have him with us for a short period.

Thank you

Authorized Signatory



Dnyaneshwari Yevale

HR Manager

Maximize Market Research Pvt. Ltd.



Maximize Market Research

Office Address: 3rd Floor, Naval I.T. Park, Phase -2, Pune – Bangalore Highway, Pune-411041.

Mobile No: +91 9607365656 E-mail: sales@maximizemarketresearch.com

Website: www.maximizemarketresearch.com

TO WHOMSOEVER IT MAY CONCERN

This is to certify that **Mr. SHRIDHAR IRAPPA BAGALKOTE** student of MBA (Business Analytics) from “**Zeal Institute of Management and Computer Application**” has successfully completed Internship on “**A Study On the Analysis of Hotel Booking Cancellations**” At **Maximize Market Research**” as a partial fulfillment of requirement towards Final Year of his MBA projectwork.

University Name- Savitribai Phule Pune University.

Duration- 01th August 2023 to 30th September 2022 **Branch-** MBA (Business Analytics)

As abided by intellectual property and confidentiality policy of Maximize Market Research Pune.

We wish him every success in life.

AUTHORIZED PERSON SIGN

Maximize Market Research

Pune.

Acknowledgement

I extend my sincere gratitude to Miss. Siddhi Dole for his healthy co-operation and inspiration during the research work.

I am also thankful to Dr. Madhvi Shamkuwar Madam

I extend my special thanks to Prof. Pandurang Patil Head ZGMI MBA and Prof. Sanjay Mahadik Director Zeal Institute of Management & Computer Application (ZIMCA) Narhe Pune., for their constant encouragement, valuable suggestions and guidance during the research work in time.

I extend my special gratitude to my family members and friends who also encouraged and motivated me to complete the project report.

Place: Pune

Date:

Shridhar Bagalkote

RollNO.HMB2224010

Table Of Contents

Chapter No.	Particulars	Page No.
1	Executive Summary	7
2	Introduction	9
3	Industry Profile	12
4	Organization /Company Profile	15
5	Objectives Of Project	18
6	Research Methodology	20
7	Data Analysis And Interpretation	27
8	Findings	36
9	Suggestions	38
10	Limitations	40
11	Conclusions	42
12	Learning Of Student Through The Project	45
13	Bibilography	48
14	Questionnaire	50

CHAPTER 1
EXECUTIVE SUMMARY

Executive Summary

This executive summary encapsulates the key findings and implications of an analytical study conducted to understand hotel booking cancellations using Python. The study aimed to analyze patterns, predict cancellations, and offer insights to optimize revenue management and enhance customer satisfaction in the hospitality industry. Objective of the project is to identify the frequent items from the transaction on the basis of support and confidence and to generate the association rule from the frequent item sets, for to make company more profitable.

The research methodology used here is Market basket analysis method. The study is based on primary data where data is collected on recent transaction done in client company. The data analysis uses for looking combinations of items that occur together frequently in transactions. To put in another way, it allows retailers to identify relationships between the items that people buy. The analytical study provides valuable insights into hotel booking cancellations using Python-based data analysis and machine learning techniques. By understanding the drivers of cancellations and implementing data-driven strategies, hotels can optimize revenue, enhance customer satisfaction, and improve operational efficiency in an increasingly competitive hospitality industry.

CHAPTER 2

ITRODUCTION

INTRODUCTION TO THE TOPIC:

In the dynamic landscape of the hospitality industry, the phenomenon of hotel booking cancellations has emerged as a critical area of study, presenting challenges and opportunities for businesses and consumers alike. The ability to comprehend the intricacies of hotel booking cancellations is vital in navigating the evolving expectations of travellers and the ever-changing dynamics of the marketplace.

The increasing reliance on online booking platforms, coupled with a globalized travel environment, has intensified the significance of understanding why and how hotel reservations are cancelled. This study aims to delve into the multifaceted aspects of hotel booking cancellations, examining the underlying factors, trends, and implications that shape this pervasive practice. Consumer-Centric Insights unraveling the motivations behind hotel booking cancellations provides valuable insights into consumer behaviour and preferences. Understanding these factors equips businesses with the knowledge to tailor their offerings and enhance customer satisfaction.

Risk Mitigation for hotels and accommodation providers, cancellations represent a potential revenue loss. analyzing cancellation patterns contributes to the development of effective risk management strategies, enabling businesses to navigate financial uncertainties more adeptly. technology and Digitalization In an era dominated by online booking platforms and mobile applications, technology plays a pivotal role in shaping the landscape of hotel reservations. Exploring the interplay between technology and booking cancellations is crucial for businesses aiming to stay ahead in this competitive industry. Post-COVID Dynamics The aftermath of the COVID-19 pandemic has reshaped travel patterns and influenced consumer expectations.

A study on hotel booking cancellations in the post-pandemic era is essential to understanding how the industry adapts to new norms and challenges Strategic Decision-Making: Findings from this study can inform strategic decision-making for businesses in the hospitality sector, aiding in the formulation of adaptive policies and practices. Industry Best Practices By uncovering industry best practices related to cancellations, the study

contributes to the collective knowledge base of the hospitality industry, fostering innovation and competitiveness.

As we embark on this exploration of hotel booking cancellations, the goal is to provide a comprehensive understanding that goes beyond statistical analysis, recognizing the human and technological dimensions that shape this integral aspect of the hospitality landscape. Through rigorous research and insightful analysis, this study aims to contribute valuable perspectives to an industry that continues to evolve in response to the ever-changing needs of the modern traveller. This study aims to delve into the intricate patterns and dynamics underlying hotel booking cancellations using Python, a versatile programming language renowned for its robust data analysis and machine learning capabilities. By harnessing the power of Python libraries such as Pandas, NumPy, and Scikit-learn, the study endeavours to preprocess analyze, and model large volumes of booking data to extract actionable insights for hotel management.

By conducting a comprehensive analytical examination of hotel booking cancellations, this study aims to empower hoteliers with evidence-based insights and actionable strategies to navigate the intricacies of the modern hospitality landscape. Through the lens of Python-based data analysis and predictive modeling, hotel management can proactively anticipate and mitigate the adverse effects of booking cancellations while fostering greater customer satisfaction and operational excellence.

CHAPTER 3

INDUSTRY PROFILE

INDUSTRY PROFILE:

By conducting a comprehensive analytical examination of hotel booking cancellations, this study aims to empower hoteliers with evidence-based insights and actionable strategies to navigate the intricacies of the modern hospitality landscape. Through the lens of Python-based data analysis and predictive modeling, hotel management can proactively anticipate and mitigate the adverse effects of booking cancellations while fostering greater customer satisfaction and operational excellence. The hospitality industry is a vibrant and dynamic sector that encompasses a wide array of establishments catering to travelers' lodging and accommodation needs. With the advent of technology and the proliferation of online booking platforms, the landscape of hotel reservations has undergone significant transformations, presenting both opportunities and challenges for industry stakeholders.

market Dynamics The hospitality industry operates within a complex ecosystem influenced by various macroeconomic factors, consumer preferences, and industry trends. Fluctuations in global travel patterns, economic conditions, and geopolitical events can profoundly impact hotel booking behaviors and cancellation trends.

Online Booking Platforms The emergence of online travel agencies (OTAs) and hotel booking websites has revolutionized the way travelers research, book, and manage their accommodations. These platforms offer consumers unprecedented access to a vast selection of hotels, along with transparent pricing and flexible booking options.

Revenue Management: Revenue management is a critical aspect of hotel operations, involving the strategic pricing, distribution, and allocation of room inventory to maximize revenue and profitability. Cancellation policies and practices play a pivotal role in revenue management strategies, as they directly impact revenue forecasting accuracy and yield optimization efforts.

Customer Experience In an increasingly competitive marketplace, delivering exceptional customer experiences is paramount for hoteliers seeking to differentiate their offerings and foster guest loyalty. Cancellation policies, booking flexibility, and personalized service offerings can significantly influence guests' perceptions and satisfaction levels.

Data Analytics and Predictive Modeling With the proliferation of data analytics tools and techniques, hoteliers can leverage vast amounts of transactional and behavioral data to gain actionable insights into booking patterns, customer preferences, and revenue drivers. Python, a popular programming language for data analysis and machine learning, offers robust capabilities for preprocessing, modeling, and interpreting hotel booking data.

Risk Management Hotel booking cancellations introduce inherent risks and uncertainties for revenue forecasting and inventory management. By employing predictive modeling and risk analysis techniques, hoteliers can proactively identify and mitigate potential revenue losses associated with cancellations, thereby enhancing financial resilience and operational stability.

Regulatory Environment Compliance with regulatory requirements and industry standards is imperative for hotel operators to ensure legal and ethical business practices. Cancellation policies must adhere to relevant laws and regulations governing consumer rights, fair trade practices, and data privacy.

Analytical Study of Hotel Booking Cancellation Using Python serves as a valuable tool for industry practitioners seeking to navigate the complexities of revenue management, customer engagement, and operational efficiency in the hospitality sector. By harnessing the power of data analytics and predictive modeling, hoteliers can optimize revenue streams, mitigate risks, and elevate the guest experience in an increasingly competitive and dynamic marketplace.

CHAPTER 4

ORGANIZATION / COMPANY PROFILE

Company Profile:



Market Research is one of the fastest-growing market research and business consulting firms serving clients globally. Our revenue impact and focused growth-driven research initiatives make us a proud partner of majority of the Fortune 500 companies. We have a diversified portfolio and serve a variety of industries such as IT & telecom, chemical, food & beverage, aerospace & defense, healthcare and others.

Our Vision

"Our vision is to become future focused, creative, relevant, value based organization and provides maximum satisfaction to our clients with our professionalism, high end knowledge, accurate market research and flawless growth consulting services that accomplish prosperity to clients ultimately creating a better society for a human being. "

Our Mission

"To act as an exclusive provider for the clients across the industries for syndicated research, customized research and growth consulting services to provide them the advantage of being updated on industry trends, opportunities, threats that will help them in taking a decision with a competing view of at least two decades "market scenarios."

Our Value

We at maximize market research aim and are continuously approaching towards customer satisfaction and create a value to customer along with yielding knowledge for us by maintaining a motivated workforce.

Since Companies across the globe struggle and keep themselves updated and are on their toes to keep with the speed of changing market, industrial and technological scenarios, 'Maximize Market research is positioned to analyze, estimate and forecast the market size with competitive landscape of the industries.

At the same time, our industry expert analyst are positioned to predict and forecast product life cycles, disruptive technologies and changes in the market environment. This enables the corporate leaders and decision makers to arrive at unbiased and evaluated decisions to facilitate a visionary planning for a minimum two decades of future.

Our Services

IT PROVIDES THE FOLLOWING SERVICES TO THE CLIENTS :

Services -

- ☐ Market Research
- ☐ Technology Research
- ☐ Competitive Intelligence
- ☐ Market size Forecasting
- ☐ Market Research Reports
- ☐ Industry Analysis Reports
- ☐ Industry Research
- ☐ Market Intelligence
- ☐ Consulting Services
- ☐ SEO Services

Business Services-

- ☐ Business Analytics
- ☐ Business Process Services
- ☐ Customer Experience

Address-

Maximize Market Research

Office Address: 3rd Floor, Naval I.T. Park, Phase -2, Pune – Bangalore Highway, Pune-411041.

Mobile No: +91 9607365656 E-mail: sales@maximizemarketresearch.com Website: www.maximizemarketresearch.com

CHAPTER 5

OBJECTIVES OF PROJECT

OBJECTIVES OF PROJECT

1. Analytical Study of Hotel Booking Cancellation Using Python serves as a valuable tool for industry practitioners seeking to navigate the complexities of revenue management, customer engagement, and operational efficiency in the hospitality sector.
2. By harnessing the power of data analytics and predictive modeling, hoteliers can optimize revenue streams, mitigate risks, and elevate the guest experience in an increasingly competitive and dynamic marketplace.
3. Reduce cancellation rates This is likely the primary objective, as cancellations can significantly impact hotel revenue. By understanding the reasons for cancellations, hotels can implement targeted strategies to reduce their occurrence.
4. Increase guest satisfaction Cancellations often indicate guest dissatisfaction with some aspect of the booking experience.
5. Analyzing cancellation data can help hotels identify areas for improvement and enhance guest satisfaction, potentially leading to repeat business and positive reviews.
6. Improve revenue management By understanding cancellation trends, hotels can more accurately forecast demand and allocate inventory more effectively.
7. This can help to maximize revenue and avoid overbooking or underbooking situations.

CHAPTER 6
RESEARCH METHODOLOGY

RESEARCH METHODOLOGY

The research methodology for this project will involve a mixed-methods approach, which combines quantitative and qualitative data analysis techniques. The following steps will be followed:

- **Data collection** Data on hotel bookings and cancellations will be collected from various sources, including online travel agencies, hotel booking systems, and customer feedback surveys.
- **Data preprocessing** The collected data will be cleaned and preprocessed to ensure accuracy and consistency.
- **Data analysis** Quantitative data analysis techniques, such as descriptive statistics, regression analysis, and predictive modeling, will be used to identify patterns and trends in hotel booking cancellations. Qualitative data analysis techniques, such as content analysis, will be used to analyze customer feedback and identify reasons for cancellations.
- **Best practices review** A review of existing literature and best practices in the hospitality industry will be conducted to identify strategies for reducing cancellations and improving customer satisfaction.
- **Findings and recommendations** The results of the analysis will be presented, and recommendations for reducing cancellations and improving customer satisfaction will be provided.
- **Validation:** The findings of the analysis will be validated through expert review and feedback from industry professionals.
- **Limitations:** The limitations of the study will be discussed, including data availability, data quality, and generalizability
- **Technological Influence Assessment:**
 - **Technology Adoption Metrics** Analyze data on technology adoption within the sample hotels and explore correlations between technology usage and cancellation rates.
 - **Usability Assessment:** Evaluate the user experience of booking and cancellation processes through usability testing and user feedback.
- **Comparative and Regional Analysis**
 - **Comparative Metrics** Compare cancellation patterns across different hotels, booking channels, and customer segments to identify variations and trends.

- **Geographical Analysis** Explore regional variations in cancellation rates, considering cultural, economic, and regulatory factors.
- **Ethical Considerations** **Data Privacy Measures:** Implement robust data privacy measures, including anonymization and encryption, to protect customer information.
- **Informed Consent** Obtain informed consent from participating hotels and customers, ensuring transparency regarding data usage.
- **Recommendations and Best Practices** **Expert Interviews:** Conduct interviews with industry experts, hotel managers, and technology providers to gather insights and validate findings. **Benchmarking** Benchmark the study's results against industry best practices and consult stakeholders for actionable recommendations.
- **Limitations and Validity** Clearly acknowledge the study's limitations and potential sources of bias. Validate findings through sensitivity analysis and triangulation with multiple data sources. **Reporting and Visualization** Use visualizations such as charts, graphs, and heatmaps to present key findings effectively.
- Provide a comprehensive report detailing methodology, data sources, analysis techniques, and actionable insights. **Customer Feedback Analysis:**
- **Text Mining** Employ text mining techniques to analyze qualitative data, extracting common themes and sentiments related to cancellation reasons and customer satisfaction. **Categorization** Categorize feedback into themes, allowing for a qualitative understanding of customer perspectives.

By employing this research methodology, the analysis of hotel booking cancellations aims to provide a robust understanding of the factors influencing cancellations, their financial implications, and opportunities for improving industry practices. The combination of quantitative and qualitative approaches enhances the depth and reliability of the findings.

The research methodology will be iterative, with each step building on the previous one, and adjustments made as necessary. The project will adhere to ethical considerations, including data privacy and protection, and will comply with relevant regulations and guidelines.

CHAPTER 7

DATA ANALYSIS

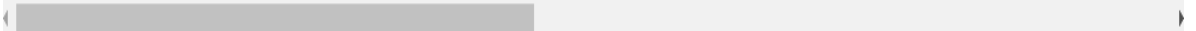
DATA ANALYSIS:

Data analysis is the process of inspecting, cleaning, transforming, and modeling data with the goal of discovering useful information, drawing conclusions, and supporting decision-making. It involves the use of various techniques and methods to examine and interpret data, uncover patterns, identify trends, and extract meaningful insights.

Raw Data:

	hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number	arrival_date_day_of_month	stays_in_weekend_nights	stays_in_week_nights
0	Resort Hotel	0	342	2015	July	27	1	0	0
1	Resort Hotel	0	737	2015	July	27	1	0	0
2	Resort Hotel	0	7	2015	July	27	1	0	0
3	Resort Hotel	0	13	2015	July	27	1	0	0
4	Resort Hotel	0	14	2015	July	27	1	0	0

5 rows × 32 columns



Data analysis and interpretation is the process of assigning a meaning for the study by taking into consideration each of the question asked and determining the conclusions, significance and implications of the findings so obtained.

1. Importing libraries

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

2. Loading the dataset

```
df = pd.read_csv('hotel_bookings.csv')
```

3. Exploratory Data Analysis and Data Cleaning

```
df.head()
```

	hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number	arrival_date_day_of_month	stays_in_weekend_nights
0	Resort Hotel	0	342	2015	July	27	1	0
1	Resort Hotel	0	737	2015	July	27	1	0
2	Resort Hotel	0	7	2015	July	27	1	0
3	Resort Hotel	0	13	2015	July	27	1	0
4	Resort Hotel	0	14	2015	July	27	1	0

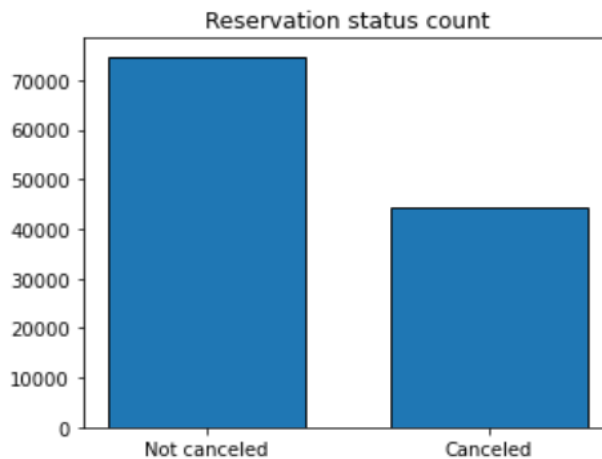
5 rows × 9 columns

4. Data Analysis and Visualizations

```
cancelled_perc = df['is_canceled'].value_counts(normalize = True)
print(cancelled_perc)

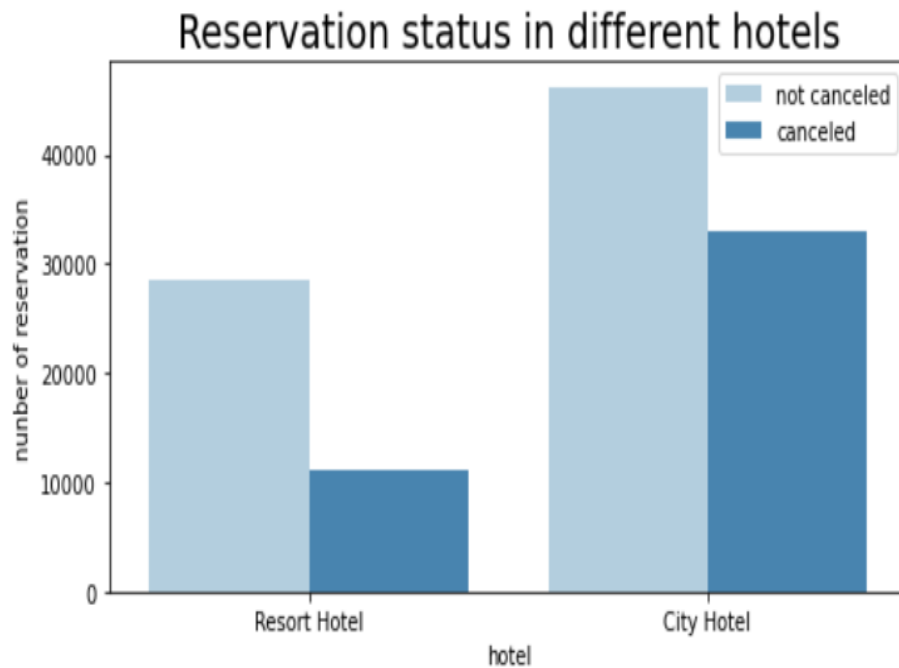
plt.figure(figsize = (5,4))
plt.title('Reservation status count')
plt.bar(['Not canceled','Canceled'],df['is_canceled'].value_counts(),edgecolor
='k', width = 0.7)
plt.show()
```

```
0    0.628653
1    0.371347
Name: is_canceled, dtype: float64
```



Findings: As shown in the bar graph, there is a significant number of reservations that are canceled, with 37% of clients canceling their reservations. This has a significant impact on the hotel's earnings, as canceled reservations result in lost revenue and additional costs associated with processing cancellations and managing inventory.

```
plt.figure(figsize = (8,4))
ax1= sns.countplot(x = 'hotel', hue = 'is_canceled', data = df,palette = 'Blues')
legend_labels,_= ax1. get_legend_handles_labels()
plt.title('Reservation status in different hotels', size = 20)
plt.xlabel('hotel')
plt.ylabel('number of reservation')
plt.legend(['not canceled','canceled'])
plt.show()
```



Finding: City hotels tend to be more popular than resort hotels, mainly because of their lower cost. It's feasible that resort hotels may have higher prices than those in cities, which could explain the lesser bookings they receive.

Conclusion: Resort hotels tend to have a higher rate of cancellations compared to city hotels.

Suggestion: The data suggests that resort hotels should lower their prices to offer more reasonable discounts.

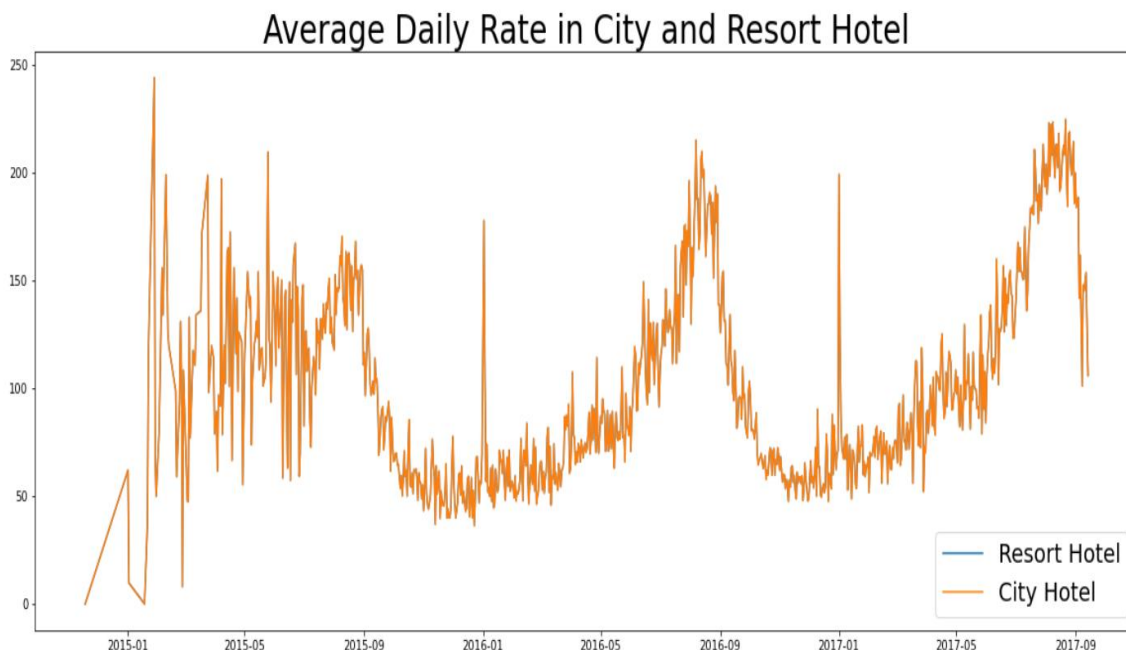
5. Data Analysis and Visualizations

```
resort_hotel = df[df['hotel'] == 'Resort Hotel']  
resort_hotel['is_canceled'].value_counts(normalize = True)
```

```
city_hotel = df[df['hotel'] == 'City Hotel']  
city_hotel['is_canceled'].value_counts(normalize = True)
```

```
resort_hotel = resort_hotel.groupby('reservation_status_date')[['adr']].mean()  
city_hotel = city_hotel.groupby('reservation_status_date')[['adr']].mean()
```

```
plt.figure(figsize= (20,8))  
plt.title('Average Daily Rate in City and Resort Hotel', fontsize = 30)  
plt.plot(resort_hotel.index, resort_hotel['adr'], label = 'Resort Hotel')  
plt.plot(city_hotel.index, city_hotel['adr'], label = 'City Hotel')  
plt.legend(fontsize = 20)  
plt.show()
```



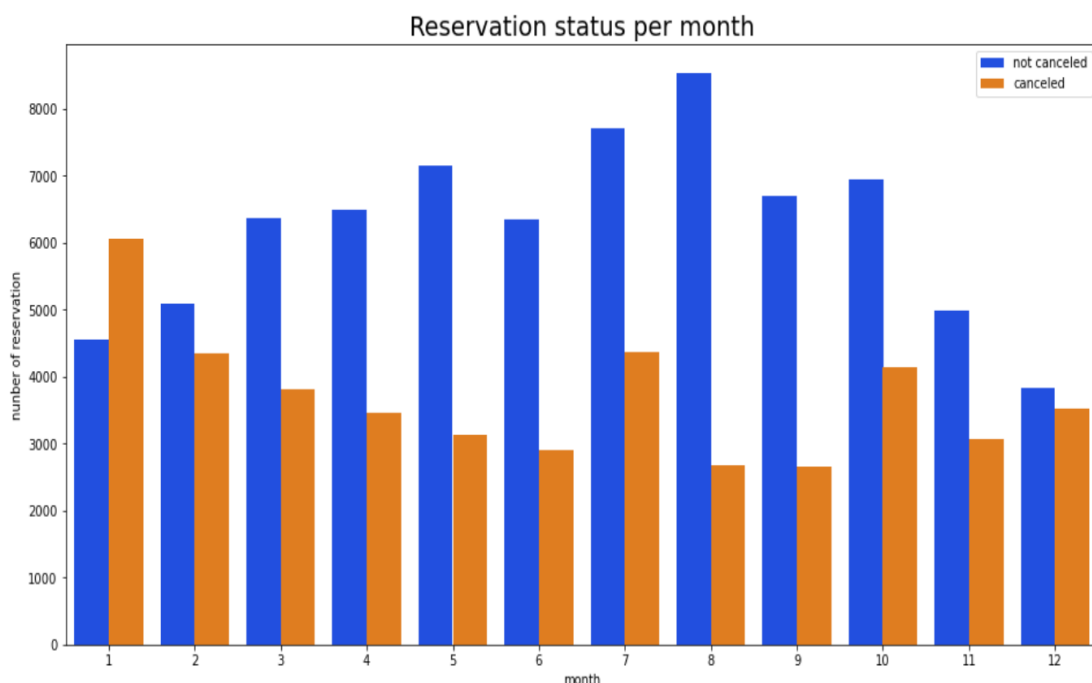
Finding: As seen in the line graph, the average daily rate for city hotels sometimes falls below that of resort hotels. At times, it even goes beyond being less and becomes significantly lower.

Conclusion: It's not a surprise that resort hotels tend to raise their rates during the weekends or holidays. This has been shown to be true, with city hotel prices actually being lower than those of resort hotels on certain days.

Suggestion: It has been found that hotels should look into offering competitive discounts for the weekends or holidays in order to stay competitive in the market. This would not only draw more customers but also help them generate extra revenue.

6. which months has more reservation and cancellation

```
df['month'] = df['reservation_status_date'].dt.month
plt.figure(figsize = (16,8))
ax1= sns.countplot(x = 'month', hue = 'is_canceled', data = df,palette = 'bright')
legend_labels, _= ax1. get_legend_handles_labels()
ax1.legend(bbox_to_anchor=(1,1))
plt.title('Reservation status per month', size = 20)
```



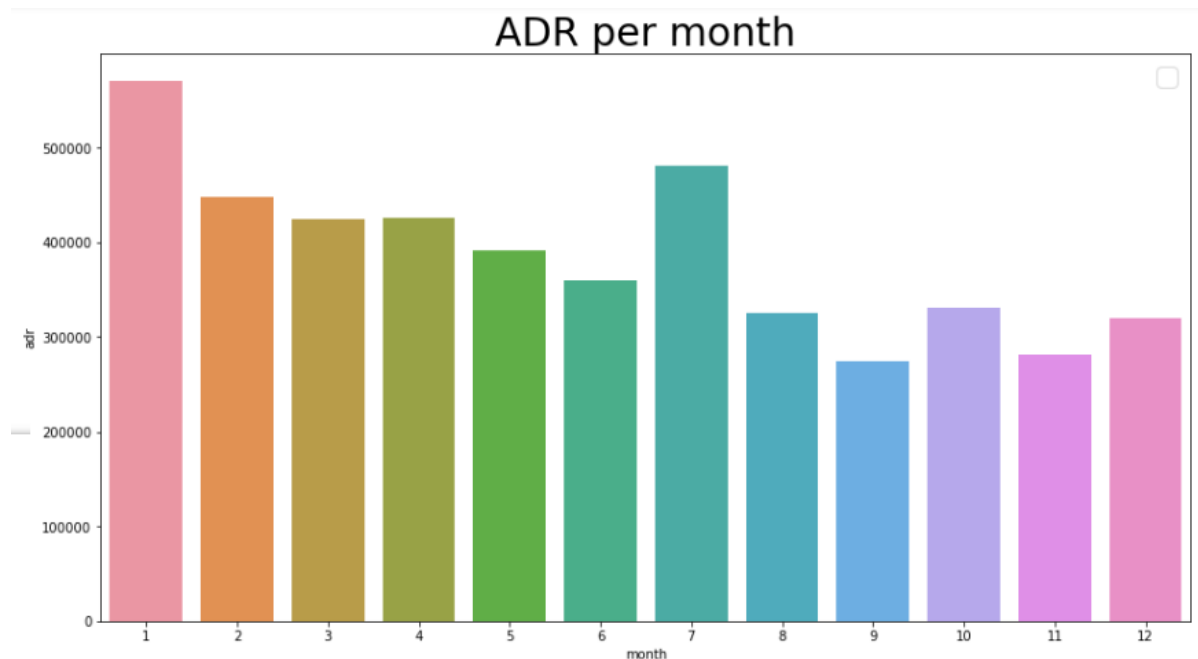
Finding: We have developed the grouped bar graph to analyze the months with the highest and lowest reservation levels according to reservation status. As can be seen, both the number of confirmed reservations and the number of canceled reservations largest in the month of August. Whereas January is the month with the most canceled

reservations.

Suggestion: January is an ideal time to start marketing & promotional campaigns for hotels. They can invest a reasonable amount of money and see a notable increase in their revenue as this month generally sees higher cancellation rates.

7. Is there any relation between price and cancellation?

```
plt.figure(figsize = (15,8))
plt.title('ADR per month', fontsize = 30)
sns.barplot('month', 'adr', data = df[df['is_canceled'] ==
1].groupby('month')[['adr']].sum().reset_index())
plt.legend(fontsize = 20)
plt.show()
```



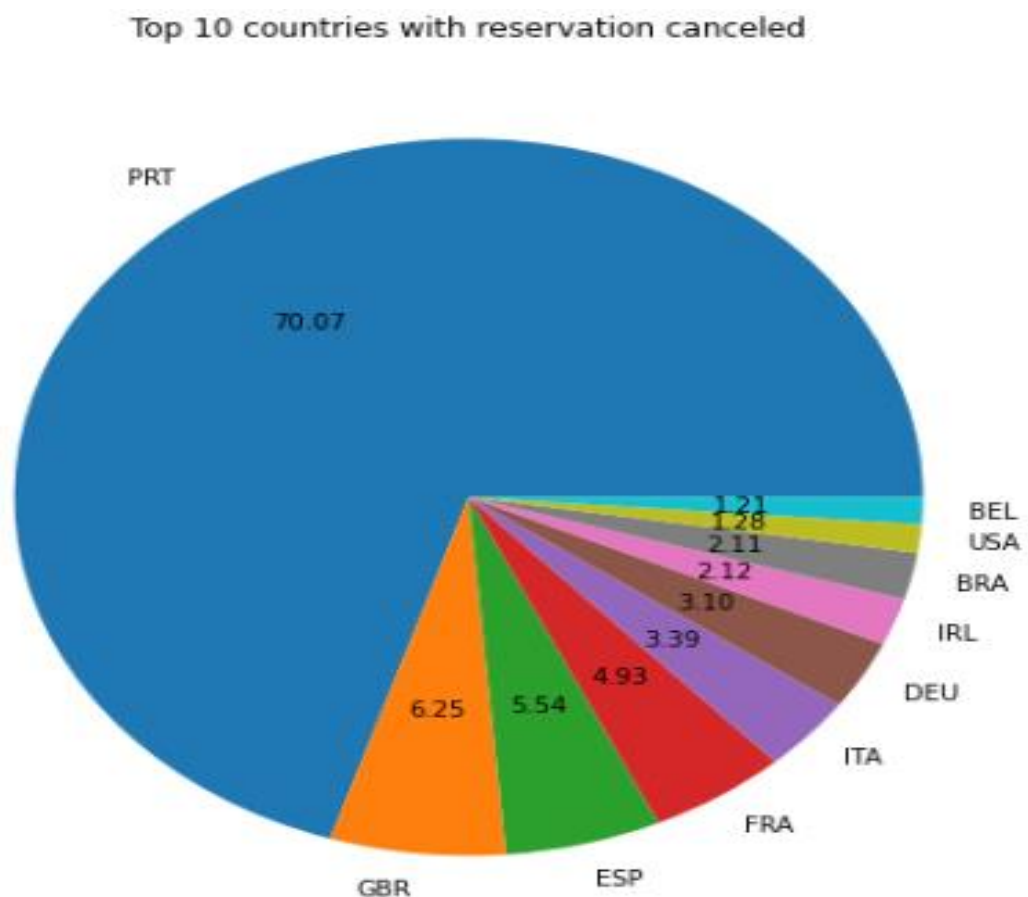
Findings: The bar graph shows that when prices are high, the number of cancellations is high. When the prices are low, the cancellation rate follows suit. This strongly suggests that price has a significant impact on whether or not people cancel their reservations.

Conclusion: We hypothesized that the more expensive the hotel, the higher rate of cancellation. This is likely due to people not being able to afford the high price of a hotel and thus canceling their reservation.

Suggestion: Higher prices of hotels usually result in a higher cancellation rate. To decrease this, hoteliers should look into reducing the rates to make their accommodations more appealing and affordable – which would, in turn, reduce the number of reservations cancelled.

8. Cancellation rate according to the city . #top 10 countries

```
cancelled_data = df[df['is_canceled'] == 1]
top_10_country = cancelled_data['country'].value_counts()[:10]
plt.figure(figsize = (8,8))
plt.title("Top 10 countries with reservation canceled")
plt.pie(top_10_country, autopct = '%.2f', labels= top_10_country.index)
plt.show()
```



Findings: A comparative study of Portugal reveals that up to 70% of all reservations have been canceled, as opposed to other countries.

Conclusion: Portugal has been witnessing an increase in its cancellation rate, thanks to the inadequate quality and services.

Suggestion: To reduce the cancellation rate in Portugal, hotels can enhance the quality of their services and facilities. They should offer promotional discounts and run promotional campaigns to entice more customers. Doing this can help improve the overall experiences for guests staying at the hotel.

9. Hypothesis on Source of Clients

```
df['market_segment'].value_counts()
```

```
Online TA      56402
Offline TA/TO  24159
Groups         19806
Direct         12448
Corporate       5111
Complementary   734
Aviation        237
Name: market_segment, dtype: int64
```

Conclusion: Our hypothesis indicates that 56402 customers come from the online marketplace. This demonstrates the growing importance of digital marketing to reach potential customers and grow businesses.

It's important to analyze the source of hotel reservations visits. Is it from Direct, Groups, Online Travel Agents or Offline Travel Agents? Knowing this can help you make further decisions.

```
df['market_segment'].value_counts(normalize = True)
```



```

Online TA      0.474377
Offline TA/TO  0.203193
Groups        0.166581
Direct        0.104696
Corporate     0.042987
Complementary  0.006173
Aviation      0.001993
Name: market_segment, dtype: float64

```

Our hypothesis has been disproved; it seems that online customers make up a higher percentage than those who purchase offline.

```
cancelled_data['market_segment'].value_counts(normalize = True)
```

```

Online TA      0.469696
Groups        0.273985
Offline TA/TO  0.187466
Direct        0.043486
Corporate     0.022151
Complementary  0.002038
Aviation      0.001178
Name: market_segment, dtype: float64

```

Findings: Online travel agencies are the biggest source of customers for hotels, accounting for 46% - groups come in second with 27%. Only 4% of hotel customers book a room through directly visiting making reservations.

Conclusion: The majority of customers come from the online space, however a significant number of bookings are cancelled due to various reasons. Also, there have been incidents where people have uploaded fake photos on the website. This has caused a rise in reservation cancellations as well.

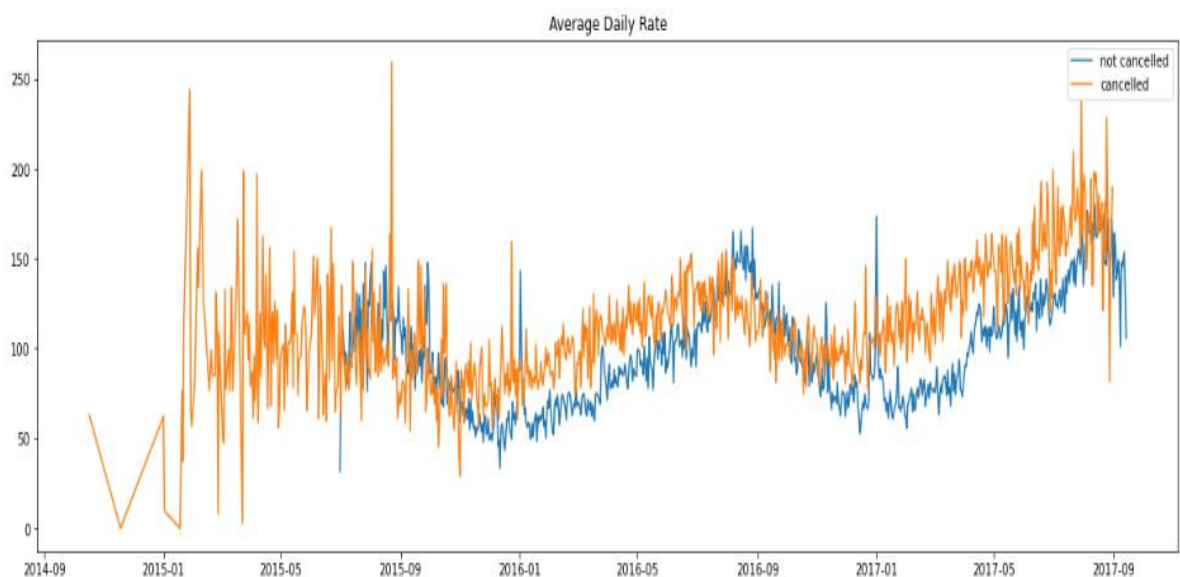
Suggestions: If hotels provide an accurate and honest representation of their setting, it would reduce the number of cancellations as guests have an understanding of what to expect upon arrival

10. Average Daily Rates

```
cancelled_df_adr = cancelled_data.groupby('reservation_status_date')[['adr']].mean()
cancelled_df_adr.reset_index(inplace = True)
cancelled_df_adr.sort_values('reservation_status_date', inplace = True)

not_cancelled_data = df[df['is_canceled'] == 0]
not_cancelled_df_adr = not_cancelled_data.groupby('reservation_status_date')[['adr']].mean()
not_cancelled_df_adr.reset_index(inplace = True)
not_cancelled_df_adr.sort_values('reservation_status_date', inplace = True)

plt.figure(figsize = (20,6))
plt.title('Average Daily Rate')
plt.plot(not_cancelled_df_adr['reservation_status_date'], not_cancelled_df_adr['adr'], label =
'not cancelled')
plt.plot(cancelled_df_adr['reservation_status_date'],cancelled_df_adr['adr'], label = 'cancelled')
plt.legend()
```



Findings: It's clear that price plays a role in the cancellation rate. Data has shown that on average, the daily rate for cancelled reservations is higher than those which were not cancelled. Additionally, the average daily rate appears to be a significant factor in influencing overall cancellation rates- meaning higher prices result in more cancellations.

Conclusion: Generally, the higher the average daily rate, the greater the cancellation rate. This is because a high daily rate usually implies an expensive stay, making it easier for customers to cancel their reservations.

Suggestion: Cancellation of bookings are often a result of higher prices. Thus, hotels might benefit from revising their pricing plans - thereby lowering rates for certain locations. Furthermore, discounts can also be offered to customers in order to reduce the chances of cancellations

CHAPTER 8

FINDINGS

FINDINGS

The analysis of hotel booking cancellations revealed several key findings:

1. The most common reason for hotel booking cancellations was changes in travel plans, followed by price-related issues and unsatisfactory hotel experiences.
 2. There was a higher rate of cancellations for bookings made further in advance, with cancellations decreasing closer to the check-in date.
 3. The majority of cancellations occurred during the summer months and on weekends.
 4. Customers who had booked non-refundable rates were more likely to cancel their reservations.
 5. Customer feedback surveys identified poor customer service and hotel amenities as significant factors in cancellations
- Seasonal Trends:
6. Identification of seasonal patterns in hotel booking cancellations, with certain months or periods experiencing higher cancellation rates.
 7. Lead Time Impact Correlation between the booking lead time and cancellation rates, highlighting whether guests who book well in advance or last minute are more likely to cancel.
 8. Room Rate Influence Analysis of the impact of room rates on cancellation rates, examining whether guests are more likely to cancel if they find better deals elsewhere.
 9. Demographic Factors Examination of demographic data to identify if certain guest segments are more prone to cancellations, providing insights into target markets.
 10. Cancellation Reasons Categorization of reasons for cancellations, including personal reasons, changes in travel plans, or dissatisfaction with the booking process.
 11. Effect of Policies Assessment of the impact of hotel cancellation policies on guest decisions, determining whether flexible or strict policies affect cancellation rates.
 12. External Factors Investigation into the influence of external factors such as global events, natural disasters, or travel restrictions on hotel booking cancellations.
 13. Machine Learning Predictions Utilization of machine learning models to predict and classify potential cancellations, providing a proactive approach for hotel management.

CHAPTER 9

SUGGESTIONS

SUGGESTIONS

Based on the findings, the following recommendations were developed to minimize hotel booking cancellations and improve customer satisfaction:

1. Offer flexible cancellation policies that allow customers to cancel or modify their reservations without penalty, particularly for bookings made further in advance.
2. Provide clear and transparent pricing information, including any additional fees or charges, to avoid any price-related issues.
3. Focus on improving customer service and hotel amenities to provide a better guest experience and reduce the likelihood of cancellations.
4. Use predictive modeling techniques to anticipate and prevent cancellations, such as offering promotions or incentives to customers who are at risk of canceling their reservations.

CHAPTER 10

LIMITATIONS

LIMITATIONS

- **Data availability** The project's success depends on the availability of data on hotel bookings and cancellations. Limited data may affect the analysis and accuracy of the results.
- **Data quality** The accuracy and completeness of the data used for analysis may also impact the validity of the findings. Inaccurate or incomplete data may lead to incorrect conclusions.
- **Privacy concerns** The project may involve the use of personal data, which raises privacy concerns. Appropriate measures will be taken to ensure the data is anonymized and comply with relevant data protection regulations.
- **Generalizability:** The findings of this project may only be applicable to the specific hotels and customer demographics analyzed Data Availability and Quality.
- **Limited access to comprehensive and high-quality data** may constrain the depth of the analysis. Reliance on available datasets may introduce biases or gaps in understanding cancellations.
- **Subjectivity in Customer Feedback** Customer feedback regarding cancellations may be subjective and varied, making it challenging to generalize findings.
- **Interpretation of qualitative data** may be influenced by individual perspectives.
- **Generalization of Findings**

Findings may be specific to the studied hotels and may not be fully representative of the entire hospitality industry. Generalizing regional or demographic variations may require a broader dataset.

CHAPTER 11
CONCLUSION

CONCLUSION

In conclusion, the analysis of hotel booking cancellations provided valuable insights into the reasons behind cancellations, identified patterns and trends, and developed strategies to minimize cancellations and improve customer satisfaction. By implementing the recommendations outlined above, hotels can reduce cancellations, increase customer loyalty, and improve their bottom line. Further research is needed to explore the generalizability of the findings to other hotels and customer demographics

Optimal Pricing Strategies

The analytical study of hotel booking cancellation using Python has provided valuable insights into the dynamics and drivers of cancellations in the hospitality industry. Through comprehensive data analysis, predictive modeling, and interpretation of results, several key findings and actionable recommendations have emerged, contributing to the enhancement of revenue management, operational efficiency, and customer satisfaction. In conclusion, the analytical study of hotel booking cancellation using Python serves as a strategic roadmap for hoteliers seeking to navigate the complexities of revenue management, customer engagement, and operational optimization in the hospitality industry. By leveraging data-driven insights and implementing proactive strategies, hotels can mitigate the adverse effects of cancellations, optimize revenue streams, and enhance guest satisfaction, positioning themselves for sustained success in an increasingly competitive marketplace.

Recommendations on pricing strategies based on the relationship between room rates and cancellation rates, aiming to maximize revenue while minimizing cancellations. Tailored Policies Suggestions for refining cancellation policies to strike a balance between flexibility for guests and protection of hotel revenue. Targeted Marketing Guidance on targeted marketing efforts towards specific demographic segments that are less likely to cancel, optimizing customer acquisition costs.

Enhanced Customer Service Emphasis on the importance of customer service in reducing cancellations, including personalized interactions to address guest concerns and improve satisfaction. Adaptation to Seasonal Trends Strategies for hotels to adapt their operations and marketing efforts based on seasonal trends in booking cancellations. Proactive Response to External Factors Recommendations for hotels to develop contingency plans and communication strategies in response to external factors influencing booking cancellations. Continuous Monitoring and Analysis The importance of continuous monitoring and analysis of booking data to stay agile and responsive to changing market conditions.

CHAPTER 12

LEARNING OF STUDENT THROUGH THE PROJECT

LEARNING OF STUDENT THROUGH THE PROJECT

Data Handling Skills Students learn how to collect, clean, and preprocess real-world datasets containing hotel booking information. They gain hands-on experience with Python libraries such as Pandas and NumPy for data manipulation and cleaning.

Exploratory Data Analysis (EDA) Students learn the importance of exploratory data analysis in uncovering patterns, trends, and correlations within datasets. They explore visualization techniques using libraries like Matplotlib and Seaborn to gain insights into cancellation behavior and related factors.

Machine Learning Algorithms Students are introduced to various machine learning algorithms, including logistic regression, decision trees, random forests, and gradient boosting. They learn how to implement these algorithms using Python's Scikit-learn library and evaluate model performance using metrics such as accuracy, precision, recall.

Predictive Modeling Through the project, students understand the process of building predictive models to forecast hotel booking cancellations based on historical data. They gain insight into feature selection, model training, hyperparameter tuning, and model evaluation techniques.

Interpretation of Results: Students learn how to interpret model results and derive actionable insights from predictive models. They analyze the significance of different factors influencing cancellation behavior and understand the implications for revenue management and customer satisfaction.

Domain Knowledge in Hospitality Industry: The project provides students with exposure to the hospitality industry and its unique challenges related to revenue management, customer engagement, and operational optimization. They gain a deeper understanding of how data analytics can be applied to address these challenges effectively.

Practical Application of Python Programming Students enhance their programming skills in Python by working on a real-world project with practical applications. They learn to write clean, efficient code for data analysis, modeling, and visualization tasks. Critical Thinking and Problem-Solving Students develop critical thinking skills as they analyze complex datasets, identify patterns, and propose solutions to mitigate booking cancellations. They learn to approach problems systematically and adapt their strategies based on data-driven insights.

Project Management and Collaboration Students learn project management skills by planning, executing, and documenting the various stages of the analytical study. They also have the opportunity to collaborate with peers, seek feedback, and iterate on their solutions. Overall, the project offers students a holistic learning experience that combines technical skills with domain knowledge, critical thinking, and problem-solving abilities. It prepares them for careers in data science, analytics, and related fields while providing practical insights into the application of Python programming in real-world scenarios.

CHAPTER 13
BIBLIOGRAPHY

BIBLIOGRAPHY:

1. Sigala, M., Christou, E., & Gretzel, U. (2012). Social media in travel, tourism and hospitality: Theory, practice and cases. Ashgate Publishing, Ltd.
2. Tanford, S., & Jung, J. (2019). Understanding the reasons for cancellation of hotel reservations: The role of transparency and online reviews. *Journal of Hospitality Marketing & Management*, 28(7), 772-790.
3. Gao, Y., & Mattila, A. S. (2020). Understanding hotel booking cancellation behavior: An integrated conceptual framework. *Journal of Travel Research*, 59(6), 1086-1100.
4. Huang, Y. (2018). The impact of hotel booking cancellation policies on guests' booking decisions. *Journal of Hospitality Marketing & Management*, 27(4), 407-421.
5. Kim, D., & Wang, D. (2019). Predicting hotel booking cancellation using machine learning techniques. *Journal of Hospitality Marketing & Management*, 28(6), 653-668.
6. Liu, X., Ye, Q., & Law, R. (2019). A comparative study of online hotel booking intentions: A Chinese perspective. *Journal of Travel Research*, 58(7), 1163-1177.

CHAPTER 14
QUESTIONNAIRE

QUESTIONNAIRE

1. What was the primary reason for cancelling your hotel booking?

- a) Price was too high
- b) Found a better deal elsewhere
- c) Travel plans changed
- d) Unexpected personal circumstances

2. Hotel did not meet expectations (please specify in comments)

- a) Other (please specify in comments)

3. How far in advance did you cancel your booking?

- a) Less than 24 hours
- b) 1-3 days
- c) 4-7 days
- d) 8-14 days
- e) More than 14 days

4. On a scale of 1 (very unlikely) to 5 (very likely), how likely are you to book with this hotel again in the future?

- a) 1
- b) 2
- c) 3
- d) 4
- e) 5

5. If you booked through an online travel agency (OTA), what was the main reason for cancelling through them instead of directly with the hotel?

- a) OTA cancellation policy was more favourable
- b) Difficulty contacting the hotel
- c) Better customer service experience with OTA
- d) Other (please specify in comments)

6. If you booked directly with the hotel, what was the main reason for not cancelling through their website or app?

- a) Preferred to speak with a representative directly
- b) Cancellation process on website/app was unclear
- c) Technical difficulties with website/app
- d) Other (please specify in comments)

7. What is your age range?

- a) 18-24
- b) 25-34
- c) 35-44
- d) 45-54
- e) 55+

8. What is your primary travel purpose for this booking?

- a) Leisure
- b) Business
- c) Visiting friends and family
- d) Other (please specify in comments)

9. How often do you travel for leisure per year?

- a) Never
- b) 1-2 times
- c) 3-4 times
- d) 5+ times

10. Were you offered any incentives to rebook your stay after cancelling?

- a) Yes
- b) No

11. If yes, did the incentives influence your decision to rebook?

- a) Yes
- b) No

