**CCT College Dublin**

**Assessment Cover Page**

|  |  |
| --- | --- |
| **Module Title:** | Strategic Thinking/ Higher Diploma Data Analytics for Business |
| **Assessment Title:** | Individual / Practical |
| **Lecturer Name:** | James Garza |
| **Student Full Name:** | Miroslava Slavikova |
| **Student Number:** | sba22382 |
| **Assessment Due Date:** | 15 Dec 2023 |
| **Date of Submission:** |  |

**Declaration**

|  |
| --- |
| By submitting this assessment, I confirm that I have read the CCT policy on Academic Misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source. I declare it to be my own work and that all material from third parties has been appropriately referenced. I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution. |

**Hospitality Management and importance of Data Analysis**

Subject area: Hospitality Management with focus on Revenue Management

Table of Contents

**Type chapter title (level 1)1**

Type chapter title (level 2)2

Type chapter title (level 3)3

**Type chapter title (level 1)4**

Type chapter title (level 2)5

Type chapter title (level 3)6

**Testing testing 1**

**Type chapter title (level 1)2**

**Type chapter title (level 1)1**

**Type chapter title (level 1)1**

Introduction

Hotel Revenue Management requires a solid understanding of business data, be it managing room rates, negotiating agreements about commissions through a variety of distribution channels, or scouting the competition’s room rates in order to remain competitive in the market. In today’s data driven world, it is increasingly important to track, collect and analyse hotel’s data to transform it into actionable results in order to target the audience effectively, to streamline operations and to retain competitive advantage.

Business Analysis and Project Plan

Historical data help us to predict future guests’ behaviour, predict demand and identify new trends. They also demonstrate cycles, seasonality, patterns and any anomalies that can impact the sales and future demand. Hotel data are typically gathered from a diverse source, including industry reports, market research, sales records, online analytics, and guests’ surveys.

We’re going to look at the historical data of the two different properties in Portugal, one being hotel is located in the city of Lisbon and the second property is a holiday resort in the region of Algarve and identify key areas of revenue enhancement.

Use of technologies

We have used two different supervised machine learning models, Random Forest and Linear regression, that may often use in regression and classification problems.

Models

We have used two different machine learning algorithms including linear regression and random forest algorithm. Both the models have focused on to resolve the regression and classification prediction of featured variable. We have done different steps to make the data useful. Such as removing unnecessary variables and making the dummy variables of symptoms using age of the patients. After encoding there were 4239 rows and 44 feature variables in this dataset.

Libraries

We have downloaded all the libraries, panda for data manipulation and analysis library, matplotlib.pyplot and seaborn for data visualisation, numpy for numerical computing, sklearn libraries for future scaling for ML models, PCA, encoding, training and testing sets, linear regression model for predictive modelling. We also included a Folium, very useful Python Library used for visualizing geospatial data. Libraries are now imported and have been assigned the abbreviated formats. The abbreviated format makes recalling and using these libraries more efficient. Lastly, we have uploaded a csv file that we renamed as a hotel.df for easy reference.

Dataset and Data Understanding

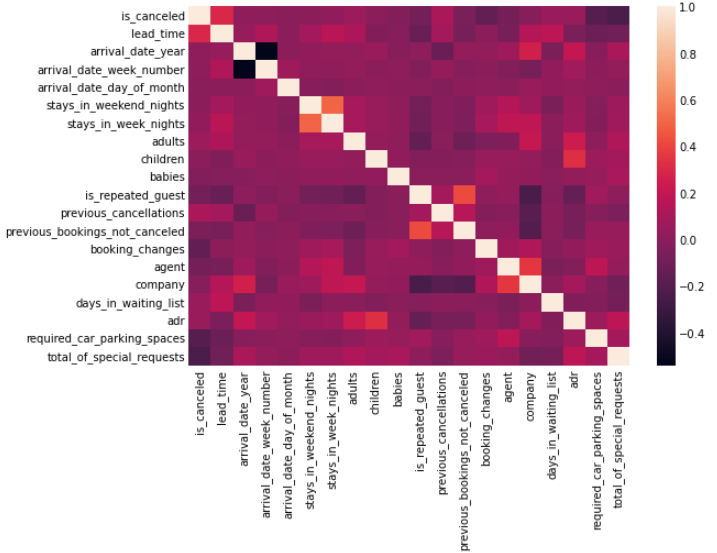
Data was acquired from Property Management Systems SQL databases. The dataset contains actual bookings due to arrive from 01st July 2015 to 31st Aug 2017 and includes bookings that have materialized and bookings that were cancelled. The data set has 119390 observations and 32 features, where each observation represents a hotel booking.

Data Prep

Data Vis

Heatmap

1. From above heatmap total special request highly correlated to adr, adults & arrival date year.
2. So total special request recieve with high adr,adults & arrival year.



When it comes to hotel revenue management, the following are essential business knowledge areas:

- Evaluate historical data to identify Average Daily Rate (ADR)

- Revenue per Available Room (RevPAR)

- % occupancy rates

- Direct bookings and distribution channels

- Meeting, Incentive, Conferences and Events (MICE) - average group sizes and even activity

arrival month

To predict data with the goal of maximising revenue, it is crucial to take few factors into consideration. Understanding the world’s holiday seasons, global conditions and special events in the region will help to understand when is an excellent opportunity to boost hotel rates because there’ll be a higher demand for the rooms due to the large number of reservations. for the off-peak season, when there’s a lower demand, hotel can create special deals to attract more traffic and foster consumer’s loyalty. creating a hotel package will additionally diversify the revenue streams by upselling a fine dining experience, spa treatments or variety upgrades, be it a better room or an executive floor usually reserved for premium guests.

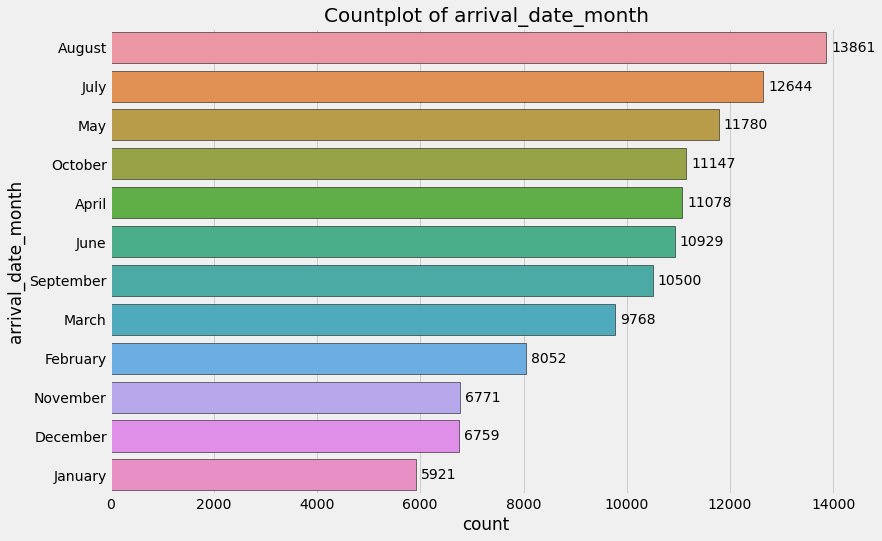


Figure x: Countplot of guest arrivals per calendar month

market segment  (online, groups)

hotel market segmentation will help us identify different categories of guest based on the travel and reservation patterns. it helps the hotel detect new business in certain areas and the hoteliers can target travellers with tailored marketing materials while focusing on the specific pricing plans.

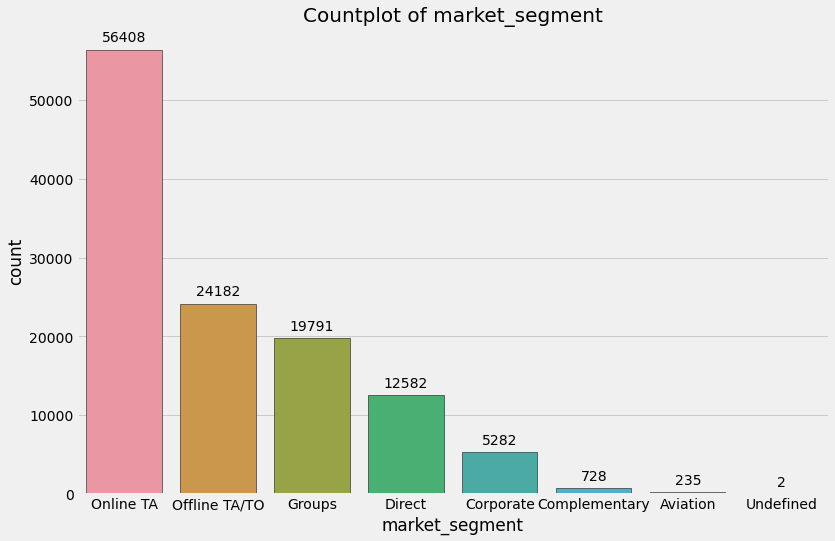


Figure x: Countplot of market segment

cancellations

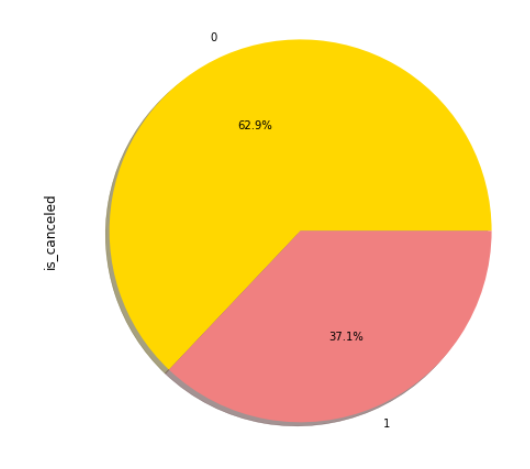
Cancellations directly impact revenue, particularly last minute cancellations lead to loss of revenue. Even, if the hotel is able to sell the room again, it’ll likely be offered at a lower room rate or it’ll be a loss-making entity for the day. The data analytics can help and understand the reasons for cancellations and how to avoid it, since it leads to extra work and preparations for guests arrival. There’d could be a genuine reasons for cancellations, such as event or conference is re-rescheduled, family holidays not materialised due to some personal emergency but it could also be fraudulent bookings that guest or agency make in order to obtain visa for the customer coming from a country with stringent visa policy into the respective destination. It is crucial for hotel

Figure % of room cancellations

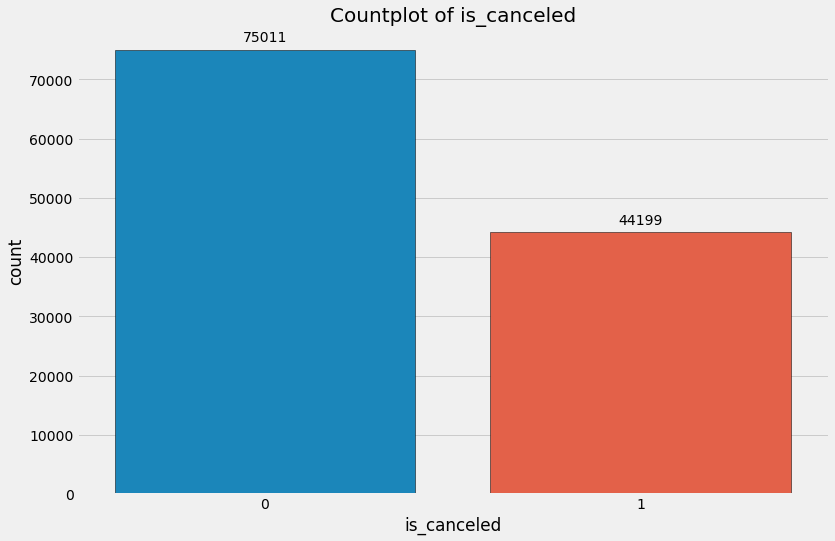
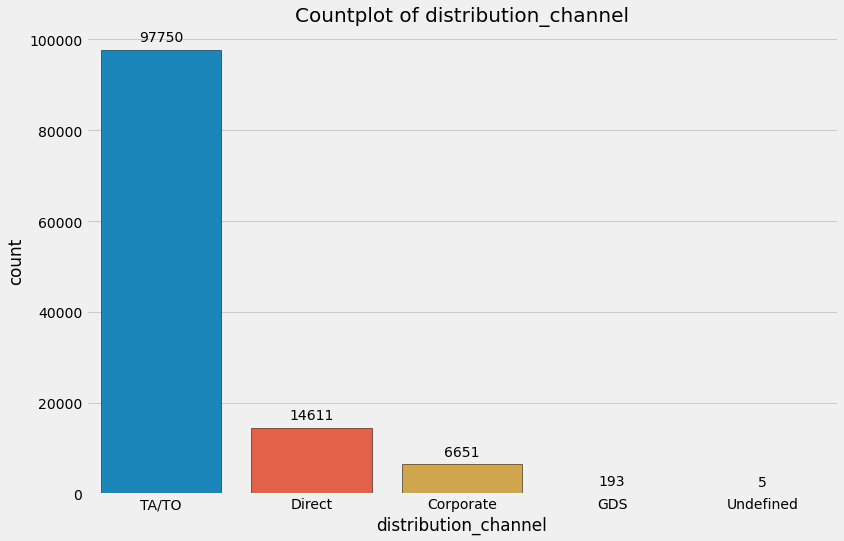


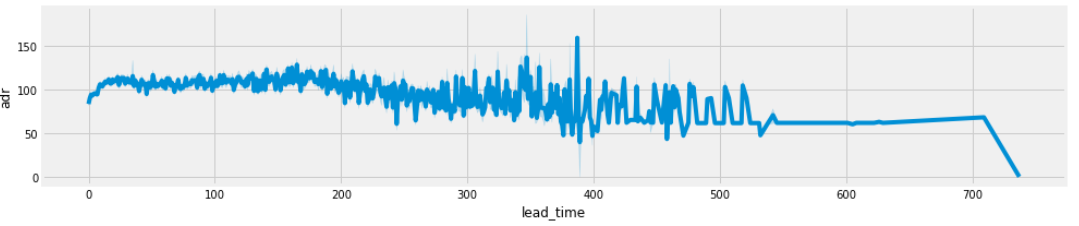
Figure Countplot of total cancellations

Distribution channels

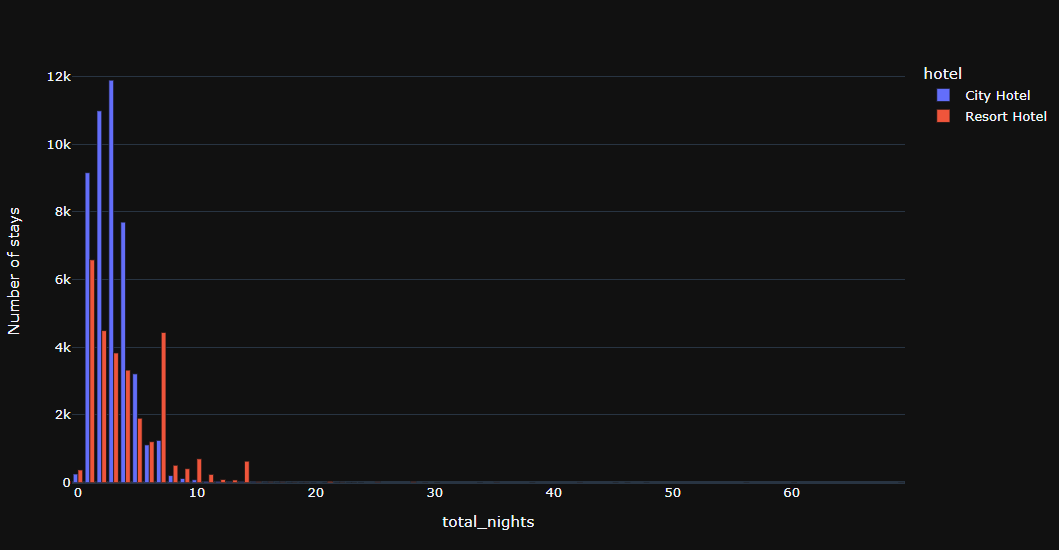


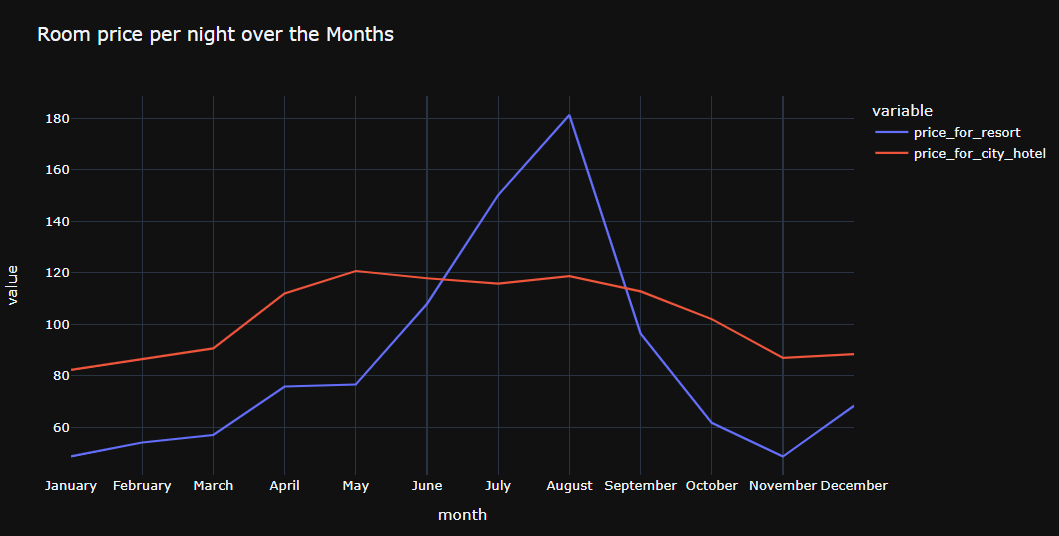
Figure

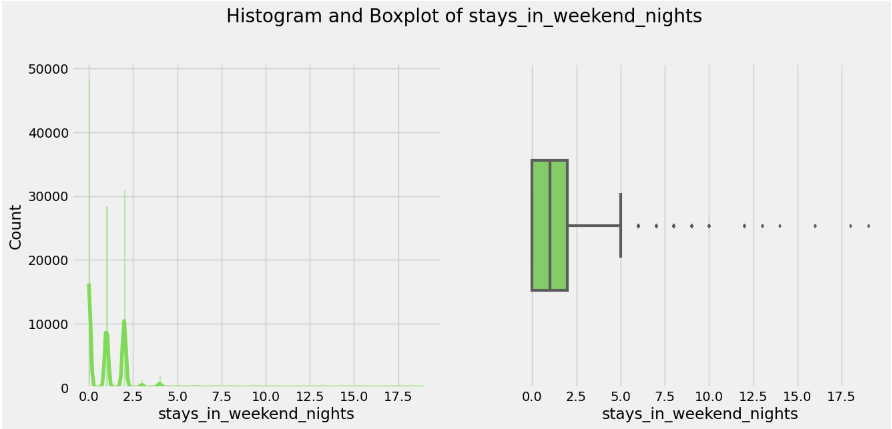
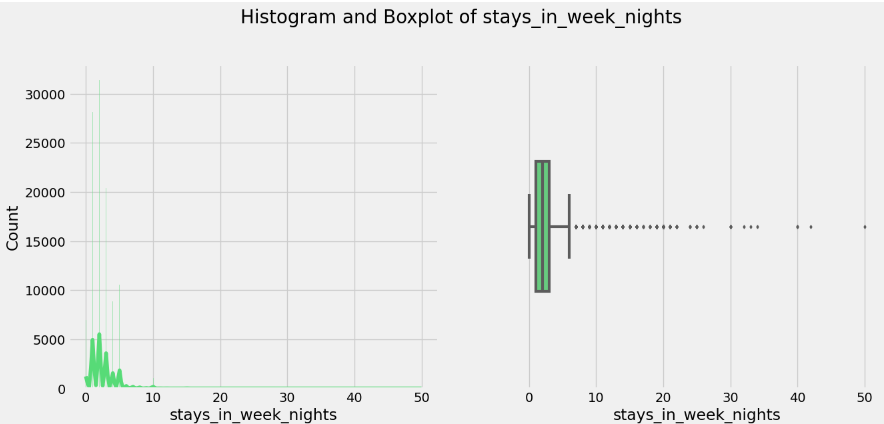
Prices of room per night**s**



Number of stays - city hotel more more nights.. resorts loaners







## Challenges encountered

Different challenges have been encountered in working with this dataset and applying models on it. These challenges may include the big size of dataset, categorizing the descriptive text of symptoms. There were a lot of missing values, cleaning and pre-processing of the data, sparse data after encoding, finding the perfect model that may give greater accuracy.

Challenges

* Cleaning
* Lots of missing values
* Identifying the target value - (ADR, vs is-cancelled)

Analysis of Results

These are my findings ..

Conclusion

Timeline

References

<https://www.sciencedirect.com/science/article/pii/S2352340918315191>

<https://www.kaggle.com/>

<https://towardsdatascience.com/how-to-clean-your-data-in-python-8f178638b98d>

<https://www.analyticsvidhya.com/blog/2020/06/guide-geospatial-analysis-folium-python/#:~:text=Folium%20is%20a%20Python%20library,library%20for%20plotting%20interactive%20maps>.

<https://realpython.com/>

<https://www.geeksforgeeks.org/>

* project management methodology to develop and execute a capstone project.
* You will select a dataset
* conduct exploratory data analysis
* pre-process the data
* implement at least one machine learning algorithm,
* and present your findings effectively through a comprehensive report

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

the project plan,

business understanding,

data understanding,

data preparation,

machine learning implementation,

and an artefact of code.

summary of the findings,

conclusions

any future recommendations.

**overview of the project's timeline, milestones achieved, and any challenges faced during the implementation phase, key insights gained from analysing the data and present any significant trends or patterns observed, the report should address any limitations or constraints encountered during the project and propose potential solutions for future improvements.**

Fig

1. City hotels are the most preferred hotel type by the guests. We can say City hotel is the busiest hotel.
2. 27.5 % bookings were got cancelled out of all the bookings.
3. Only 3.9 % people were revisited the hotels. Rest 96.1 % were new guests. Thus retention rate is low.
4. The percentage of 0 changes made in the booking was more than 82 %. Percentage of Single changes made was about 10%.
5. Most of the customers (91.6%) do not require car parking spaces.
6. 79.1 % bookings were made through TA/TO (travel agents/Tour operators).
7. BB( Bed & Breakfast) is the most preferred type of meal by the guests.
8. Maximum number of guests were from Portugal, i.e. more than 25000 guests.
9. Most of the bookings for City hotels and Resort hotel were happened in 2016.
10. Average ADR for city hotel is high as compared to resort hotels. These City hotels are generating more revenue than the resort hotels.
11. Booking cancellation rate is high for City hotels which almost 30 %.
12. Average lead time for resort hotel is high.
13. Waiting time period for City hotel is high as compared to resort hotels. That means city hotels are much busier than Resort hotels.
14. Resort hotels have the most repeated guests.
15. Optimal stay in both the type hotel is less than 7 days. Usually people stay for a week.
16. Almost 19 % people did not cancel their bookings even after not getting the same room which they reserved while booking hotel. Only 2.5 % people cancelled the booking.

Conclusion: Create flexible price plans depending on competition analysis, seasonality, and demand.

Strategically use promotions and discounts to increase demand.

Working together and communicating:

Encourage cooperation between the marketing, sales, and revenue management departments.

Align price tactics with marketing initiatives and sales targets.

Integrating operations:

Make sure that decisions about pricing and inventory are in line with the hotel's capacity and service capabilities by closely collaborating with operations.

Loop of Feedback and Adjustment:

Create a routine evaluation procedure to evaluate how well revenue management tactics are working.

Pricing and distribution strategies can be iteratively adjusted by using feedback.

Instruction and Growth:

Staff members should receive training on new tools and technology.

Ascertain that the group is prepared to adjust to modifications in price and distribution tactics

Arrival per month

Market segment  (online, Groups)

Disctributions channels

Cancelatoins

To predict data with the goal of maximizing revenue, it is important to take special events, holiday seasons and global condition into consideration.

A pair plot of the first ten variables facilitates a comprehensive exploration of their interdependencies. It reveals patterns, trends, and potential outliers through scatterplots and helps assess correlations. This visual inspection aids in understanding the data's structure, guiding further analysis, and informing data-driven decision-making in diverse applications.