CS 571 Final Project Deliverable #2

Dhruvi Chauhan, Anushka Neerav Doshi, Genglin Liu, Vinitha Maheswaran

Introduction

For this project, we are making medium-fidelity prototypes for the hotel booking demand visualization dashboard.

Intended Users

We believe that this visualization dashboard could provide useful information for a traveler who is looking for a good time of the year to book a hotel in the city or a resort. The user may be interested in the price of the hotel booking since no one would like to pay for overpriced stays. This visualization could also be useful for hotel management. As they observe the booking demand across a two-year span, they could analyze the seasonality of the demand, forecast the future booking trend, and adjust their pricing and marketing strategies accordingly.

Dashboard Type

In class, we have discussed two kinds of dashboard visualizations, for data monitoring and data exploration respectively. For our visualization project, different users of various interests would be working with the hotel booking demand dataset, and more interactivity is warranted. Therefore, a dashboard for data exploration appears more suitable for our tasks.

Features

We would like to enable the users to be able to inspect the hotel booking demand and the rate (price) of the booking on an interactive scale. For this project, we are making medium-fidelity prototypes for the hotel booking demand visualization dashboard.

Hotel types

We would like to enable the users to be able to switch between the two hotels, `city` and `resort`. For both of the hotels, we will provide the same set of information and data visualization so a user could compare them directly. For now, the two quantities that we are interested in are the `demand` and `booking rate`.

Interactive Scale

For each hotel, we would let the user decide which time scale they want to view. For example, the "all-time" yearly scale contains data from July 2015 to July 2017; and a "monthly" scale would provide demand data for each month in a year. If the user is interested in an even more refined scale, they can choose the "daily" scale and inspect the specific hotel booking information for each day in a given month. The data will be represented in a mix of histogram and line plots in the prototype.

A naive implementation of the interactive scale could consist of a simple drop-down menu that prompts the user to select a specific scale and then display it accordingly. Or we could let the user select their preferred view in a form and then update the visualization accordingly.

Booking Rate (time permitting)

If we have enough resources to visualize more data, we could also show the trend of hotel booking prices (directly linked to hotel revenue) along the number of guests (booking demand)

Data Extraction

For our interests, we would like to extract a few specific columns from the dataset that will serve for our visualizations.

In short, demand data comes from the following columns:

- adults
- children
- babies

For each booking (row in the dataset), these three attributes are given, and we can sum up the numbers to get the `total number of guests` in a given booking.

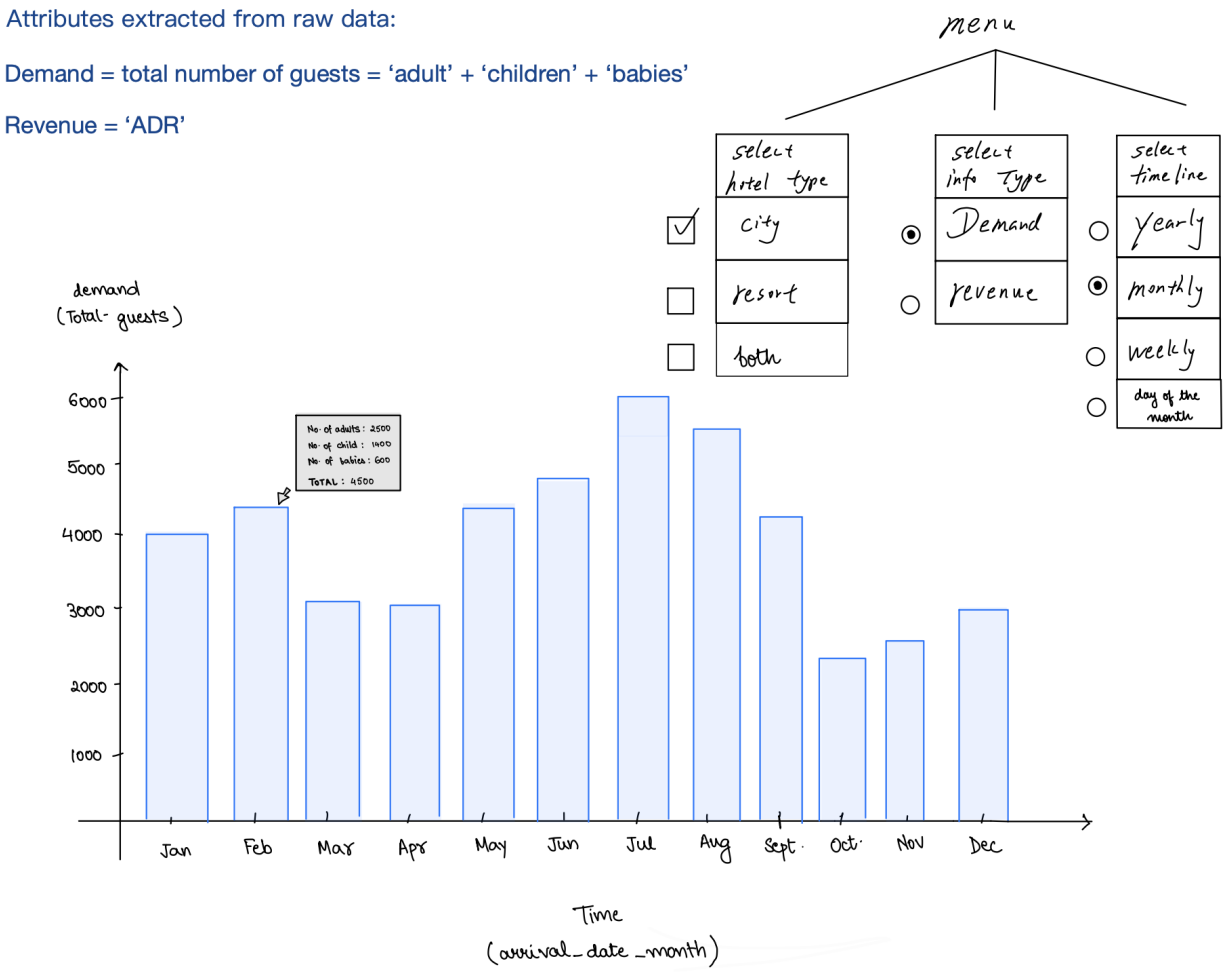
For the `rate`, the attribute of our interest is `ADR`, or average daily rate. This attribute tells how much a given booking costs per day.

If we are visualizing the demand or hotel booking revenue from a larger scale, then we would not have enough space on the screen to include every single data point. In that case, we will aggregate the corresponding values across different rows grouping them by

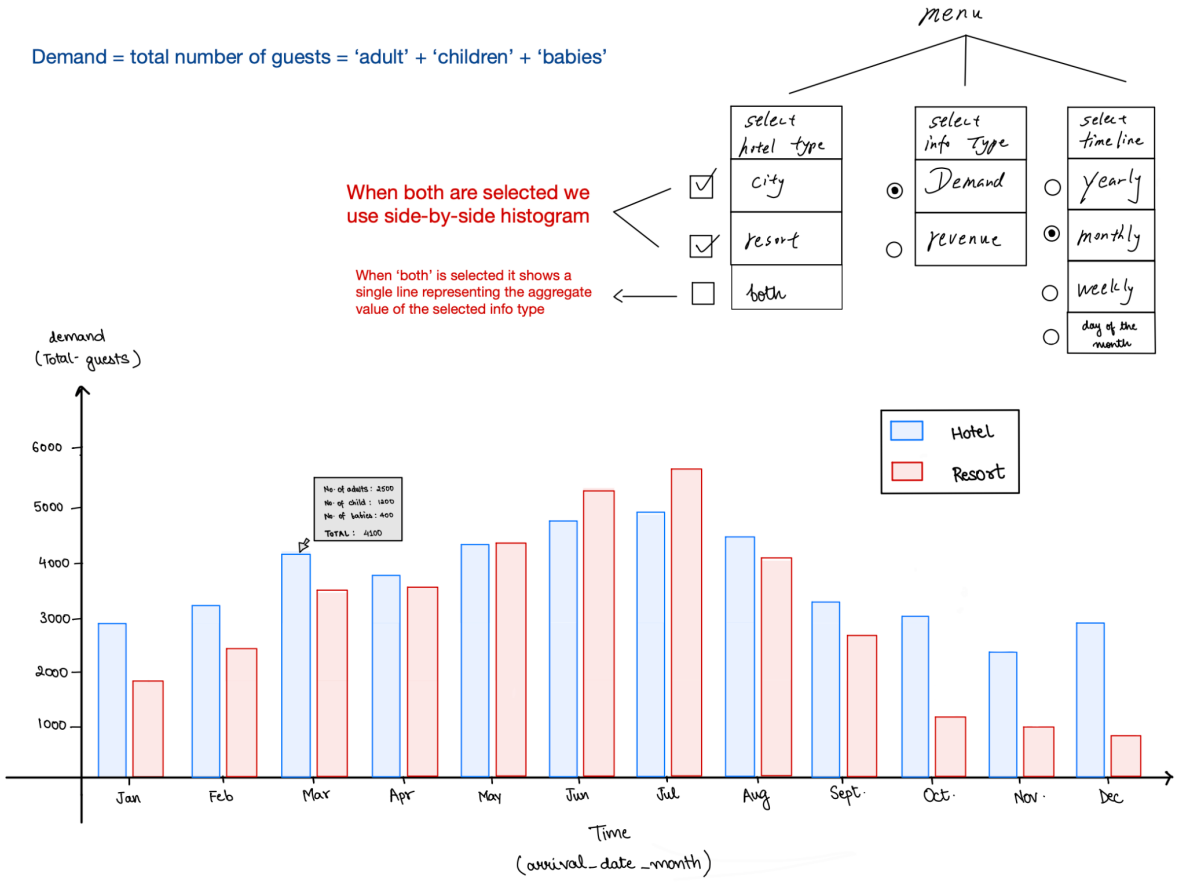
the arrival week/month/year, etc. We will try to implement the data manipulation in D3 as well.

This preprocessing step could also be done in Python or R, and ultimately we need to provide a clean `.csv` file that contains all the data that we need for our dashboard, and then read and visualize the data in Javascript D3.

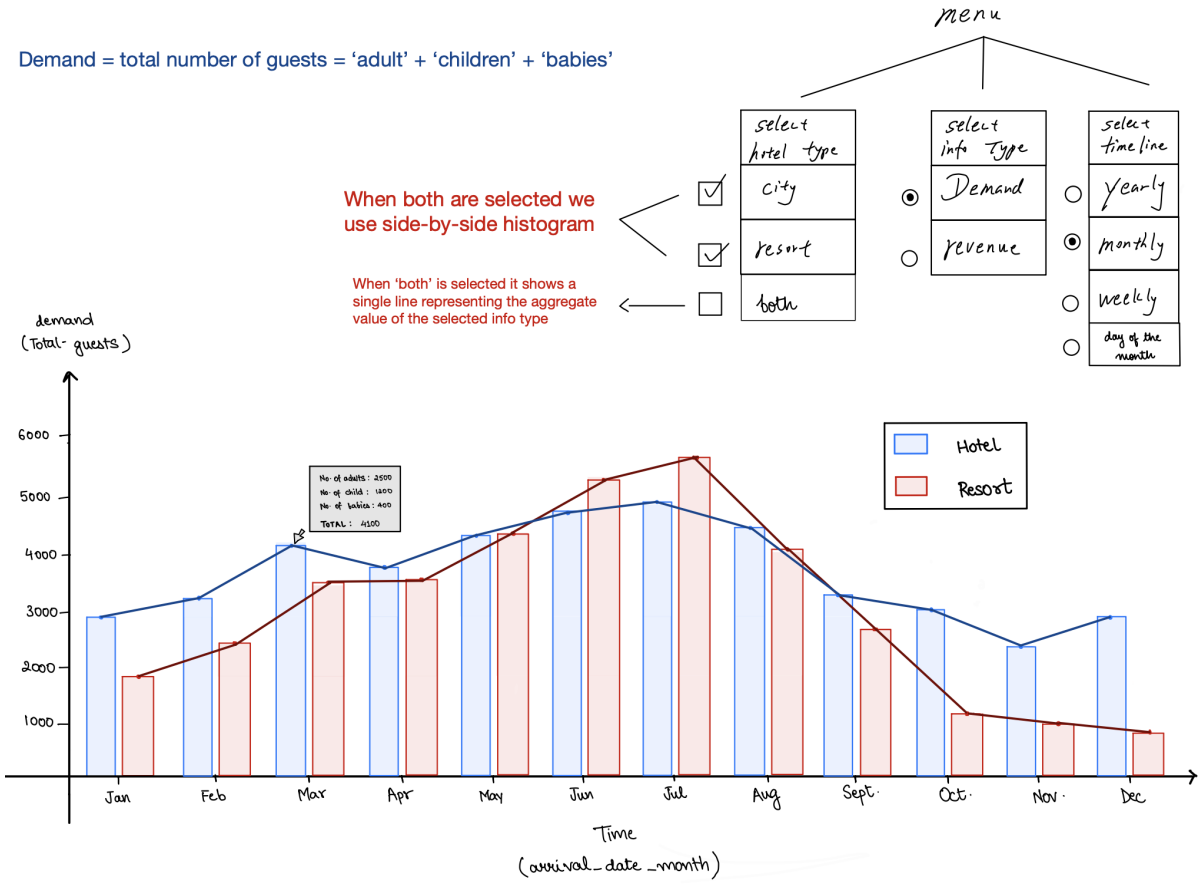
Rough sketches of what we intend the visualizations to look like:



Demand = total number of guests = 'adult' + 'children' + 'babies'



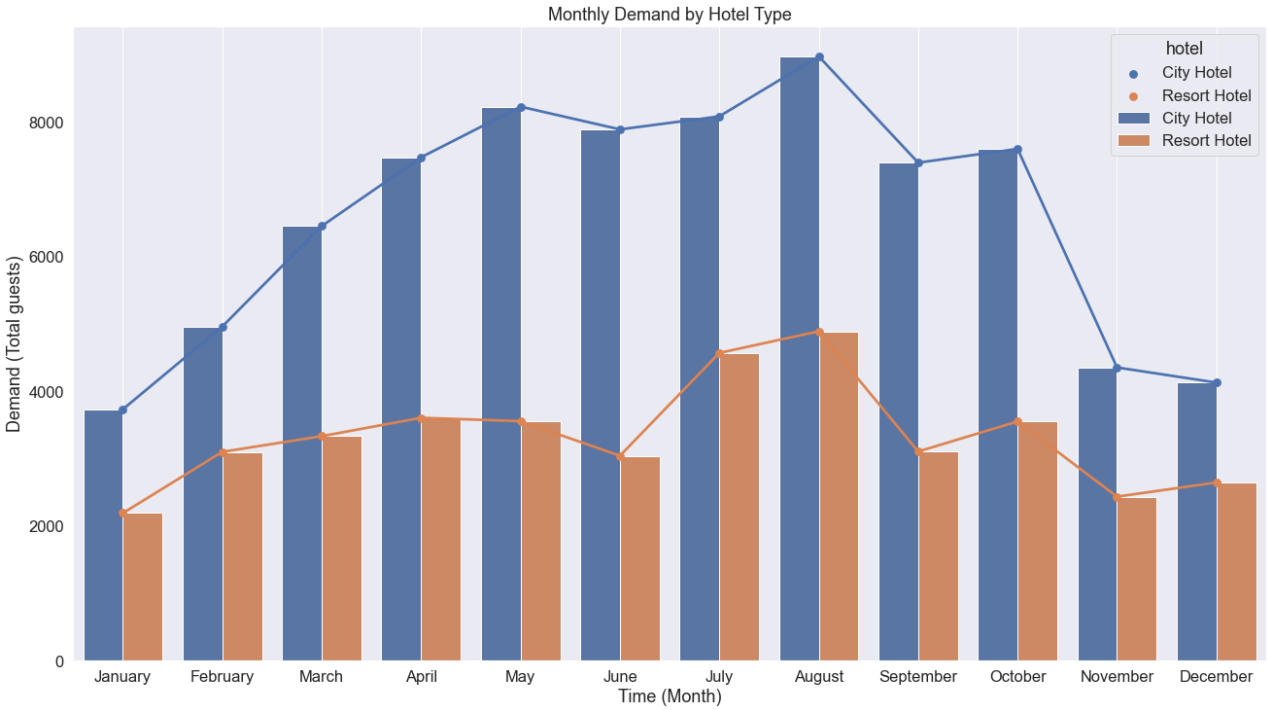
Demand = total number of guests = 'adult' + 'children' + 'babies'



Some visualizations generated by implementing python (seaborn and matplotlib) on the Hotel Demand dataset:

A.

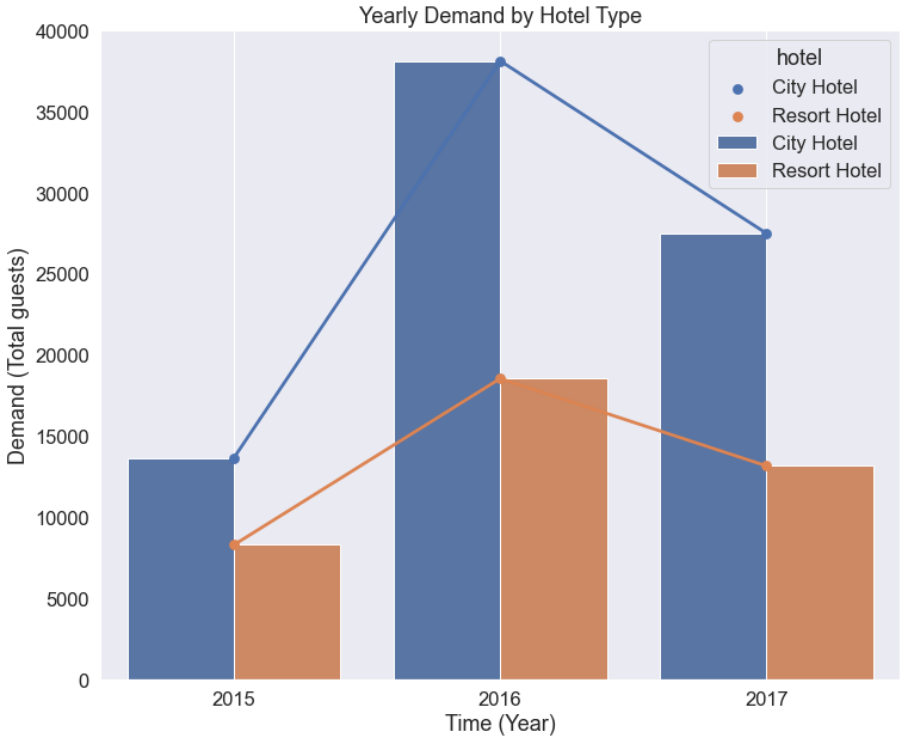
The following plot will be shown on the dashboard when the user checks the menu options:
1) select hotel type: city & resort
2) select information type: demand
3) select timeline: monthly



The visualization above shows the trend in the demand for the city hotel and resort hotel for every month. Both types of hotels are represented in different colors (blue for city and orange for resort). The bar chart allows the user to compare the demand for each hotel every month, and the line plot helps the user see the trend in demand for each hotel over time. The demand for both hotels was highest in August.

B.

The following plot will be shown on the dashboard when the user checks the menu options:
1) select hotel type: city & resort
2) select information type: demand
3) select timeline: yearly

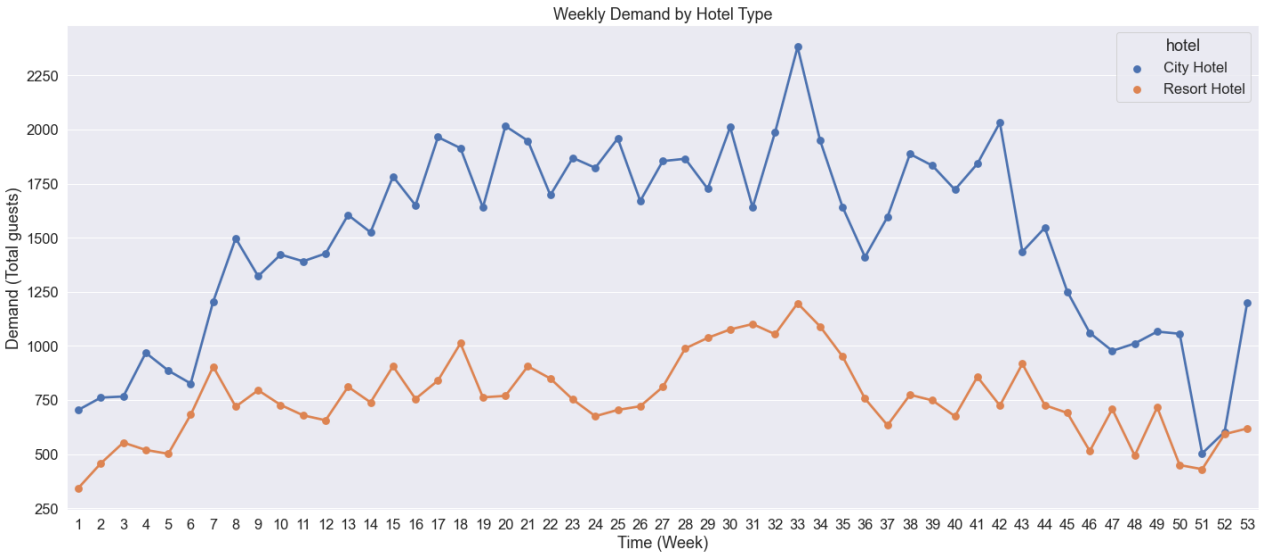


This plot represents the yearly demand for each hotel. The year 2016 had the highest demand and 2015 had the lowest demand for both hotels. The demand for the city hotel was consistently higher than that for the resort hotel.

C.

The following plot will be shown on the dashboard when the user checks the menu options:

1) select hotel type: city & resort
2) select information type: demand
3) select timeline: weekly

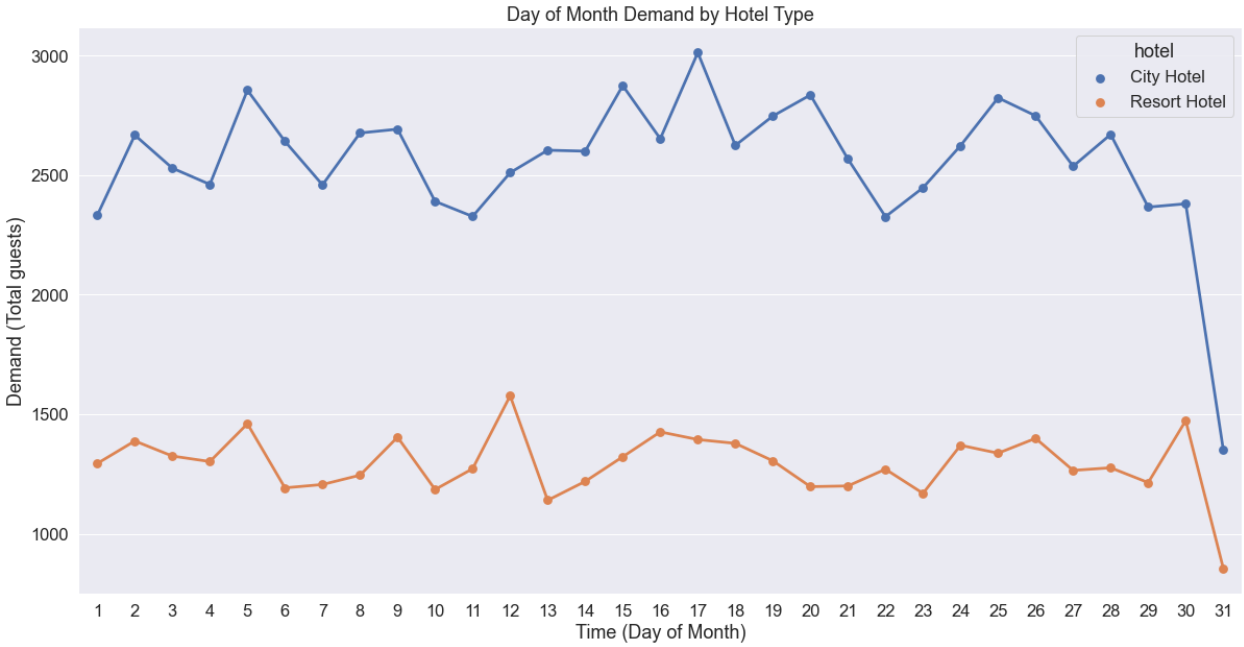


This line chart compares the weekly demand for each hotel. Every week, the demand for city hotels was higher compared to resort hotels.

D.

The following plot will be shown on the dashboard when the user checks the menu options:

1) select hotel type: city & resort
2) select information type: demand
3) select timeline: day of month

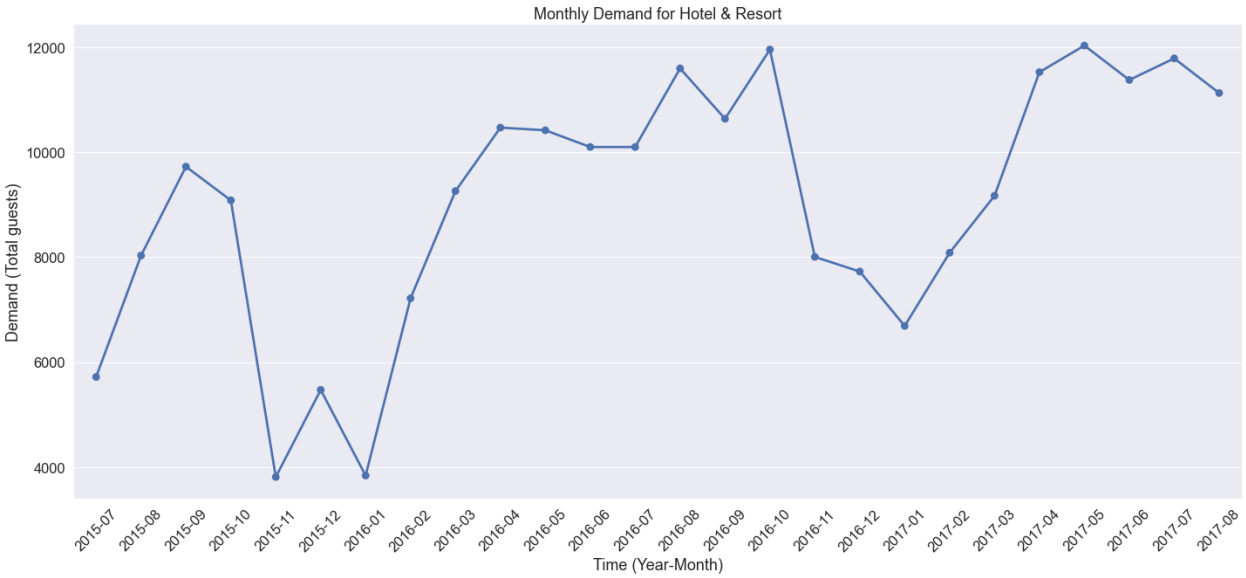


These line charts compare the demand for each hotel on the days of month for the years 2015 to 2017. From this line chart we can observe that the highest demand is during the middle of the month and lowest during the end of the month.

E.

The following plot will be shown on the dashboard when the user checks the menu options:

1) select hotel type: both
2) select information type: demand
3) select timeline: monthly



This visualization shows the total monthly demand for both hotels combined, for every month over a period of three years. This aggregate value representing the total demand for both hotels would be useful for tourist/hospitality services and local government/independent data analysts to understand the demand for hotels in general.

Dashboard Usage

This dashboard will enable the user to achieve their task by

- Displaying the trend of hotel booking demand (and revenue) in a defined time period with a bar chart and/or line plot.
- Supports side-by-side visualization so that a user can compare the booking demand from both city and resort hotels easily.
- The dashboard shows more detailed data when the cursor hovers over the bars, making the visualization more informative.
- User: Tourists
    - The above plots will prove significant to a tourist in planning trips and making decisions about when to start booking their reservations because if it's a popular month where the demand is high, it's likely the rooms will be booked in no time.
- User: Hotel Management team
    - This dashboard will enable the hotel management team to achieve their task by displaying the demand and revenue trends given a specific timeline (yearly, monthly, weekly) so as to understand their own and the competitor's standings in the market.
    - The information about the number of adults, children, and babies will help the team understand the demographics of their bookings so they can arrange for more amenities that will appeal to that demographic.

NOTE: The sketches shown in this document are different frames of the same plot based on the selections made in the menu. We picked this set-up for our visualization because it targets different tasks from the many users we described in deliverable 1 by taking into account the same attributes from the dataset. It is also worth mentioning that there will be other visualizations on our dashboard which aren't mentioned in this document.