Take-Home Exercise

Pace Revenue Science Team

October 20, 2022

Instructions

In addition to this file, you should have received several others containing the data to analyse:

- `hotel_A_bookings.csv`, containing historical booking information for a single hotel. This is used in questions 1, 2, and 5.

- `competitor_prices.csv`, containing prices for five hotels at a single point in time for one year. This is used in question 4.

- `competitor_bookings.csv`, containing historical bookings for the same hotels as in `competitor_prices.csv` over the same period as in `hotel_A_bookings.csv`. This is used in question 5.

- `capacities.csv`, containing the number of rooms of each type for the hotels considered above. This information is available for all questions, if you wish to use it.

Please attempt all questions. This exercise should take about 8 hours. Each question has a time to help guide your work, but these should be considered suggestions only. The questions here do not generally have a single "right" answer. Rather, we want to see how you approach analysing the complex data set we provide, what conclusions you can draw, and how you test those conclusions.

1. *(1 hour)* This question refers to the data in file `hotel_A_bookings.csv`. Compute the final number of bookings for each night in the data. Describe any patterns you observe in the results. How might final bookings

be related to demand? What might that tell us about general pricing strategies?

2. *(2 hours 30 minutes)* This uses the same data as the previous question. We want a simple model to describe the arrival over time of bookings for a night. To do this, convert each booking time into a lead time: the difference between the booking time and the end of the night for which the booking was made. Plot the fraction of final bookings as a function of lead time. Can you make any other observations about the data?

3. *(30 minutes)* Briefly outline how a prediction of the fraction of final bookings as a function of lead time might be used in a pricing strategy.

4. *(1 hour 30 minutes)* In file `competitor_prices.csv` we have prices at one point in time as a function of date for a number of hotels. Assume that these hotels are in the same city, competing with each other. What features can you see in this data? How can you relate it to your results from question 1?

5. *(2 hours 30 minutes)* In file `competitor_bookings.csv` we have bookings for the same hotels considered in question 4, over the same period as the bookings for quesiton 1. Based on these bookings, which hotels do you think are most/least similar to hotel A? Why? Using these similarities, construct a predictive model for either bookings or price (or both) for hotel A, given the data from its competitors