

Case Study

Business question

Having observed a high rate of customers churning (i.e. failing to rebook again within 6 months from their last booking), Hcom.com leadership team would like to have a better understanding on which are the factors which contributes most to a churning behaviour and implement some actions to reduce customer attrition.

Task

Using the data provided can you:

1. Identify the importance of the factors provided with regards to churn rate behaviour, ideally highlighting which of the drivers contribute more than the others
2. Build a statistical model to identify which customers are more vs. less likely to churn
3. Summarize your insights in a Power Point presentation aimed at the Hotel.com leadership team, including some recommendation on how they can use the results retrieved in the previous two points to reduce their churn rate

Case study submission:

At the time of submission, please provide the presentation and all the supporting materials created as part of the case study (e.g. programming code, ...). The PowerPoint slides should be targeted to a senior executive audience for a 20-minute presentation.

Data provided - PIP_case_study_data.csv

The .csv table contains (fictional) historic customer information and attributes related to their past purchasing behaviour.

- Each record is at booking level, so there might be multiple rows related to the same customers
- For each record, selected information (i.e. factors) related to the customer and booking are reported, along with a flag that identifies whether the customer churned.

Below is the detailed column explanation:

(*For confidentiality reasons, the data provided have been simulated and the email address and booking id provided has been encrypted. While representative of customer behaviour recorded, Hotel.com trading performance cannot be inferred)

Column name	Notes
email_address	Unique id for each customer
booking_id	Unique id for each booking
bk_date	Booking date, format YYYY-MM-DD
coupon_flag	1 = use coupons when booking

pay_now_flag	1 = payment choice 'pay now', 0 = 'pay later'
cancel_flag	1 = canceled
cancel_date	Cancel date, format YYYY-MM-DD
customer_type	Indicates whether the customer has booked another accommodation through Hotels.com in the 3 months prior the booking in question (booking_id).
loyalty_tier	<p>Indicates customer's loyalty level (more info on this can be found here).</p> <ul style="list-style-type: none"> • Loyalty_tier = 0 → Customer who made the booking didn't sign in to the loyalty program (= Not a member) • Loyalty_tier = 1 → Customer who made the booking was signed into the loyalty program and was base member (=Base member) • Loyalty_tier = 2 → Customer who made the booking was signed into the loyalty program and at the time of booking was a very loyal member (=Silver/Gold member)
platform	Indicates which platform the booking came (e.g. desktop, mobile website, ...)
marketing_channel	Indicates how the visit came to the website (e.g. Google, TripAdvisor, ...). <i>Direct</i> marketing channel identifies customers who reached the website by directly typing the www.hotels.com URL in the address bar
total_visit_minutes	Total minutes of visits to the website
total_visit_pages	# of total pages visited on the website
landing_pages_count	# of landing pages visited on the website
search_pages_count	# of search pages visited on the website
property_pages_count	# of property pages visited on the website
bkg_confirmation_pages_count	# of booking confirmation pages visited on the website
bounce_visits_count	Total number of single page visits to the website
searched_destinations_count	Total number of different destinations searched on the website
hotel_star_rating	Star rating score of the booked hotel
churn_flag	1 = No repeat booking 6 months after check out