NOVA IMS
Information Management School

4

BUSINESS CASE 1: HOTEL CUSTOMER SEGMENTATION

Business Cases with Data Science

Acreditações e Certificações

Instituto Superior de Estatística e Gestão da Informação
Universidade Nova de Lisboa

Summary

1. Business situation
2. Key problems
3. Data
4. Assignment

NOVA IMS Information Management School

4.1

Business situation

Case 1: Customer segmentation

Business situation (1/3)

Finding new customers is vital in every industry.

The process of finding new customers begins by learning as much as possible from the existing customers.

Understanding current customers allow organizations to identify groups of customers that have different product interests, different market participation, or different response to marketing efforts.

Business situation (2/3)

Market segmentation, the process of identifying customers' groups, makes use of geographic, demographic, psychographic, and behavioral characteristics of customers.

By understanding the differences between the different segments, organizations can make better strategic choices about opportunities, product definition, positioning, promotions, pricing, and target marketing (1, 2).

Business situation (3/3)

Like most hotels, hotel H, a hotel located in Lisbon, Portugal, a member of the independent hotel chain C, uses a hospitality standard market segmentation based on the origin of the customer.

However, A, the new marketing manager of hotel H, recognized that this type of segmentation, as is today well-known (3, 4), is not useful for the hotel marketing department.

The name of the individual and the company name is anonymized to protect confidentiality. The referenced data are real.

NOVA IMS Information Management School

4.2

Key problems

Case 1: Customer segmentation

Key problems (1/3)

Until 2015 hotel chain C operated 4 hotels, however, with the acquisition of new hotels, the hotel chain board decided to invest more in marketing.

However, it was not until 2018 that the hotel chain created a marketing department and hired a new marketing manager, A.

A realized that the current customer segmentation was not adequate, as it **only reflected one only customer characteristic**, its sales origin. It **did not reflect** geographic characteristics, such as the country of origin, demographic characteristics, such as age, or behavioral characteristics, such as the number of stays.

Key problems (2/3)

Without proper customer segmentation, it is difficult for A to define a strategy to reach new customers and to continue to captivate the current customers.

Taking into consideration the multiple distribution channels that hotels operate nowadays (travel agencies, travel operators, online travel agencies – OTA, brand websites, meta searchers websites, among others). For example, corporate customers tend to make reservations very near the arrival date, book directly with the hotel, and be willing to pay more for a better-equipped room, while a customer on holiday tends to make reservations more distant from the arrival date, book with a travel operator or OTA, and to look for better price opportunities.

Key problems (3/3)

Therefore, products “creation”, pricing definitions, and other marketing tasks, such as advertising, must take into consideration the targets of its efforts according to the different channels and groups of customers.

NOVA IMS Information Management School

4.3

Data

Case 1: Customer segmentation

HASH Definition

A SHA-n (n=256, 512, etc.) hash is a cryptographic function that converts any length input data into a fixed-size n-bit string of characters.

Key characteristics:

1. Fixed output length: Regardless of the input size, it always generates a n-bit hash value
2. One-way function: It's computationally infeasible to reverse-engineer the original input from the hash value
3. Deterministic: The same input will always produce the same hash output
4. Avalanche effect: Even a tiny change in the input data results in a completely different hash value

Data (1/6)

The provided dataset is composed of the following columns:

- **ID**: Customer ID
- **Nationality**: Nationality of the customer in ISO 3166-1 (Alpha 3) format
- **Age**: The age of the customer
- **DaysSinceCreation**: Number of elapsed days since the customer was created
- **NameHash**: Hash of the customer's name
- **DocIDHash**: Hash of the customer's personal document identification number (usually a passport or ID card)
- **AverageLeadTime**: Average number of days before arrival date the customer makes bookings

Data (2/6)

- **LodgingRevenue**: Total amount of lodging revenue paid by the customer so far
- **OtherRevenue**: Total amount of other revenue (e.g., food & beverage, spa, etc.) paid by the customer so far
- **BookingsCanceled**: Number of bookings the customer made but subsequently canceled
- **BookingsNoShowed**: Number of bookings the customer made but subsequently made a "no-show"
- **BookingsCheckedin**: Number of bookings the customer made, which actually ended up staying

Data (3/6)

- **PersonNights**: Total person/nights the customer has stayed at the hotel so far. Persons/Nights are the sum of *Adults* and *Children* in each booking, multiplied by the number of *Nights* (Length-of-stay) of the booking
- **RoomNights**: Total of room/nights the customer has stayed at the hotel so far. Room/Nights are the multiplication of the number of rooms of each booking by the the number of *Nights* (Length-of-stay) of the booking
- **DistributionChannel**: Distribution channel normally used by the customer to make bookings at the hotel
- **MarketSegment**: Current market segment of the customer
- **SRHighFloor**: Indication if the customer usually asks for a room in a higher floor (0: No, 1: Yes)

Data (4/6)

- **SRLowFloor**: Indication if the customer usually asks for a room in a lower floor (0: No, 1: Yes)
- **SRAccessibleRoom**: Indication if the customer usually asks for an accessible room (0: No, 1: Yes)
- **SRMediumFloor**: Indication if the customer usually asks for a room in a middle floor (0: No, 1: Yes)
- **SRBathtub** Indication if the customer usually asks for a room with a bathtub (0: No, 1: Yes)
- **SRShower**: Indication if the customer usually asks for a room with a shower (0: No, 1: Yes)
- **SRCrib**: Indication if the customer usually asks for a crib (0: No, 1: Yes)

Data (5/6)

- **SRKingSizeBed**: Indication if the customer usually asks for a room with a king size bed (0: No, 1: Yes)
- **SRTwinBed** Indication if the customer usually asks for a room with a twin bed (0: No, 1: Yes)
- **SRNearElevator**: Indication if the customer usually asks for a room near the elevator (0: No, 1: Yes)
- **SRAwayFromElevator**: Indication if the customer usually asks for a room away from the elevator (0: No, 1: Yes)
- **SRNoAlcoholInMiniBar**: Indication if the customer usually asks for a room with no alcohol in the mini bar (0: No, 1: Yes)

Data (6/6)

SRQuietRoom: Indication if the customer usually asks for a room away from the noise (0: No, 1: Yes)

NOTE: All time-based columns (e.g., *Age* or *DaysSinceCreation*) were calculated at the dataset extraction date.

NOVA IMS Information Management School

4.4

Assignment

Case 1: Customer segmentation

Assignment

1. Explore the data and identify the variables that should be used to segment customers
2. Use a clustering algorithm to identify customer segments
    1. Justify your selection of the number of clusters (taking into consideration the business use)
3. Suggest business applications for the findings

Business Cases with Data Science

Acreditações e Certificações

Instituto Superior de Estatística e Gestão da Informação
Universidade Nova de Lisboa

References

1. MILLER, Thomas W. Marketing Data Science: Modeling Techniques in Predictive Analytics with R and Python. Old Tappan, NJ, USA : Pearson FT Press, 2015. ISBN 978-0-13-388655-9.
2. CHAPMAN, Chris. R for marketing research and analytics. New York, NY : Springer Science+Business Media, LLC, 2015. ISBN 978-3-319-14435-1.
3. RONDAN-CATALUÑA, Francisco Javier and ROSA-DIAZ, Isabel M. Segmenting hotel clients by pricing variables and value for money. Current Issues in Tourism. 2 January 2014. Vol. 17, no. 1, p. 60–71. DOI 10.1080/13683500.2012.718322.
4. GUILLET, B., GUO, Y., and LAW, R. Segmenting hotel customers based on rate fences through conjoint and cluster analysis. Journal of Travel and Tourism Marketing. August 2015, p. 835-851. DOI 10.1080/10548408.2015.1063825