

241127회의록

1. 데이터셋 소개

배경

호텔 객실을 예약하기에 가장 좋은 시기가 언제인지 궁금했던 적이 있나요? 혹은 일일 요금을 가장 저렴하게 얻기 위한 최적의 숙박 기간이 얼마나 되는지 알고 싶으신가요? 아니면, 특정 호텔이 비정상적으로 많은 특별 요청을 받을 가능성을 예측하고 싶으신가요?

이 호텔 예약 데이터셋은 이러한 질문들을 탐구하는 데 도움을 줄 수 있습니다!

내용

이 데이터셋은 도시 호텔과 리조트 호텔의 예약 정보를 포함하고 있으며, 다음과 같은 다양한 정보를 제공합니다:

- 예약이 이루어진 시점
- 숙박 기간
- 성인, 어린이, 유아의 수
- 이용 가능한 주차 공간의 수
- 그 외 기타 정보

모든 개인 식별 정보는 데이터에서 제거되었습니다.

출처

이 데이터는 **Nuno Antonio, Ana Almeida, Luis Nunes**가 작성한 논문 "*Hotel Booking Demand Datasets*" (2019년 2월, *Data in Brief*, Volume 22)에서 처음 제공되었습니다.

이후, **Thomas Mock**와 **Antoine Bichat**가 2020년 2월 11일 주간 *#TidyTuesday* 이벤트를 위해 데이터를 다운로드하고 정리하였습니다.

컬럼명

- hotel
호텔 유형 (H1 = 리조트 호텔, H2 = 시티 호텔)
- is_canceled
예약 취소 여부를 나타내는 값 (1 = 취소됨, 0 = 취소되지 않음)
- lead_time
예약이 호텔 관리 시스템(PMS)에 입력된 날짜와 도착일 사이의 경과 일수
- arrival_date_year
도착 날짜의 연도
- arrival_date_month
도착 날짜의 월
- arrival_date_week_number
도착 날짜가 속한 연도의 주 번호
- arrival_date_day_of_month
도착 날짜의 일
- stays_in_weekend_nights
투숙객이 호텔에서 숙박한 주말(토요일 또는 일요일) 밤의 수
- adults
성인 투숙객 수

hotel_bookings.csv (16.86 MB)

Detail

Compact

Column

Δ hotel	# is_canceled	# lead_time	# arrival_da...	Δ arrival_da...	# arrival_da...	# arrival_da...	# stays_in_...	# stays_in_...	# adults
Resort Hotel	0	342	2015	July	27	1	0	0	2
Resort Hotel	0	737	2015	July	27	1	0	0	2
Resort Hotel	0	7	2015	July	27	1	0	1	1
Resort Hotel	0	13	2015	July	27	1	0	1	1
Resort Hotel	0	14	2015	July	27	1	0	2	2

호텔 선호 데이터셋

<https://www.kaggle.com/datasets/jessemostipak/hotel-booking-demand/data>

2. 역할 분담

- 임준용: EDA+전처리
- 박세민: 데이터 분석
- 문장훈: 회귀 모델
- 백인서: 발표자료 + 발표

3. 기한

- 임준용: 12월 2일 까지
- 박세민: 12월 3일 까지
- 문장훈: 12월 6일 까지
- 백인서: 12월 8일 까지