

호텔 예약 취소 리스크 분석

(팀프로젝트)-팀장

1. 프로젝트 개요

본 프로젝트는 호텔 예약 데이터를 기반으로
예약 조건에 따라 발생하는 취소 및 매출 손실 위험 요인을 분석하고,
관리 가능한 고위험 예약 조합을 도출하는 것을 목표로 진행되었습니다.

단순 취소를 분석이 아닌, 실제 손실 규모와 연결된 위험 구조를 파악하고,
추가적으로 **예약 취소 가능성을 사전에 예측하는 모델을 구축하여 운영 개선에 활용 가능한 분석**을 수행했습니다.

2. 문제 정의

호텔 운영 과정에서 다음과 같은 문제가 반복적으로 발생합니다.

- 취소율은 높지만 실제 손실로 이어지는 예약 유형이 명확하지 않음
- 예약 조건에 따른 위험 관리 기준 부재
- 취소 발생 이전에 위험 예약을 예측하기 어려움

따라서 본 프로젝트에서는 다음 핵심 질문을 설정했습니다.

“어떤 예약이 취소될 때 가장 큰 손실을 만들며, 이를 사전에 예측할 수 있을까?”

3. 데이터 설명

Kaggle의 *Hotel Booking Demand* 데이터를 활용했습니다.

데이터에는 다음과 같은 정보가 포함됩니다.

- 예약 채널
- 환불 가능 여부
- 보증금 유형
- 예약 시점 정보
- 숙박 요금
- 예약 취소 여부

이를 기반으로 예약 조건과 취소 발생 간 관계를 분석했습니다.

4. 분석 접근 방식

4-1. 기본 분석

- 예약 조건별 취소율 비교
- 채널 및 환불 조건별 분포 분석
- 단순 취소율과 실제 손실 규모 차이 분석

4-2. 핵심 가설

취소율이 높은 예약이 반드시 가장 위험한 예약은 아닐 수 있다.

이를 검증하기 위해 취소 발생 시 손실 금액 중심 분석을 수행했습니다.

5. 핵심 분석 내용

5-1. 예약 조건 조합별 손실 분석

환불 여부 × 예약 채널 × 보증금 조건 조합을 생성하고,

각 조합별로

- 취소 건수
- 총 손실 규모
- 평균 손실 규모를 비교했습니다.

그 결과, 일부 조건은 취소율이 낮음에도 불구하고 취소 발생 시 손실 규모가 매우 큰 구조를 확인했습니다.

5-2. 고위험 예약 조합 도출

취소율이 아닌 손실 규모 기준으로 고위험 예약 유형을 정의하고, 운영 관점에서 관리 우선순위를 도출했습니다.

5-3. 예약 취소 예측 모델 구축

분석 결과를 실제 운영에 활용하기 위해

예약 시점에서 취소 가능성을 예측하는 모델을 추가 구축했습니다.

- 예약 조건, 시점 정보, 고객 특성을 feature로 사용
- 예약 취소 여부를 예측하는 분류 모델 구축
- 예약 단계에서 취소 가능성 사전 판단 가능 구조 설계

이를 통해 향후 다음과 같은 운영 개선이 가능함을 확인했습니다.

- 고위험 예약 사전 식별

- 보증금 정책 조정
- 객실 판매 전략 개선

6. 문제 해결 과정

문제 1. 취소율과 실제 손실 규모 불일치

취소율만으로 위험 판단 시 실제 손실 구조를 설명하지 못하는 문제 발생

→ 손실 규모 기반 분석으로 전환하여 해결

문제 2. 단일 변수 분석 한계

환불 여부, 채널, 보증금을 개별 분석 시 위험 구조가 분산됨

→ 조건 조합 기반 분석으로 전환

문제 3. 분석 결과의 실무 활용 한계

단순 분석 결과로는 운영 개선 연결이 어려움

→ 취소 예측 모델 추가 구축으로 사전 대응 구조 제시

7. 프로젝트 의의

- 단순 취소율 분석을 넘어 실제 손실 구조 분석 수행
- 고위험 예약 유형 도출 및 정책 개선 방향 제시
- 취소 예측 모델 구축을 통해 분석 결과를 운영 전략으로 연결

8. 사용 기술 및 개발 환경

- MySQL
- Excel
- Power BI (시각화 및 대시보드)
- 머신러닝 기반 예약 취소 예측 모델

GitHub Repository

https://github.com/rjaekawpxm1-netizen/portfolio/tree/main/Projects/hotel_risk_model