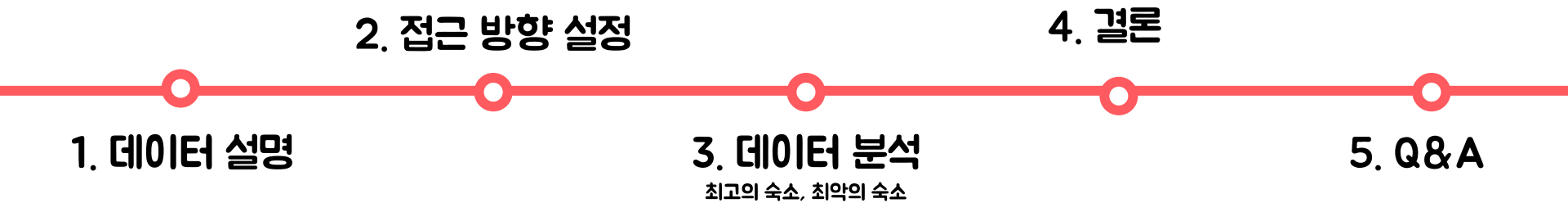




# 에어비앤비 EDA

**DSL EDA Project 5조**

김도연, 박세승, 박수빈, 박준우, 이승재





# ① 데이터 설명

# 1 데이터 설명

id		Room	Price/Reservation	Location	Host	Review
id	property_type		log_price	city	description	first_review
name	room_type	cancellation_policy		latitude	host_has_profile_pic	last_review
thumbnail_url	amenities	cleaning_fee		longitude	host_identity_verified	number_of_reviews
-	accommodates	instant_bookable		neighbourhood	host_response_rate	review_scores_rating
-	bathrooms	-		zipcode	host_since	-
-	bed_type	-		-	-	-
-	bedrooms	-		-	-	-
-	beds	-		-	-	-

약 7만개의 행, 29개의 열로 이루어진 방대한 데이터

# ID

# 객실 정보

# 가격 및 예약

# 위치

# 호스트 정보

# 리뷰



## 2 접근 방향 설정

## 이용고객의 고려사항

객실의 정보

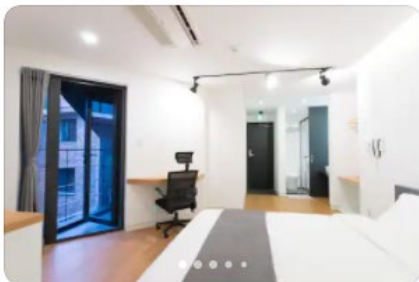


유명 관광지와의 거리



# 호스트의 고려사항

## 리뷰



최대 인원 2명 · 침실 1개 · 침대 1개 · 욕실 1개  
에어컨 · 무선 인터넷 · 세탁기

★ 4.78 (후기 83개)



최대 인원 1명 · 원룸 · 침대 1개 · 단독 사용 욕실 1개  
주방 · 에어컨 · 무선 인터넷 · 세탁기

★ 4.88 (후기 17개)

★ 4.98 · 후기 109개

청결도	_____	5.0
정확성	_____	5.0
의사소통	_____	5.0
위치	_____	4.9
체크인	_____	5.0
가격 대비 만족도	_____	4.9

## 2 접근 방향 설정

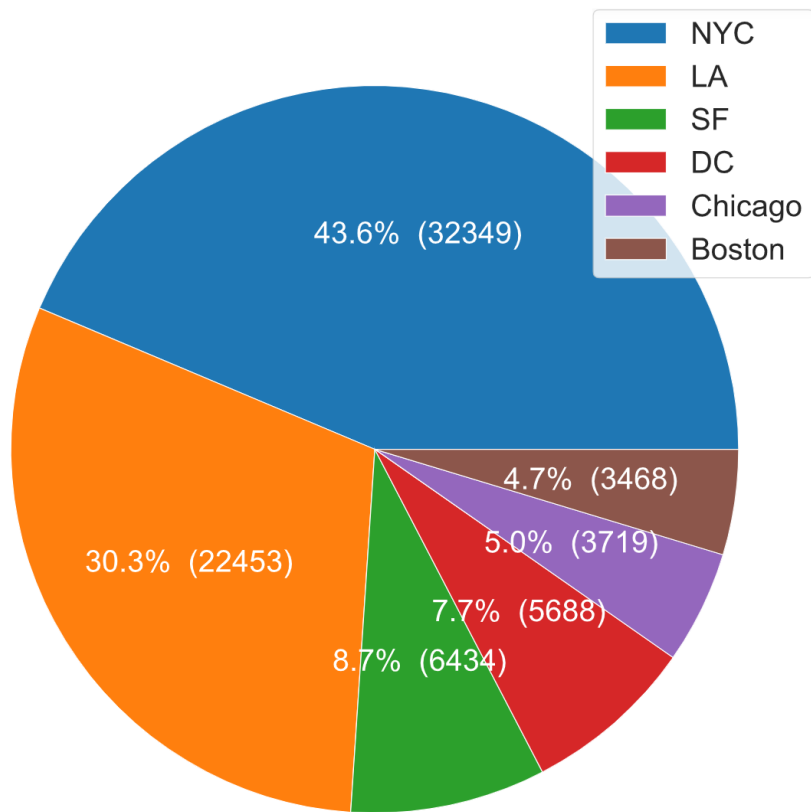
고려요소	
고객	호스트
가격	예약 수
편의시설	리뷰 점수
관광지 접근성	리뷰 수
...	...

고객 또는 호스트의 입장이 있지만  
호스트의 입장, 특히 리뷰 점수에 초점을 맞추었다.

“호스트 입장에서 좋은 리뷰를 많이 받기 위해서는 어떠한 점들을 신경써야 할까?”



## 2 접근 방향 설정



모든 도시를 분석하지 않고  
주요 도시 3군데를 설정하여 분석.

New York City

Los Angeles

San Francisco

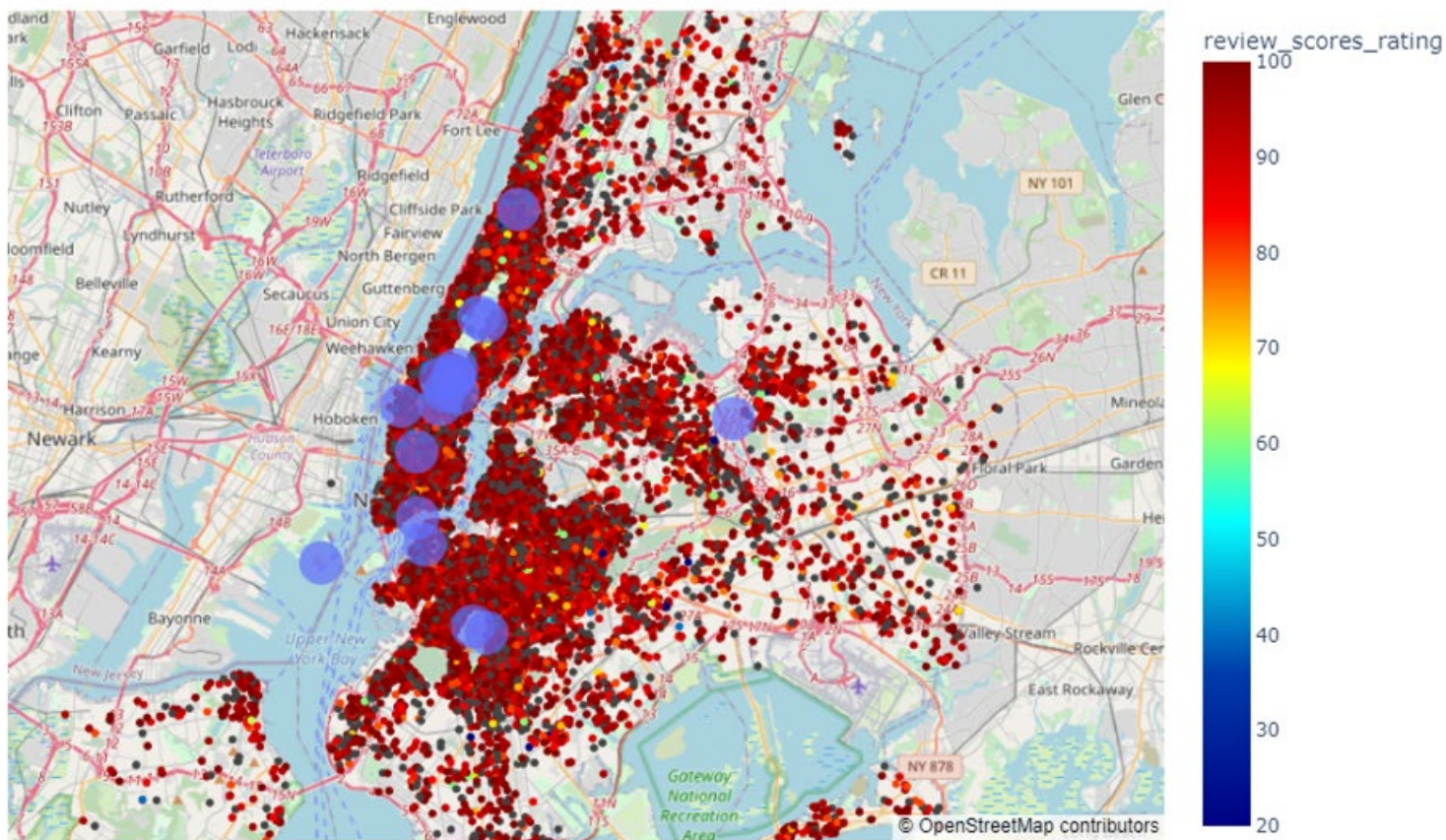
### 첫번째 접근

숙소와 랜드마크까지의 거리가 리뷰점수에 영향을 끼치는 요인일 것이다.

## 2 접근 방향 설정

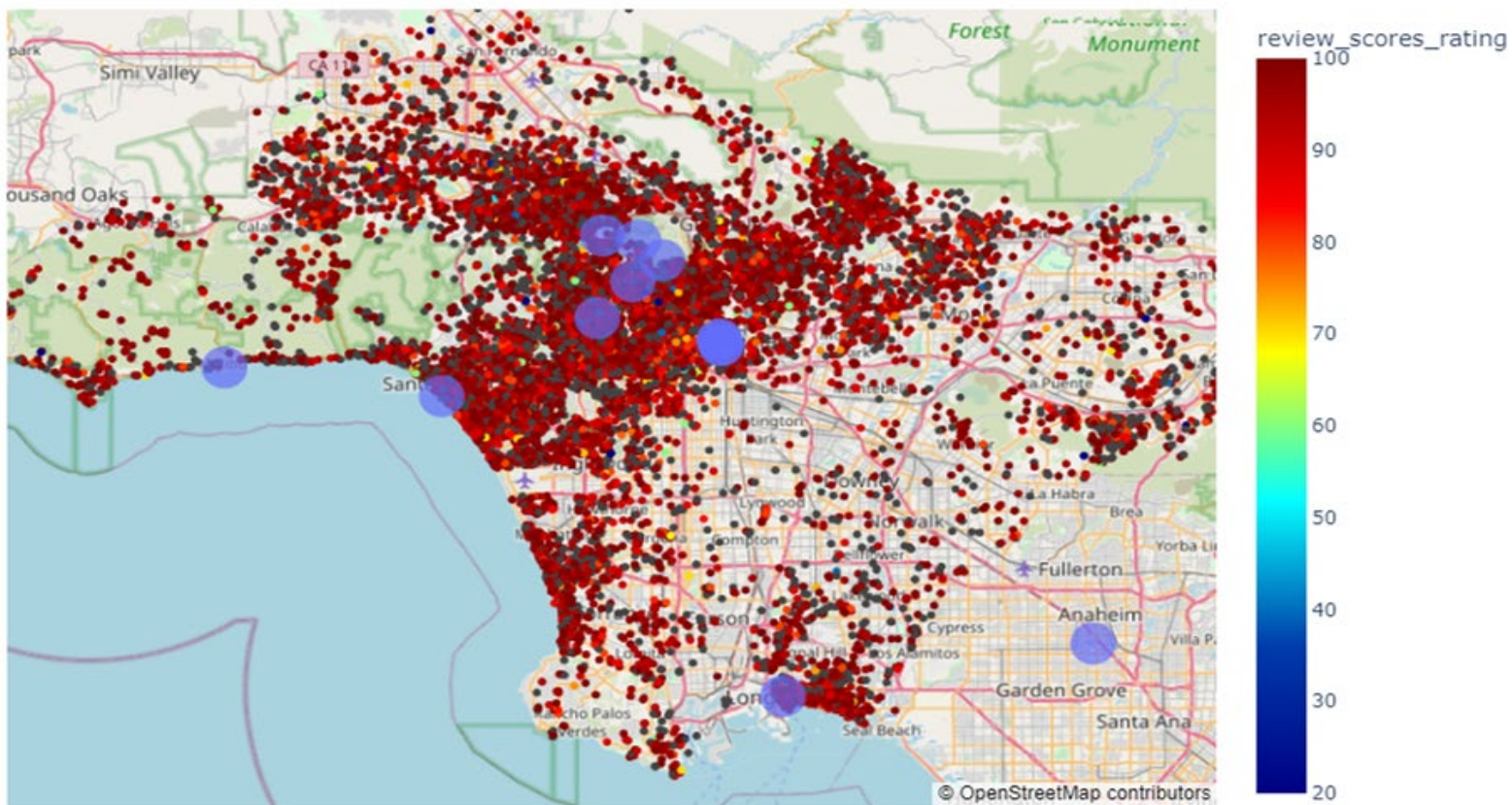
# New York City

리뷰 점수와 랜드마크와의 관계



# Los Angeles

리뷰 점수와 랜드마크와의 관계

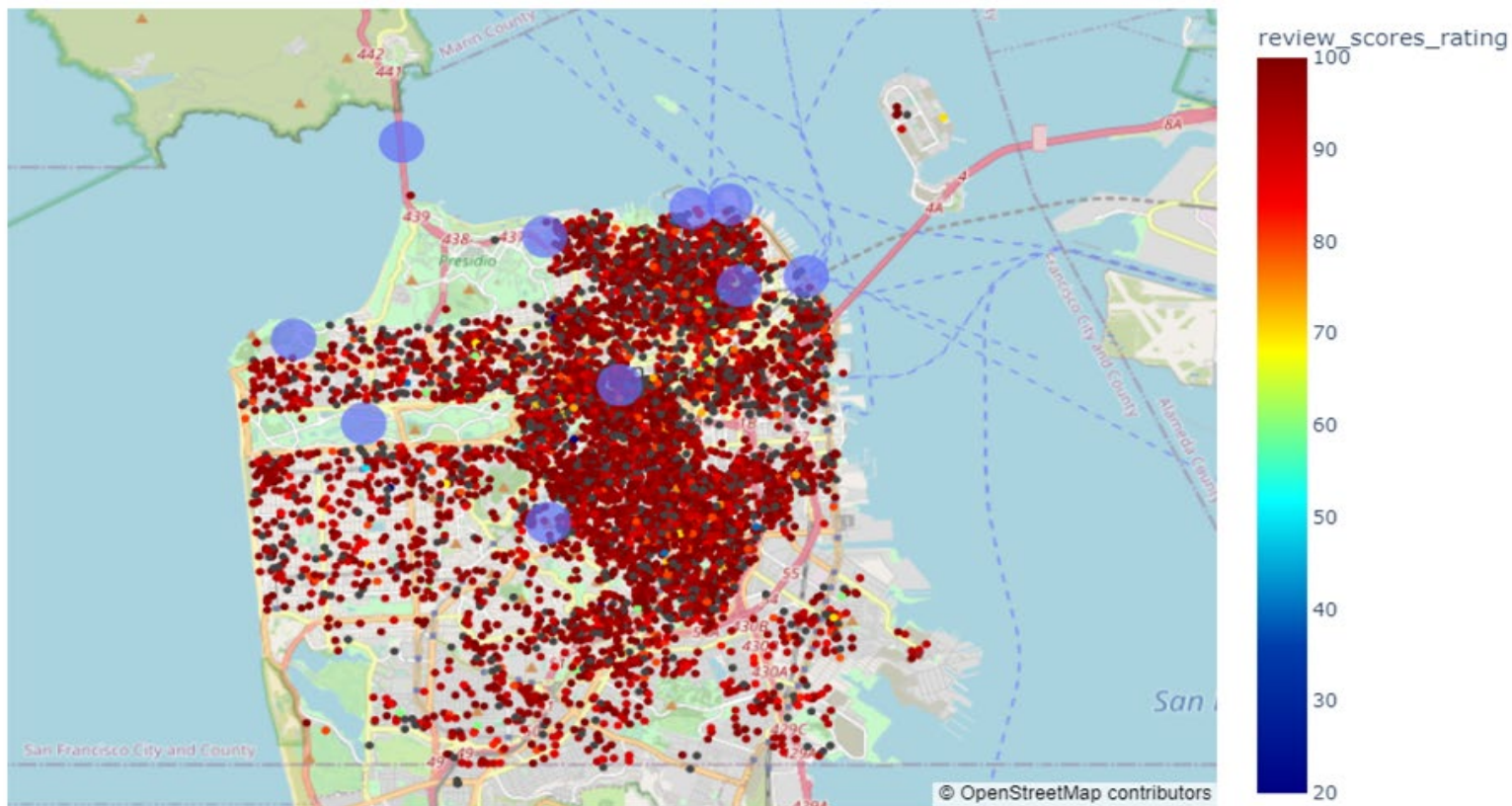




## 2 접근 방향 설정

# San Francisco

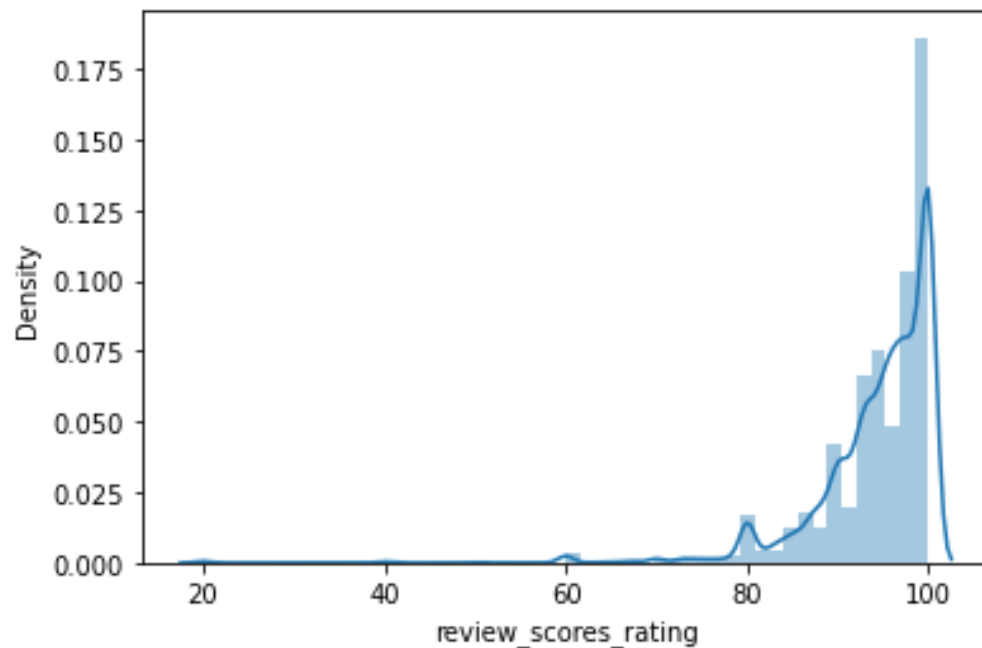
리뷰 점수와 랜드마크와의 관계



**리뷰 점수와 랜드마크와의 관계  
→ 불확실함.**



### 리뷰 점수의 분포는 Left-Skewed



### 두번째 접근

리뷰 점수를 보완할 수 있는 변수를 찾아본다.



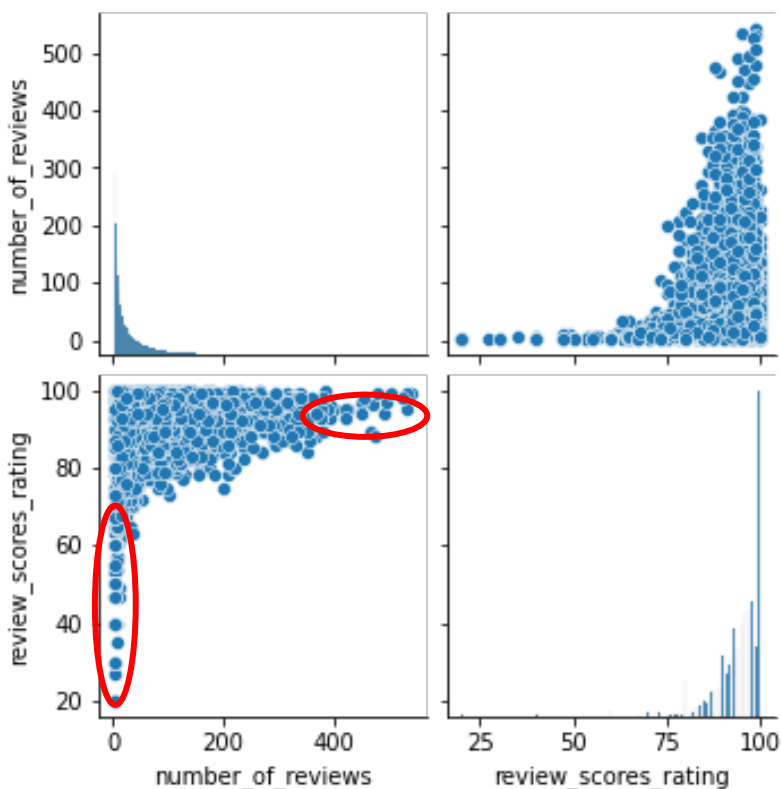
리뷰 수  $\propto$  예약 수

리뷰 수를 일종의 인기 지표로서 활용

인기가 있다  $\Rightarrow$  리뷰 수가 많다

## 2 접근 방향 설정

### 리뷰 점수와 리뷰 수의 관계



리뷰 개수가 **적은** 숙소들의 경우  
낮은 평점부터 높은 평점까지 다양하지만,

리뷰 개수가 **많은** 숙소들의 경우  
낮은 평점을 찾아보기 힘들다.

### 최고의 숙소(top):

리뷰 점수가 **높고**, 리뷰 수도 **많은** 숙소

### 최악의 숙소(bottom):

리뷰 점수가 **낮고**, 리뷰 수도 **적은** 숙소

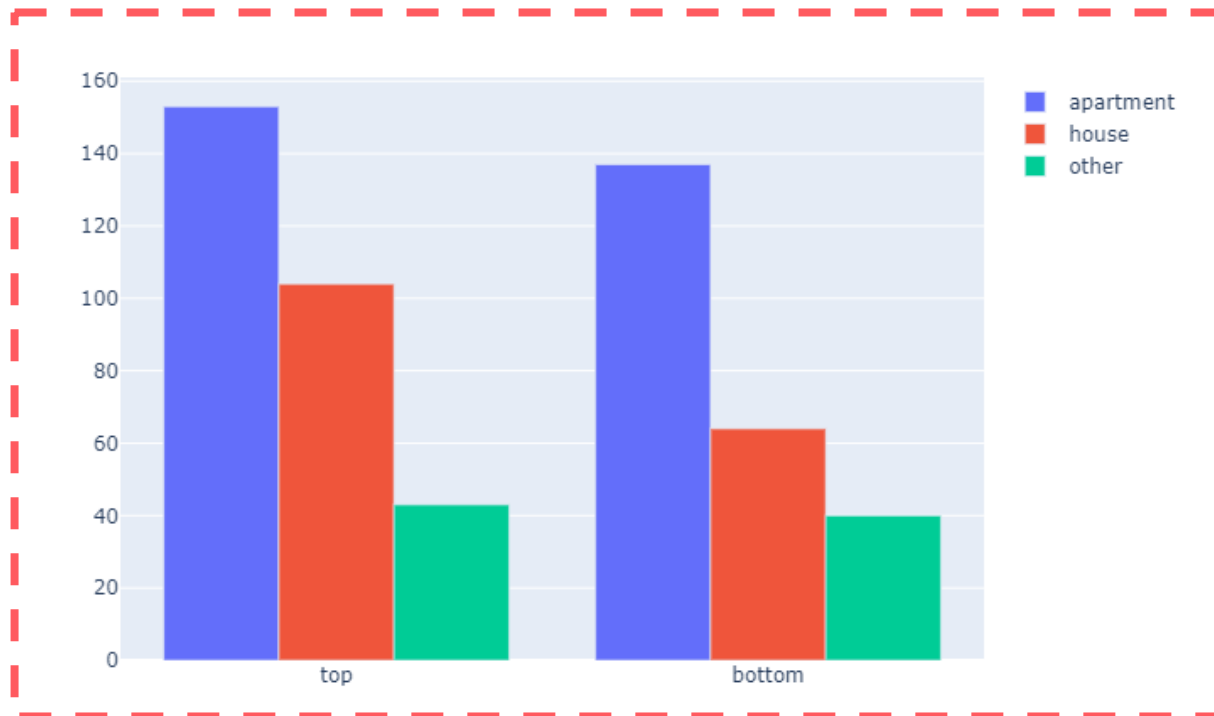


# ③ 데이터 분석

최고의 숙소, 최악의 숙소!

### 3 최고의 숙소, 최악의 숙소 !

#### 방 관련 요소 분석- 방의 type 분석 결과

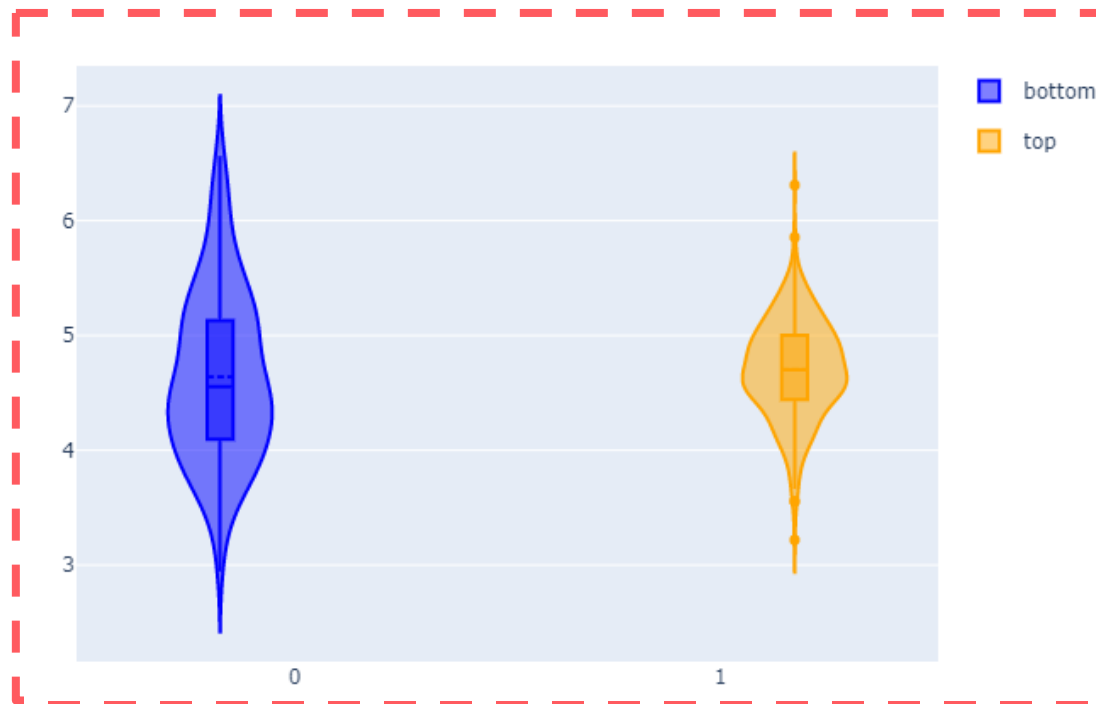


차이를 찾아보기 힘들었으며,  
객실 관련 문제에서는 호스트를 위한  
즉각적인 피드백이 어렵기에,

다른 요소에 주목해보기로 함

### 3 최고의 숙소, 최악의 숙소 !

#### (로그)가격분포에 대한 분석

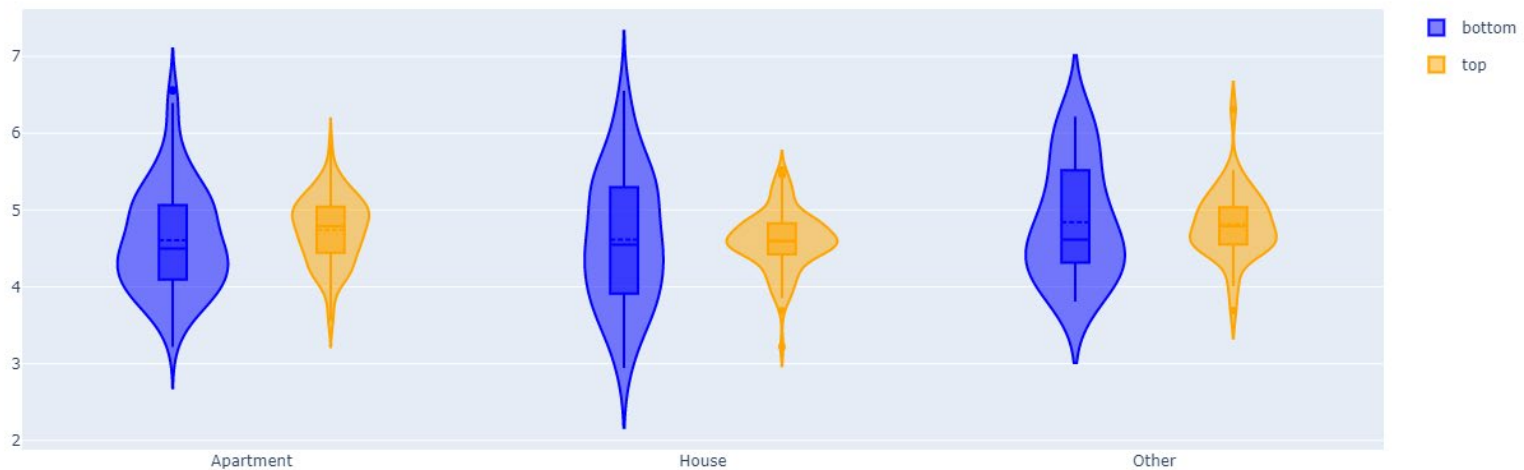
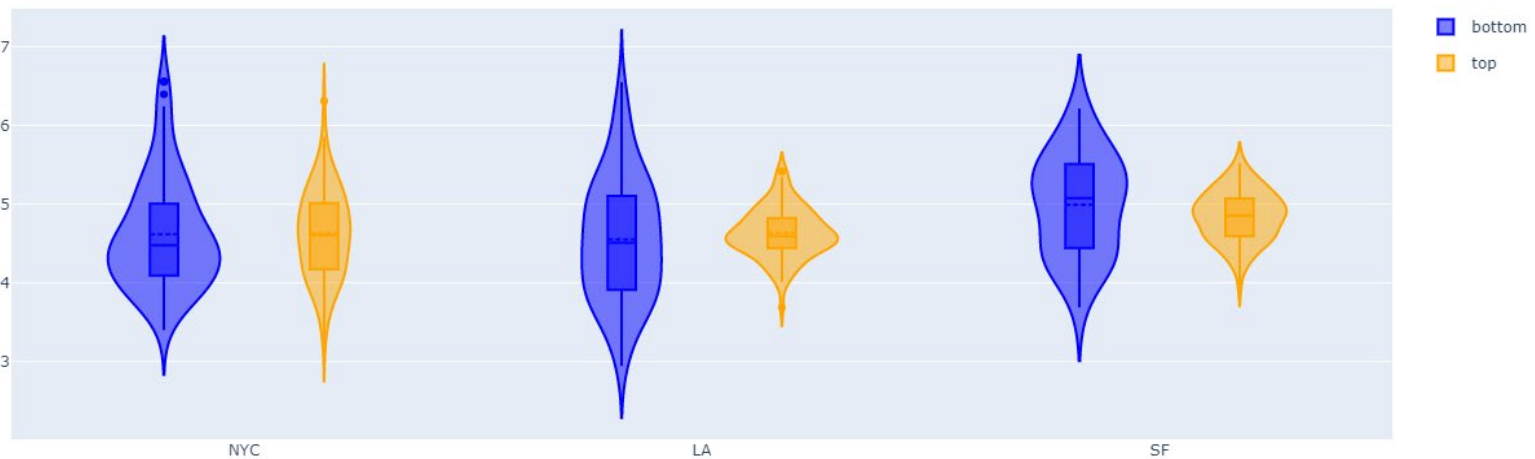


Top 숙소의 가격의 분산이  
더욱 작다는 것을 알 수 있음

→가격이 비슷하고 이상치가 거의 없는 모습

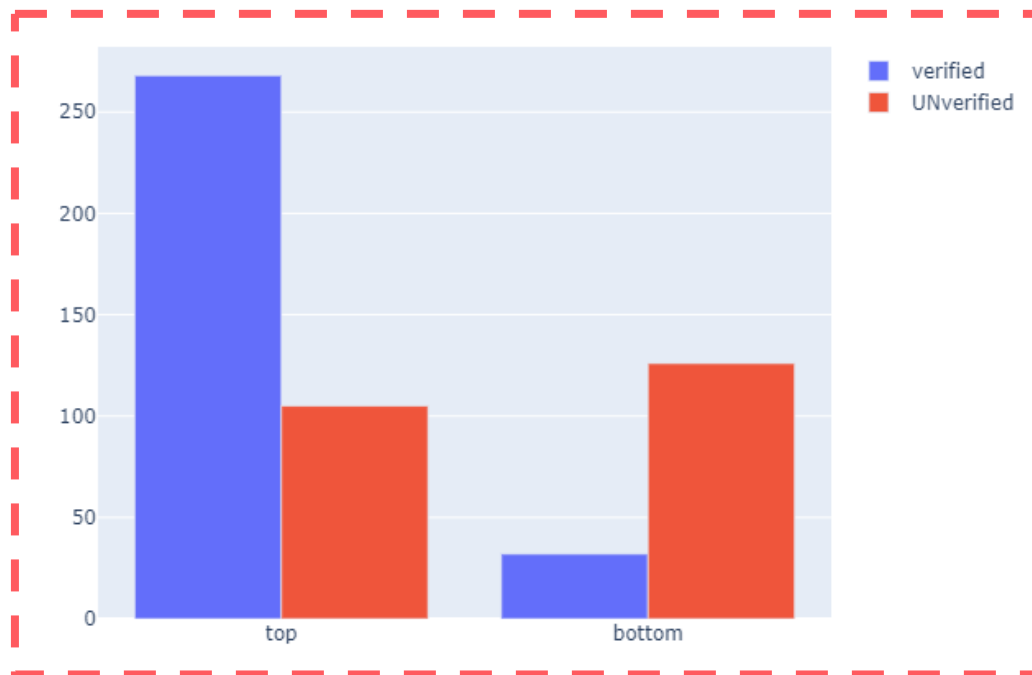
### 3 최고의 숙소, 최악의 숙소 !

#### 가격분포에 대한 분석 - city , property type



### 3 최고의 숙소, 최악의 숙소 !

#### HOST 관련 변수 분석 - host verified



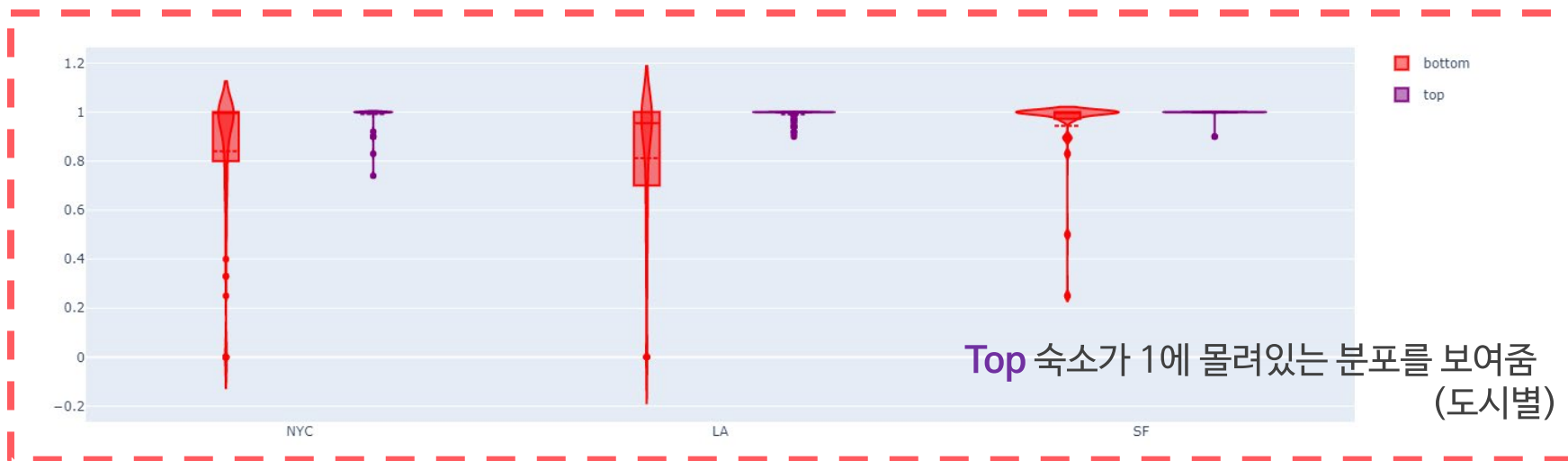
Top 숙소에서 호스트가 인증된 비율이  
80%정도로 bottom 숙소에 비해 월등히 높음

→ 손님들에게 신뢰도를 줄 수 있는 중요 요소라고 판단



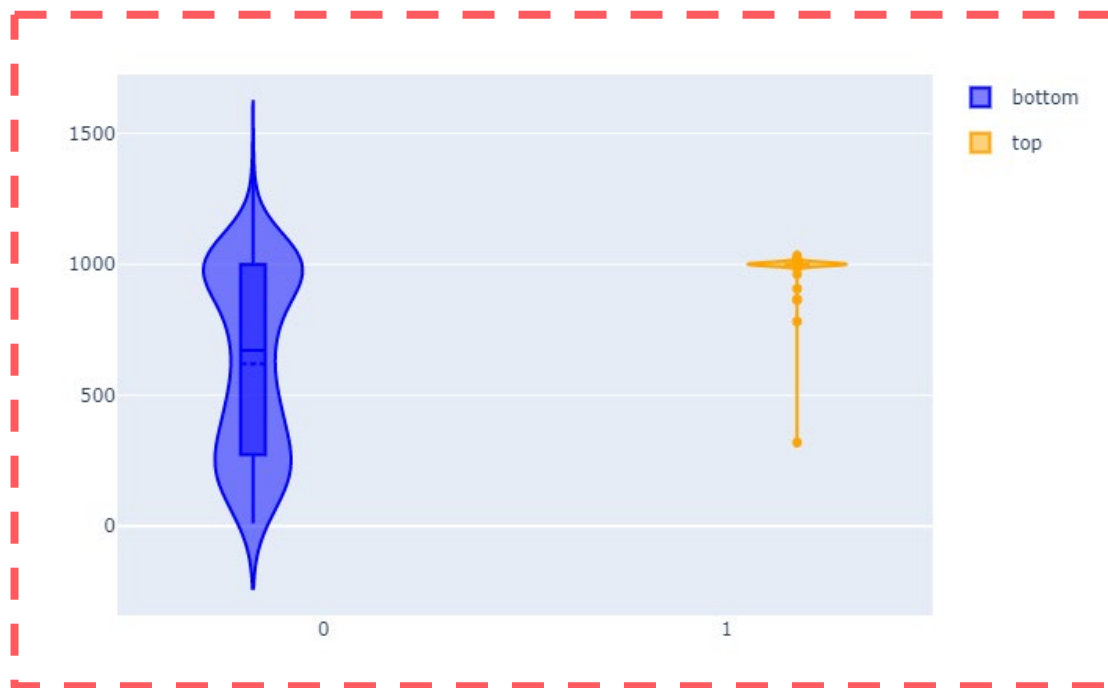
### 3 최고의 숙소, 최악의 숙소 !

#### HOST 관련 변수 분석 - host\_response\_rate



### 3 최고의 숙소, 최악의 숙소 !

## Description 분석 - length



Description 자체는 host가 작성하는 것이기에  
내용 분석은 큰 의미가 없었음

Description 길이 분석 결과, Top 숙소가 대체적으로  
길었으며 이를 호스트 성향과 연결지어 볼 수 있다고 판단



# 4 결론

## 4 결론

가격이 일정하면 좋겠어

믿을 수 있는 호스트가 좋겠어

호스트가 피드백을 주면 좋겠어



< 서울 강남구

날짜 입력



슈퍼호스트



★ 4.74 (54)

공동 주택 전체 · 강남

S2 Clean & cozy room right next to Ga...

위치가 너무 좋아요 호스트님도 친절하세요

가성비 대박이었습니다! 강남역 출구 나오자마자 있어서  
코엑스, 한강 가기 너무 좋았구요 2박을 머물러서 수건 추가  
요청 드렸는데 다음날 바로 갖다주셨어요!  
쓰레기봉투를 포함해서 면봉, 솜 등 각종 소모품도  
부족하지 않았고 돌돌이도 있어서 잘 썼어요!  
가성비 너무 좋아요 잘 머무르다 갑니다~~

깨끗하고 위치도 좋음

답장이 너무 빠르고 친절하심

요청 사항도 잘 들어주셔서 너무나 감사드립니다

편히 잘 사용했음

너무 깨끗하고 호스트분도 친절해서 다음에 또 찾을것같아요

실제 '서울시 강남구'의 슈퍼호스트 숙소의 리뷰를 분석해본 결과

54건 중 15건의 리뷰에서 호스트에 대한 피드백이 존재함을 확인

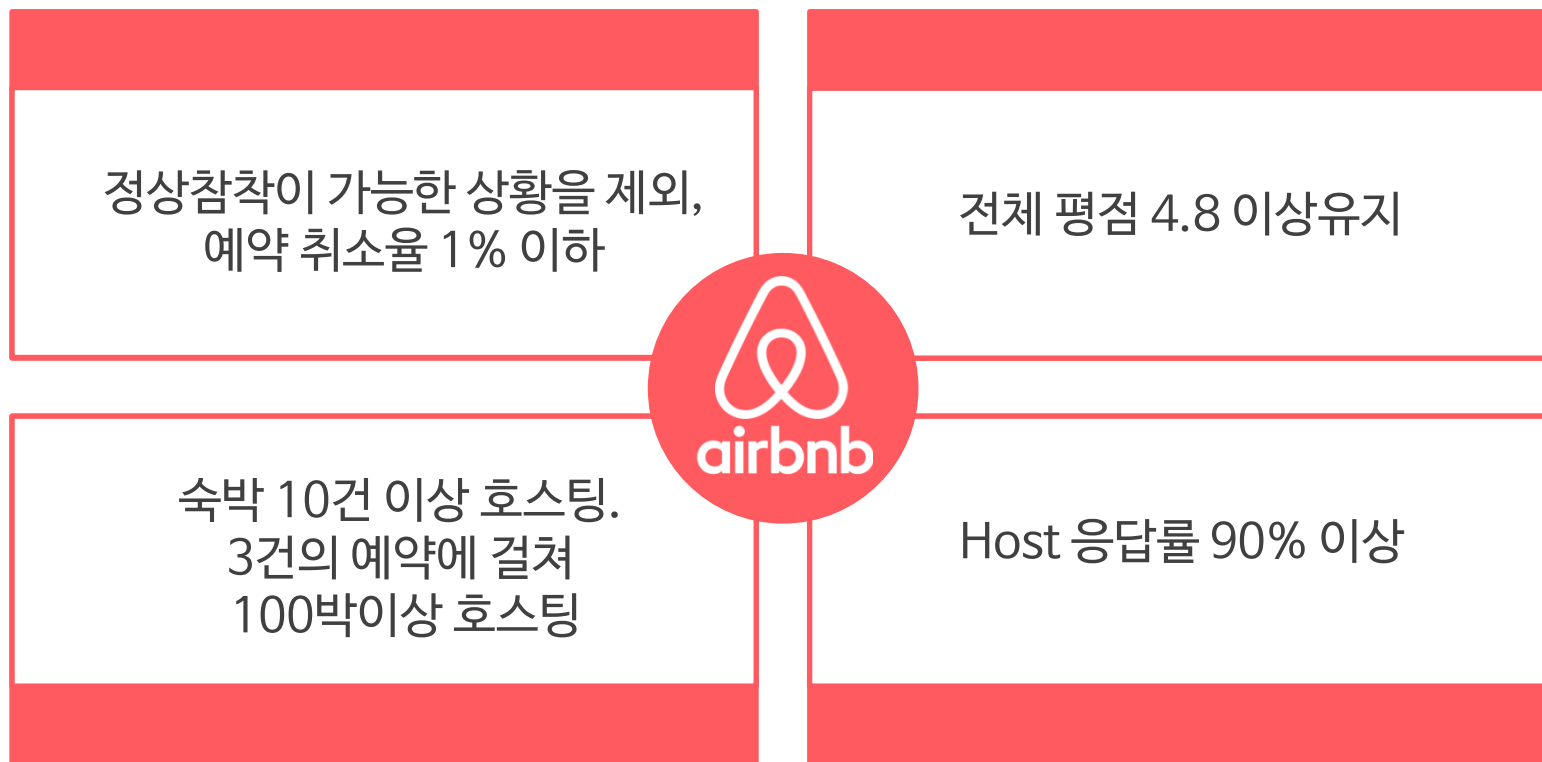
이때, 슈퍼호스트란 무엇일까?



### 에어비앤비의 슈퍼 호스트 제도

다른 호스트에게는 모범이 되고 게스트에게는 특별한 경험을 선사하는 숙련된 호스트

## 슈퍼호스트의 자격요건



호스트 관련 요소의 변화를 통해,  
슈퍼호스트 제도라는 추가적인 이점을 노릴 수 있으며  
이가 앞서 결론 지은 후기에도 영향을 미칠 것으로 판단



궁금하신 점을 질문해주세요 !



**감사합니다**