



New York 핫플 한식당 만들기 Project

‘브루클린 무브’

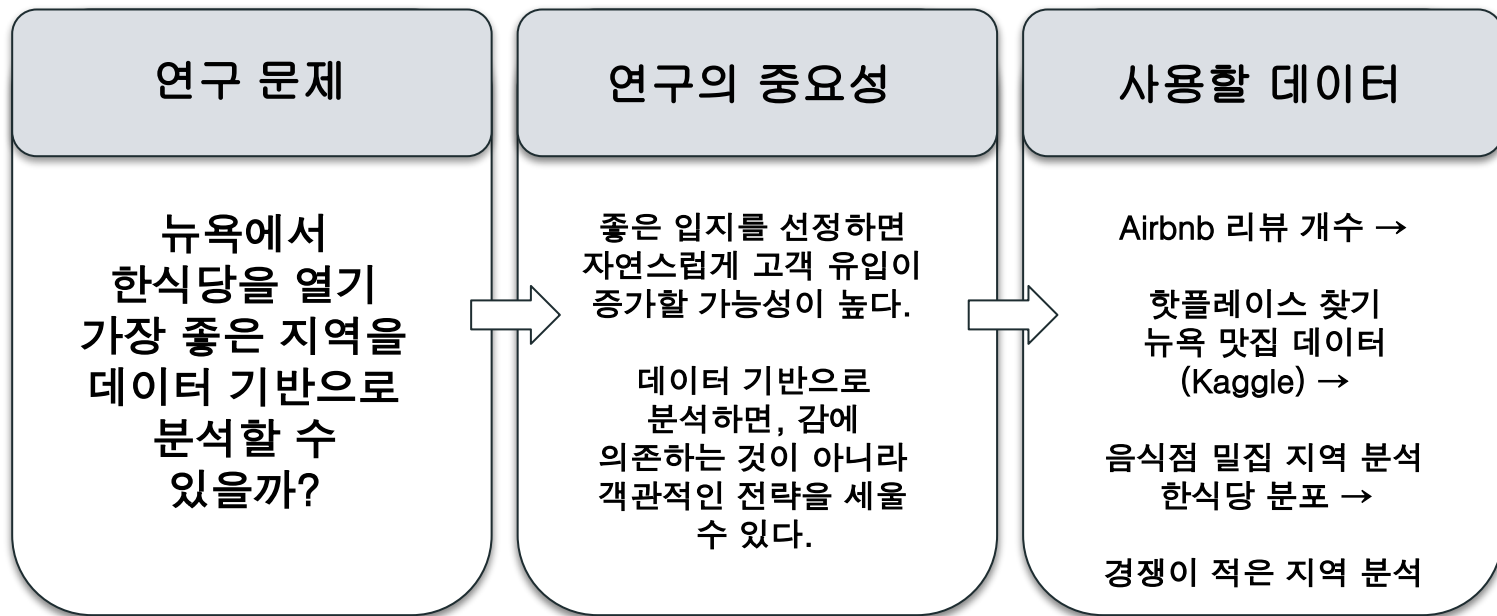
1. 개요 및 문제 정의

2. 데이터셋 소개

3. 과정

4. 결론

문제정의 |



배경 | 왜 한식당을 뉴욕에 열려는가?



경제력이 세계 최고인 도시이며,
높은 소비력을 가진 시장

한식당의 인기가 증가하면서
시장 규모가 커지고 있음

K-컬처의 확산으로 한식이
대중화되고 있음

다양한 인종과 배경을 가진
소비자들이 한식을 적극적으로 수용

입지를 전략적으로 잘 선정하면
성공할 가능성이 매우 높음

배경 | 왜 한식당을 뉴욕에 열려는가?

미국 전체에서 한식 레스토랑 산업은
연평균 4.9% 성장할 것으로 전망됨
(2025년까지)

이는 뉴욕 내 한식당 창업이 유망한
비즈니스 모델이 될 가능성을 시사

Projected Growth of Korean Restaurant Industry in the US (2020-2025)



이러한 흐름 속에서, 뉴욕의 **가장
적절한 지역**에 한식당을 개업하는
것이 중요합니다.

우리는 데이터를 기반으로
한식당을 열기 좋은 입지를
분석하여 최적의 위치를 찾고자
합니다.

배경 | 뉴욕에서 한식당이 증가하는 이유

한식에 대한 관심 증가

K-콘텐츠(넷플릭스, BTS,
블랙핑크 등)

뉴욕 타임스가 선정한 '2024년
뉴욕 최고의 레스토랑 100곳' 중
7곳이 한식당

2023년 5월 기준, 뉴욕 미술관
레스토랑 72곳 중 9곳이 한식당

프리미엄 한식당의 성장

고급 한식당이 증가하면서
고객층이 다양해지고 있음

전통 한식뿐만 아니라 퓨전 한식,
채식 한식, 고급 한식당이
늘어나고 있음

따라서 한식은 뉴욕에서 시장 성장
가능성이 있다



VRINDA KALLU · UPDATED A YEAR AGO



37

<> Code

Download



New York Airbnb Open Data 2024

Airbnb listings and metrics in NYC, NY, USA as of (05 January, 2024)



RAYHAN AHMED · UPDATED 2 YEARS AGO



29

<> Code

Download

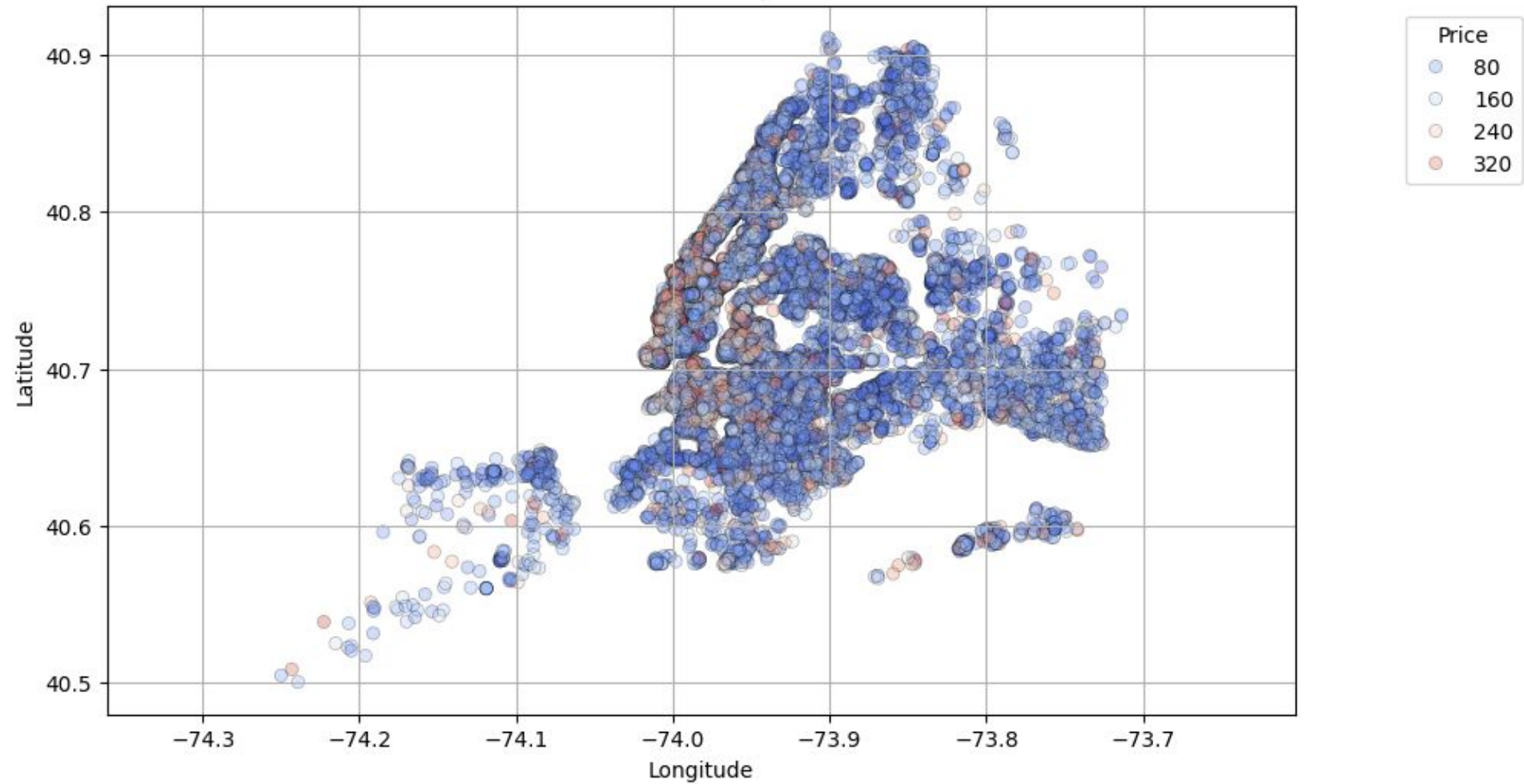


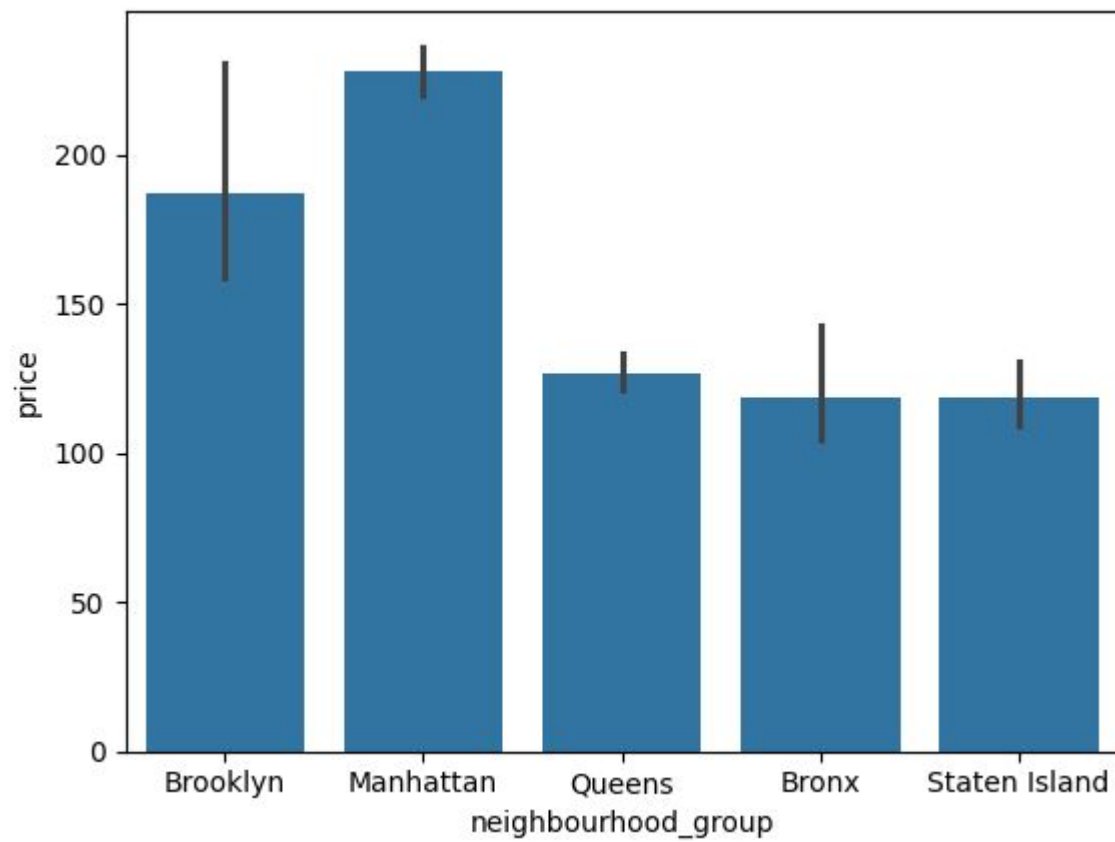
Trip Advisor Newyork City restaurants Dataset 10k+

Newyork City Dataset from Tripadvisor

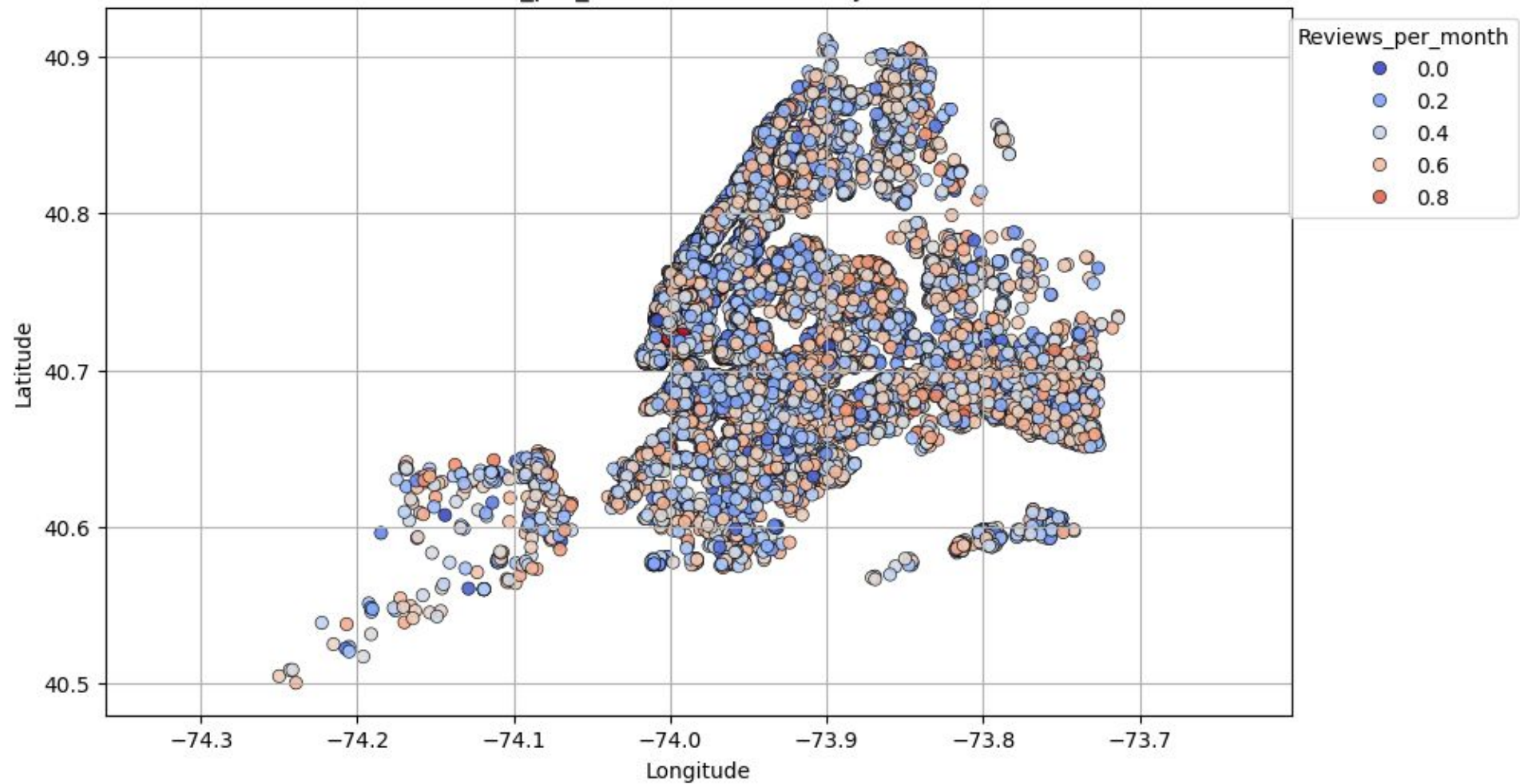


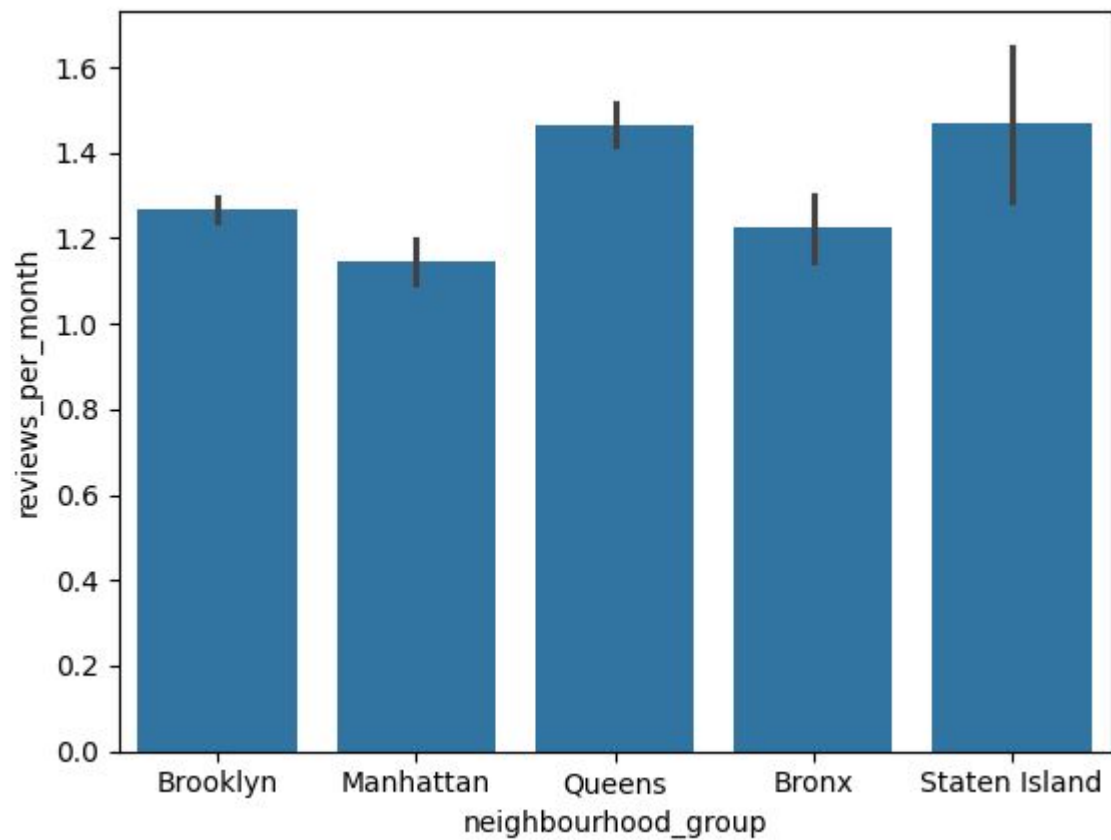
Price Distribution by Location





Reviews_per_month Distribution by Location





논리구조 | 리뷰 수가 많은 Airbnb = 핫플레이스 근처

- (1) 여행객들은 관광지과 맛집 근처의 숙소를 선호한다.
- (2) 리뷰 수가 많은 Airbnb는 사람들이 많이 이용한 숙소라는 뜻이다.
- (3) 입지가 좋은 지역에 Airbnb가 많고, 그 숙소들은 리뷰 수도 많다.
- (4) 따라서, 리뷰 수가 많은 Airbnb 근처는 핫플레이스일 가능성이 높다.

즉, Airbnb 리뷰 수를 핫플레이스의 지표로 활용할 수 있다.

전제 | 뉴욕시의 숙소와 인기맛집 지역 분포에 대한 상관관계

NYC Airbnb Analysis



Matilda Sorić

Data Analyst @ Brooklyn Nets & NY Liberty | Former NCAA
& FIBA Athlete | NBA Future Analytics Stars '24

2022년 11월 11일

When Everything Is “Nearby”: How Airbnb Listings in New York City Exaggerate Proximity

Mikael Brunila ✉️

Platial Analysis Lab, Department of Geography, McGill University, Montréal, Canada
Urban Politics & Governance Lab, School of Urban Planning, McGill University, Montréal, Canada

Priyanka Verma ✉️

Platial Analysis Lab, Department of Geography, McGill University, Montréal, Canada

Grant McKenzie ✉️

Platial Analysis Lab, Department of Geography, McGill University, Montréal, Canada

Cambridge Handbook of the Law of the Sharing Economy (forthcoming)

Available at <https://ssrn.com/abstract=3048397>

Airbnb Usage Across New York City Neighborhoods: Geographic Patterns and Regulatory Implications¹

Peter Coles² Michael Egesdal³ Ingrid Gould Ellen⁴ Xiaodi Li⁵ Arun Sundararajan⁶

October 12, 2017

결론 : 상관관계가 유의미하다.



전제 | 뉴욕시의 숙소와 인기맛집 지역 분포에 대한 상관관계



2. 코드분석


```
# 위에서 반환한 한식당 이름 리스트를 구글 api에서 좌표로 받기
# 식당 이름에 지역 단서를 달아두면 정확한 식당을 찾아낼 가능성이 높아짐
# API 키 설정
```

```
gmaps = googlemaps.Client(key="AIzaSyCX6BzKlK49z5DGMc3wRiLEyYR_aqgvW2E")
```

```
# 각 식당 이름에 "(New York)" 추가 >> 위치 힌트
```

```
restaurant_names = [restaurant + " (New York)" for restaurant in KR_title_list]
```

```
# 각 식당에 대한 위도, 경도 리스트
```

```
restaurant_coords = []
```

```
for restaurant in restaurant_names:
```

```
    try:
```

```
        # Place Search API 요청
```

```
        places_result = gmaps.places(restaurant)
```

```
        # 결과 확인
```

```
        if places_result["status"] == "OK":
```

```
            latitude = places_result["results"][0]["geometry"]["location"]["lat"]
```

```
            longitude = places_result["results"][0]["geometry"]["location"]["lng"]
```

```
            restaurant_coords.append({"name": restaurant, "latitude": latitude, "longitude": longitude})
```

```
            print(f"{restaurant} - 위도: {latitude}, 경도: {longitude}")
```

```
        else:
```

```
            print(f"{restaurant} - 식당을 찾을 수 없습니다.")
```

```
    except Exception as e:
```

```
        print(f"Error occurred while searching for {restaurant}: {e}")
```

```
# API 호출 제한을 피하기 위해 잠시 대기
```

```
# 각 식당의 좌표를 1초 간격으로 불러오기 >> API 호출 후 1초 대기
```

```
time.sleep(1)
```

```
# 숙소 반경 1km 내 한식당 개수
isnearby = []

# 각 숙소에 대해 반경 1km 내 식당 수 계산
for _, nyAirbnb_row in nyAirbnb.iterrows():
    nyAirbnb_lat, nyAirbnb_lon = nyAirbnb_row['latitude'], nyAirbnb_row['longitude']

    count = 0
    for _, res in restaurants.iterrows():
        res_lat = res['latitude']
        res_lon = res['longitude']

        # 거리 계산
        distance = haversine(nyAirbnb_lat, nyAirbnb_lon, res_lat, res_lon)

        # 1km 이내인 경우 카운트
        if distance <= 1.0:
            count += 1

    isnearby.append(count)

# 결과를 숙소 데이터에 추가
nyAirbnb['restaurant_count'] = isnearby
```

```
print(nyAirbnb.shape)  
nyAirbnb['reviews_per_month'].head() #
```

✓ 0.0s

```
0    0.03  
1    0.24  
2    1.67  
3    1.38  
4    0.24
```

Name: reviews_per_month, dtype: float64

```
nyAirbnb['reviews_per_month'].head()
```

✓ 0.0s

```
0    0.084159  
1    0.306664  
2    0.537971  
3    0.514964  
4    0.306664
```

Name: reviews_per_month, dtype: float64

```
train_data['target_value'] = train_data['reviews_per_month'] * train_data['inverse_scaled_score']
```

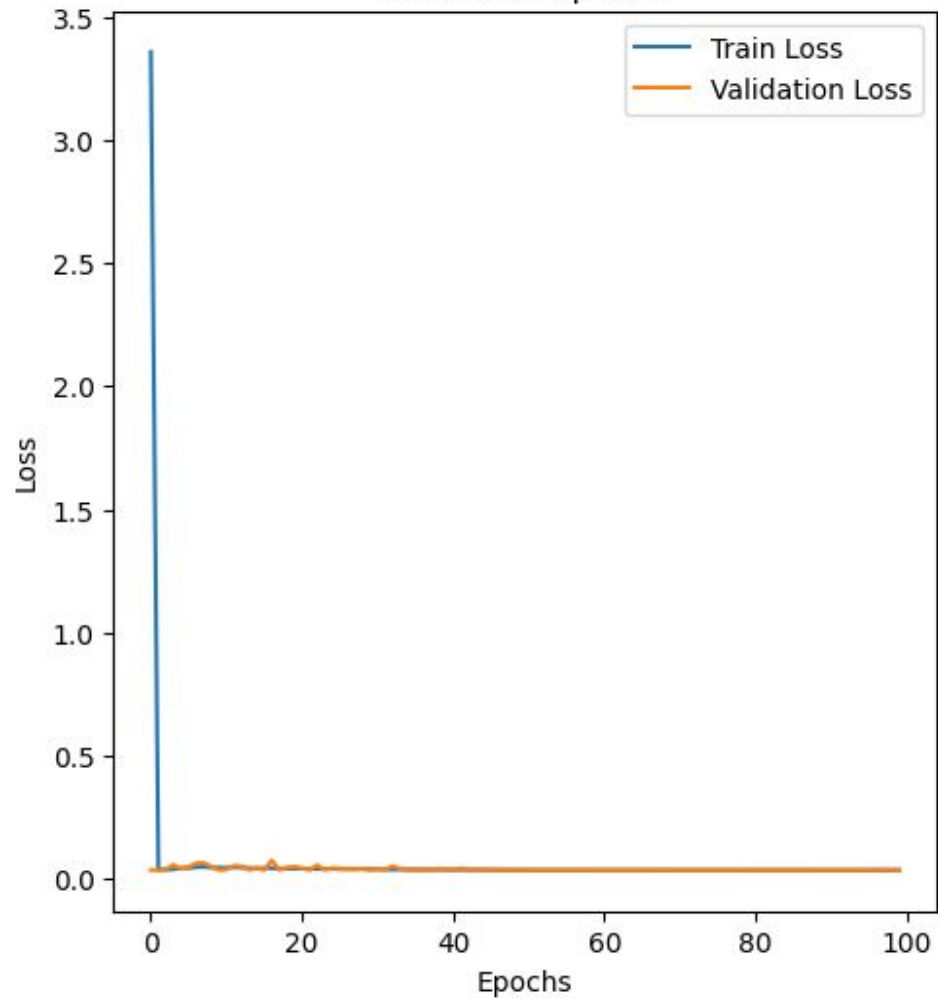
```
x_train = train_data[['latitude', 'longitude']]
```

```
y_train = train_data['target_value']
```

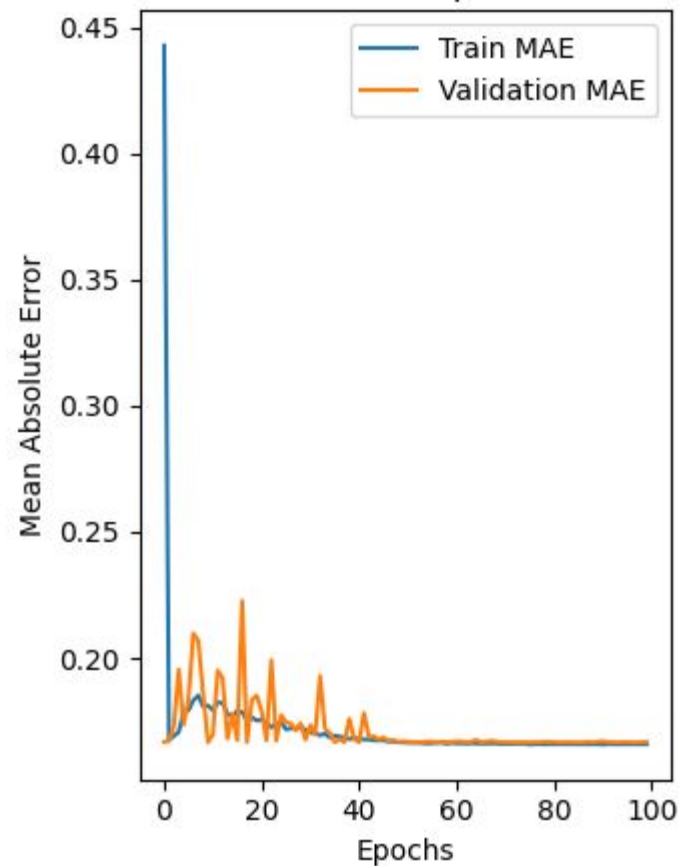
```
def build_model():  
    model = keras.Sequential([  
        layers.Dense(32, activation = 'relu'),  
        layers.Dense(32, activation = 'relu'),  
        layers.Dense(1)  
    ])  
    optimizer = keras.optimizers.Adam(learning_rate=0.001)  
    model.compile(optimizer = optimizer, loss = 'mse', metrics = ['mae'])  
    return model
```

```
model = build_model()
```

Loss Over Epochs



MAE Over Epochs

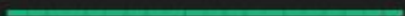


```
predictions = model.predict(test_data_selected)
print(predictions[:5])
```

[34]

✓ 0.1s

...

98/98  **0s** 801us/step

```
[0.33966503]
[0.33966503]
[0.33966503]
[0.33966503]
[0.33966503]
```

```
print(test_data_selected.head())
```

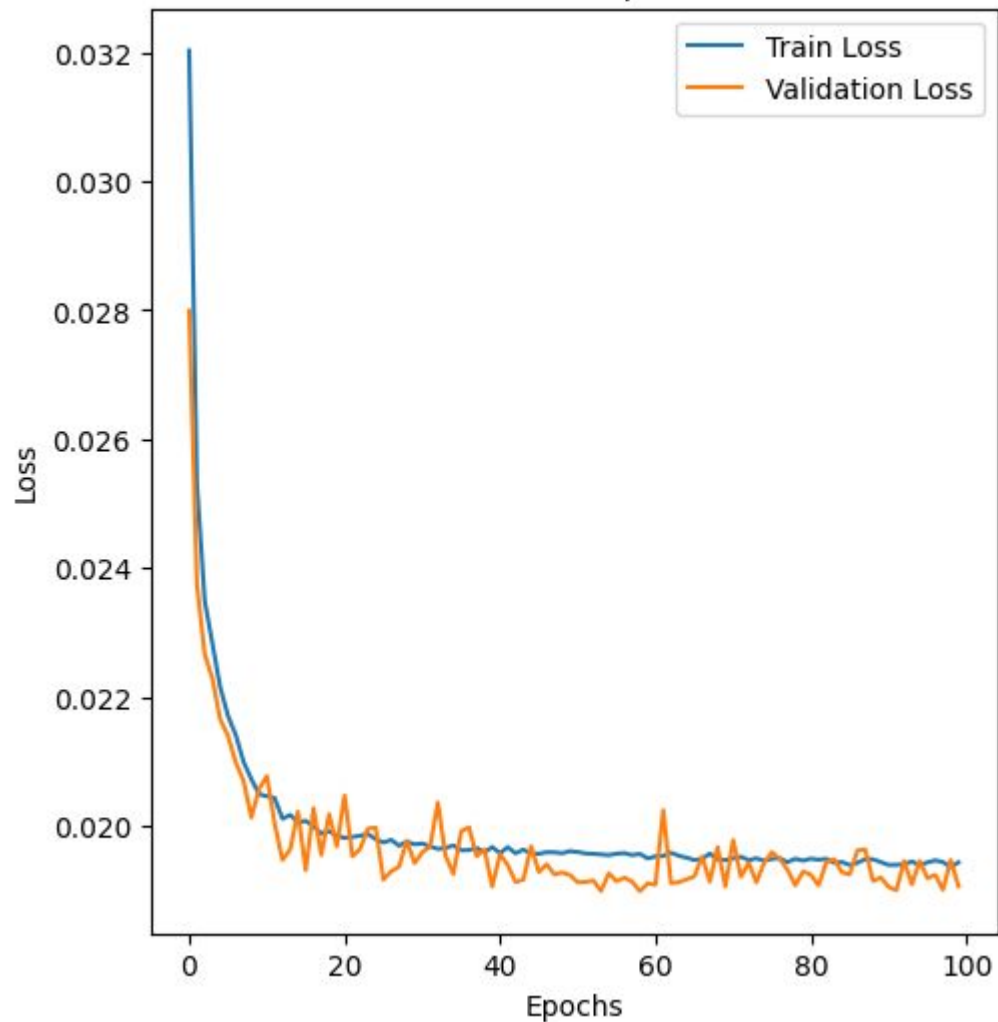
[35]

✓ 0.0s

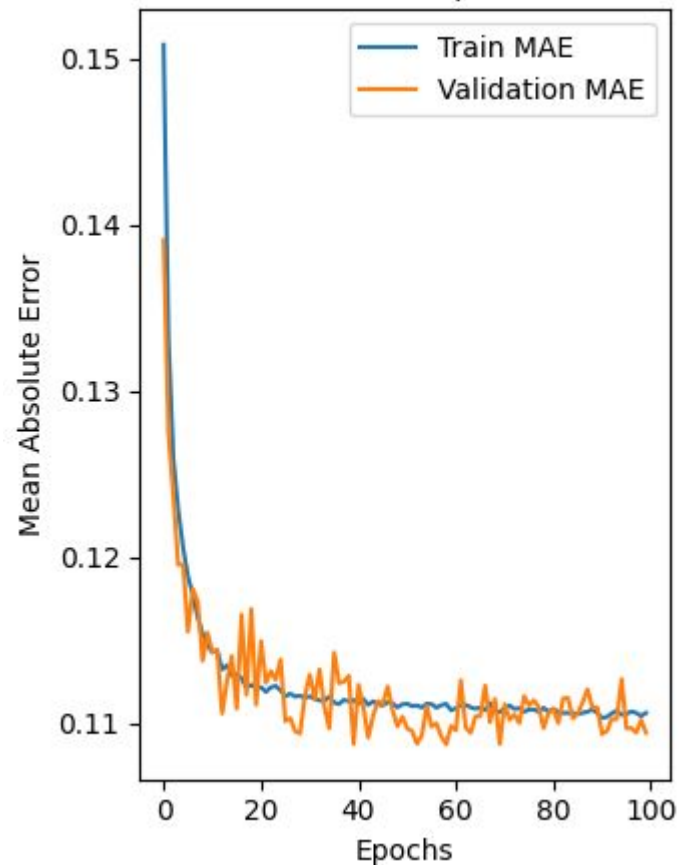
...

	latitude	longitude
7341	40.833080	-73.941030
4821	40.808540	-73.922670
1783	40.772293	-73.993665
16037	40.688680	-73.924070
20539	40.697920	-73.940570

Loss Over Epochs



MAE Over Epochs




```
test_data_selected = test_data.iloc[:, :2]
predictions = model.predict(test_data_selected)
print(predictions[:5])
```

✓ 0.1s

98/98 ————— 0s 804us/step

```
[[0.40986806]
 [0.4535644 ]
 [0.20760411]
 [0.41612947]
 [0.39812642]]
```

```
print(test_data_selected.head())
```

✓ 0.0s

	normalized_latitude	normalized_longitude
7341	0.809979	0.575934
4821	0.750246	0.610175
1783	0.662018	0.477768
16037	0.458497	0.607564
20539	0.480988	0.576792

```
top_3_reviews = test_data.sort_values(by='reviews_per_month', ascending=False).head(3)
print(top_3_reviews)
```

✓ 0.0s

	normalized_latitude	normalized_longitude	reviews_per_month	\
14767	0.607076	0.475794	1.000000	
6409	0.510455	0.444768	0.988964	
8342	0.640859	0.494433	0.951986	


	inverse_scaled_score	target_value
14767	0.184820	0.184820
6409	0.796205	0.787418
8342	0.166995	0.158977

```
example_data = {
    'normalized_latitude' : [0.607100, 0.510400, 0.640900],
    'normalized_longitude' : [0.475800, 0.444800, 0.494400]
}
example = pd.DataFrame(example_data)
```

✓ 0.0s

```
predictions = model.predict(example)
print(predictions)
```

✓ 0.0s

1/1  0s 25ms/step

```
[[0.10227874]
 [0.2758556 ]
 [0.10414751]]
```

프로젝트를 마치며

연구 결과

이번 프로젝트에서 저희는

한식당 위치가 미치는 영향을 새롭게 분석했습니다

브루클린의 발전과 에어비엔비의 성장은 뉴욕을 크게 바꾸고 있습니다

도전과 성장

뉴욕이라는 큰 도시를 완벽히 이해하긴 어려웠지만, 많이 배울 수 있었습니다

우리의 연구는 아직 부족한 점이 있지만,

여러분께 새로운 생각을 드릴 수 있었다면 좋겠습니다

이번 데이터톤은 쉽지 않았지만 즐거웠고 배움이 많았습니다





감사합니다.

Reference

<https://www.kaggle.com/datasets/rummagelabs/nytimes-best-restaurants-2024>

<https://www.kaggle.com/datasets/lorentzyeung/top-240-recommended-restaurants-in-new-york-2023/data>

<https://www.kaggle.com/datasets/kalilurrahman/newyork-subway-entry-exit-dataset-for-dataviz>

<https://www.kaggle.com/datasets/lorentzyeung/top-240-recommended-restaurants-in-new-york-2023>

<https://www.kaggle.com/datasets/vrindakallu/new-york-dataset>

https://www.kaggle.com/datasets/vrindakallu/new-york-dataset/data?select=new_york_listings_2024.csv

https://www.kaggle.com/datasets/rayhan32/trip-advisor-newyork-city-restaurants-dataset-10k?select=trip+advisor+restaurents++10k+-+trip_rest_newyork_1.csv

<https://www.kaggle.com/datasets/khashayarahmadi/usa-shooting-2023>

<https://www.kaggle.com/datasets/whisperingkahuna/gunviolence1>

<https://campaigns.do/datasets/76>

Reference

<https://www.linkedin.com/pulse/nyc-airbnb-analysis-matilda-sori%C4%87/>

https://www.academia.edu/101079511/The_Sharing_Economy_as_a_Local_Economic_Engine_The_Heterogeneous_Impact_of_Airbnb_on_Restaurant_Employment_Growth

<https://grantmckenzie.com/academics/LIPIcs.GIScience.2023.16.pdf>

<https://www.linkedin.com/pulse/new-york-city-airbnb-data-analysis-visualization-rayana-tns/>

<https://www.redfin.com/neighborhood/219258/NY/New-York/Brooklyn/housing-market>

<https://www.zillow.com/home-values/6181/new-york-ny/>

<https://www.redfin.com/city/35948/NY/Manhattan/housing-market>

<https://corenyc.com/wp-content/uploads/2023/12/Property-Club-Jessica-Silver-June-13-2019-1.pdf>

<https://blog.dwellworks.com/brooklyn-is-it-actually-cheaper-or-better2>