



Airbnb & Chicago Crime Data Study

Final Presentation

Team 3

Colleen Jung

Kevin Ko

Khushi Ranganatha

Saphir Qi

Agenda

- I. Business Case Recap
- II. Data Models
- III. Exploratory Data Analysis
- IV. Lessons Learned & Recommendations



Business Case Recap



Data Sources

- Airbnb Listing Data
- Chicago Crime Rate

Our Focus

- Analyze factors affecting Airbnb pricing in Chicago, including property type, location, license status, and user reviews
- Explore the impact of local crime rates on Airbnb listing prices in Chicago

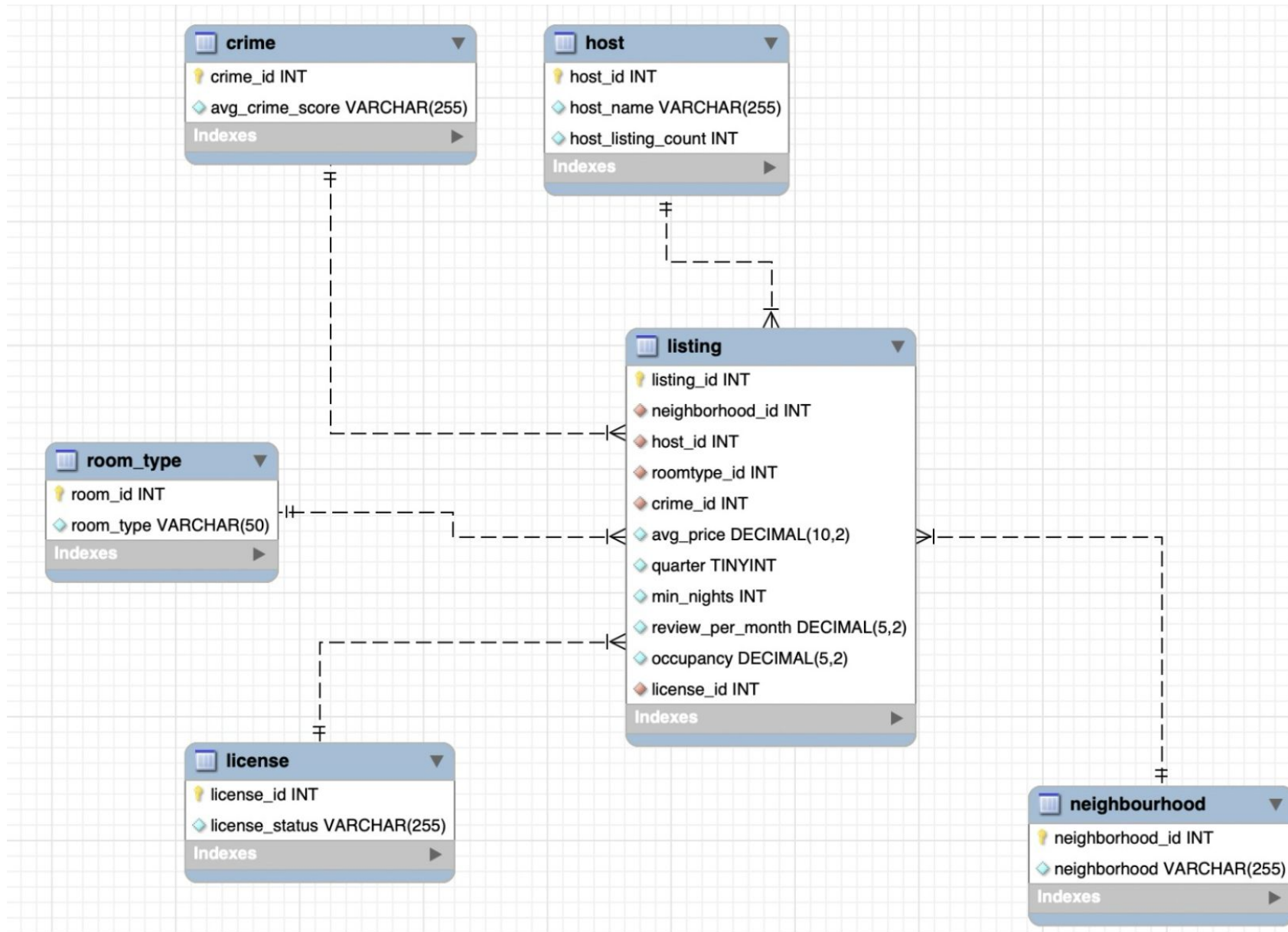


Agenda

- I. Business Case Recap
- II. Data Models
- III. Exploratory Data Analysis
- IV. Lessons Learned & Recommendations



Data Model - EER



Data Transformation - Listing Data

Example 1 - License

Original Data

license
2209984
2446868
NaN
R20000055258
R17000015589
City registration pending
398144 Chicago: R190000050938
2446867

Normalized Data

license_id	license_status
1	Licensed
2	No License
3	Pending

Licenses status are **categorized into 3 groups**:

- With a license number -> *Licensed*
- NaN -> *No License*
- City registration pending -> *Pending*

Example 2 - Room Type

Original Data

room_type
Shared room
Private room
Hotel room
Entire home/apt
Entire home/apt
Entire home/apt
Private room
Entire home/apt

Normalized Data

room_type	room_type_id
Private room	1
Entire home/apt	2
Shared room	3
Hotel room	4

All the room types are **categorized into 4 groups** and an unique `room_type_id` is assigned for each of them



Data Transformation - Crime Data

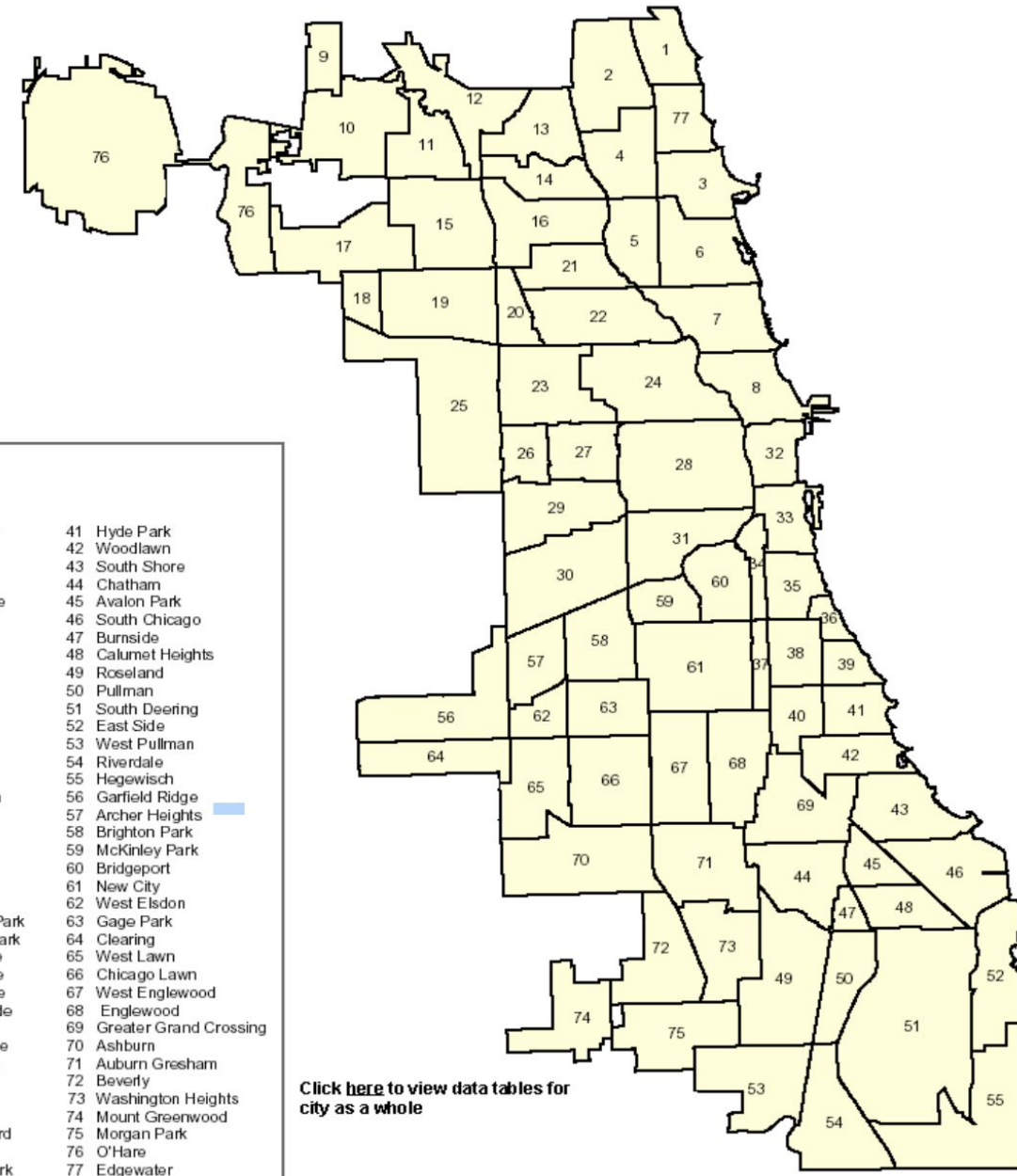
Example 1 - Neighborhood

neighborhood_id	neighborhood
1	0000X W WACKER DR
2	023XX N LOCKWOOD AVE
3	055XX S ROCKWELL ST
4	045XX N ALBANY AVE
5	087XX S SANGAMON ST

Example 2 - Community Area

community_area	name
1	Rogers Park
2	West Ridge
3	Uptown
4	Lincoln Square
5	North Center

- Defined to get a sense of 'community_area'
- Assign 'neighbourhood_id' that falls in unique community code



[Click here to view data tables for city as a whole](#)

Data Transformation -Crime Data

Crime Type

ID	Primary Type	year	quarter	neighbourhood
12876758	BATTERY	2022	4	Uptown
12876833	THEFT	2022	4	Rogers Park
12877365	CRIMINAL DAMAGE	2022	4	Auburn Gresham
12877003	THEFT	2022	4	Grand Boulevard
12877755	THEFT	2022	4	Near South Side
12879013	THEFT	2022	4	Lincoln Park
12877976	THEFT	2022	4	Chicago Lawn
12877327	BATTERY	2022	4	Lincoln Park
12878616	OTHER OFFENSE	2022	4	Washington Park
12879098	BATTERY	2022	4	Washington Heights
12876457	MOTOR VEHICLE THEFT	2022	4	Roseland
12877395	ASSAULT	2022	4	Near West Side

Primary Type	Score
DECEPTIVE PRACTICE	4
BATTERY	7
THEFT	5
CRIMINAL DAMAGE	8
OTHER OFFENSE	2
MOTOR VEHICLE THEFT	5
ASSAULT	8
CRIMINAL SEXUAL ASSAULT	8
BURGLARY	6
ROBBERY	6
PUBLIC PEACE VIOLATION	2
WEAPONS VIOLATION	7
...	...

Score	Score_avg
7	5.82
5	5.86
8	6.12
5	5.9
5	5.88
5	5.6
5	6.06
7	5.6
2	5.96
7	6.0
5	6.13
8	5.69



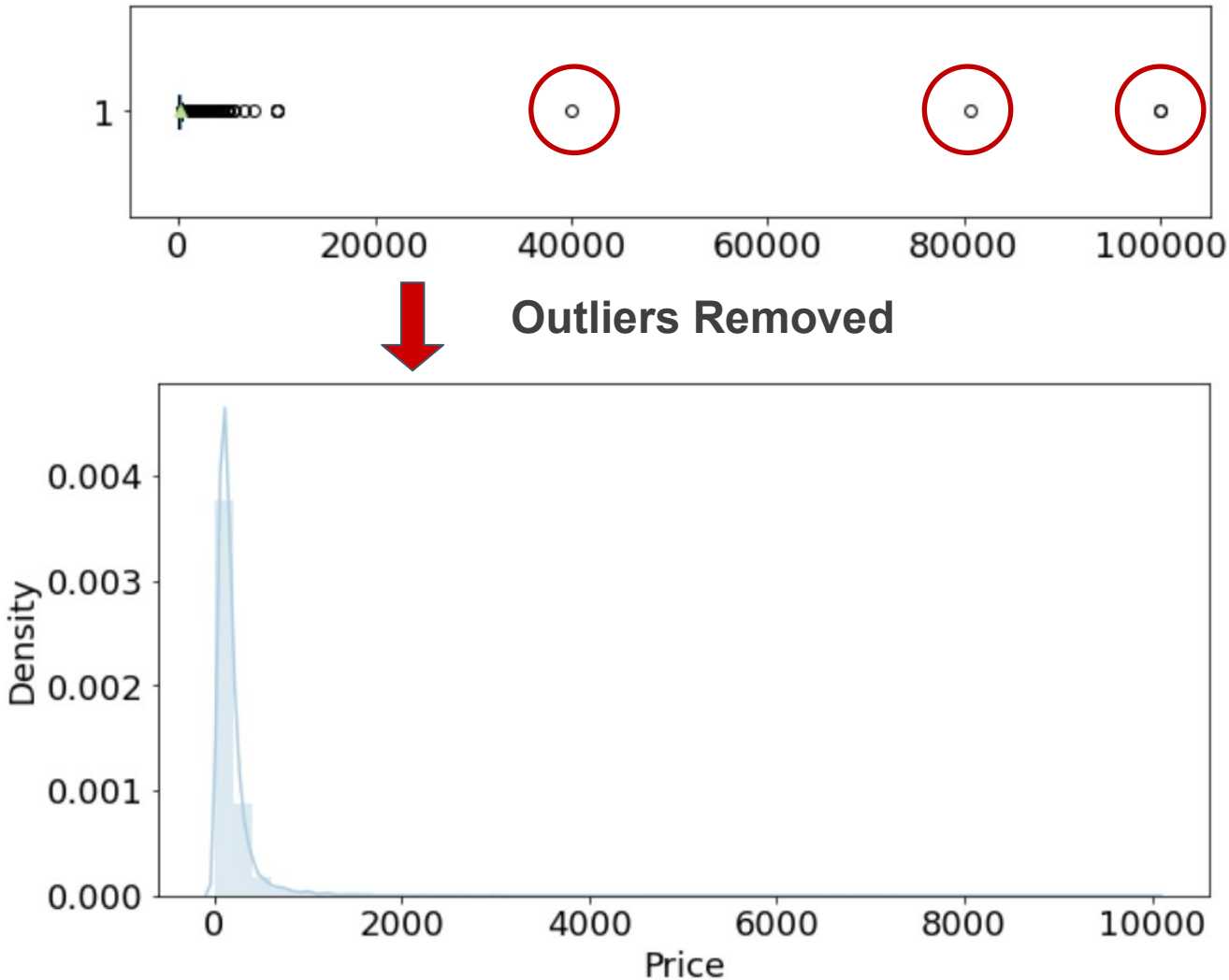
Agenda

- I. Business Case Recap
- II. Data Models
- III. Exploratory Data Analysis
- IV. Lessons Learned & Recommendations



Exploratory Data Analysis

Listing Price Distribution



Factsheet

Max Listing Price	\$99,998
Min Listing Price	\$0
Average Listing Price	\$199
Median Listing Price	\$130
Standard Deviation	\$974

Key Takeaways

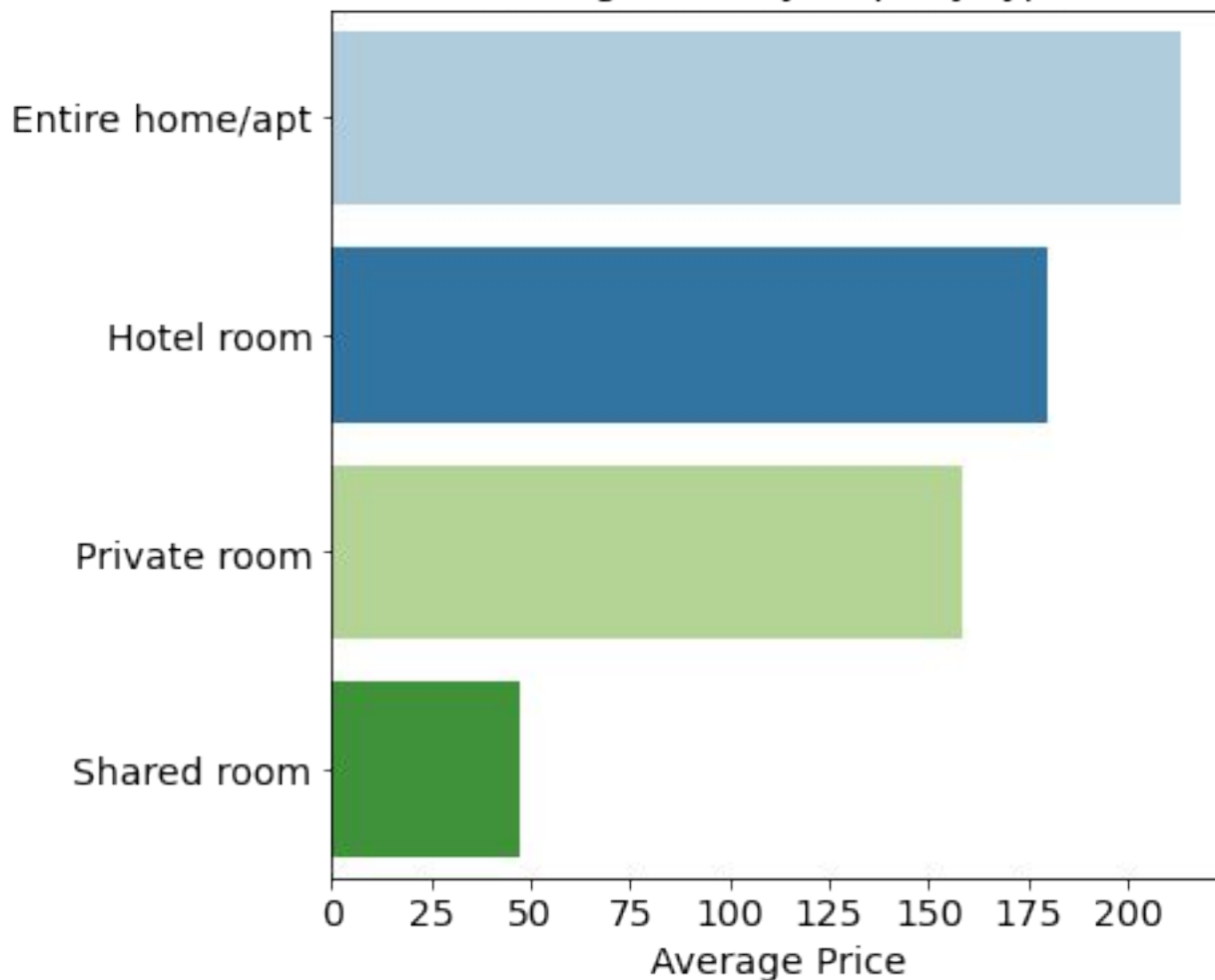
- **Price Distribution Skewness:** Right-skewed distribution with a concentration of listings at the lower price end, indicating a market dominated by more affordable rental options.
- **Variation in Listing Price:** The wide range of listing prices indicates a diverse market from budget to luxury accommodations.
- **Outliers:** The outliers on the higher end of the spectrum hint at the existence of premium listings



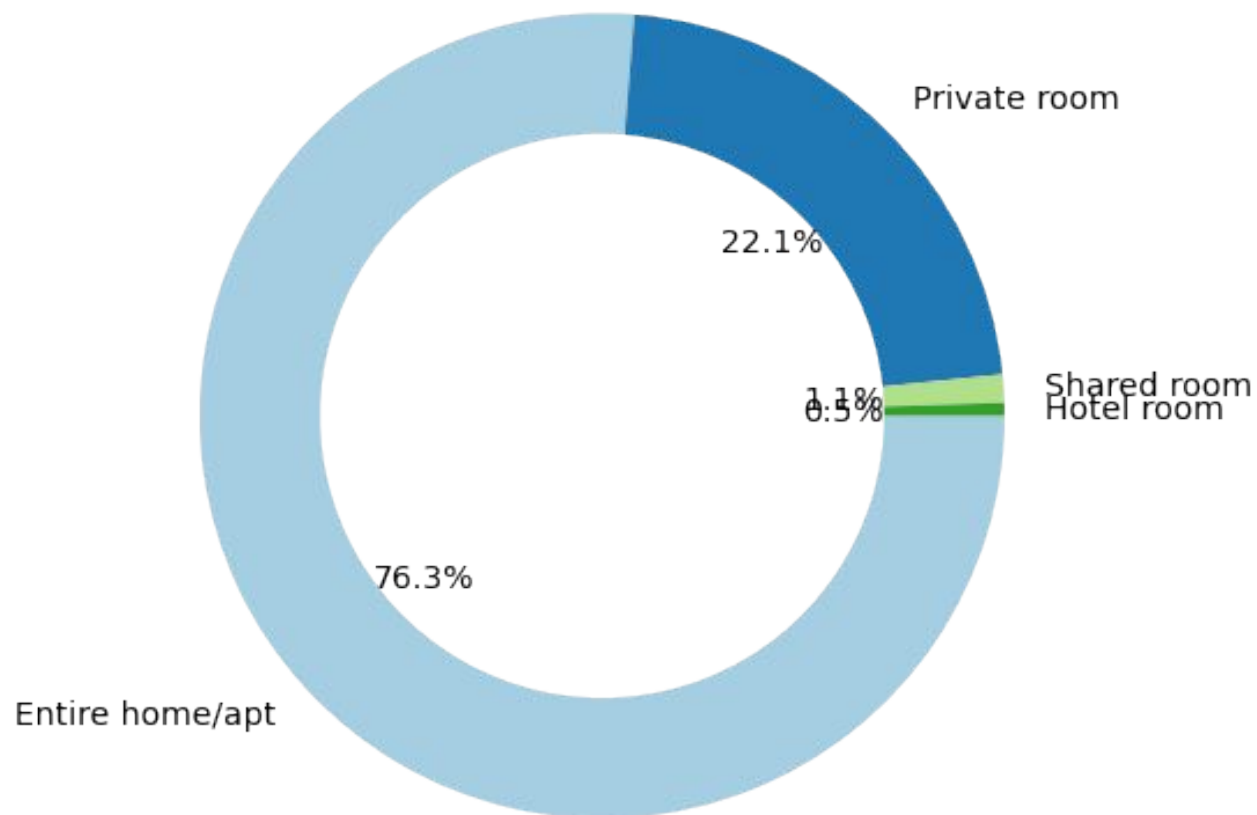
Exploratory Data Analysis

Property Type

Average Price by Property Type

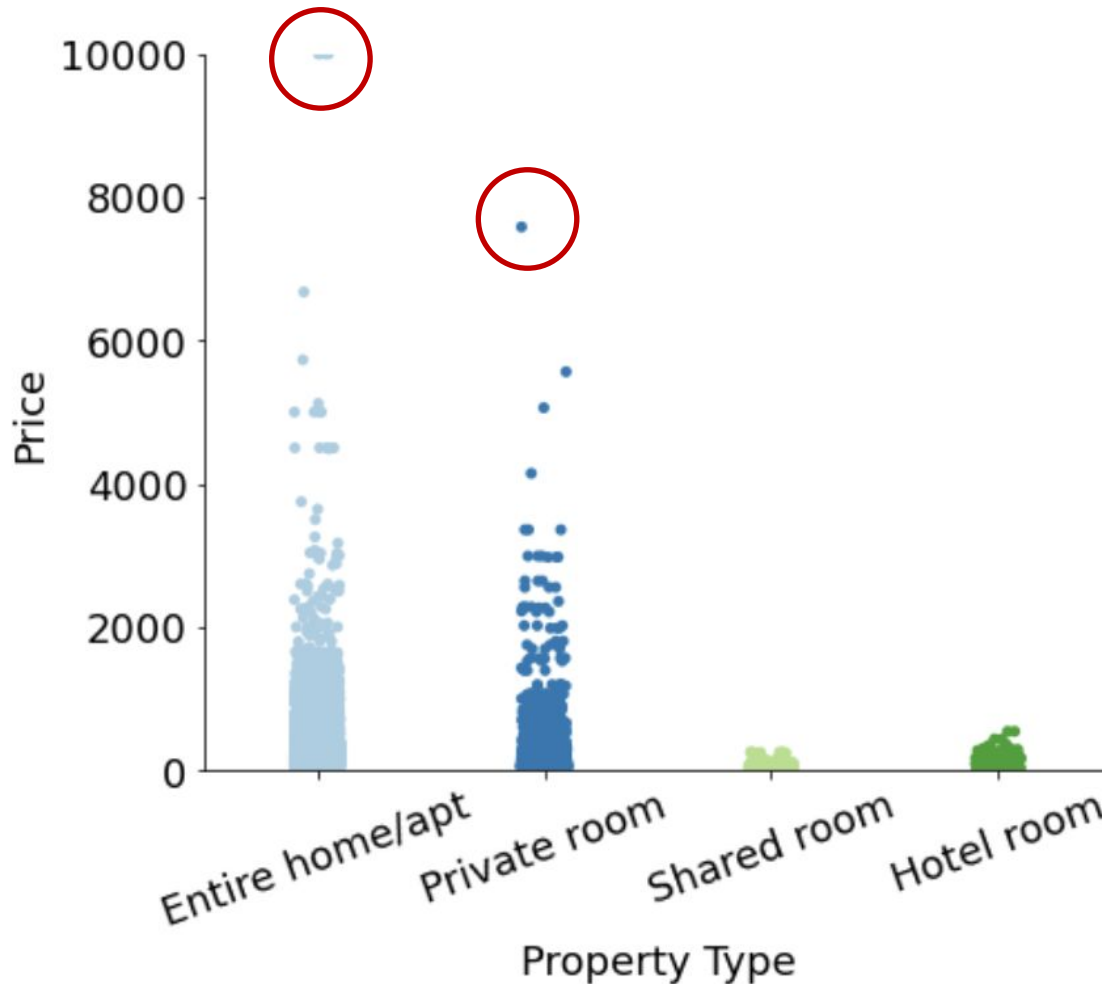


Percentage by Property Type



Exploratory Data Analysis

Property Type

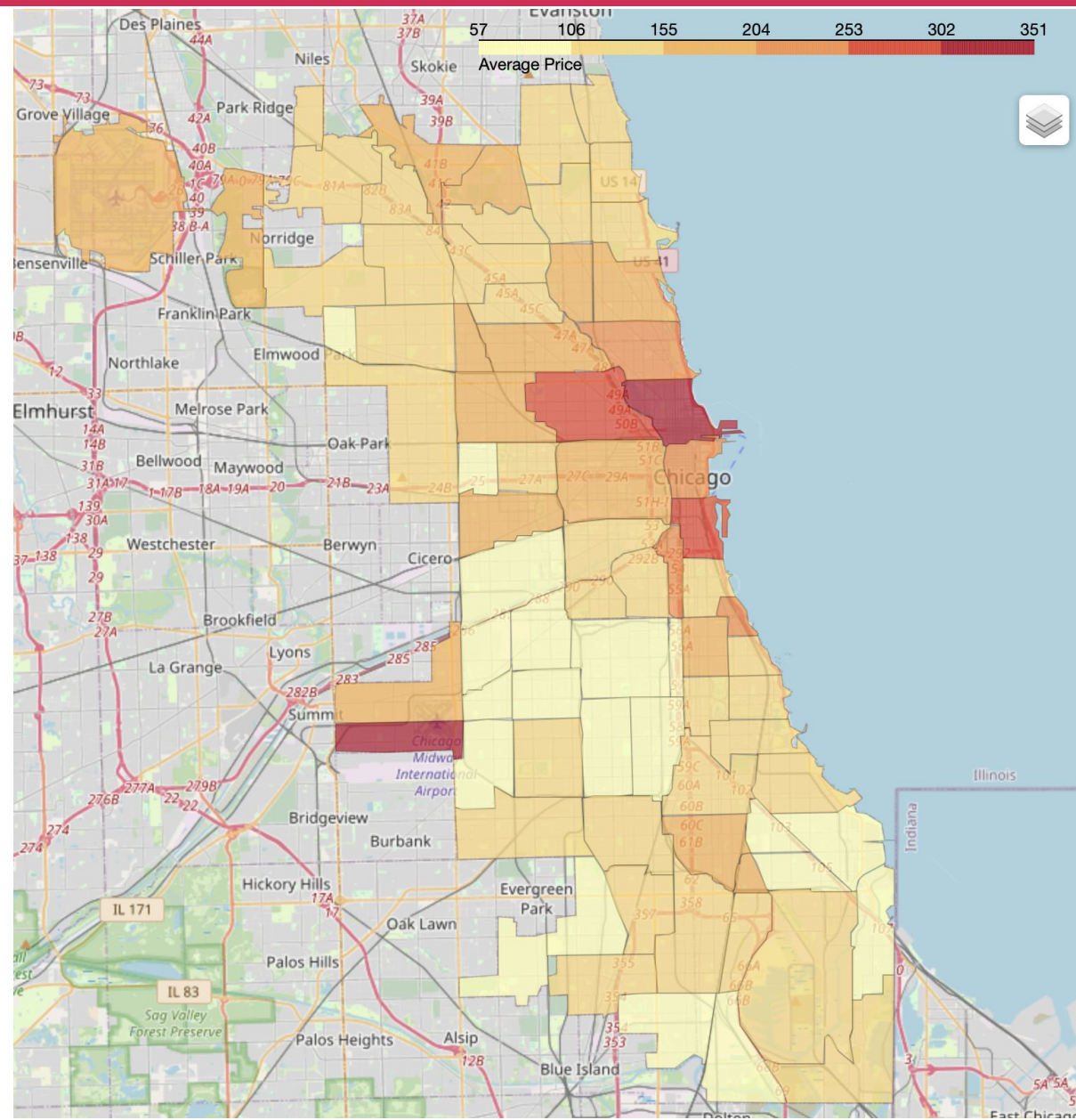
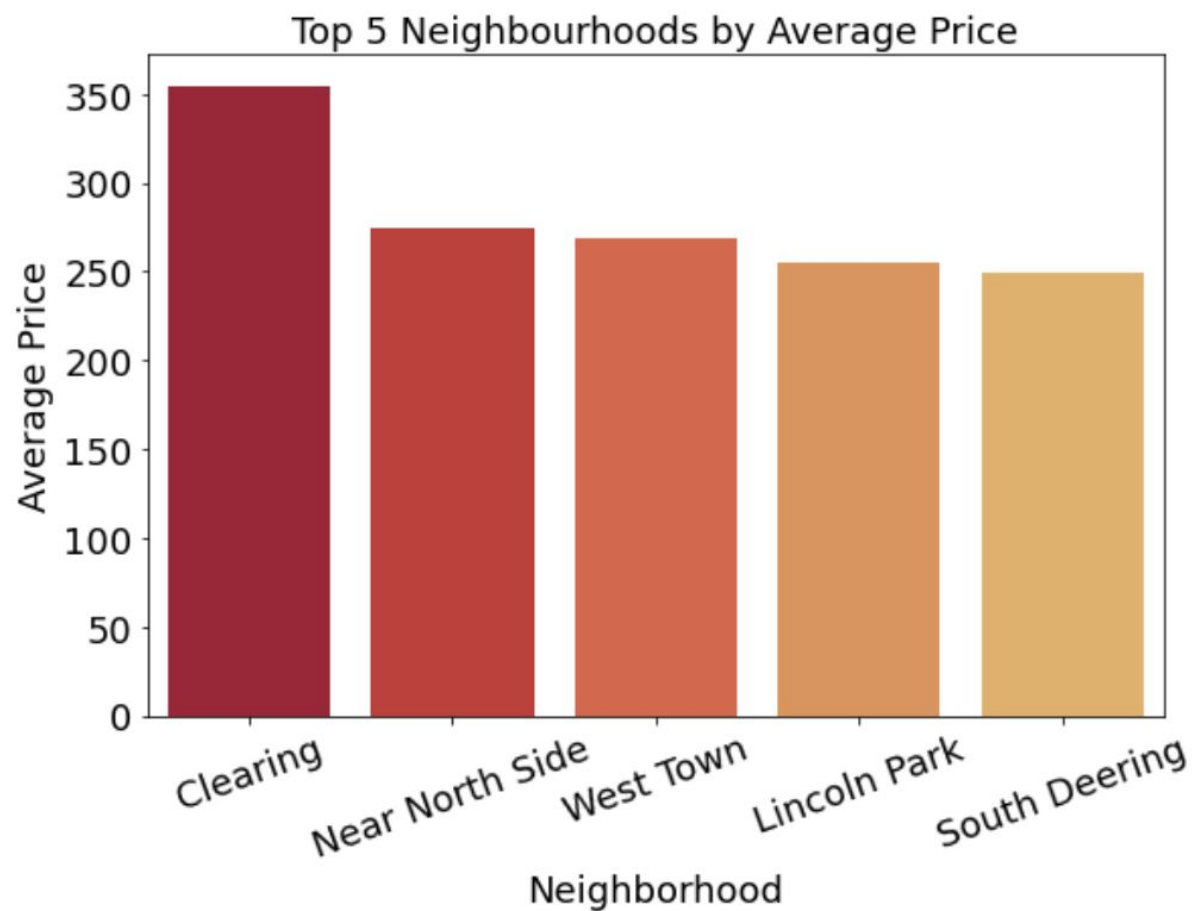


Key Takeaways

- **Property Type Preference:** The majority of listing supply are entire homes/apartments, driven by the higher demand for privacy and space.
- **Average Price by Property Type:** Entire homes/apartments have the highest average price, followed by hotel rooms and then private rooms, suggesting that space and exclusivity come at a premium.
- **Distribution of Prices:** There's a wide range of prices within each property type, especially with entire homes/apartments, as indicated by the significant outliers.

Exploratory Data Analysis

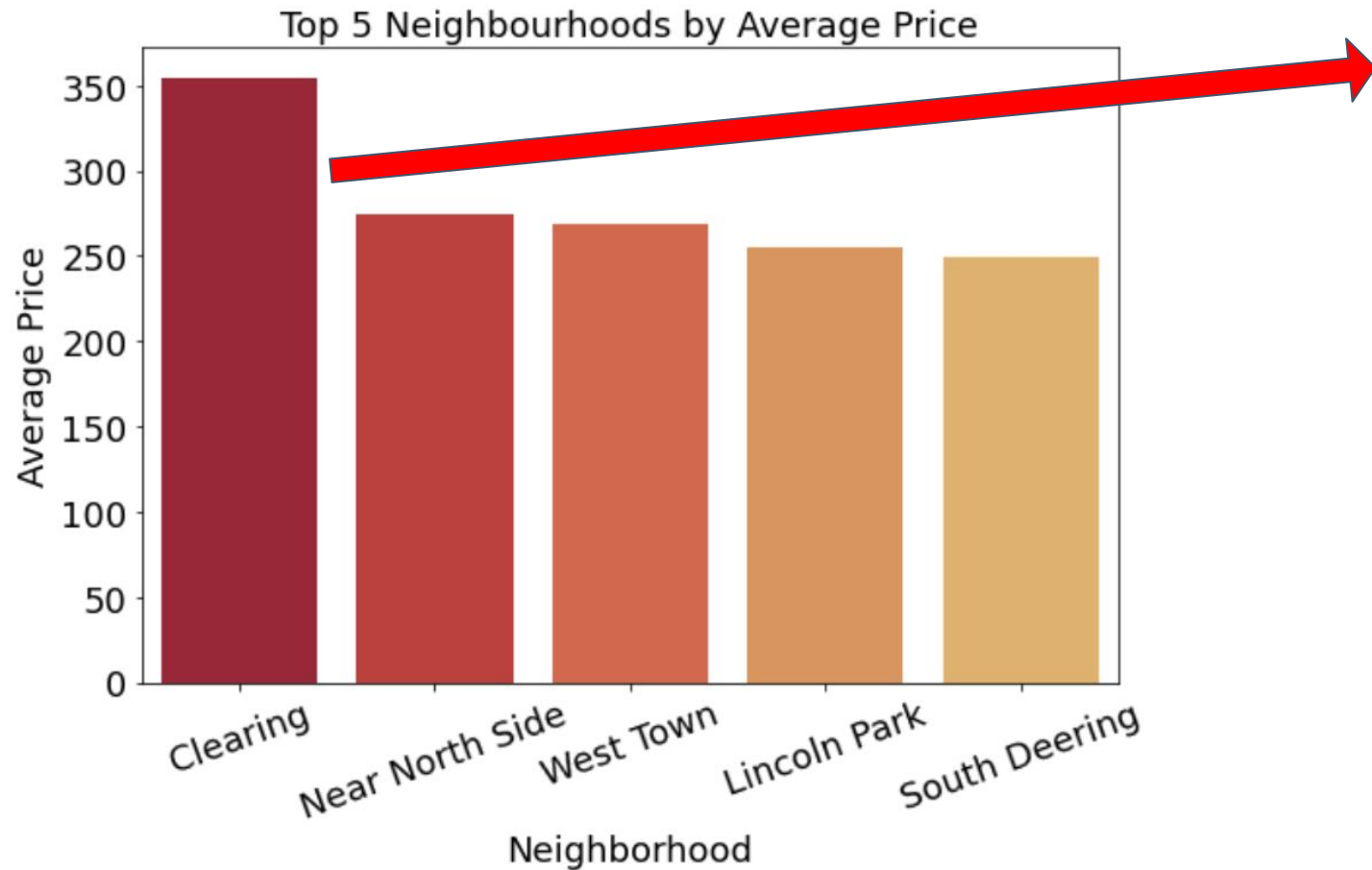
Spatial Data Analysis - Average Price By Neighborhood



Exploratory Data Analysis

Why Clearing is the most expensive neighbourhood?

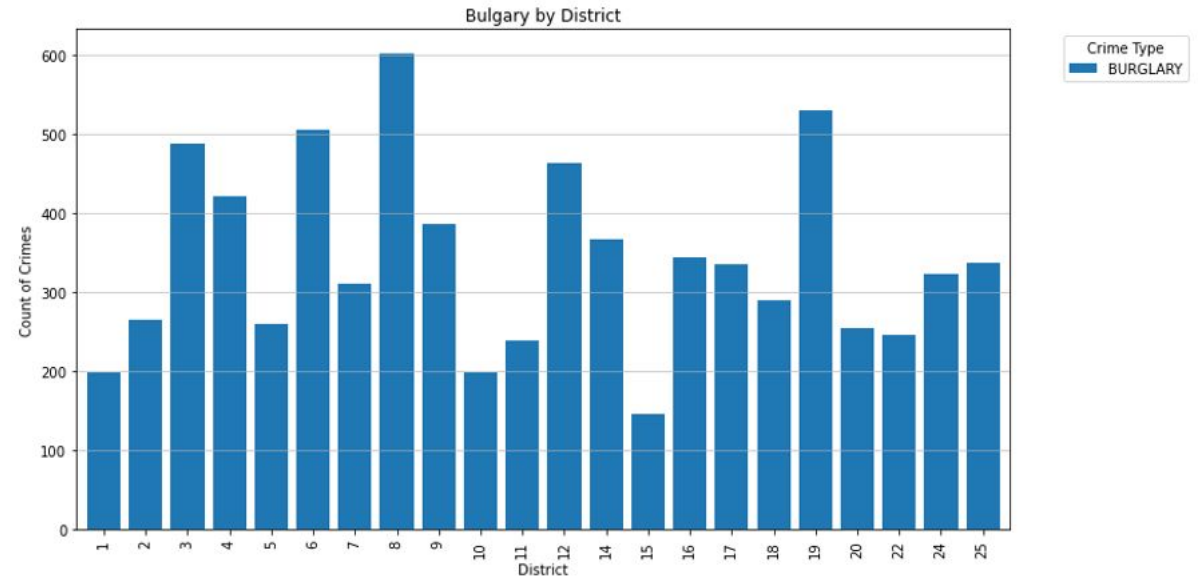
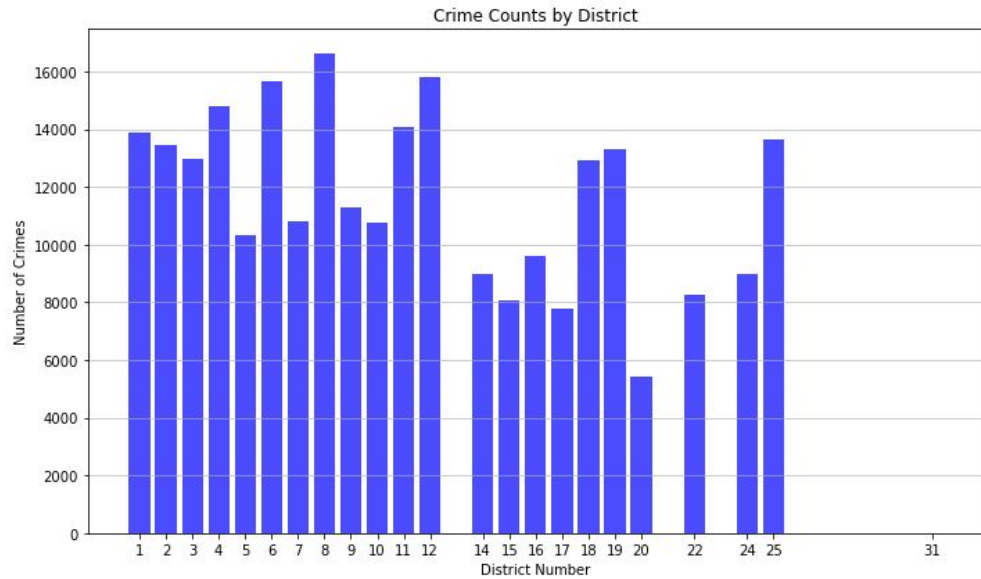
– Likely to be driven by one professional agency



name	host_name	neighbourhood	room_type	price
Home in Chicago · 5 bedrooms · 6 beds · 2 baths	Leslie	Clearing	Entire home/apt	549
Home in Chicago · 5 bedrooms · 7 beds · 2 baths	FreeStand Home Solutions	Clearing	Entire home/apt	549
Home in Chicago · 4 bedrooms · 5 beds · 2 baths	Erik	Clearing	Entire home/apt	549
Home in Chicago · ★New · 5 bedrooms · 7 beds · ...	Brandon	Clearing	Entire home/apt	549
Home in Chicago · ★New · 4 bedrooms · 5 beds · ...	Brandon	Clearing	Entire home/apt	549
Home in Chicago · 4 bedrooms · 5 beds · 2 baths	FreeStand Home Solutions	Clearing	Entire home/apt	549

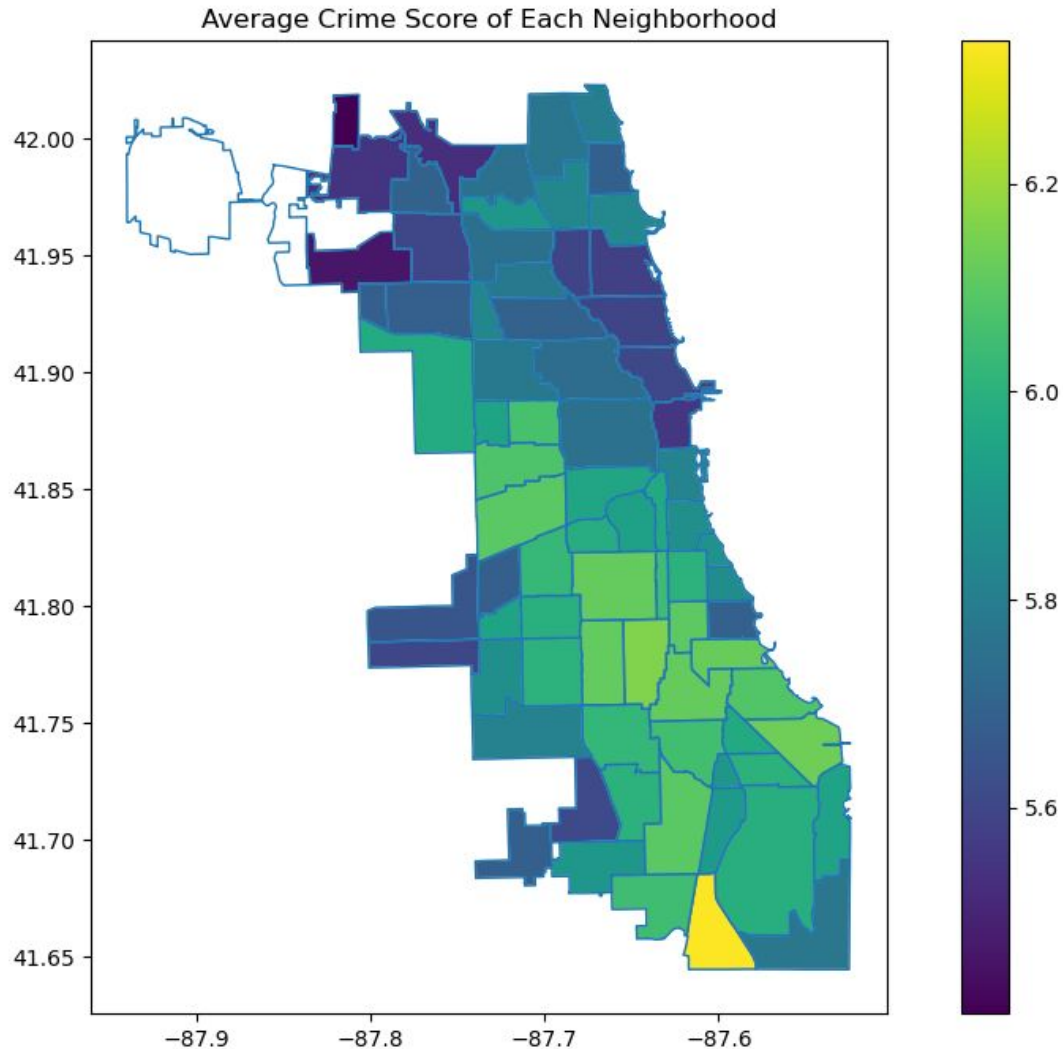


Exploratory Data Analysis



- **No significant information** between the total crime rate between the districts
- **District population data** is crucial when analyzing crime rates to gain a **more comprehensive** understanding of the local safety landscape to compare with Airbnb listing
- Create a numeric scale for crime types to calculate neighborhood-specific average crime rates relevant to Airbnb listing prices.

Exploratory Data Analysis

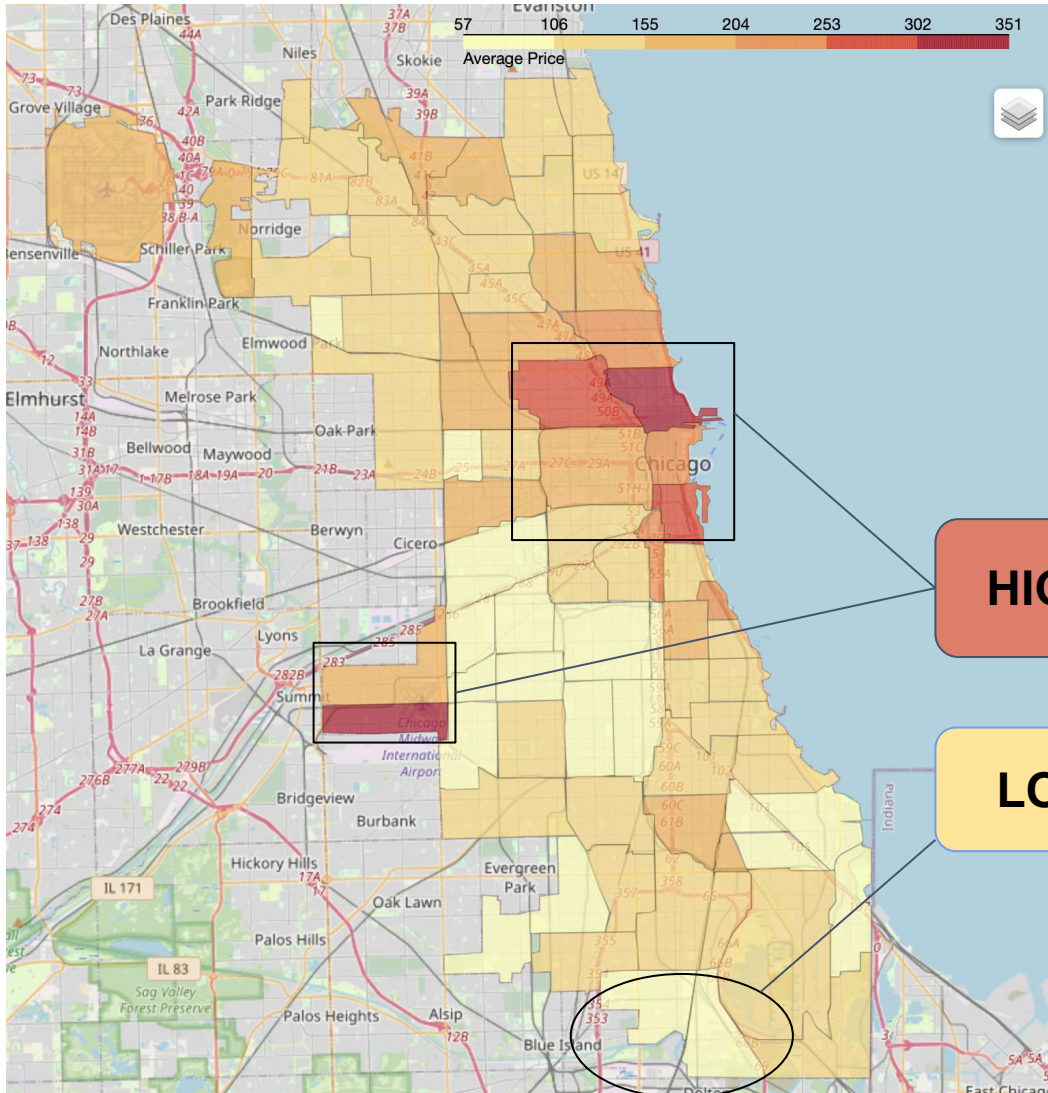


Key Observations

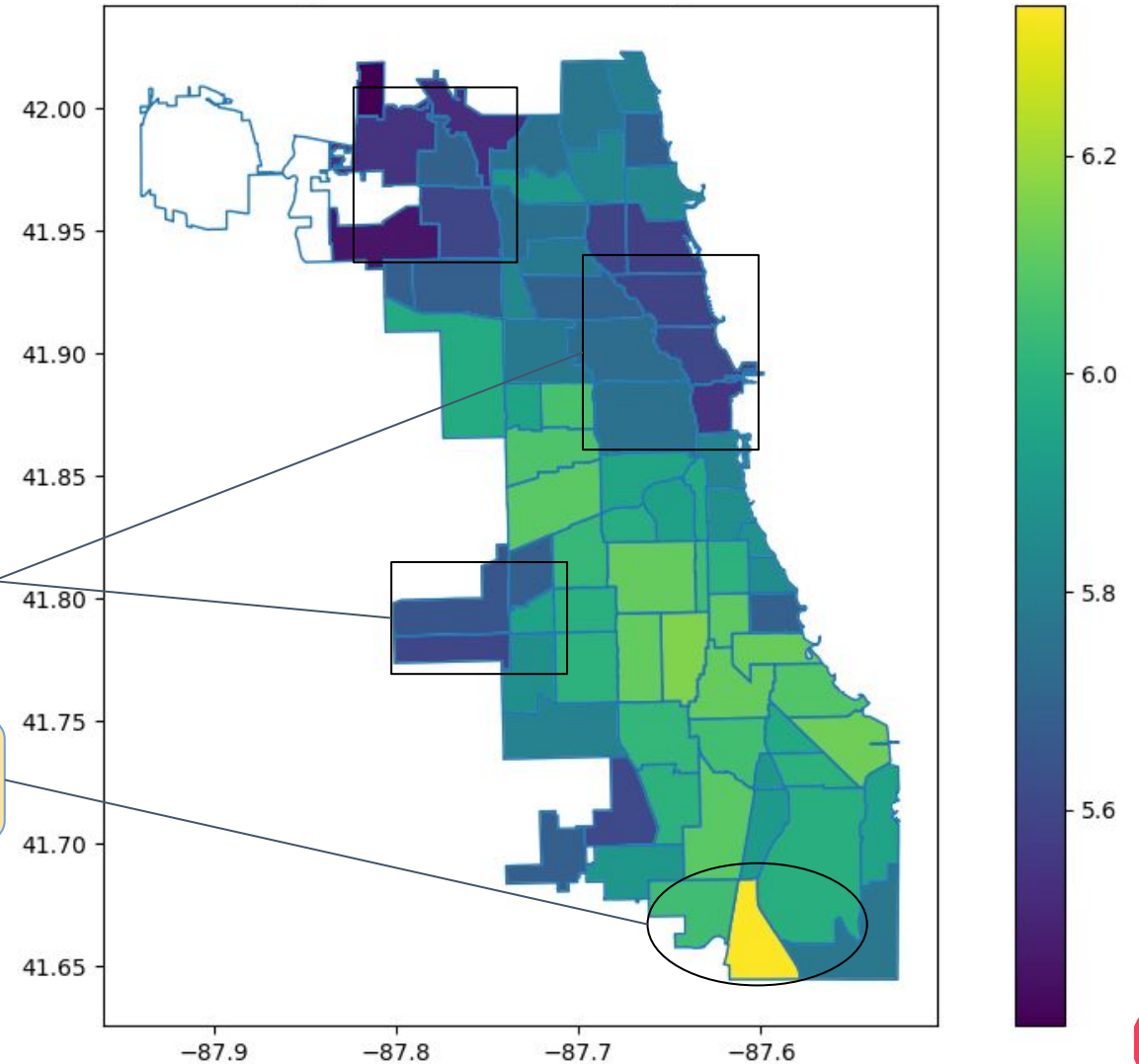
- The graph presents a heat map of Chicago's neighborhoods, color-coded by the **average crime score**
- There is a visible pattern where certain neighborhoods with **higher crime scores** are **associated with Airbnb housing pricing** trends.
- “Safest” Neighbourhood: Edison Park
“Most Dangerous” Neighbourhood: Riverdale

Summary by Neighborhood

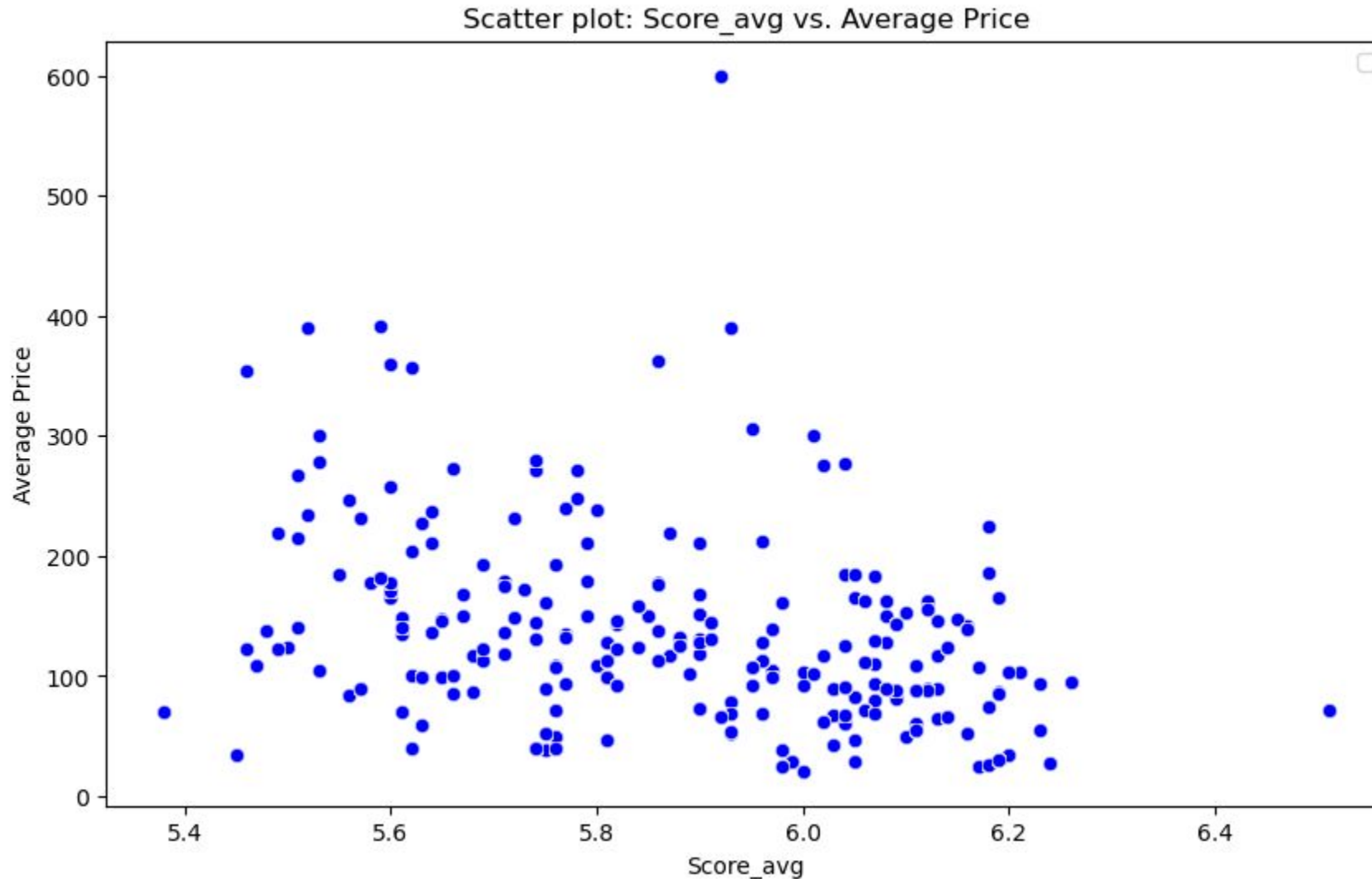
Average Price



Average Crime Score



Exploratory Data Analysis



- **Negative trend** between average cime score and average price
- Higher average listing price generally have lower average crime score
- Coefficient of -0.6

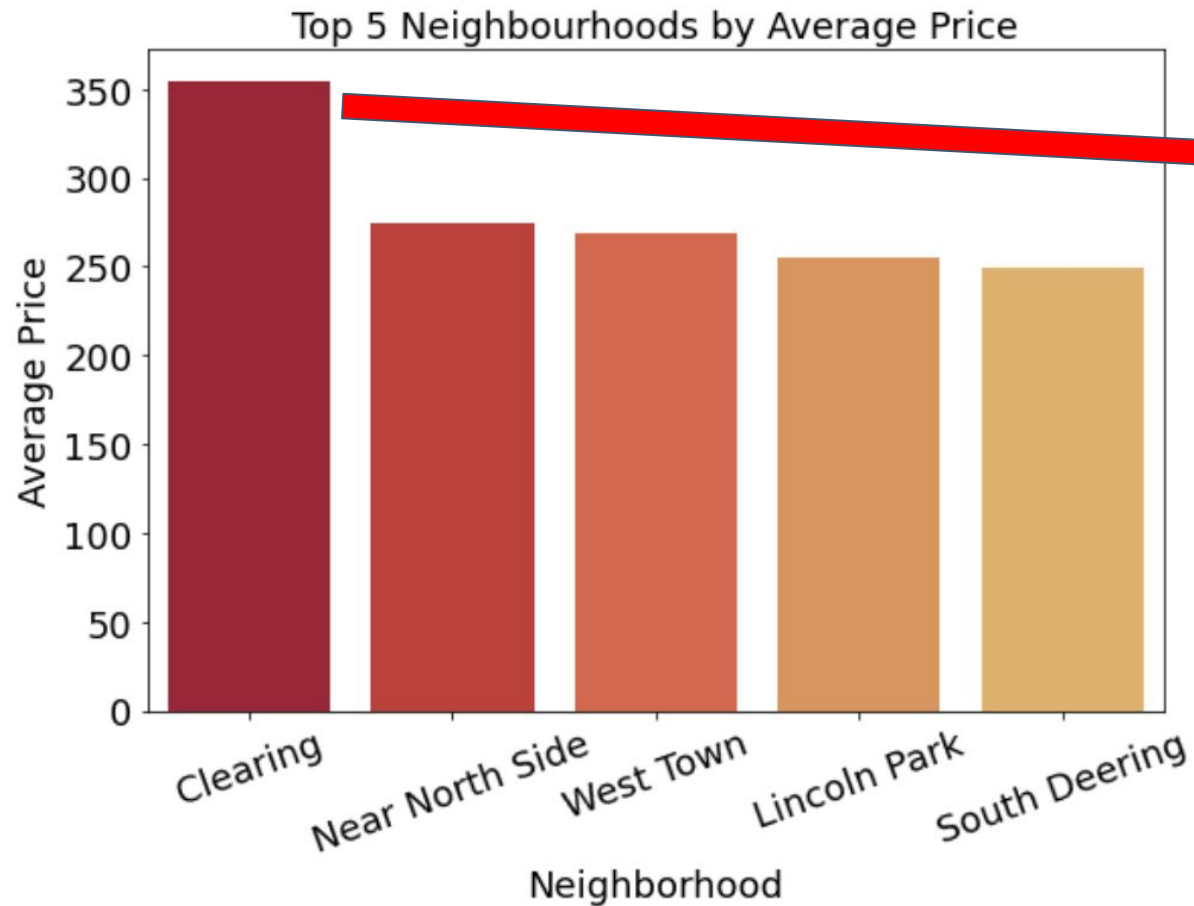
Agenda

- I. Business Case Recap
- II. Data Models
- III. Exploratory Data Analysis
- IV. Lessons Learned & Recommendations



Lessons Learned

- Data does not lie, but it can be misleading:



name	host_name	neighbourhood	room_type	price
Home in Chicago · 5 bedrooms · 6 beds · 2 baths	Leslie	Clearing	Entire home/apt	549
Home in Chicago · 5 bedrooms · 7 beds · 2 baths	FreeStand Home Solutions	Clearing	Entire home/apt	549
Home in Chicago · 4 bedrooms · 5 beds · 2 baths	Erik	Clearing	Entire home/apt	549
Home in Chicago · ★New · 5 bedrooms · 7 beds · ...	Brandon	Clearing	Entire home/apt	549
Home in Chicago · ★New · 4 bedrooms · 5 beds · ...	Brandon	Clearing	Entire home/apt	549
Home in Chicago · 4 bedrooms · 5 beds · 2 baths	FreeStand Home Solutions	Clearing	Entire home/apt	549

Recommendations

1. For Airbnb Hosts

- **Competitive Pricing:** Encourage hosts in high crime rate areas to adjust their pricing to remain competitive.
- **Safety Features:** Advise hosts to invest in additional safety features and to clearly communicate these enhancements to potential guests.

1. The City of Chicago

- **Policy Changes:** Advocate for policy changes that support increased safety measures in neighborhoods with lower Airbnb activity
- **Collaboration with Airbnb:** Work alongside Airbnb to develop community programs that address guests' safety and support local law enforcement efforts

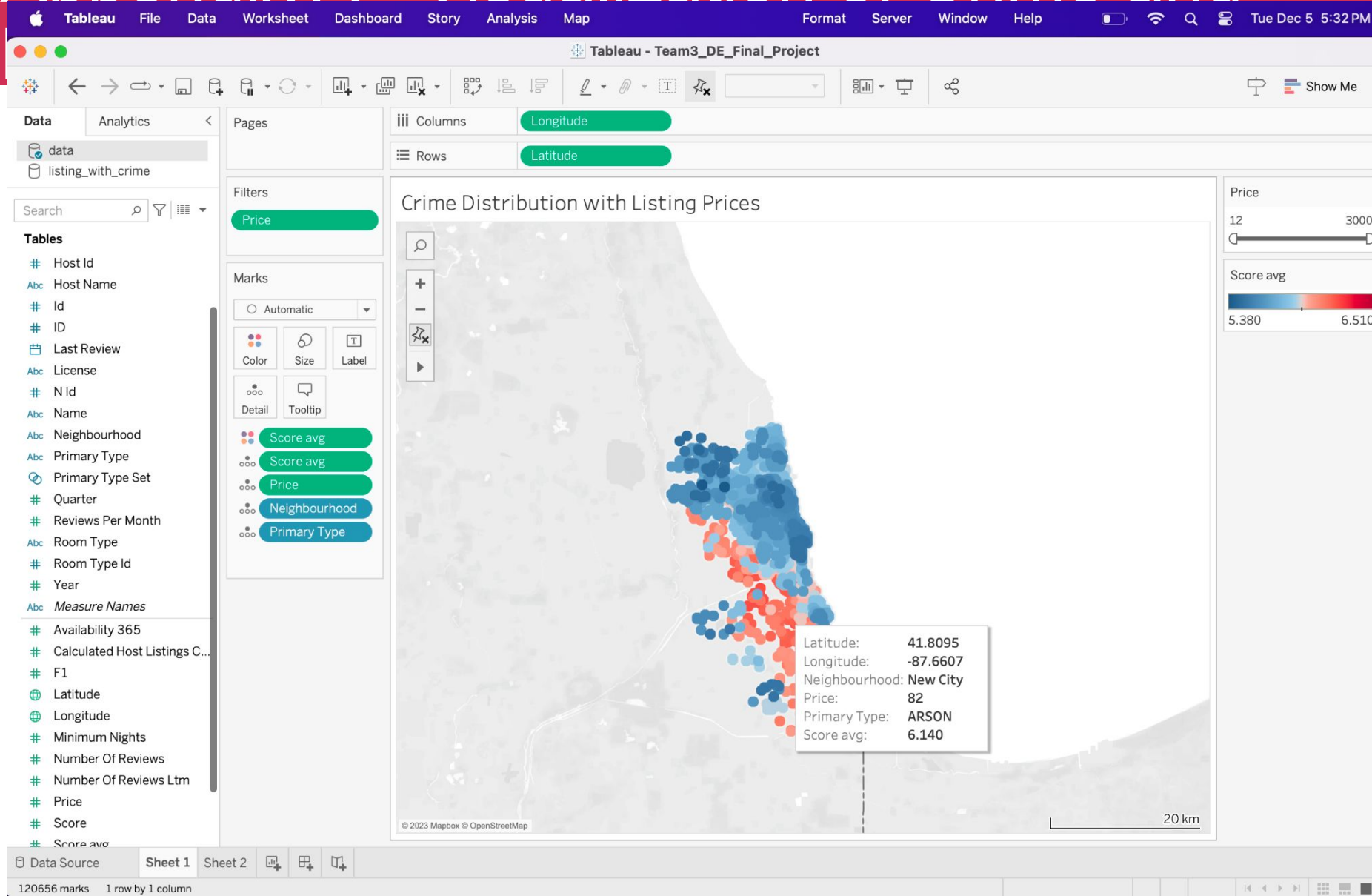
1. Further Research

- **Broader Analysis:** Expand the scope of research to consider additional factors like local amenities, seasonal trends, and economic conditions.
- **Comparative Studies:** Conduct studies in other cities to provide comparative insights and develop a more comprehensive understanding of the dynamics between crime rates and short-term rental prices.





Appendix A - Visualization of Crime and



Appendix A - Visualization of Crime and Listing Data (Listing Price at Modified Scale)

