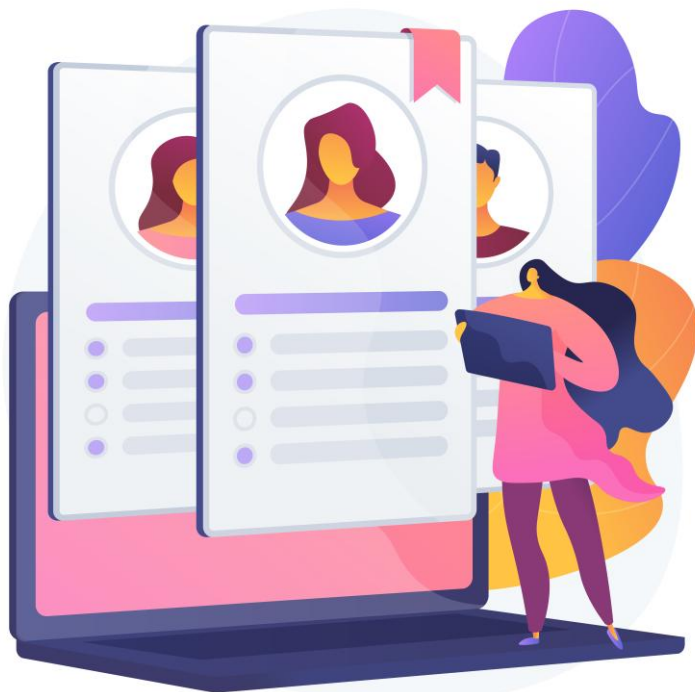# AIRBNB PRICE PREDICTION

Charlene Vu

Introduction

Short-term rental platforms like Airbnb have transformed the urban housing market.

Explore how different features such as amenities, location, and listing descriptions affect pricing.

Implications for stakeholders:

- Help hosts optimize their strategies

- Assist renters in identifying fair prices

- Offer platforms deeper insights into market behavior

Dataset

InsideAirbnb:

https://insideairbnb.com/get-the-data/

- Dataset is updated every quarter; covers the most recent 12 months (currently 36403 listings, 18 variables)

- Drop irrelevant columns: id, name, host_id, host_name, last_review, license

- Remove extreme outliers (> $1000 per night)

id	name	host_id	host_name	neighbourl	neighbourl	latitude	longitude	room_type	price	minimum_	number_of	last_review	reviews_pe	calculated	availability	number_of	license
2539	Superfast	2787	John	Brooklyn	Kensington	40.64529	-73.9724	Private roc	260	30	9	10/19/2018	0.08	6	365	0	
2595	Skylit Stud	2845	Jennifer	Manhattar	Midtown	40.75356	-73.9856	Entire hom	240	30	47	6/21/2022	0.25	3	365	0	
6848	Only 2 stop	15991	Allen	Brooklyn	Williamsbu	40.70935	-73.9534	Entire hom	98	30	195	6/9/2025	0.99	1	1	3	
6872	Uptown Sa	16104	Kahshanna	Manhattar	East Harle	40.80107	-73.9426	Private room		30	1	6/5/2022	0.03	2	24	0	
6990	UES Beaut	16800	Cynthia	Manhattar	East Harle	40.78778	-73.9476	Private roc	74	30	249	5/27/2025	1.3	1	288	6	
7064	Amazing lc	17297	Joelle	Brooklyn	Williamsbu	40.71248	-73.9588	Private room		30	13	9/12/2022	0.07	2	0	0	
7097	Perfect for	17571	Jane	Brooklyn	Fort Green	40.69194	-73.9739	Private roc	214	2	416	7/21/2025	2.2	2	8	38	OSE-STRRI
7801	Sunny Will	21207	Chaya	Brooklyn	Williamsbu	40.71881	-73.9562	Entire hom	345	30	12	7/31/2024	0.06	1	0	0	
8490	Maison de:	25183	Nathalie	Brooklyn	Bedford-St	40.68456	-73.9396	Entire hom	170	30	189	10/16/2023	0.97	2	246	0	
9357	Midtown P	30193	Tommi Lau	Manhattar	Hell's Kitcl	40.76724	-73.9866	Entire hom	175	30	58	8/13/2017	0.3	1	57	0	
10452	Radiant Oa	35935	Angela	Brooklyn	Bedford-St	40.68294	-73.9568	Private roc	90	30	82	3/31/2025	0.44	6	183	2	
11943	Country sp	45445	Harriet	Brooklyn	Flatbush	40.63702	-73.9633	Private room		30	0			1	0	0	
12192	ENJOY Dov	46978	Edward	Manhattar	East Villag	40.72296	-73.9838	Private room		30	314	1/1/2025	1.64	2	15	1	
12940	Charming	50148	Adreinne	Brooklyn	Bedford-St	40.67946	-73.9542	Entire hom	153	30	80	5/31/2025	0.42	1	327	4	
14314	Greenpoin	56246	Jeanne	Brooklyn	Greenpoin	40.73535	-73.9558	Entire hom	162	30	176	9/30/2024	0.92	1	240	1	
15341	* Lovely Nc	60049	Ted	Manhattar	Lower Eas	40.72136	-73.9937	Entire hom	180	30	42	7/1/2025	0.29	1	241	4	
15396	Sunny & Sr	60278	Petra	Manhattar	Chelsea	40.74493	-73.9952	Entire home/apt		180	5	11/3/2018	0.04	1	61	0	

EDA Visualization

Price Distribution

Distribution of Airbnb Prices

EDA Visualization
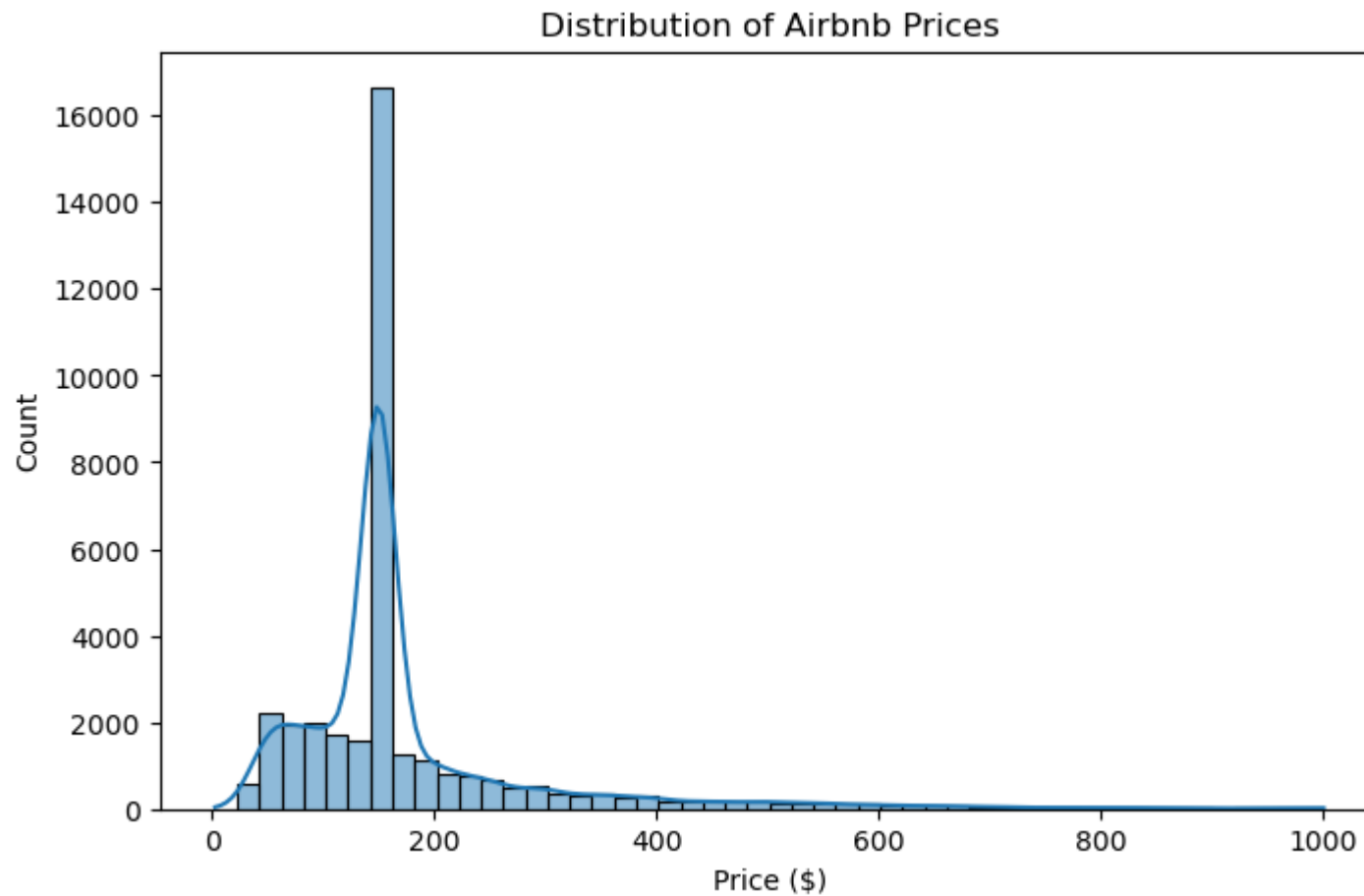
Price Distribution by Room Type

Price Distribution by Room Type
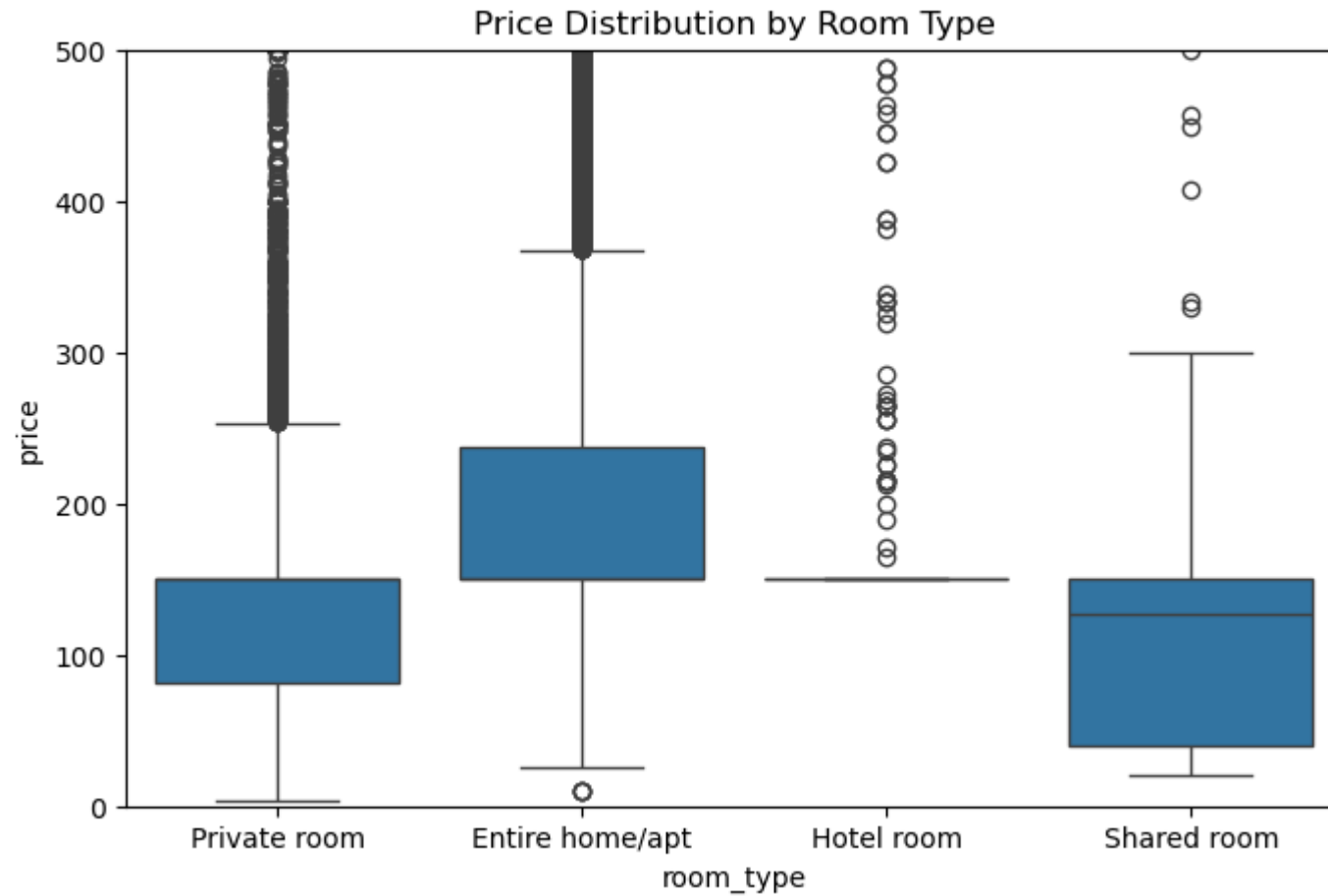
EDA Visualization

Price by Neighbourhood Group

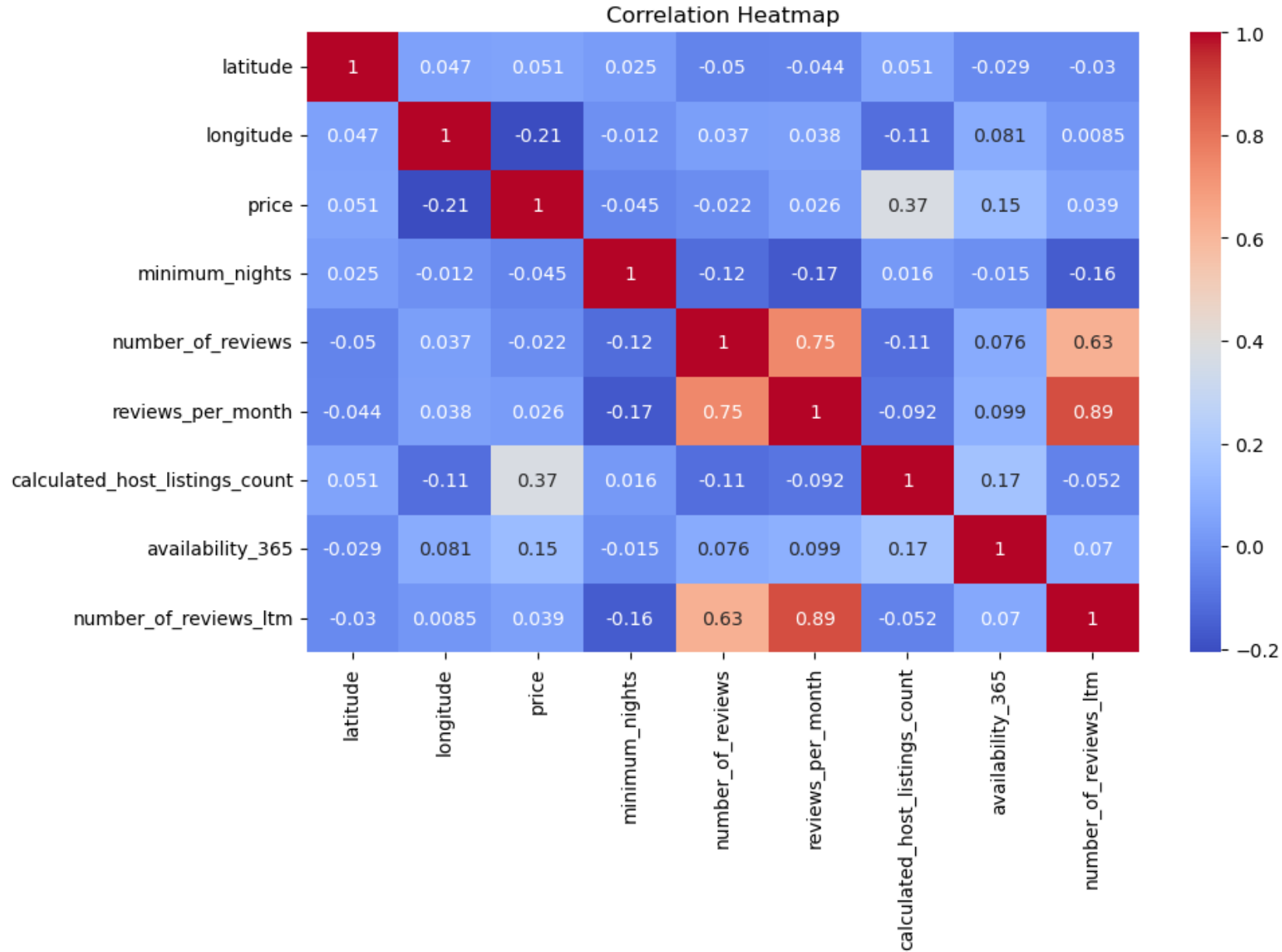Median Price by Neighbourhood Group
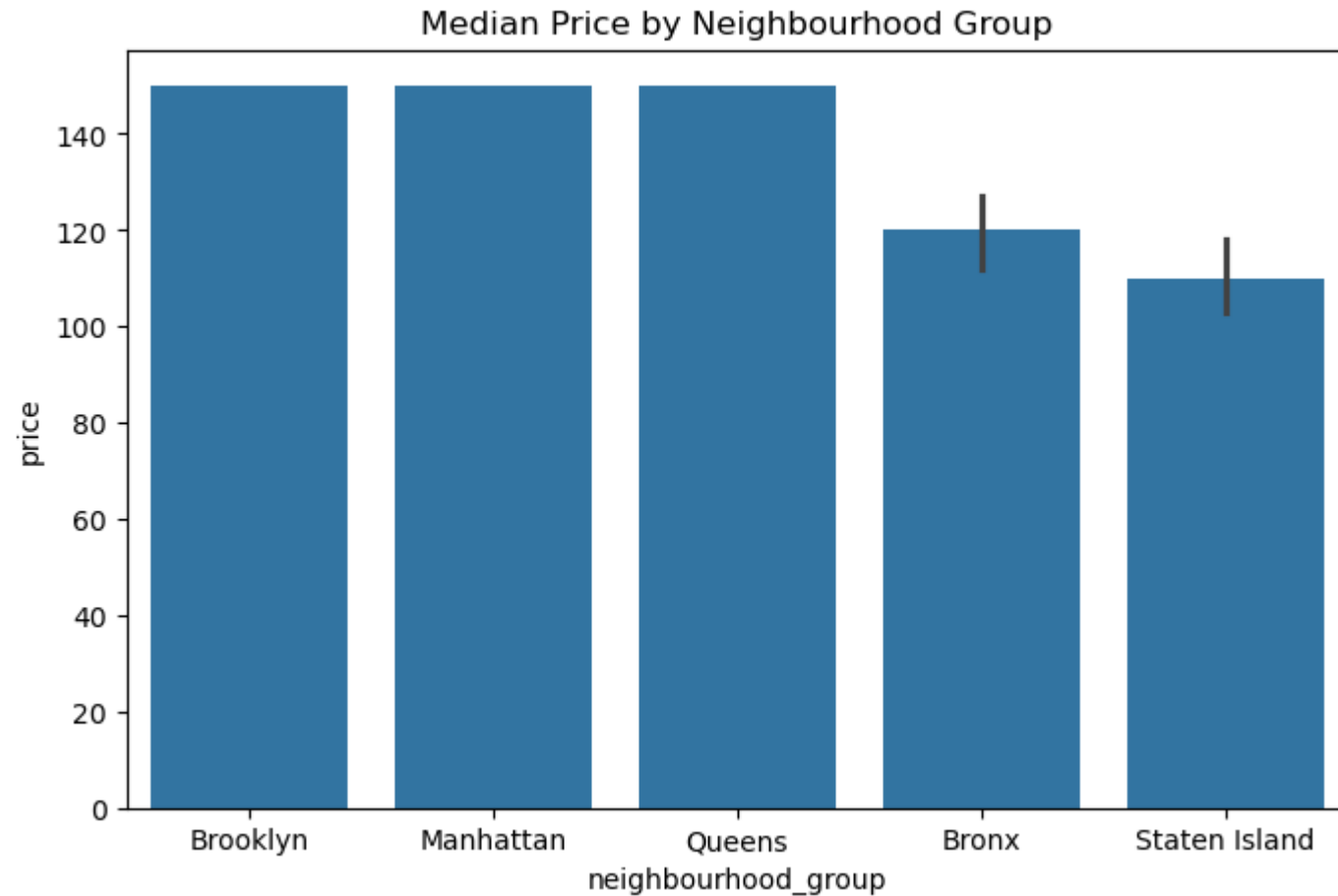
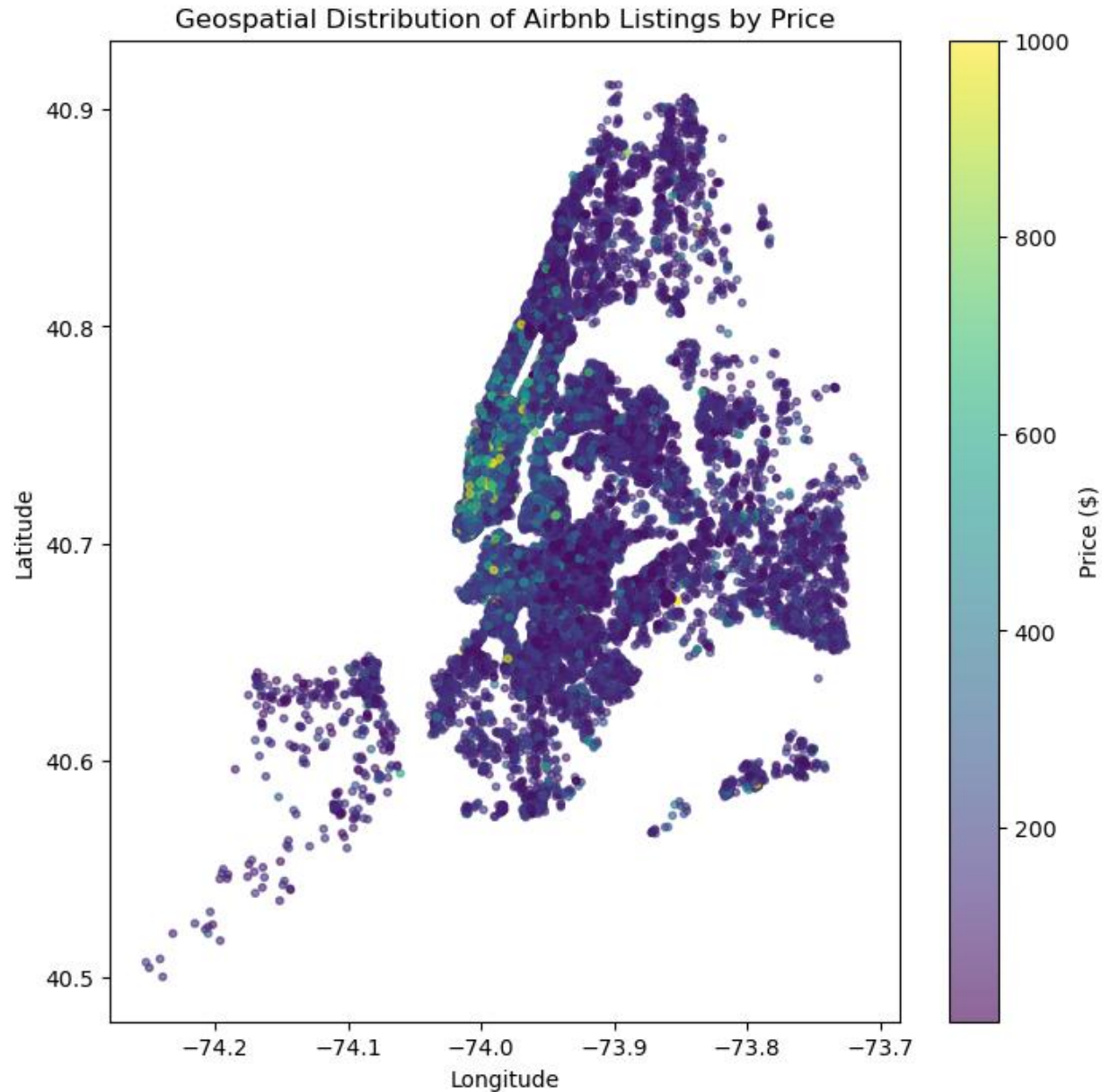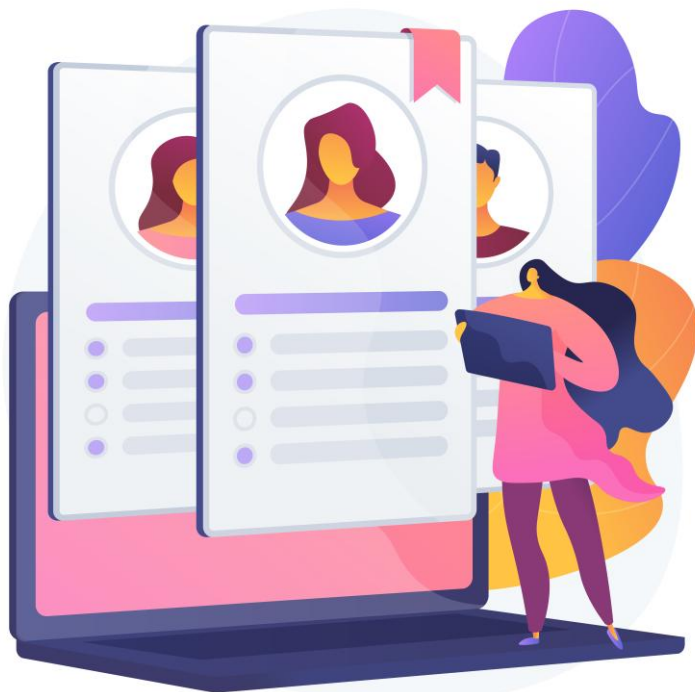# EDA Visualization

Reviews vs Price

Reviews vs Price

Correlation Heatmap

Geospatial Scatter Plot (Listings Colored by Price)


Geospatial Distribution of Airbnb Listings by Price
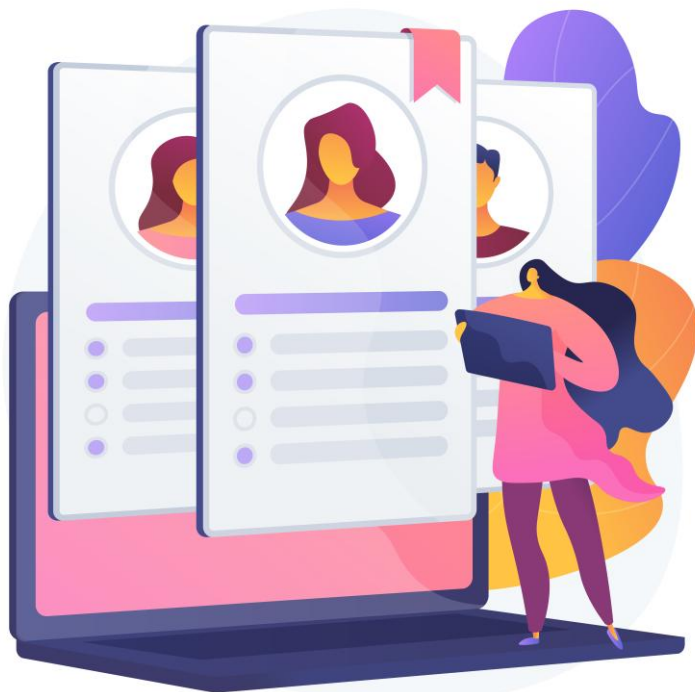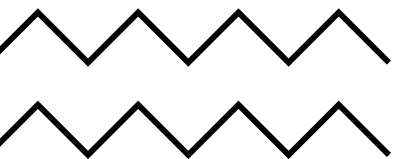
Ethical Implications

- **Privacy & personal data:** only using publicly available listing information, not private host or guest data

- **Fairness & bias:** cautious about using neighborhood or property features to infer economic or social characteristics

- Example: mapping prices by neighborhood could unintentionally reinforce stereotypes or contribute to gentrification debates
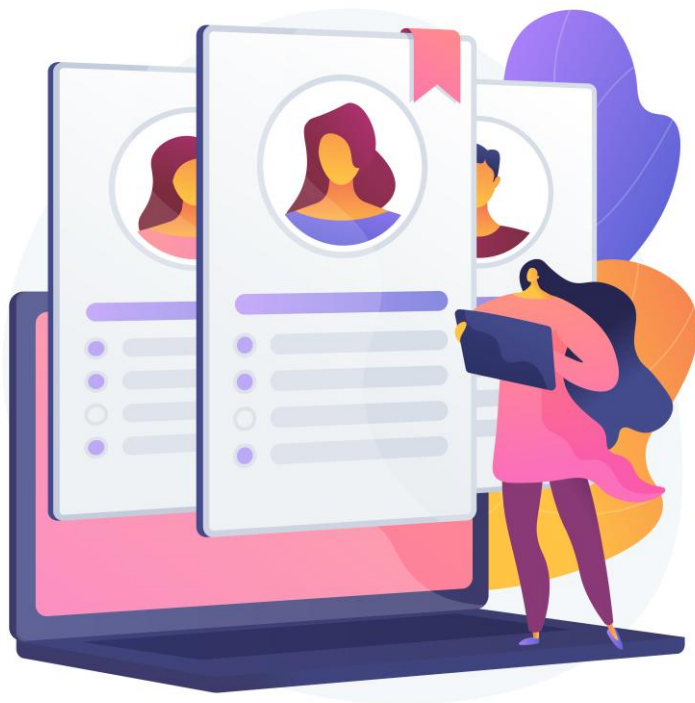
Legal Implications

- **Terms of Service:** Airbnb generally blocks automated scraping, especially at scale. Using Inside Airbnb datasets is safer from a legal perspective, because they are curated for research use

- **Data usage rights:** Use the data only for academic research purposes, not commercial gain. Properly cite external datasets (Inside Airbnb)

Societal Implications

-   **Urban housing & pricing:** inform discussions about housing affordability, neighborhood gentrification, or short-term rental regulation

-   **Bias awareness:** Airbnb data does not fully represent the rental market. Certain types of properties, neighborhoods, or hosts may be underrepresented
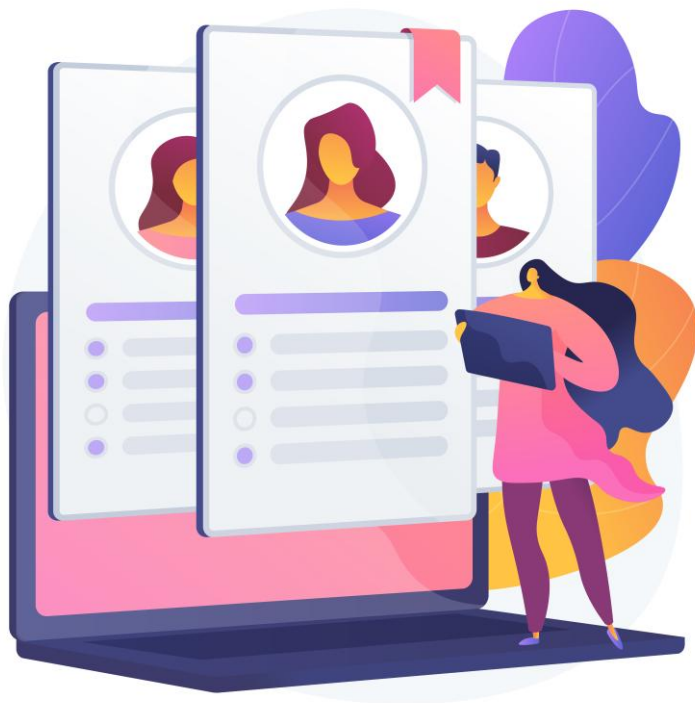
Engagement & Orginality:

- Active involvement with real-world data: actual Airbnb listings for NYC in the last 4 quarters

- Creating visualizations (maps, boxplots, heatmaps, and summary statistics) and consider relationships between multiple features (price vs. room type, neighborhood vs. price)