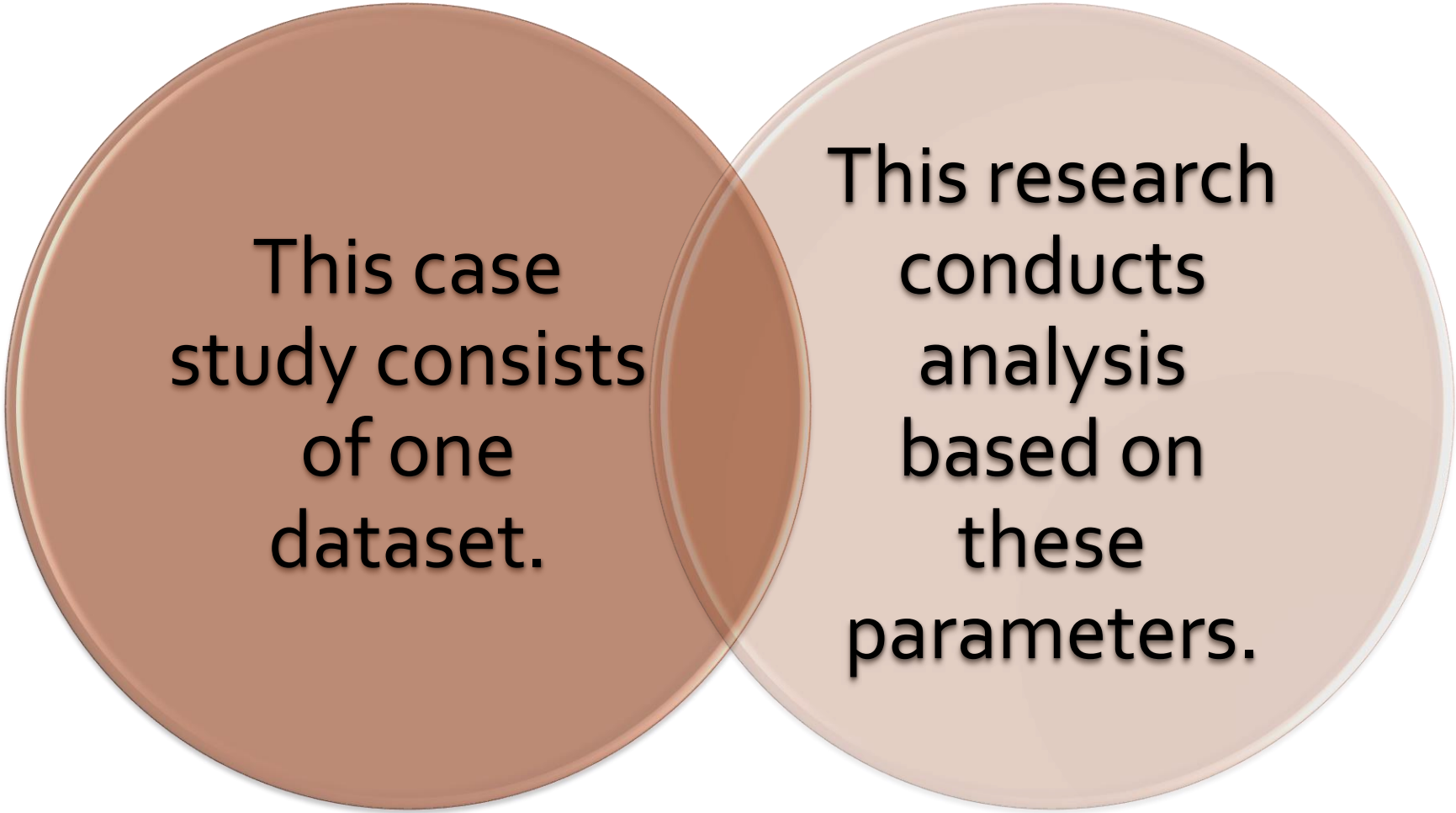
AIRBNB (NYC) SALES ANALYSIS

By Krishnakoli Datta

Contents

Case Study


This case study consists of one dataset.
This research conducts analysis based on these parameters.

Objectives

This presentation centers on the analysis of an Airbnb dataset for New York City.

Dataset consists of hotels associated with Airbnb in various locations of New York City, ranging from 2011 to 2019.

The project's primary aim is to extract valuable insights from the dataset to support data-driven decision-making.

Dataset Overview

Dataset Source

- Obtained from Kaggle, a renowned data science platform.

Dataset Size

- Includes over 50 thousand rows and 14 key columns.

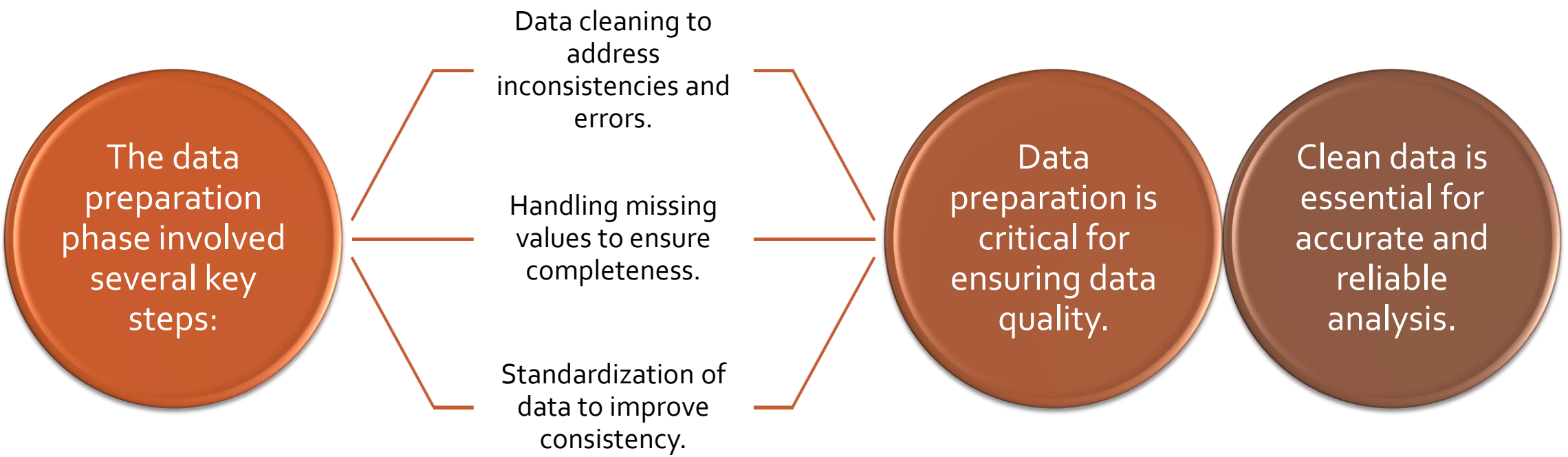
Data Types

- Consists of various data types, including text, numeric, and date.

Structure

- Contains information on Airbnb listings, hosts, neighborhoods, pricing, and availability.

Data Preparation



Research Questions

The project was designed to address several key research areas, like:

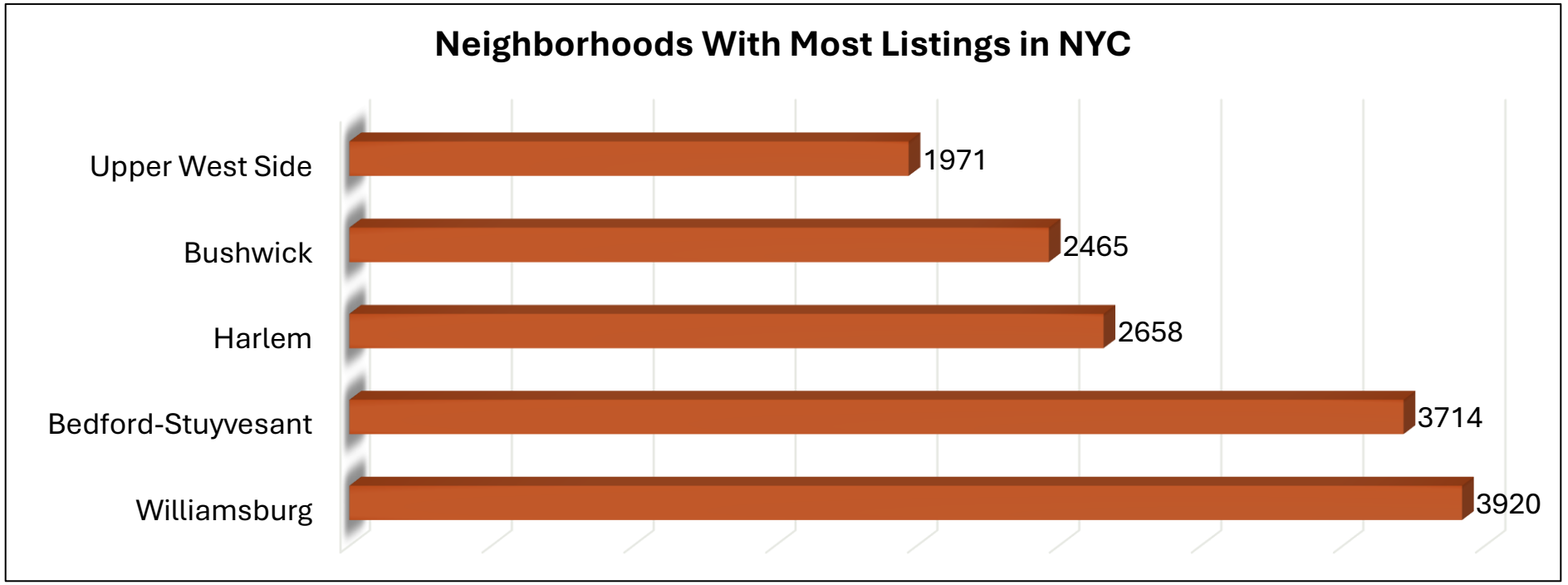
- Seasonal pricing trends.
- Analyzing availability in each district.
- Examining changes in availability over time.
- Determining the average price for each room type in various districts.
- Assessing neighborhoods with the best price-to-availability ratio.

Total Revenue Earned by Airbnb in NYC (2011-2019)

SQL QUERIES

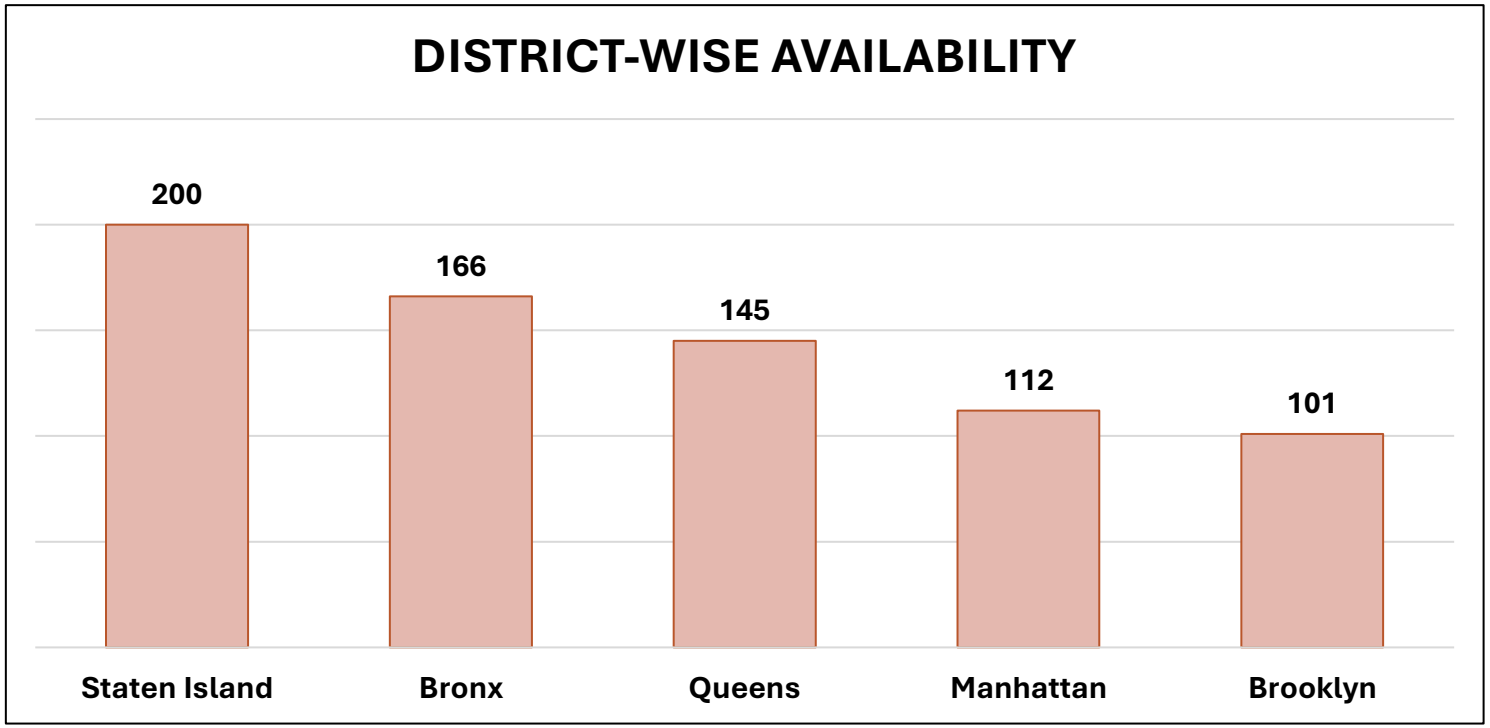

```sql
-- Is there a seasonal trend in pricing?
SELECT EXTRACT(MONTH FROM last_review_date) AS month, AVG(price) AS average_price
FROM airbnb_data
GROUP BY month
ORDER BY month;
```

SQL QUERIES



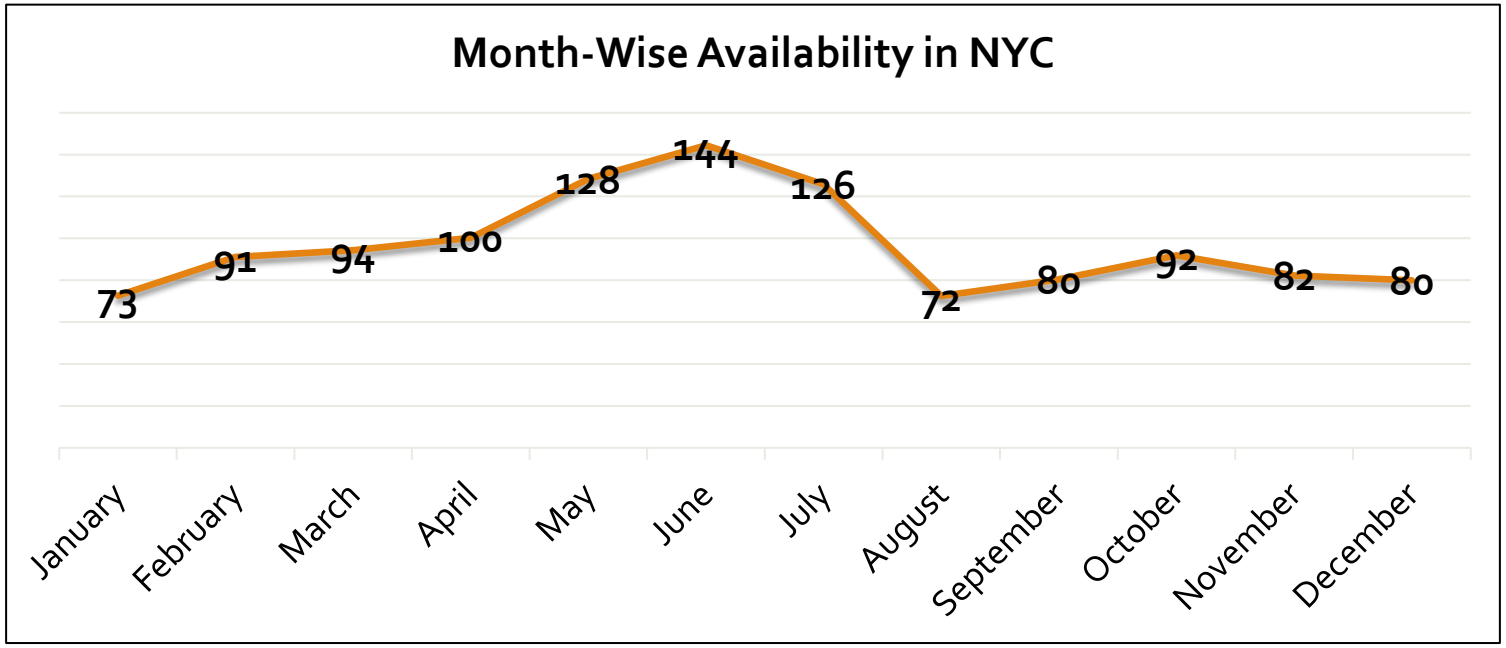

```sql
-- Which (Top 5) neighborhoods have the highest number of Airbnb listings?
SELECT neighbourhood, COUNT(*) AS listing_count
FROM airbnb_data
GROUP BY neighbourhood
ORDER BY listing_count DESC
limit 5;
```

SQL QUERIES




```sql
-- What's the average availability for listings in each district?
SELECT district, ceil(AVG(availability_365)) AS average_availability
FROM airbnb_data
GROUP BY district
ORDER BY average_availability DESC;
```

SQL QUERIES




```sql
-- How does availability change over time?
SELECT EXTRACT(MONTH FROM last_review_date) AS month,
AVG(availability_365) AS average_availability
FROM airbnb_data
WHERE last_review_date IS NOT NULL
GROUP BY month
ORDER BY month;
```

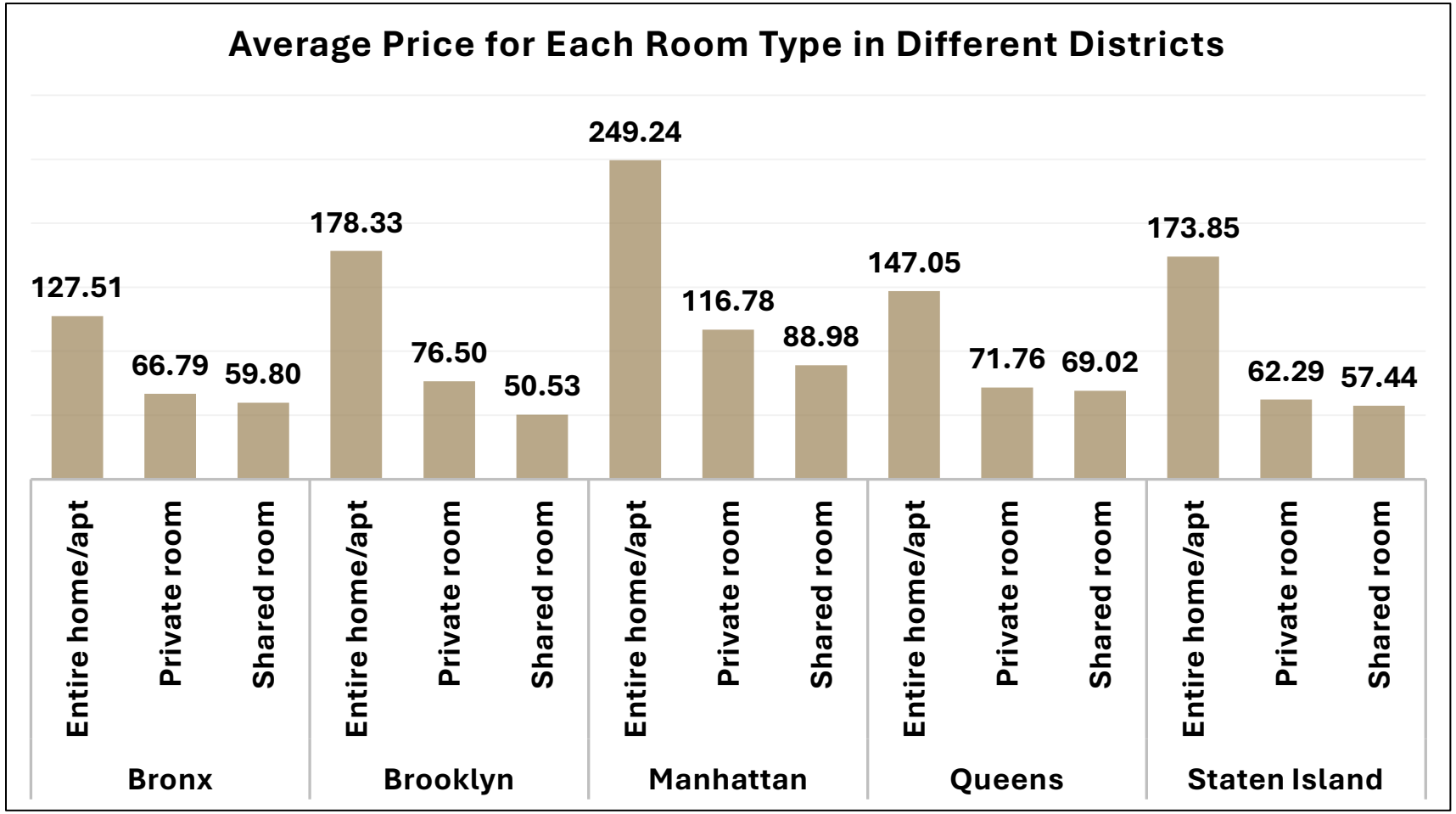
SQL QUERIES

```
-- Which listings have not been reviewed for a long time?
SELECT id, name, last_review, district
FROM airbnb_data
WHERE last_review IS NOT NULL
ORDER BY last_review asc
limit 10;
```

Name	Last Reviewed On	District
Sunlit and Cozy Williamsburg/Green point, Brooklyn	28-03-2011	Brooklyn
Luxurious Condo in DUBMO with View	25-04-2011	Brooklyn
Loft w/ Terrace @ Box House Hotel	12-05-2011	Brooklyn
Modern Apt with Spectacular Views	18-09-2011	Brooklyn
Sweet and Spacious Brooklyn Loft	28-12-2011	Brooklyn
Oversized Studio in Park Slope	02-01-2012	Brooklyn
Large Room w/ Private Entrance	29-05-2012	Brooklyn
East Village Sanctuary	10-12-2011	Manhattan
LOVELY APARTMENT IN THE HEART OF NY	03-01-2012	Manhattan
Fully Furnished Basement Apartment	19-09-2011	Queens

SQL QUERIES

```sql
-- What's the average price for each room type in different districts?
SELECT district, room_type, AVG(price) AS average_price
FROM airbnb_data
GROUP BY district, room_type
ORDER BY district, room_type;
```



SQL QUERIES

```sql
-- Can we identify neighborhoods with the best price-to-availability ratio?
SELECT district, AVG(price / availability_365) AS price_to_availability_ratio
FROM airbnb_data
WHERE availability_365 > 0
GROUP BY district
ORDER By price_to_availability_ratio DESC;
```