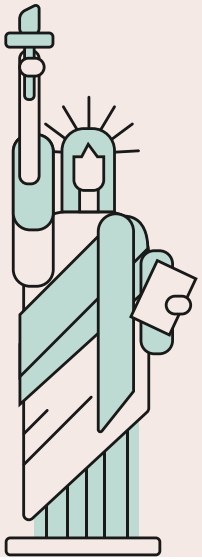
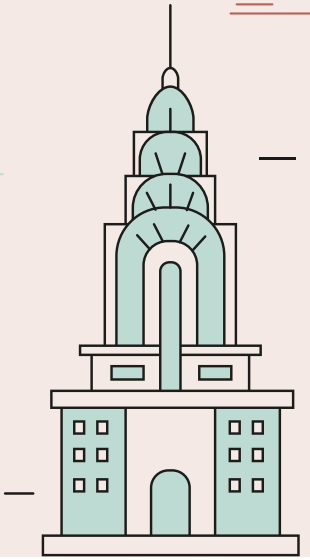


Interactive Airbnb Visualization

Name: Rayleen Ramos
Course: CPS 5745 – Interactive Information Visualization
Instructor: Dr. Huang



Project Overview

This project:

- Builds an interactive web-based data visualization system
- Allows users to explore Airbnb pricing trends across New York City
- Uses real-world data and interactive controls to reveal patterns and trends

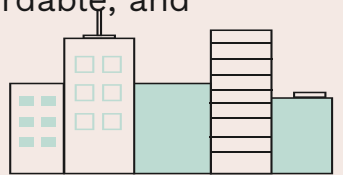
This project is important:

- Airbnb prices vary significantly by location and listing characteristics
- Viewing all data at once can hide meaningful differences
- Interactive charts allow users to filter data and clearly observe trends

Overall goal:

- Transform raw Airbnb data into clear, meaningful visualizations

Help users understand where prices are high, where they are more affordable, and why



Dataset Description

- Dataset Used: Inside Airbnb-New York City
- Public dataset collected from real Airbnb listings worldwide
- Dataset contains 79 attributes and 36,000 records
- Including:
  - Location details (neighborhood & borough)
  - Listing characteristics (room type, price, # of reviews
  - Amenities (# of rooms & baths)
- Simplified the dataset to 16 attributes for better analysis

neighbourhood_cleans...	neighbourhood_group_clean...	latitude	longitude	property_type	room_type	accommodat...	bathrooms	bedrooms	beds	price	number_of_revie...
Midtown	Manhattan	40.75356	-73.98559	Entire rental unit	Entire home/apt	1	1.0	0	1	240.00	47
Williamsburg	Brooklyn	40.70935	-73.95342	Entire rental unit	Entire home/apt	3	1.0	2	1	96.00	195
East Harlem	Manhattan	40.80107	-73.94255	Private room in condo	Private room	1	1.0	1	1	59.00	1
East Harlem	Manhattan	40.78778	-73.94759	Private room in rental unit	Private room	1	1.0	2	2	73.00	249
Williamsburg	Brooklyn	40.71248	-73.95881	Private room in loft	Private room	2	1.0	1	0	0.00	13
Fort Greene	Brooklyn	40.69194	-73.97389	Private room in guest suite	Private room	2	1.0	1	2	216.00	423
Williamsburg	Brooklyn	40.71880722	-73.95617676	Entire place	Entire home/apt	2	1.0	0	0	0.00	12
Bedford-Stuyvesant	Brooklyn	40.68455584	-73.93963415	Entire loft	Entire home/apt	5	1.0	1	4	170.00	189
Hell's Kitchen	Manhattan	40.76724	-73.98664	Entire rental unit	Entire home/apt	2	1.0	0	1	175.00	58
Bedford-Stuyvesant	Brooklyn	40.68294	-73.95682	Private room in rental unit	Private room	2	1.0	1	2	90.00	82

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

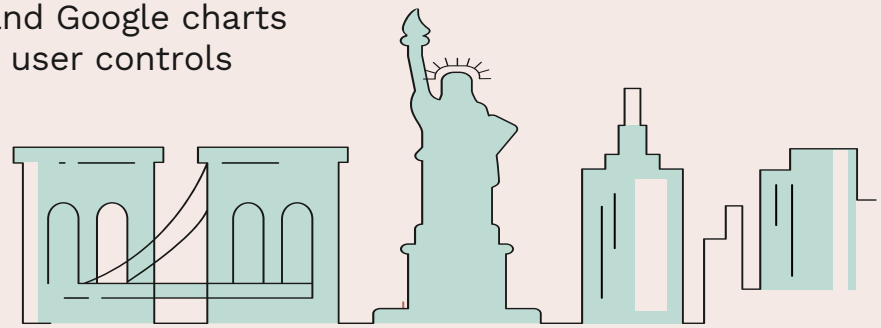
System Architecture

This project follows a three-layer architecture

Data Layer:
- stores raw Airbnb data and aggregated data marts in MySQL

Application Layer
- Uses PHP and SQL to process user requests
- Retrieves filtered and aggregated data from database

Presentation Layer
- Uses HTML/CSS, JavaScript, Chart.js and Google charts
- Displays interactive charts, tables and user controls



Data Marts

01

```sql
CREATE TABLE airbnb (
    id BIGINT primary key,
    name VARCHAR(300),
    neighbourhood_cleansed VARCHAR(100),
    neighbourhood_group_cleansed VARCHAR(100),
    latitude DOUBLE,
    longitude DOUBLE,
    property_type VARCHAR(150),
    room_type VARCHAR(100),
    accommodates INT,
    bathrooms DECIMAL(3,1),
    bedrooms INT,
    beds INT,
    price DECIMAL(10,2),
    number_of_reviews INT,
    review_scores_rating DECIMAL(3,2),
    review_scores_location DECIMAL(3,2)
);
```

02

```sql
CREATE TABLE price_borough AS SELECT
    neighbourhood_group_cleansed AS borough,
    SUM(price) AS total_price,
    COUNT(price) AS listing_count,
    SUM(price) / COUNT(price) AS avg_price
FROM airbnb
WHERE price > 0
  AND neighbourhood_group_cleansed IS NOT NULL
GROUP BY neighbourhood_group_cleansed
ORDER BY avg_price DESC;
```

03

```sql
CREATE TABLE price_neighborhood AS SELECT
    neighbourhood_cleansed AS neighborhood,
    neighbourhood_group_cleansed AS borough,
    SUM(price) AS total_price,
    COUNT(price) AS listing_count,
    (SUM(price) / COUNT(price)) AS avg_price
FROM airbnb
WHERE price > 0
  AND neighbourhood_cleansed IS NOT NULL
  AND neighbourhood_group_cleansed IS NOT NULL
GROUP BY neighbourhood_cleansed, neighbourhood_group_cleansed;
```

04

```sql
CREATE TABLE price_reviews AS SELECT
    number_of_reviews,
    SUM(price) AS total_price,
    COUNT(price) AS listing_count,
    SUM(price) / COUNT(price) AS avg_price
FROM airbnb
WHERE price > 0
  AND number_of_reviews IS NOT NULL
GROUP BY number_of_reviews
ORDER BY number_of_reviews;
```

Data Flow

Presentation layer

User logs in (all buttons are disabled until user logs in)

Presentation layer

Once user logs in, user can click Load DB Data

Application layer

JavaScript sends an AJAX request to PHP to retrieve the data

Data layer

PHP queries the appropriate data mart in MySQL

Data layer

Aggregated results are returned as a JSON

Presentation layer

Charts and tables update dynamically in the browser

Source Code Integration

Presentation Layer:

- *Index.html* – page layout, controls and chart containers
- *Style.css* – visual styling and layout

Application Layer:

- *App.js* – handles user events, filters and AJAX requests
- Dynamically updates charts when slider or dropdown changes

Data Layer:

- *Login.php* & *logout.php* – user authentication
- *Load_data.php* – retrieves aggregated data from data marts
- *Saved_settings.php* & *get_settings.php* –saves and retrieves user slider settings
- *Dbconfig.php* – database connects to MySQL

Chart #1 – Avg. Price by Neighborhood

Chart Type:

- Bar Chart

Attributes Used:

- X-axis: Neighborhood
- Y-axis: Average Airbnb Price
- Data Marts: price_borough & price_neighborhood

Goal

- Compare average Airbnb prices between neighborhoods within a selected borough and price range

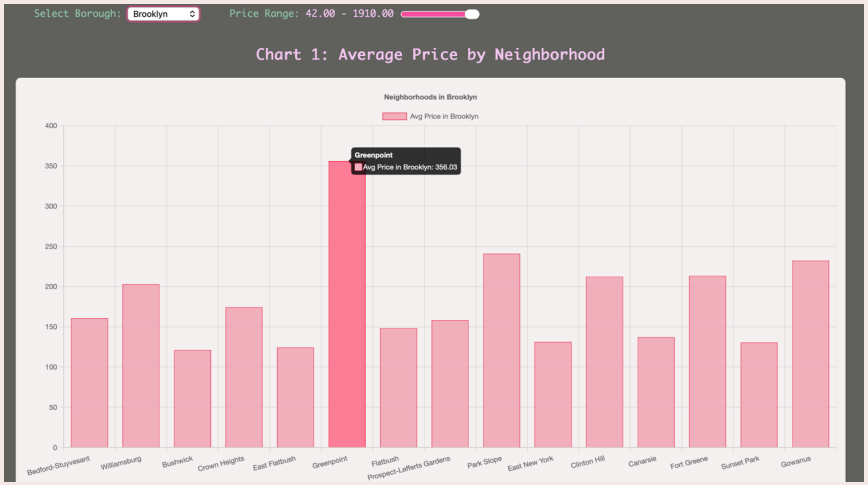


Chart #2 – Avg. Price by Neighborhood

Chart Type:

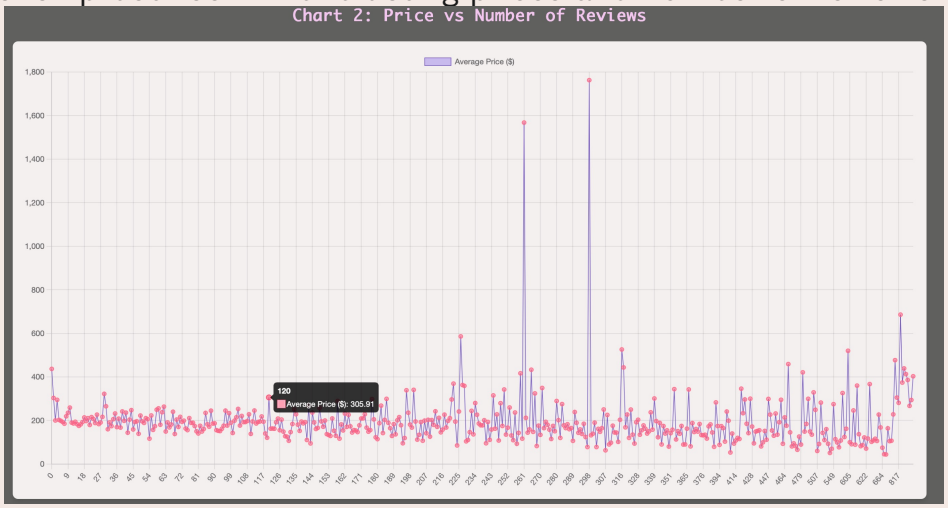
- Line Chart

Attributes Used:

- X-axis: Number of Reviews
- Y-axis: Average Airbnb price
- Data Marts: price_reviews

Goal

- Explore relationship between Airbnb listing prices and number of reviews

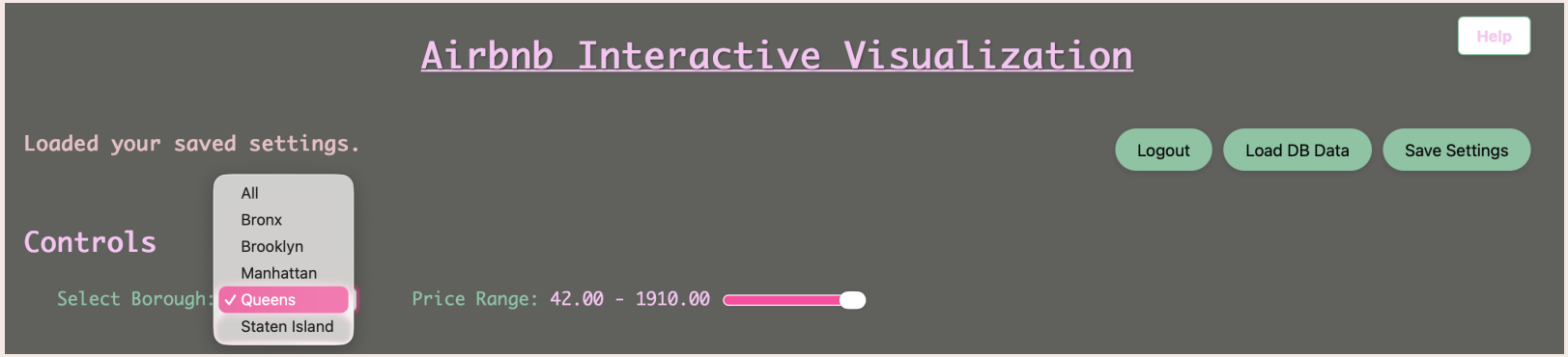


Loaded Data & Outliers

- Raw data is displayed using google charts table
- Displays the first 500 rows of the Airbnb and data marts table
- Rows containing price outliers are highlighted yellow
- Outliers are detected using the quartile (IQR) method
- The threshold values are displayed below the tables

Interactive Features

- Username and password required before accessing data and charts
- Load DB data button retrieves data from database
- Borough dropdown filters data by selected borough
- Price range slider filters listings by minimum to maximum price
- Charts and tables are updated dynamically
- User settings can be saved and restored for future sessions



Insights & Conclusion

- Airbnb prices vary significantly across neighborhoods, even within the same borough
- High Airbnb prices are driven by specific neighborhoods, not the entire city
- Listings with a high number of reviews are not always the most expensive
- Interactive filtering makes pricing patterns and outliers easier to identify


Live Demo


Thank You!