



Airbnb stay

WITH

CAPECASA FINDERS

A wide-angle photograph of a beach at sunset. The sky is filled with warm, orange, and yellow hues, transitioning into cooler blues and purples at the top. The ocean waves are breaking onto the sandy shore in the foreground. In the distance, a few small figures of people are visible in the water.

TAGLINE

Recommending the best price
&
Stay Options in Cape Town

CAPECASA FINDERS

TEAM

Peter Mbuthia
Kelvin Murithi
Festus Muchemi
Eric Muema

Business Understanding

- The project develops a recommender system for Cape Town Airbnb listings to assist hosts in optimizing pricing, occupancy, and guest satisfaction using Inside Airbnb data.
- It suggests optimal prices based on listing features, guest sentiment, and occupancy patterns, helping hosts set competitive and profitable prices.



DATA DESCRIPTION

- Key data files include:
 - **listings.csv**: Contains details about each listing, such as property type, location, amenities, and host information.
 - **reviews.csv**: Provides guest feedback, which is used to derive sentiment scores.
- ‘Price’ is our target column.
- 81 unique columns



Key Columns

- **Property Details:** property_type, accommodates, bathrooms, bedrooms, and beds.
- **Host Information:** Fields like host_id, host_response_rate, host_is_superhost, and host_listings_count.
- **Pricing and Occupancy:** Columns such as price, availability, and number_of_reviews.
- **Guest Sentiment:** Derived from guest comments using sentiment analysis to score the emotional tone of each review.



Data Cleaning

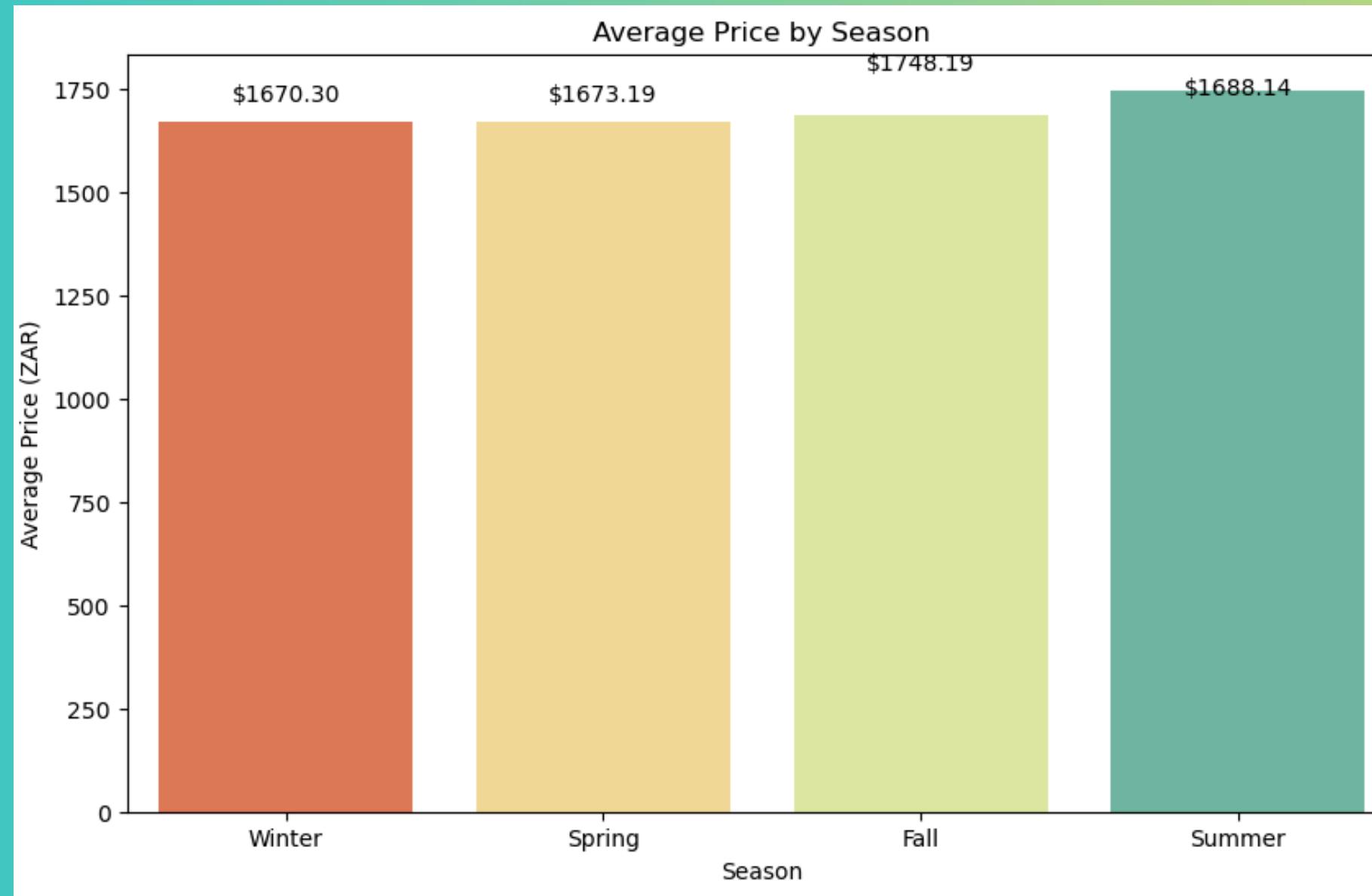
- Dropped null values.
- Checked for duplicates. There were none.
- Merged our 2 datasets

Feature Engineering

- Created new `avg_rating` using the mean of 7 existing rating columns.
- Created new features such as `sentiment_score`, extracted from guest reviews using sentiment analysis.
- Encoded categorical variables, applying frequency encoding to `neighbourhood_cleansed` and `property_type` columns for improved model performance.



Exploratory Data Analysis

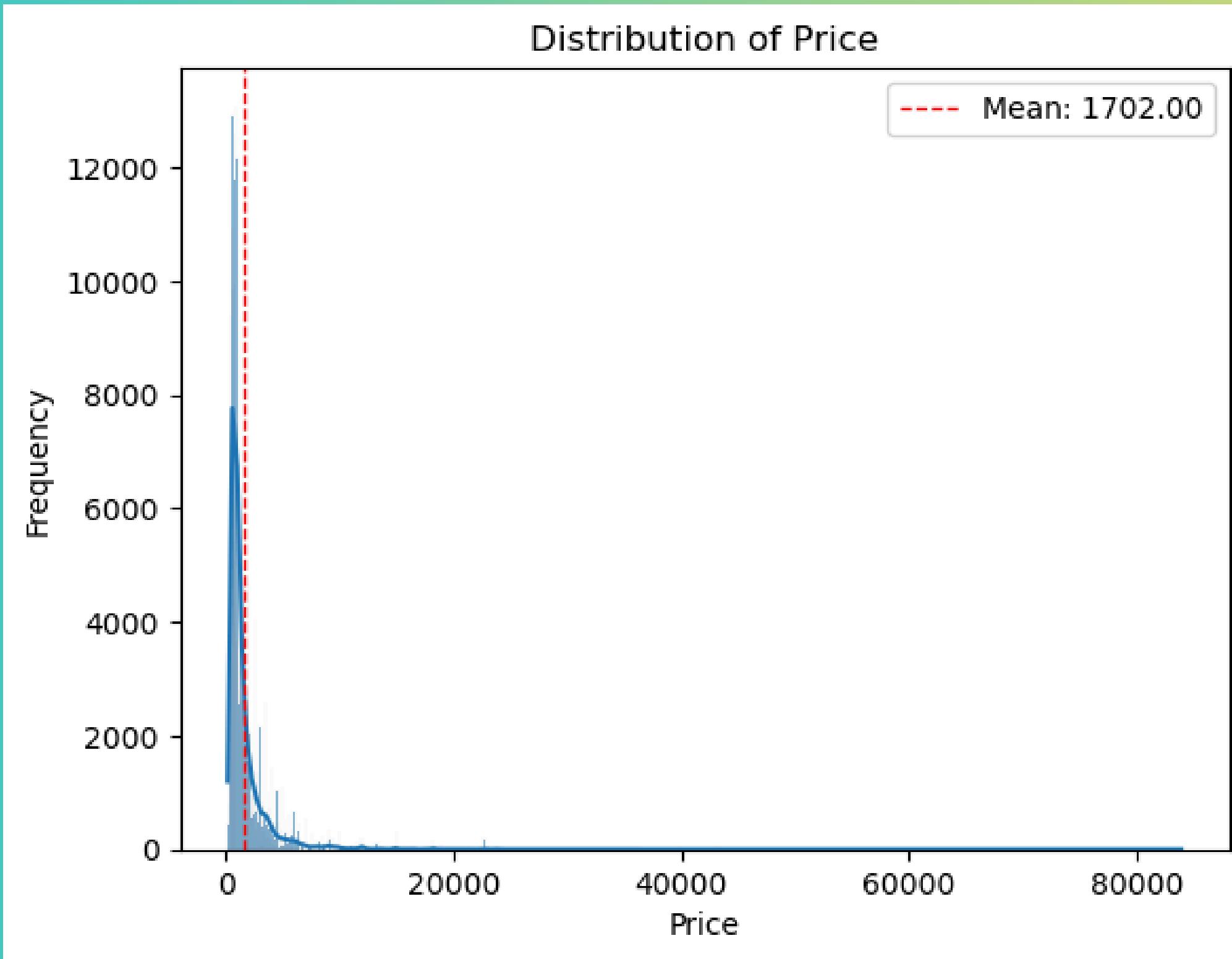


Key insights:

- Summer: Highest average price, around 1748 ZAR, indicating peak tourist season and high demand.
- Fall: Average price of 1688 ZAR, suggesting continued strong interest.
- Spring: Close to Fall with an average of 1673 ZAR, possibly due to milder weather and events.
- Winter: Lowest average price, around 1670 ZAR, reflecting the off-peak season with fewer visitors.

The chart highlights that Summer is the most expensive time for Airbnb stays, likely driven by high tourist influx.

Exploratory Data Analysis

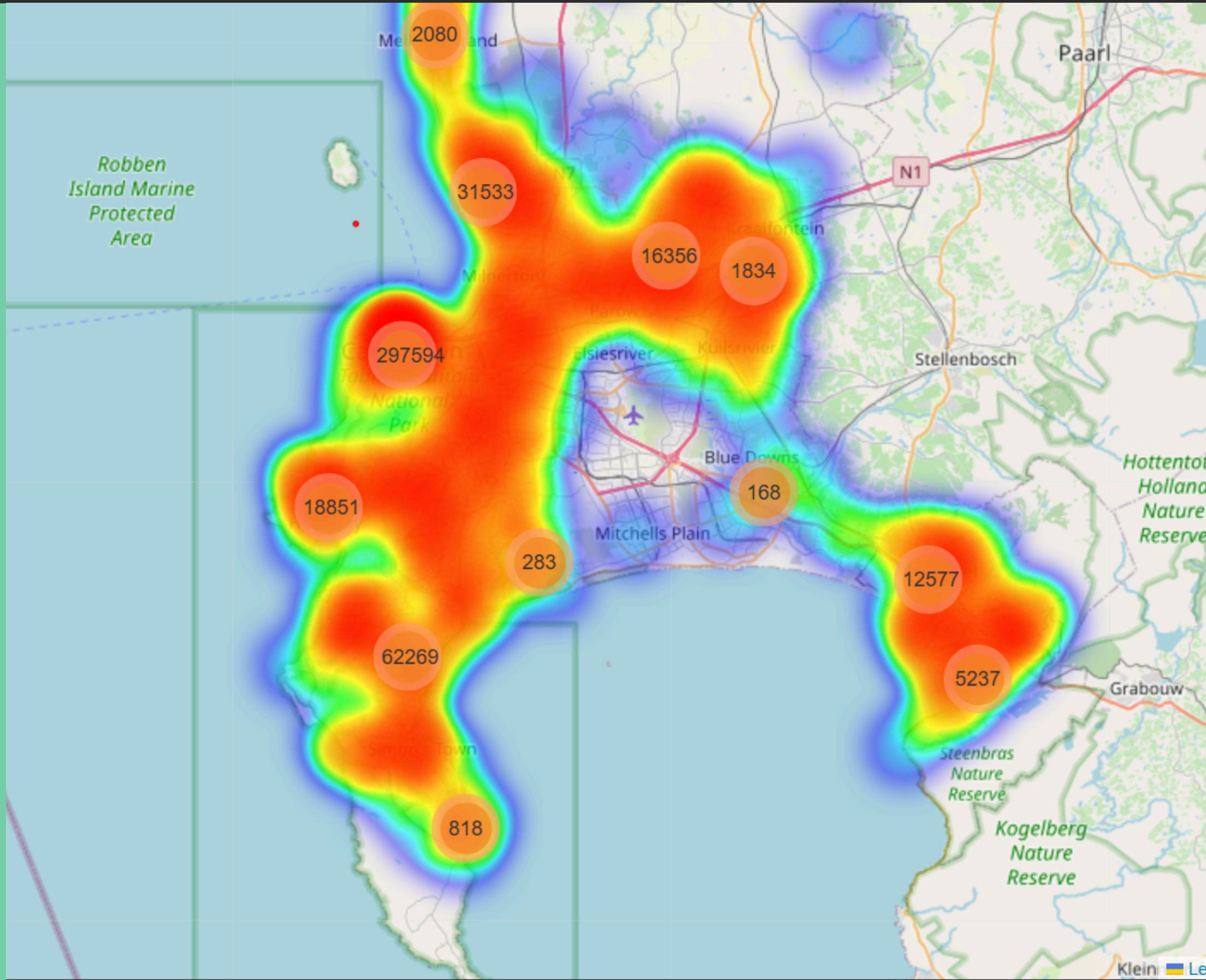


The KDE of the histogram above shows that the average price of Airbnbs in CapeTown is 1702 ZAR.

Exploratory Data Analysis

Insights

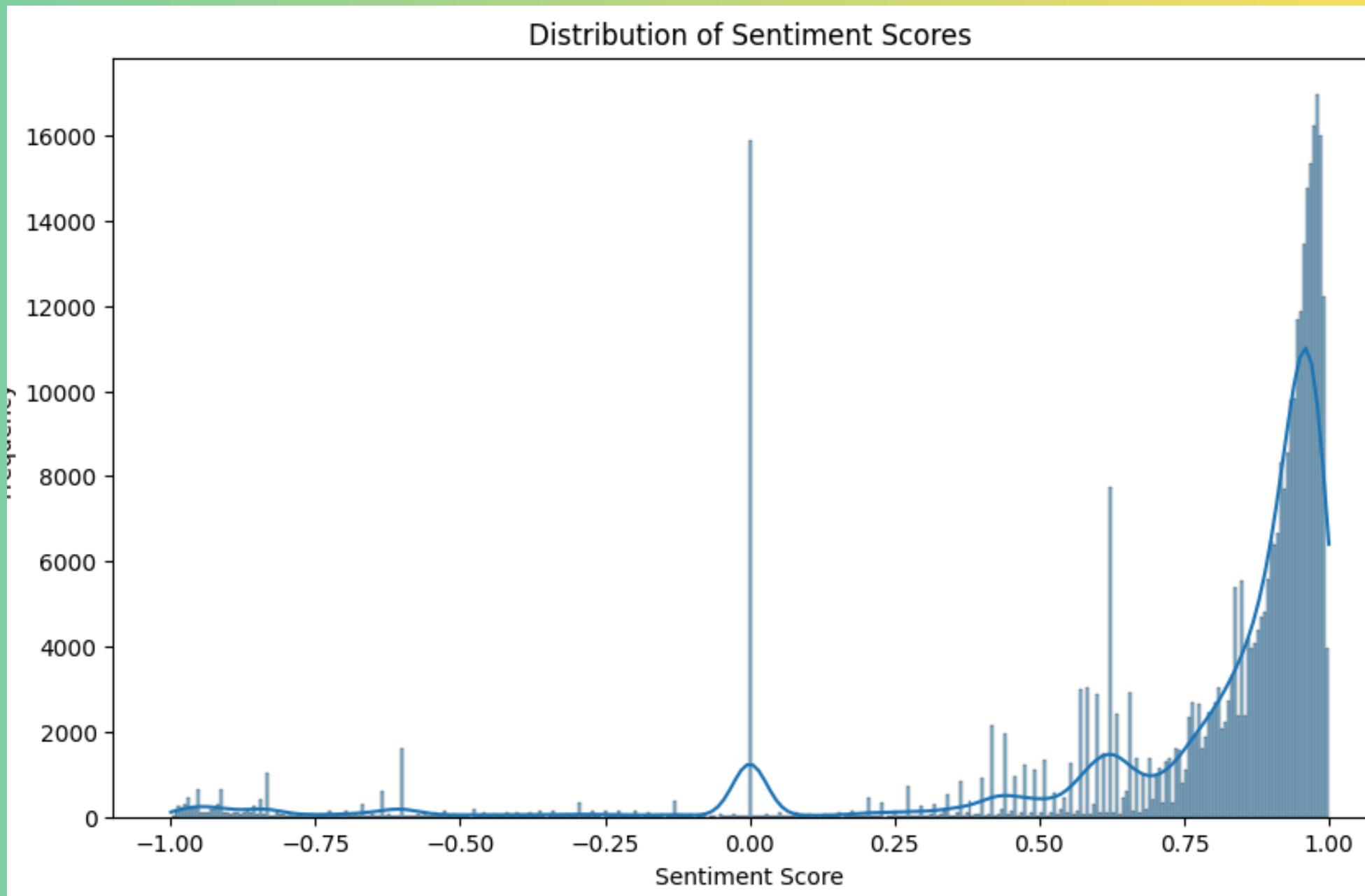
- The highest concentration of listings (297,594) is in central Cape Town.
- Airbnb listings are primarily located in tourist-friendly coastal areas and the city center, with much less activity in inland regions.



Exploratory Data Analysis

Insights

- There's a strong positive skew, with most comments ranging from 0.5 to 1, indicating general positivity.
- A spike around the 0 score suggests many neutral comments.
- Negative sentiments are infrequent, reflecting few strongly negative remarks.



Modelling

	RMSE	R ²
Linear Regression	0.77	0.59
Linear Regression (with PCA)	0.79	0.58
Decision Tree	0.71	0.64
Random Forest	0.57	0.75
KNN Regression	0.68	0.6616
Tuned KNN(with Grid Search)	0.68	0.66
Polynomial Regression	0.64	0.69
Neural Network	0.64	0.69



- **BaselineOnly** model is the most effective model for the collaborative filtering-based recommender system.

FUTURE IMPROVEMENTS

- Incorporate Additional Model Types
- Dynamic Price Optimization



LINK

The following CAPECASA recommender system link helps you find the best price for a stay in Cape Town

<https://capecasa-airbnb.streamlit.app/>



Recommendation

1. Optimize pricing based on model recommendations considering location, amenities, and demand.
2. Enhance guest satisfaction by maintaining quality and providing excellent customer service to boost sentiment scores.
3. Use data insights to market listings with high occupancy, positive sentiment, and competitive pricing to attract guests.



An aerial photograph of Cape Town, South Africa, showing the city nestled at the foot of the Twelve Apostles mountain range. The foreground is filled with the dense urban sprawl of Camps Bay, with numerous white houses built on rocky hillsides overlooking the deep blue Atlantic Ocean. A winding coastal road follows the curve of the bay. In the background, the iconic peaks of Signal Hill and Lion's Head rise prominently against a clear blue sky.

Welcome to
Cape Town