AirBnB Rooms in Copenhagen, Denmark:

Activities, Prices, and Availability

Background

- Scandinavian capital cities like Copenhagen in Denmark receive millions of tourist every year, catalysed by globalisation and megatrends such as sustainability and minimalism.
- Thanks to budget airlines and affordable accommodation, such as AirBnB, tourists on a tight budget are now able to experience these notoriously expensive countries.

Challenge

- Copenhagen, Denmark is not yet an established urban tourism capital such as Paris, London, or Rome.
- As such, tourists who wish to visit need extensive research to get the bang for their buck.
- This report aims to get a better insight on the activity, availability, and price of AirBnB in Copenhagen's surrounding neighbourhood.

Data Sources

- The dataset (df_listing, 8525 rows) is a list of AirBnB listing in Copenhagen, downloaded from this link. Essential information was taken in df_listing such as:
  - Neighbourhood
  - Geolocation(Latitude, Longitude)
  - Price
  - Room Type
  - Reviews per month
  - Availability of listing in a year
- Venue dataset (cph_venues) is a list of first 100 venues within 500 meters of a particular AirBnB listing in Copenhagen. API Frousquare was used in this dataset.

Data Treatment

Data Acquisition

- Downloading the dataset: (df_listing)
- Focusing in Private Room – Room Type listing

Data Treatment and Transformation

- Adding new features such as utilization factor.
- Removing Outliers.
- Combining df_listing dataset and cph_venues

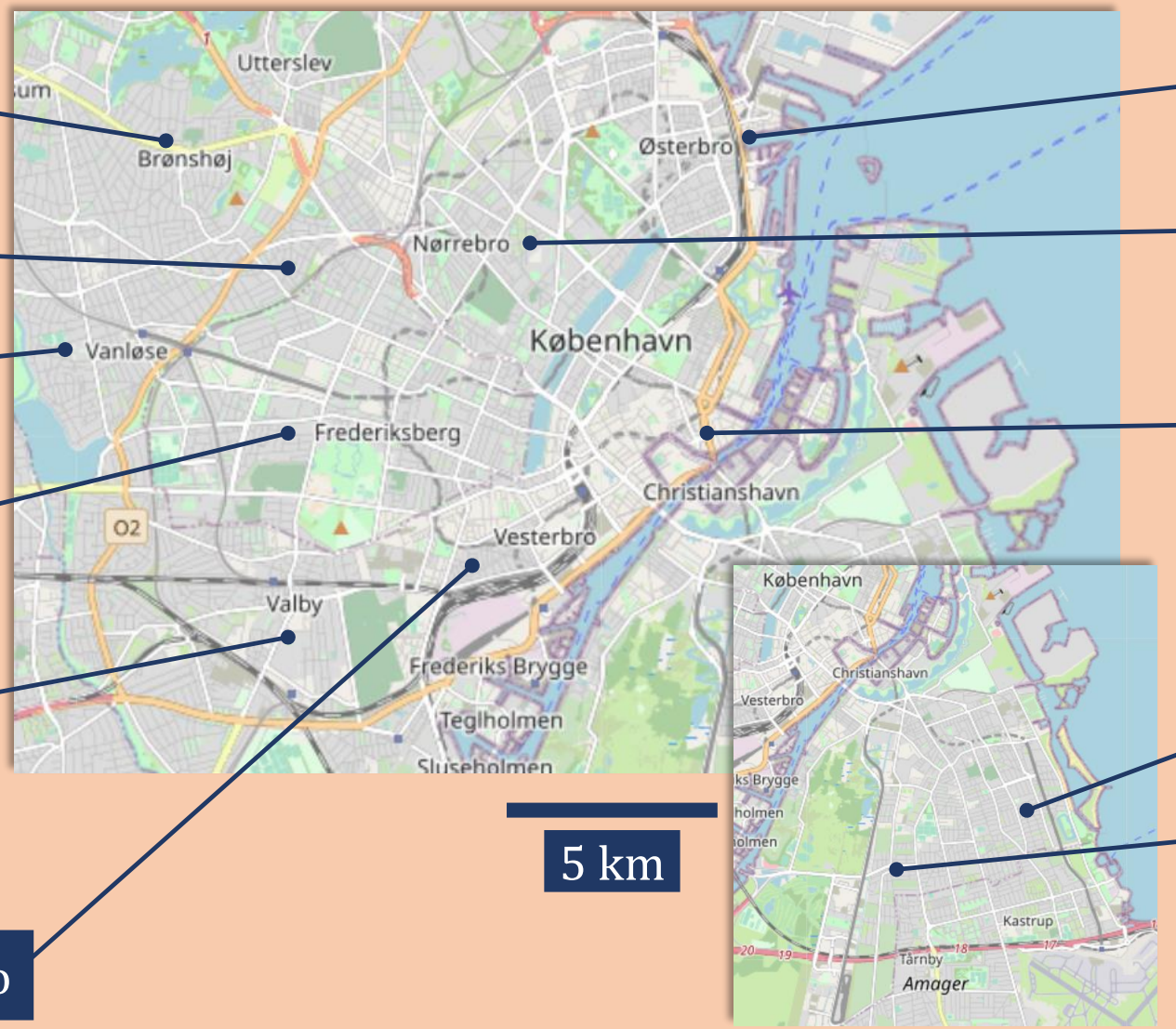
Exploratory Analysis

- Comparing different neighbourhood characteristics
- Comparing room types
- Reducing the dataset to be analysed

Data Insights and Conclusion

- Clustering the dataset using Kmeans
- Gaining insights for incoming tourists in Copenhagen, Denmark.

Neighbourhoods in Copenhagen



Bronshoj

Bispebjerg

Vanlose

Frederiksberg

Valby

Vesterbro

Osterbro

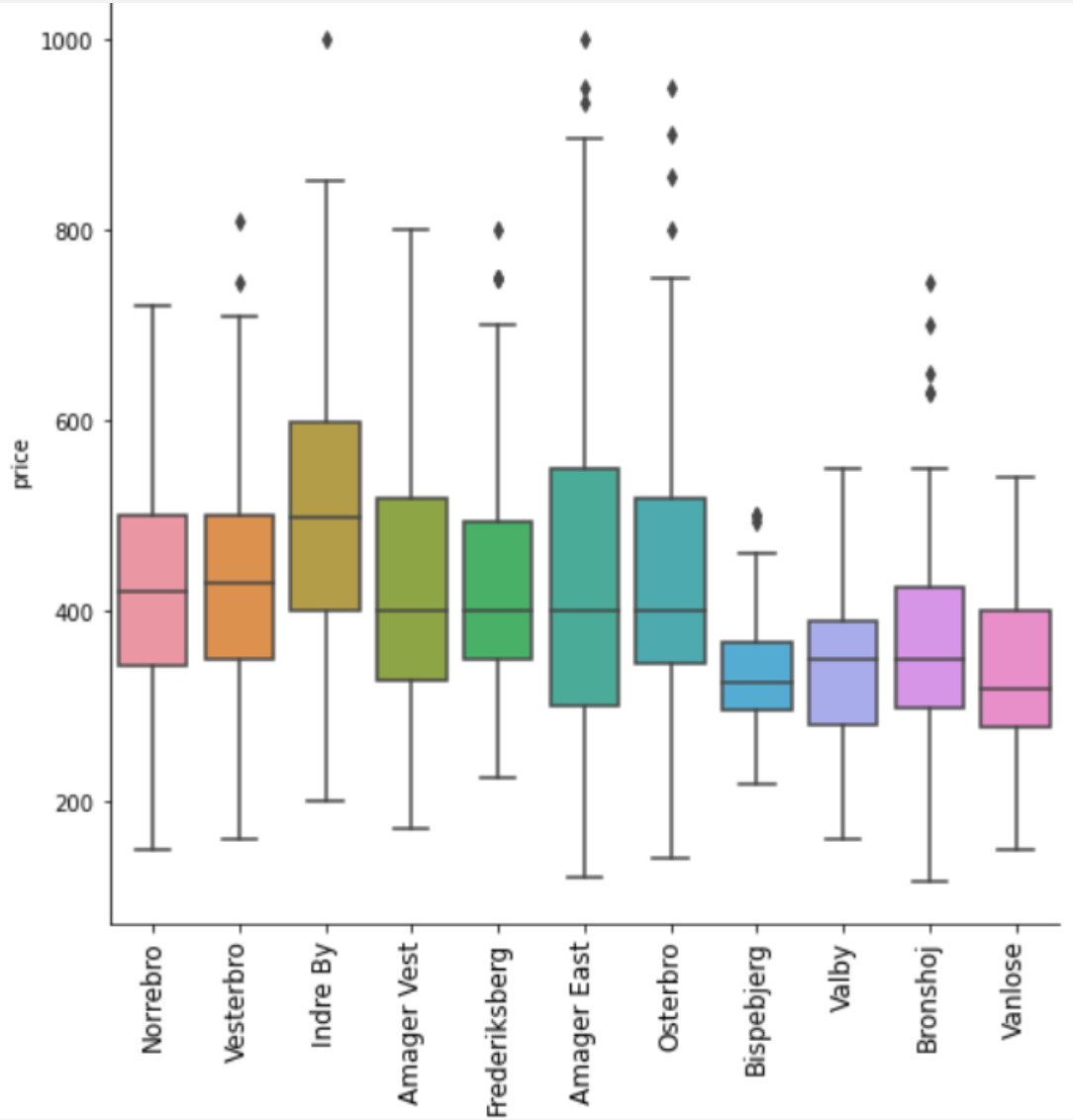
Norrebro

Indre By (City Center)

Amager East

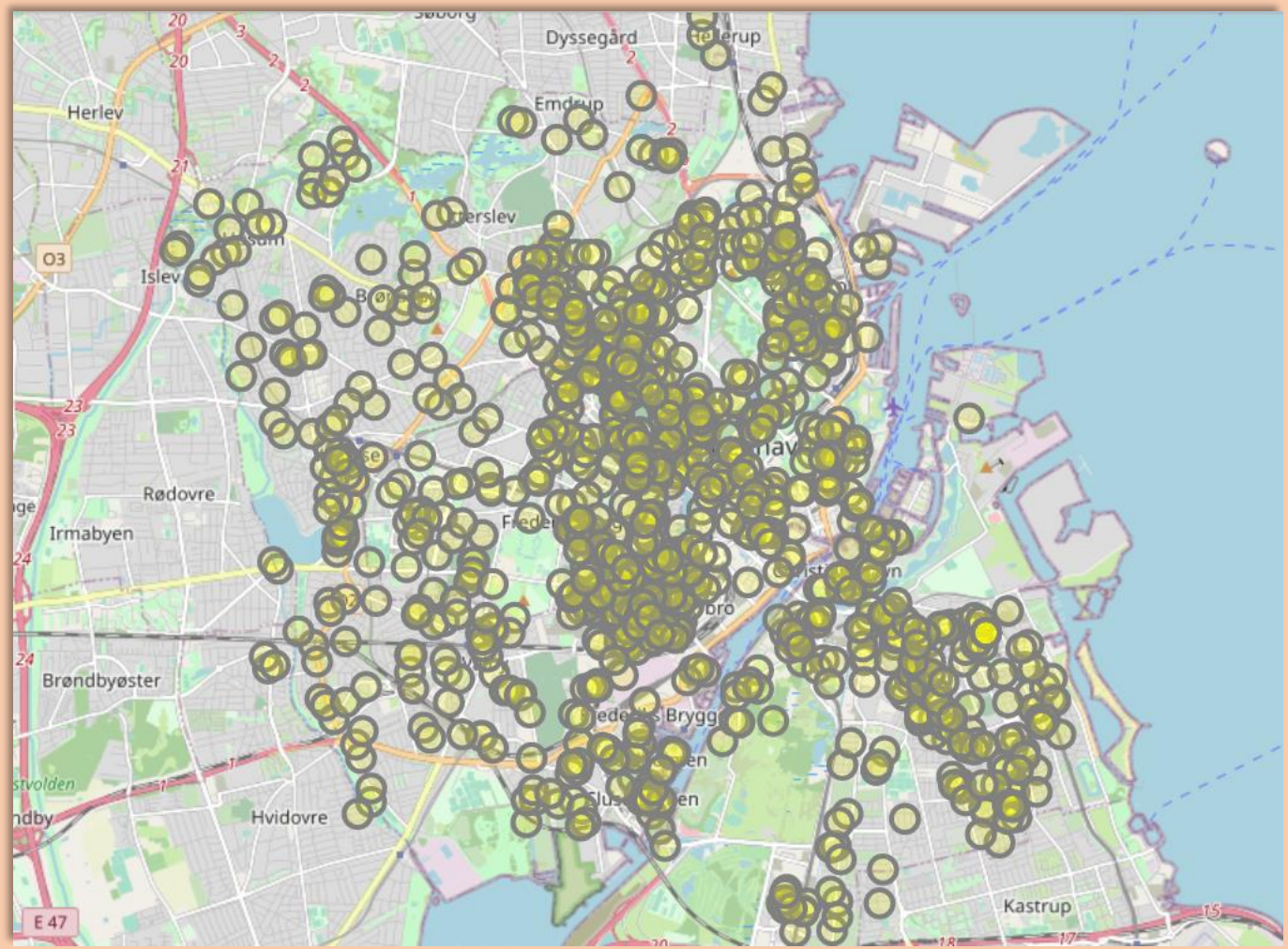
Amager West

Data Exploration: Price



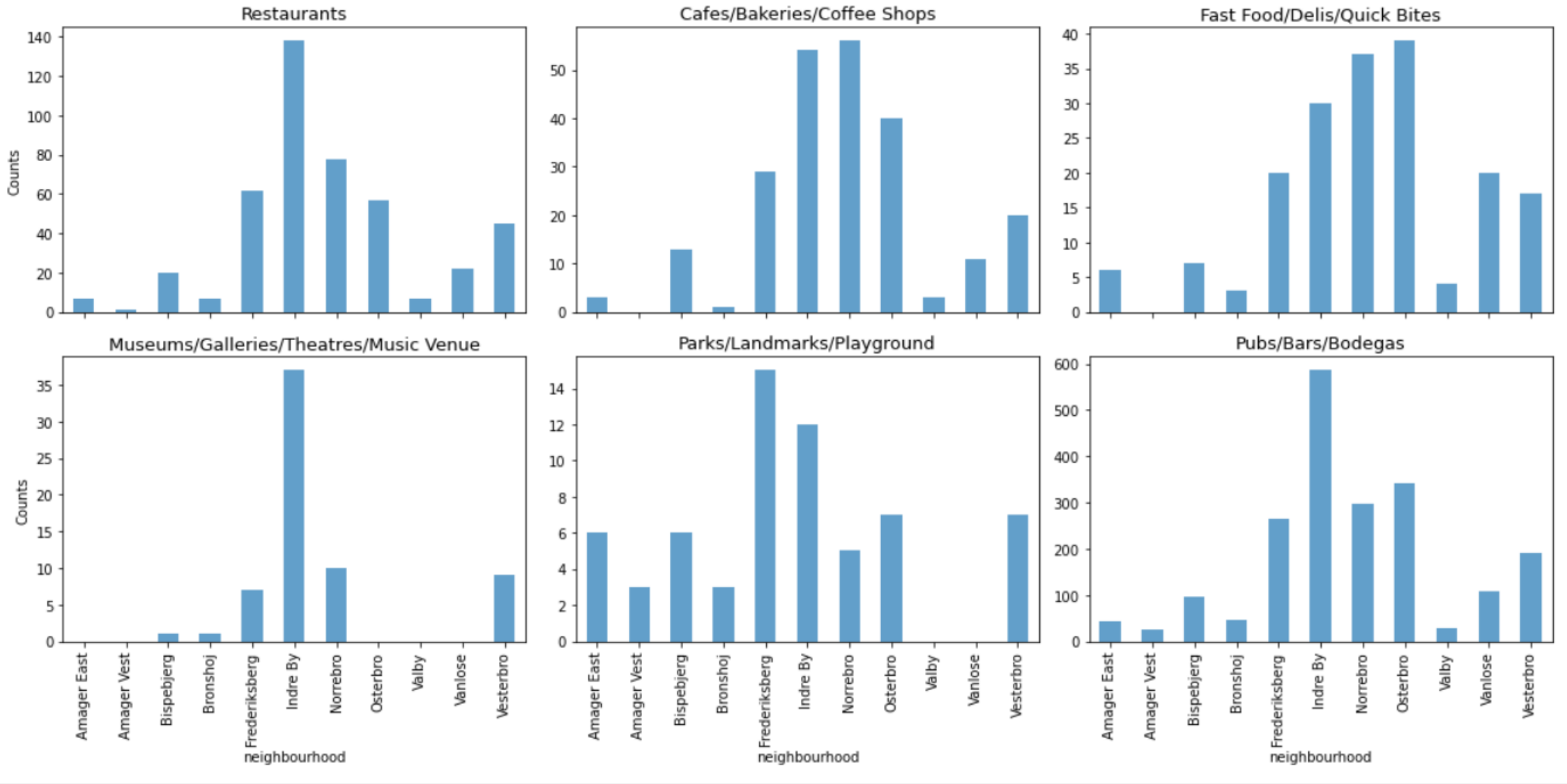
- Indre By (City Center) has the highest median nightly price rate for a private room in Copenhagen, by DKK 100 (14 EUR) more than in other nearby neighbourhoods.
- The highest variance for the median price rate is at Amager East. This suggest that this neighbourhood has high and low price points for private room listings in AirBnB.
- Tourist should expect to pay about DKK 420 in Norrebro, Vesterbro, Frederiksberg, Amager East, and Osterbro. Cheaper neighbourhoods such as Bispebjerg, Valby, Bronshog and Vanlose has a median price of DKK 380.

Data Exploration: Location

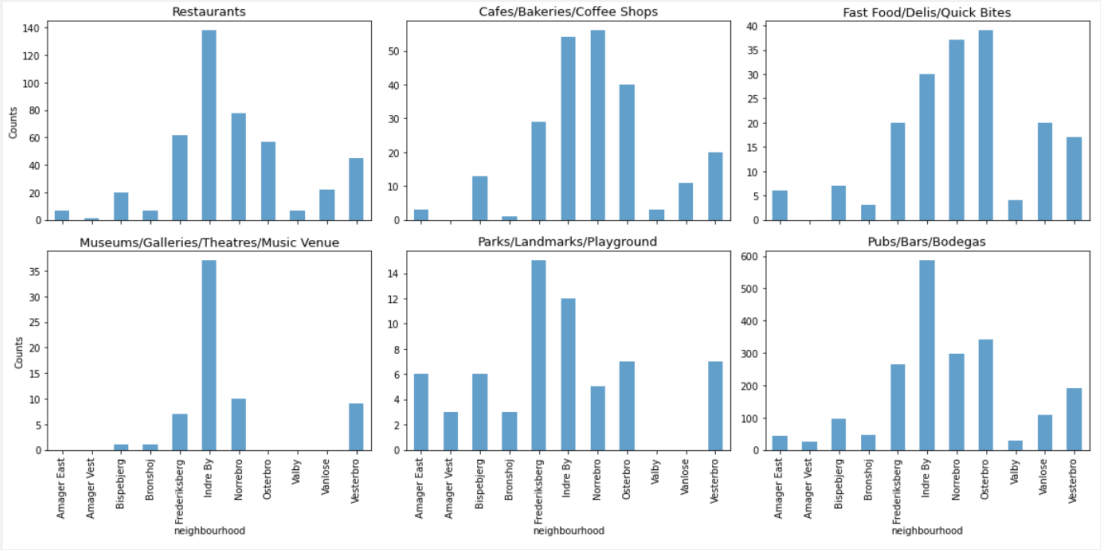


- The yellow mark on the map shows the location of room listings from AirBnB in Copenhagen.
- Most of the Room Listings are concentrated at the inner part of the City – Indre By, Norrebro, Frederiksbreg, Osterbro, and Vesterbro.
- The neighbourhoods surrounding the inner city exhibit room listings that are less and spread out.

Data Analysis: Comparing Neighbourhoods

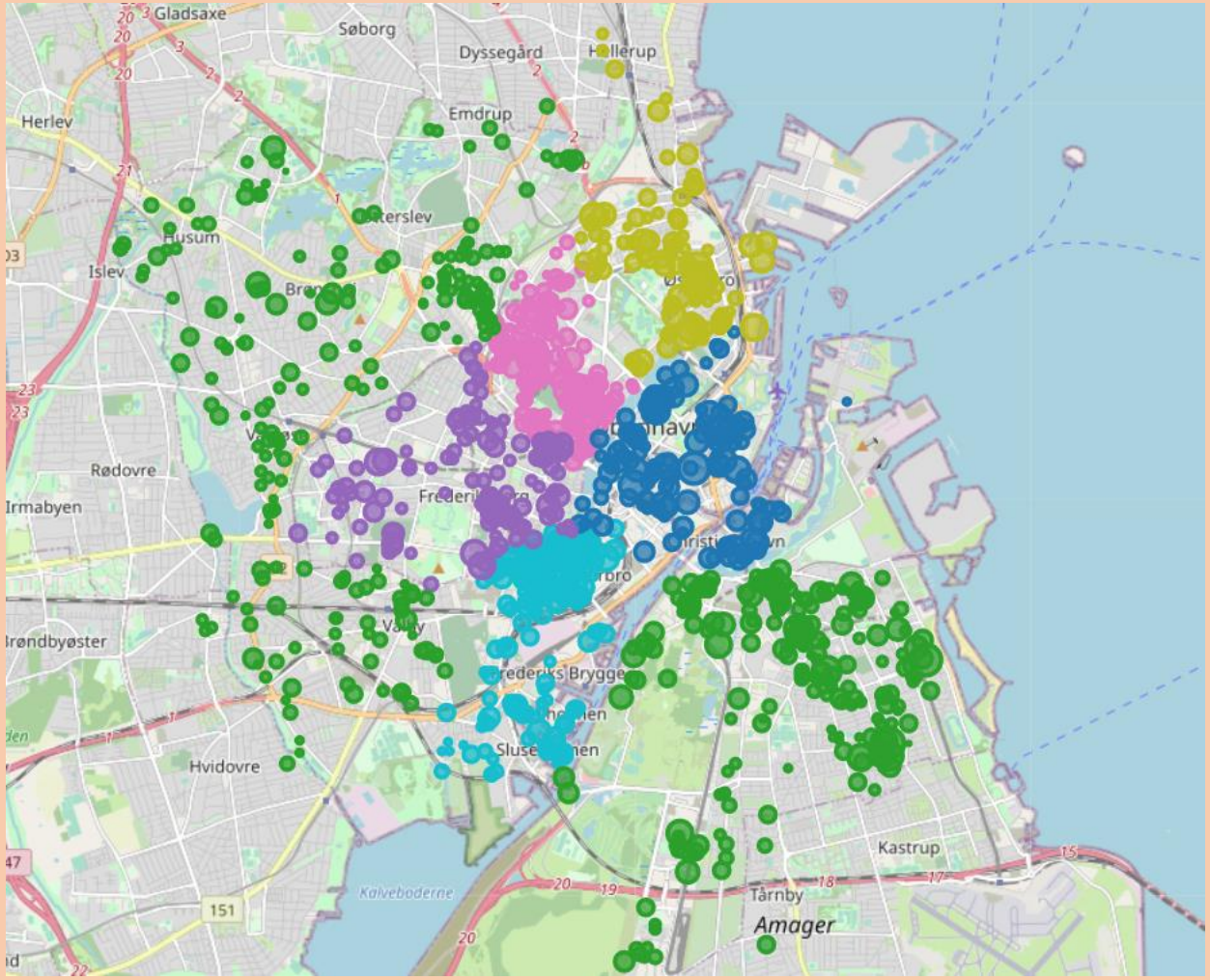


Data Analysis: Comparing Neighbourhoods



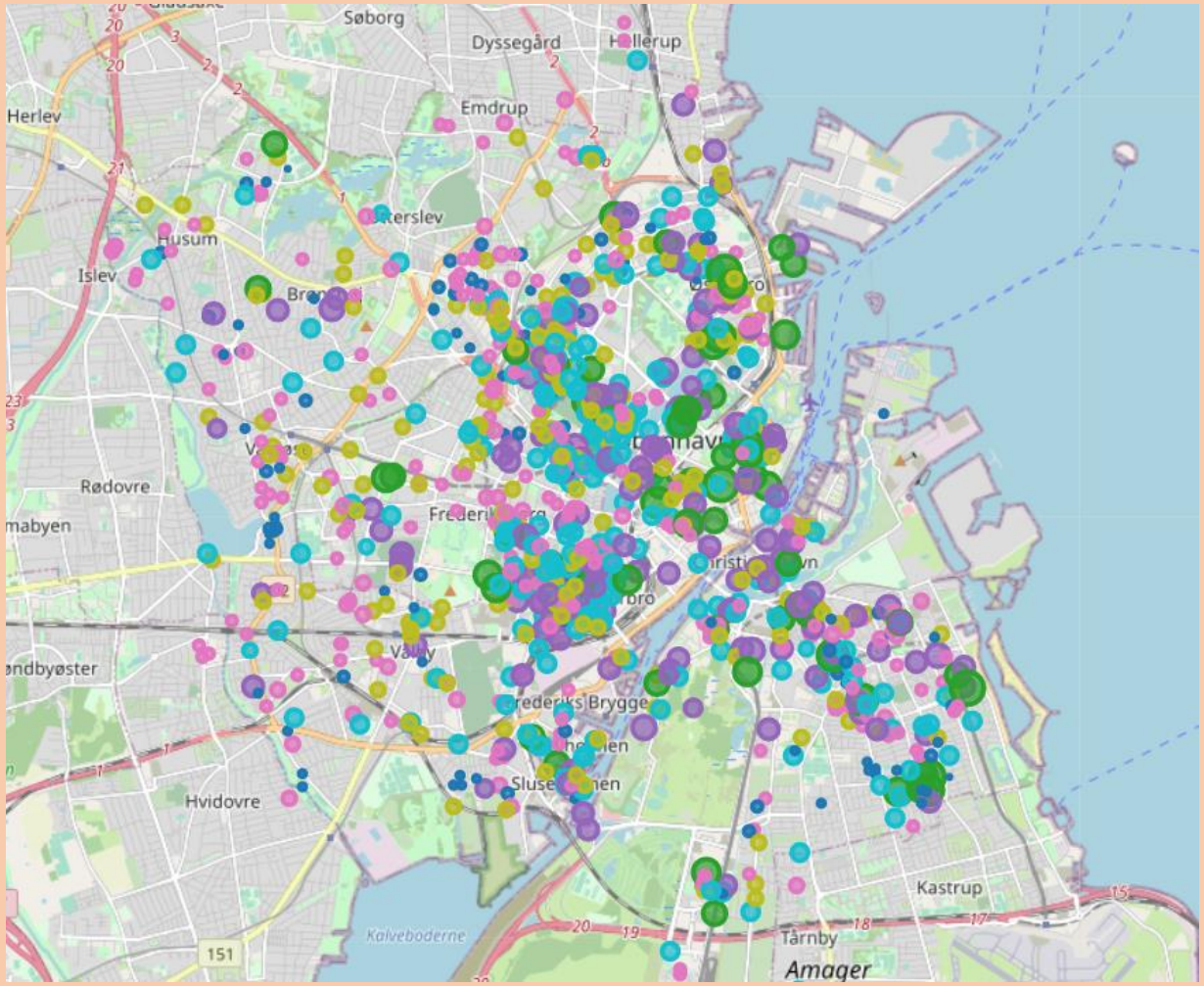
- Using Foursquare API, we collected nearby venues surrounding AirBnB room listings.
- We further grouped these listings according to their venue categories and counted their number of occurence in the neighbourhoods.
- Data shows that restaurants, cafes, museums, and bars/pubs are fairly more frequent in Indre By, parks are more frequent in Frederiksberg, while quick bites/street food are more frequent in Norrebro and Osterbro.
- As expected outer neighbourhoods such as Valby, Vanlose, and Amager did not fair well, relative to inner neighbourhoods.

K-Means Clustering: API Foursquare Features



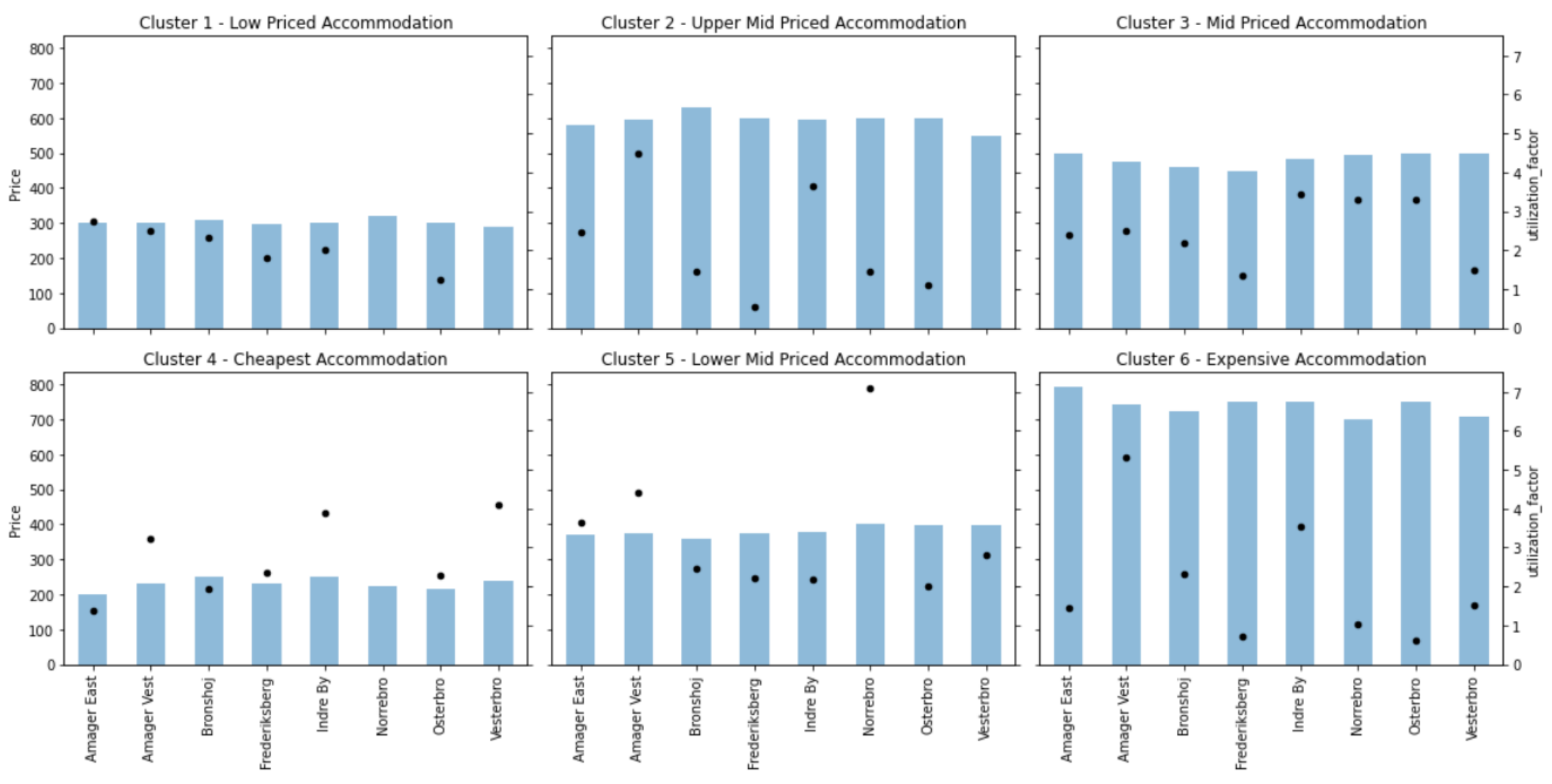
- Using K Means clustering with 6 clusters as parameter, we grouped the airbnb listing using the information obtained with API Foursquare as features.
- The clustering effectively identified the distinct neighbourhoods in the inner city, namely, Norrebro, Osterbro, Vesterbro, Frederiksberg, and Indre By.
- The neighbourhoods, located outside the city (green markers) were collectively clustered, suggesting similarity in features.

K-Means Clustering: Price and Availability

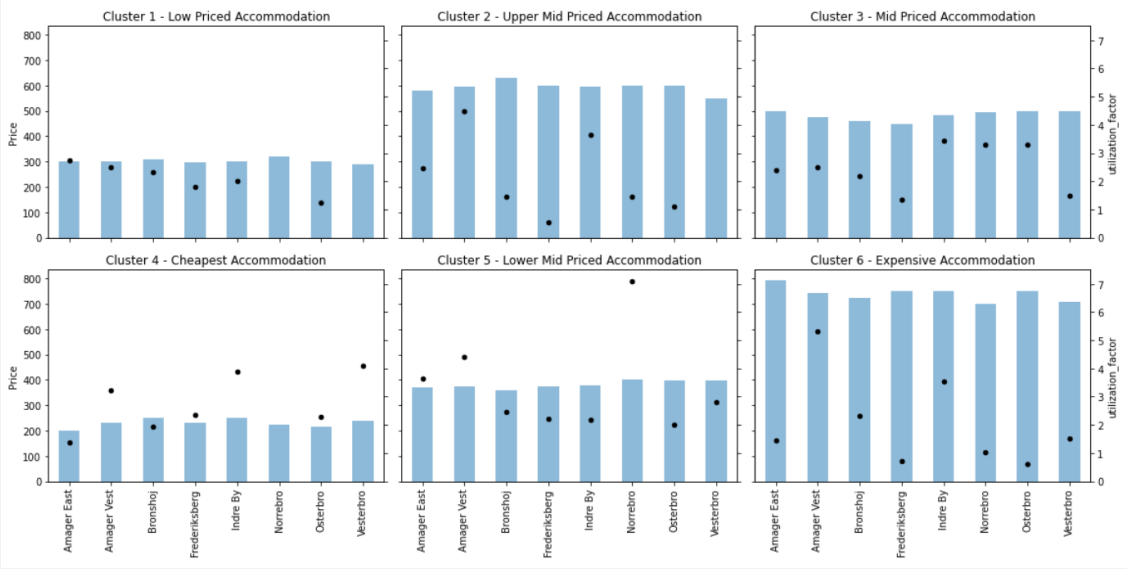


- Another similar routine was performed in the dataset but this time, addition of price and utilization factor was added in the feature set.
- The clusters did not follow the location of the neighbourhoods, but rather seemed to be dependent to price and utilization factor.

K-Means Clustering: Price and Availability



K-Means Clustering: Price and Availability



- K-Means clustering followed the price points of the individual private room listings.
- The cheapest accommodations, which were Cluster 4 and Cluster 1 has a median price of DKK 200 and DKK 300, respectively. Their utilization factor is fairly high at 3.5-2.
- The most expensive accommodations are in cluster 6 and Cluster 2 with a median price of DKK 600 and DKK 700. Their utilization is quite high in Indre By and Amager, while relatively low in other neighbourhoods.
- Interestingly, Amager East is fairly utilized irrespective of the price points and being outside the inner city.

Conclusion and Future Directions

- We explored a dataset describing the AirBnB private room listings in Copenhagen and the associated venue and venue categories around them.

- We identified that Indre By (City Center) is more than DKK 100 in median price point, despite the fact that the characteristics (i.e., venue categories) are quite similar to the surrounding neighbourhoods – Norrebro, Vesterbro, Osterbro, and Frederiksberg. Therefore, toursits should not be afraid in venturing in these areas.

- Conversely, the outer neighbourhoods such as Vanlose, Valby, Amager East and West, Bronshoj and Bispebjerg exhibited similar characteristics. These were confirmed using Kmeans clustering. As such, tourists should expect similar activities in these neighbourhoods when booking in these areas.

- Different price points are present in all of Copenhagen's neighbourhood. This means that tourists can score various price points irrespective of the neighbourhood. Interestingly, however, outer neighbourhood areas with higher price points are utilized effectively. This could be due to newly built areas in these neighbourhood.

- This analysis could be further by looking at the API Foursquare on a granular level and utilizing tourist traffic data in Copenhagen.