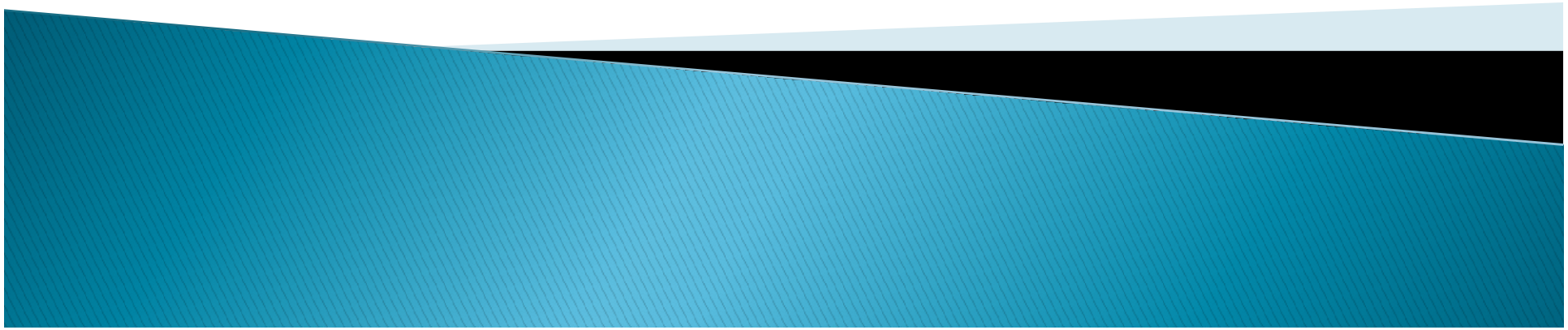
# Analysis of similarities between cities



# Introduction

- More people are relocating to major cities around the world as global mobility increases
- However, people might be worried about the quality of life in the new city, especially if the city is unfamiliar to them
- Hence, this project is to do a comparison of major cities located in North America, Europe and Asia-Pacific to find clusters of similar cities

# Description of data (1/2)

- A total of 30 major cities are selected, with 10 from each region:
  - North America: New York, Washington DC, Boston, Houston, Chicago, Los Angeles, Seattle, San Francisco, Toronto, Vancouver
  - Asia-Pacific: Tokyo, Osaka, Shanghai, Beijing, Seoul, Kuala Lumpur, Jakarta, Melbourne, Singapore, Bangkok
  - Europe: London, Paris, Berlin, Frankfurt, Amsterdam, Madrid, Brussels, Copenhagen, Vienna, Zurich.

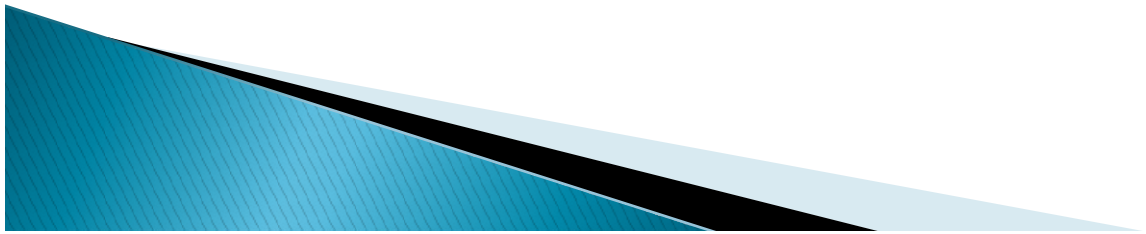
# Description of data (2/2)

The features used are:

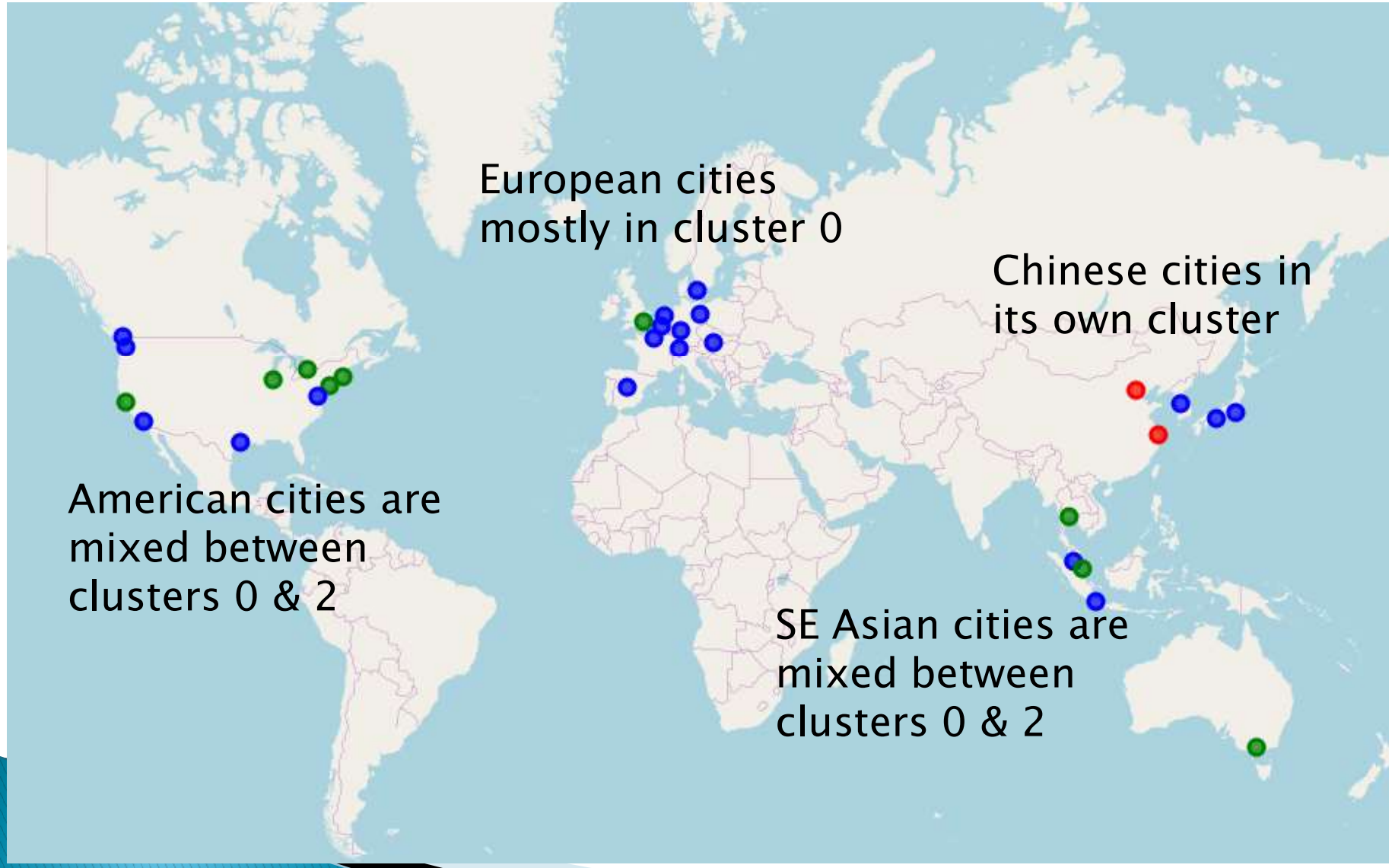
- Number of venues within 1km radius of the city centre in 8 different categories (restaurant, bar, gym/fitness center, park, movie theatre, supermarket, laundromat, metro station);
- Average number of likes for all the restaurants given in the search for number of venues;
- Variety score for restaurants within 1km radius of the city centre (at least 5 restaurants in each major cuisine type: Asian, French, Italian, Vegetarian, Steakhouse).

# Methodology

- Pre-processing to condition the feature vectors individually so that their distribution is normalised to zero mean and unit variance.
- k-means algorithm to cluster the cities into 3 groups.



# Results & Discussion



# Conclusion

- This study offers an insight to the broad similarities of the cities studied and as well as the distribution of each category of city in the different regions.