

LONDON  
CITY OF  
OPPORTUNITIES

# INTRODUCTION

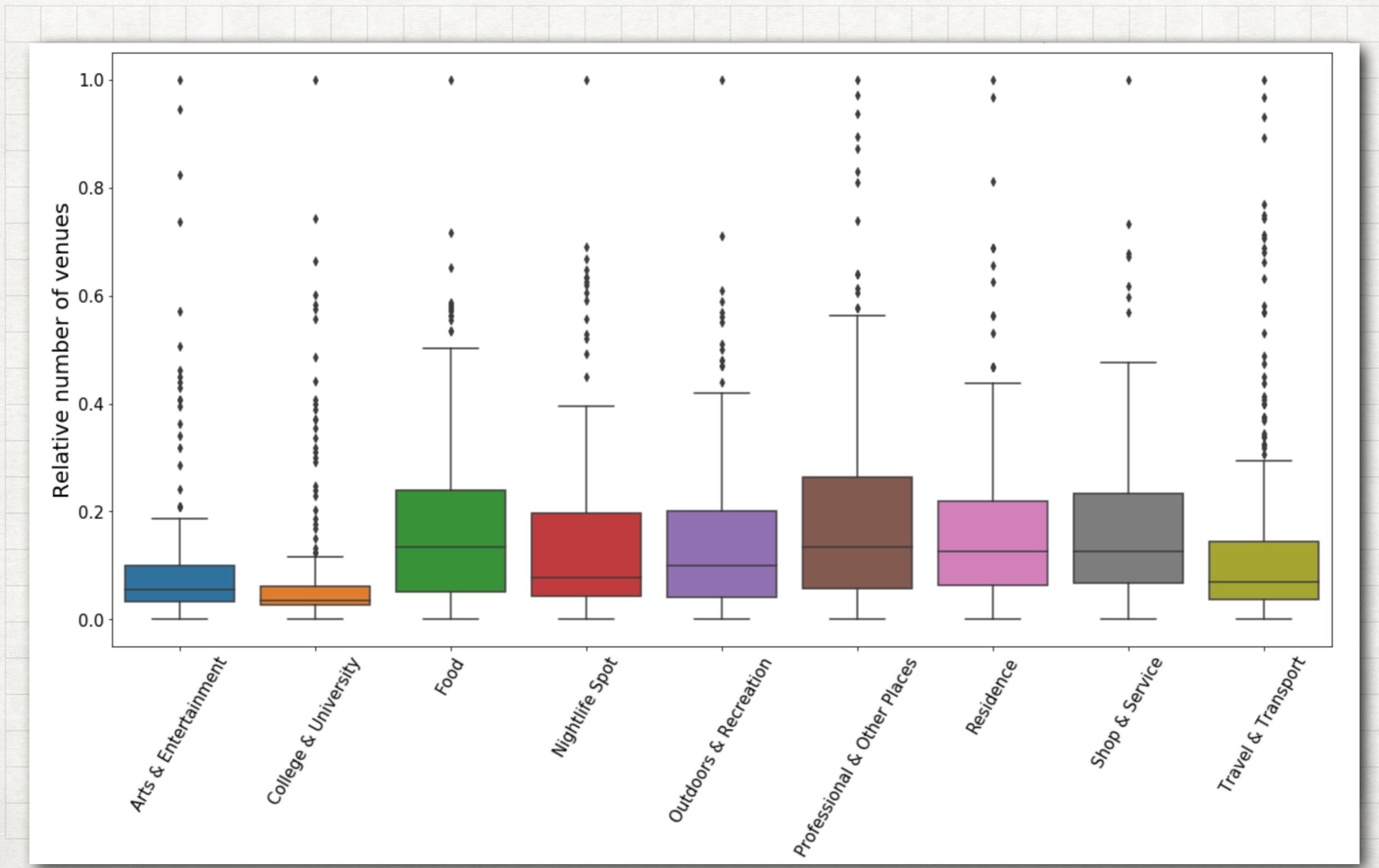
- London is one of the largest and busiest city in Europe. It also one of the world's capitals of finance, culture, art and entertainment. With all of its cultural diversity it provides many opportunities for new businesses and working professionals. Greater London Authority published population projection for the city and updated it annually at both local authority and ward level for year up to 2050.
- In the modern worlds of cosmopolites and fast changing trends it becomes increasingly important to examine and understand all these changes quantitatively. City government and urban planners, entrepreneurs and investors - all have an interest in identifying early opportunities and growing urban footprint in prospective districts.
- My idea here is that using population projection data and current venues information (from extensive Foursquare database) I can do following analysis:
  - - classify neighbourhoods to following categories: underdeveloped, business center or downtown, highly developed;
  - - explore underdeveloped neighbourhoods looking at significant population growth;
  - - identify opportunities for new businesses in urbanizing neighbourhoods;

# DATA

FOR THIS ANALYSIS I WILL USE FOLLOWING SOURCES OF DATA

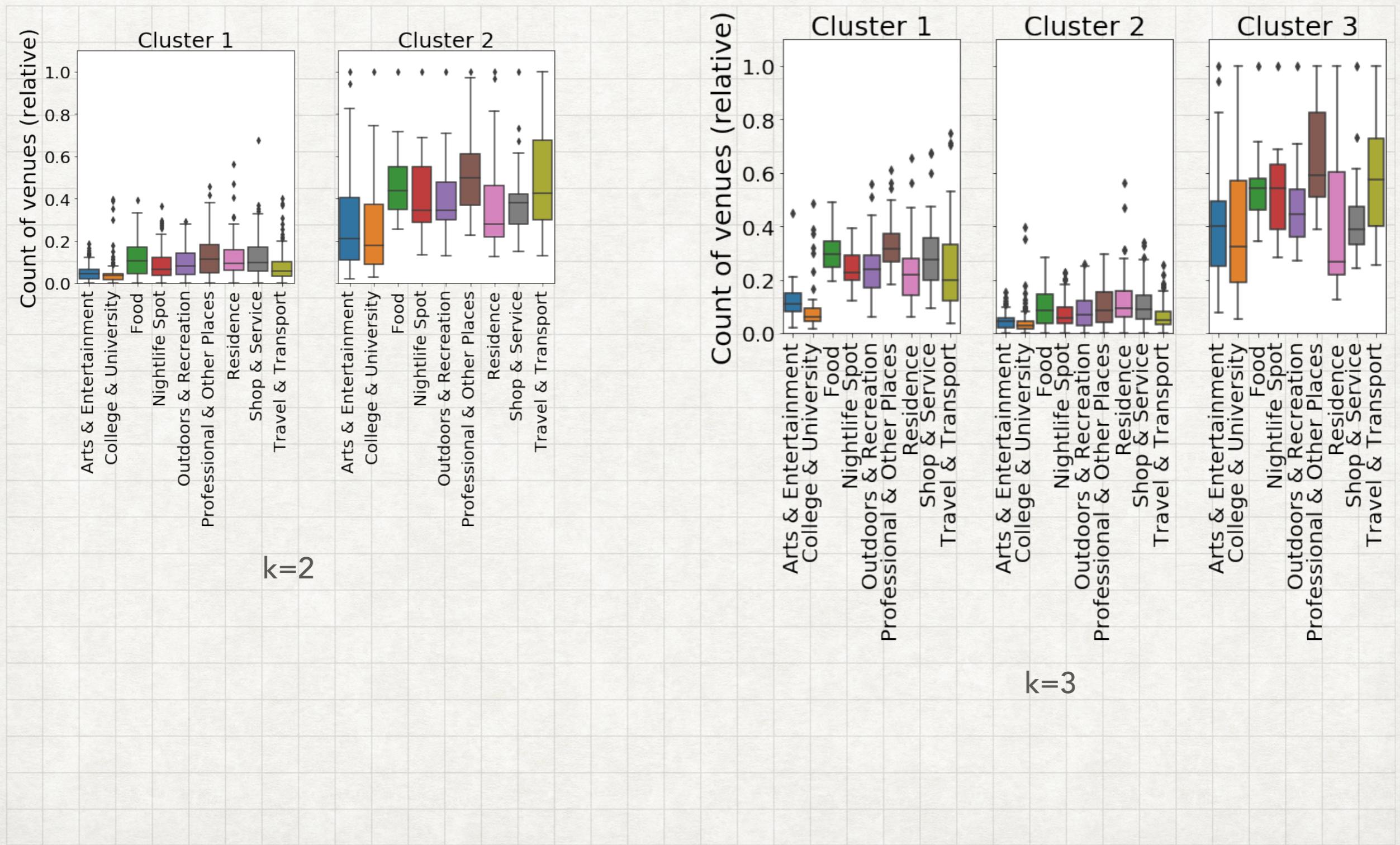
- Statistical GIS Boundary Files for London, published on [London Datastore](<https://data.london.gov.uk/dataset/statistical-gis-boundary-files-london>). It contains National Statistics data © Crown copyright and database right [2015]" and "Contains Ordnance Survey data © Crown copyright and database right [2015]".
- Greater London Authority Population and Household Projections dataset, published on [London Datastore](<https://data.london.gov.uk/dataset/land-area-and-population-density-ward-and-borough>). From this source we can take following data: ward name, ward location, ward current population and population projection for years 2020-2050.
- London Ward Profiles and Atlas by [London Datastore](<https://data.london.gov.uk/dataset/ward-profiles-and-atlas>). We will take all the additional information about wards from here: borough it belongs to, what part of city it belongs to, etc
- Venues data using [Foursquare API](<http://developer.foursquare.com>). From this source we can take following data: list of venues names and locations in each ward of interest, also venue category (we will only need high level category, which will extract also from Foursquare, but as list of categories, available [here](<https://developer.foursquare.com/docs/resources/categories>)).

# GENERAL DISTRIBUTION OF VENUES



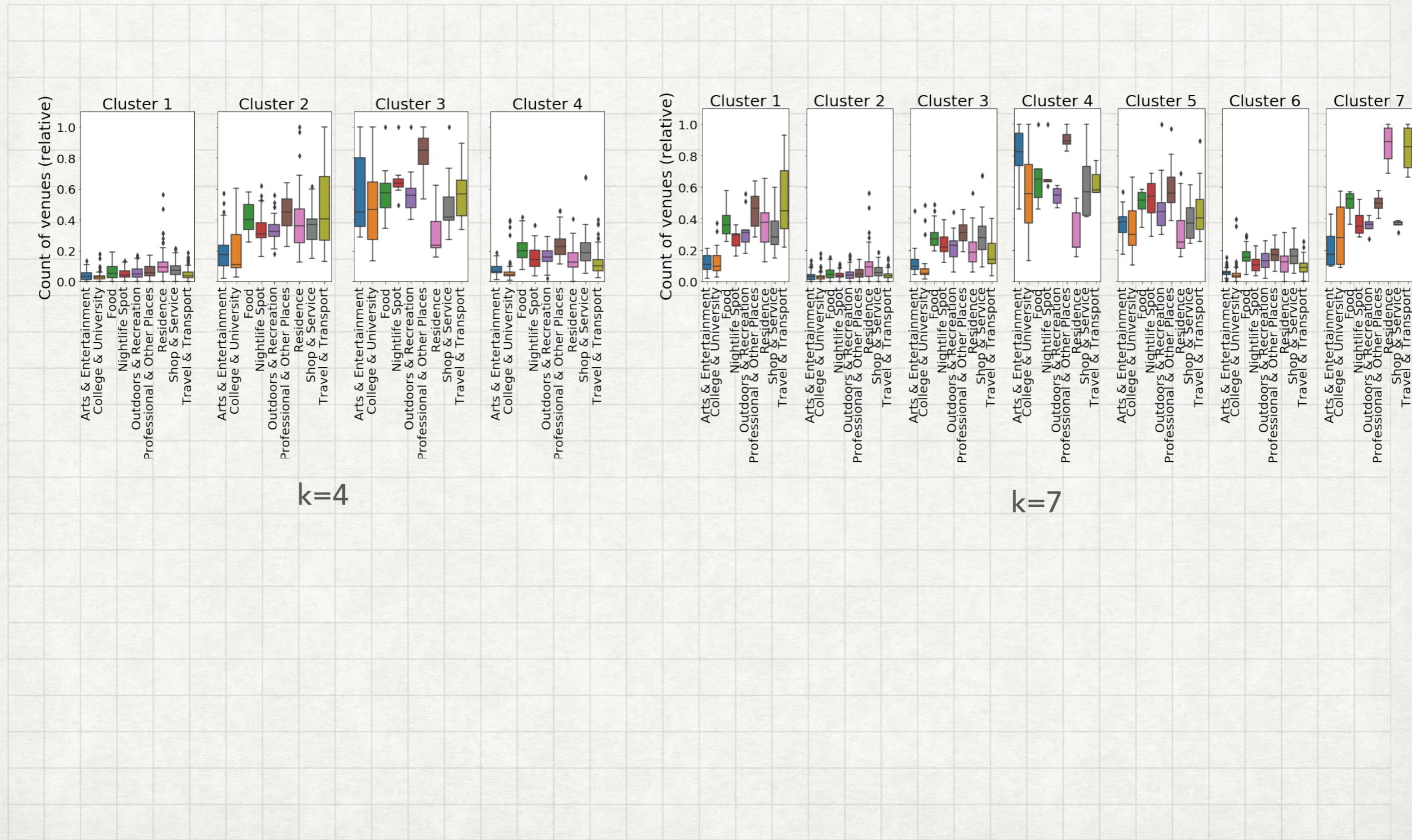
# K-MEANS CLUSTERING

## EXPERIMENTING WITH NUMBER OF CLUSTERS



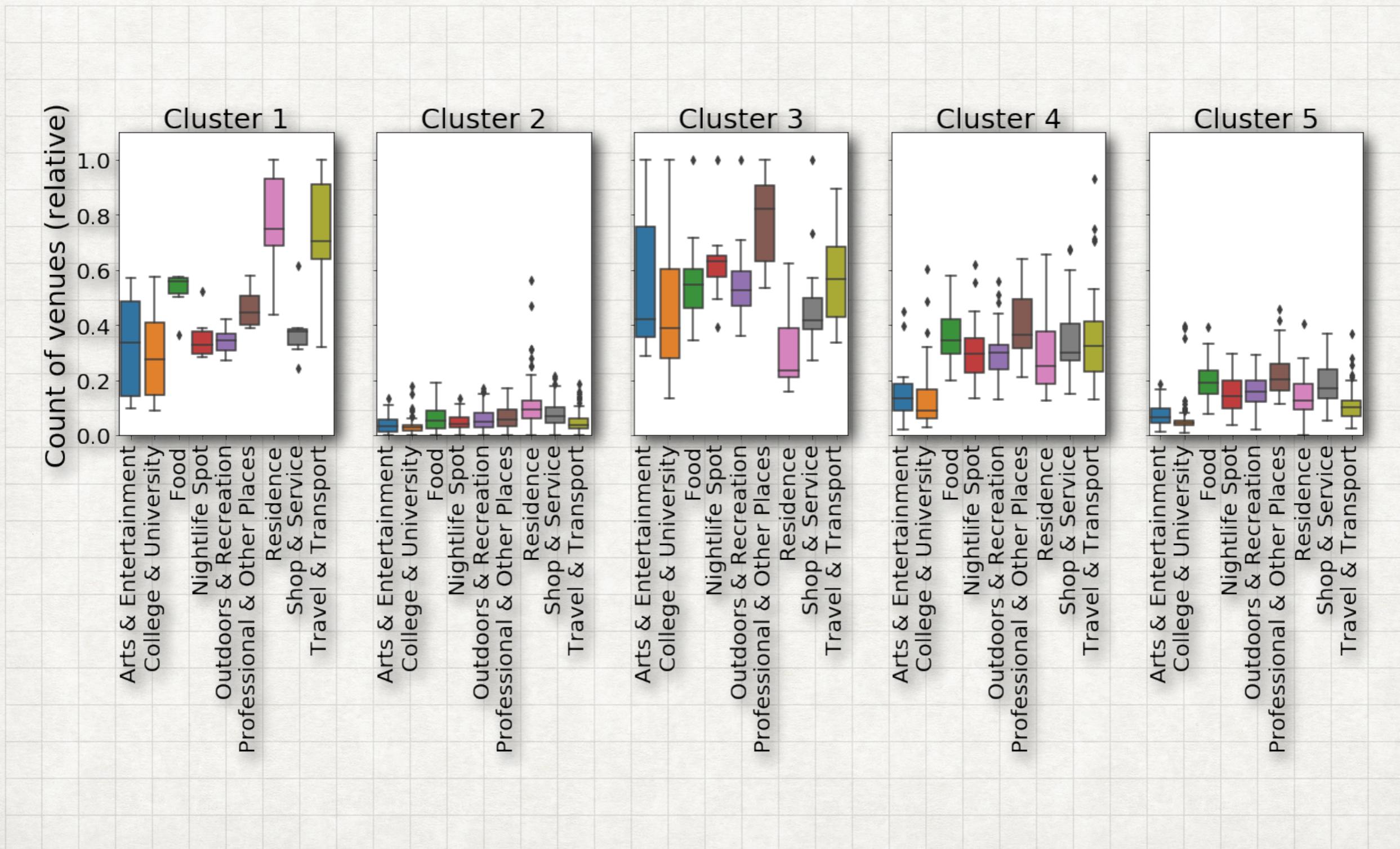
# K-MEANS CLUSTERING

## EXPERIMENTING WITH NUMBER OF CLUSTERS



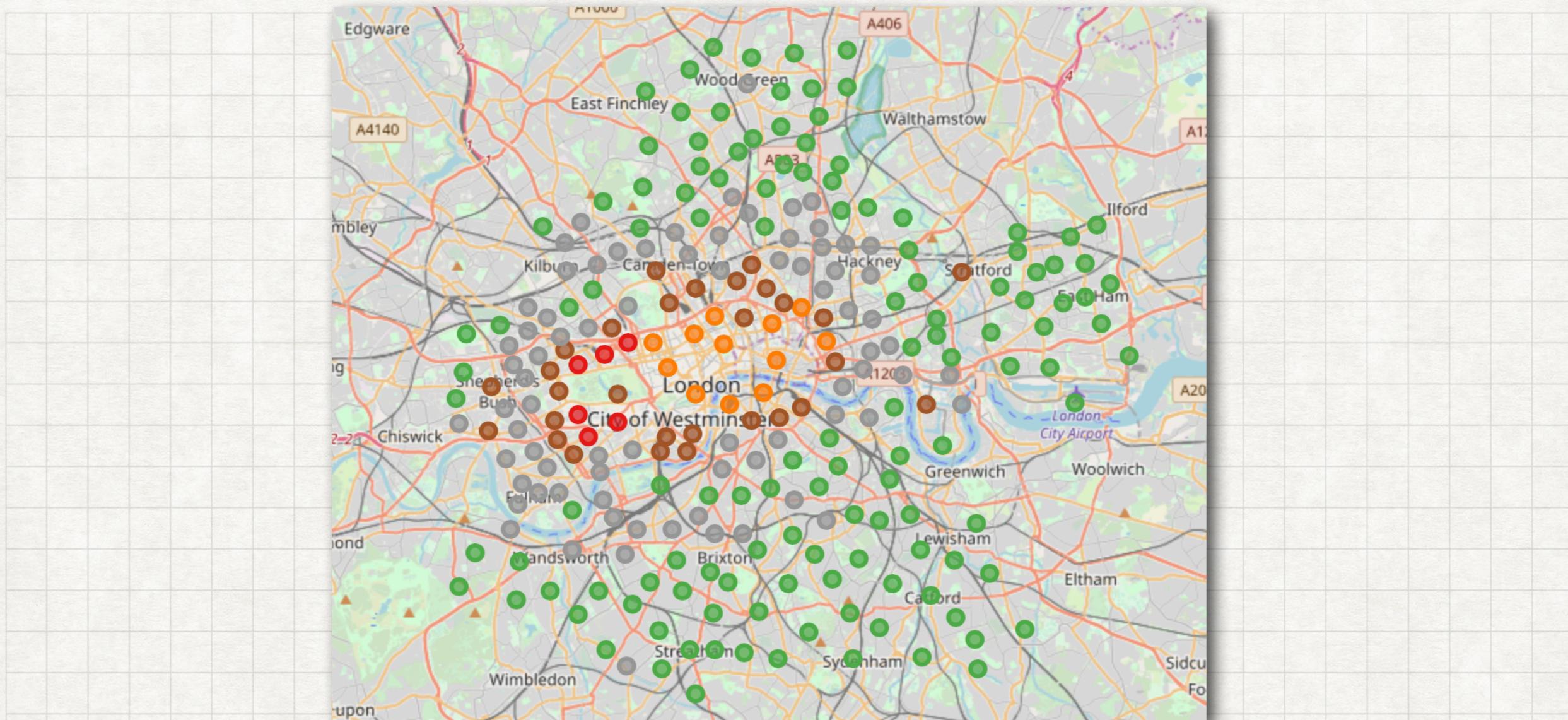
# K-MEANS CLUSTERING

## BEST K FOUND



# MAPPING RESULTS

## 5 CLUSTERS



RED(1) CLUSTER - HAS MODERATE SCORES WITH RESIDENTIAL AND TRANSPORT BEING THE MOST POPULAR. THESE ARE HIGHLY-DEVELOPED RESIDENTIAL SUBURBS.

GREEN(2) CLUSTER - UNDERDEVELOPED, HAS LOW FREQUENCIES OF ALL CATEGORIES OF VENUES.

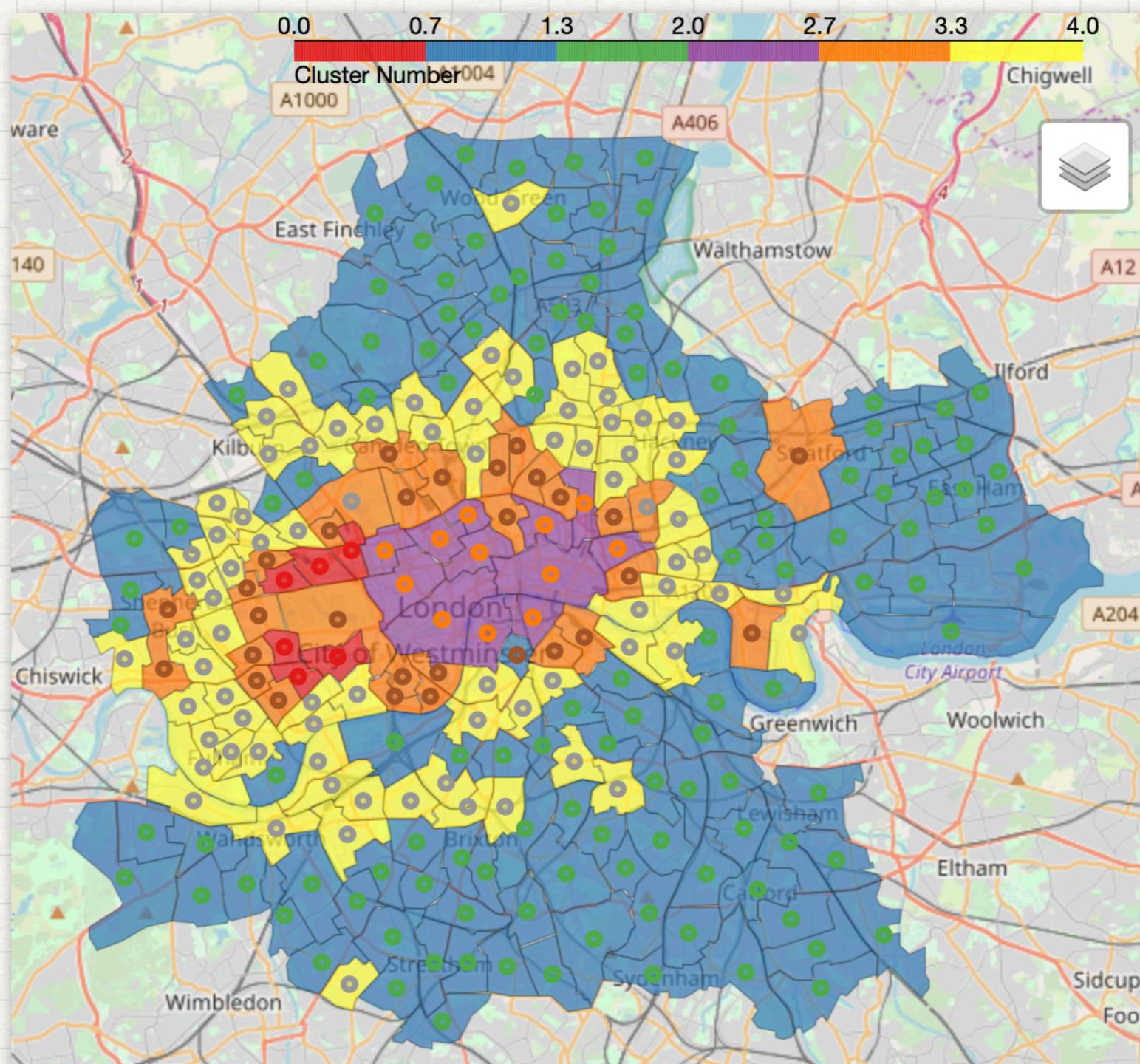
ORANGE (3) CLUSTER - BUSINESS OR DOWNTOWN CLUSTER. HIGHLY DEVELOPED BUSINESS PART OF THE CITY, WHICH HAS VERY HIGH FREQUENCIES FOR ALL VENUE CATEGORIES EXCEPT RESIDENTIAL. IT IS PART OF THE CITY WHERE PEOPLE MOSTLY WORK AND ENTERTAIN, BUT NOT LIVE HERE.

BROWN(4) CLUSTER - HIGHLY DEVELOPED AREA, HAS HIGH FREQUENCIES WITH LESS RESIDENTIAL PLACES AND MORE PROFESSIONAL ONES. DEVELOPED PROFESSIONAL AND INDUSTRIAL AREAS.

GREY(5) CLUSTER - MID-DEVELOPED, HAS MODERATE FREQUENCIES OF ALMOST ALL CATEGORIES WITH SMALL PREVALENCE OF PROFESSIONAL PLACES

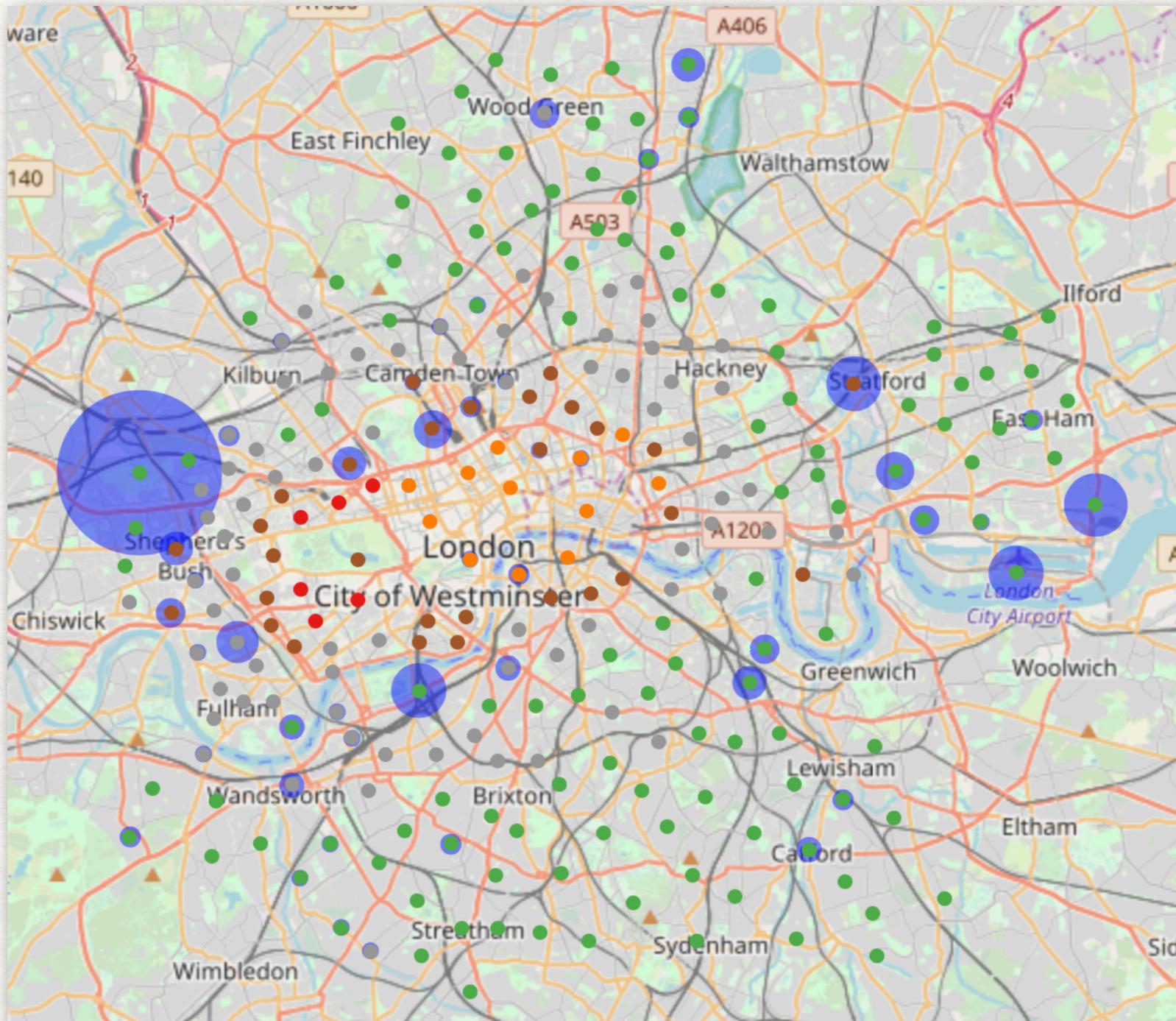
# MORE MAPS PLEASE

## CHOROPLETH WITH CLUSTERS



# POPULATION PROJECTION

## 2020-2040



## GREEN(2) CLUSTER - WINNER IN POPULATION GROWTH

BROWN(4) CLUSTER - 2ND PLACE

# RESULTS

OUR ANALYSIS SHOWS THAT THERE ARE DISTINCT AREAS IN INNER LONDON WHICH CAN BE GROUPED TOGETHER AS SIMILAR. THERE ARE BIG NUMBER OF UNDERDEVELOPED AREAS ACCORDING TO PRESENCE IN THEM DIFFERENT VENUE CATEGORIES WHICH CAN HELP IDENTIFY PLACES FOR NEW BUSINESSES. POPULATION PROJECTION DATA SHOWS THAT THERE WILL NOT BE A HUGE GROWTH IN POPULATION IN ANY OF THE IDENTIFIED CLUSTER. NEVERTHELESS THE MOST "OPORTUNISTIC" AREAS WILL STILL SHARE MORE THAN A HALF OF CITY POPULATION.

IDENTIFYING CURRENTLY UNDERDEVELOPED AREAS CAN GIVE BIG ADVANTAGE TO EARLIER BUSINESSES AND SERVICE PROVIDES IN THESE AREAS.

# CONCLUSION

OUR MAIN AIM WAS TO IDENTIFY LESS DEVELOPED AND UNDERDEVELOPED AREAS IN LONDON. WE SUCCESSFULLY ACHIEVED THIS GOAL I BELIEVE.

DESPITE THESE RESULTS IT WOULD BE INTERESTING TO FURTHER CONDUCT CLUSTERING ANALYSIS CONSIDERING OTHER SOCIO-ECONOMIC MEASURES LIKE DEMOGRAPHICS, RELIGION, HOUSING TYPES, EMPLOYMENT RATES, INCOME PER HOUSEHOLD AND MANY, MANY OTHER. FORTUNATELY, LONDON DATASTORE HAS A HUGE AMOUNT OF OPEN DATA FOR FURTHER ANALYSIS AND GETTING DEEPER INSIGHT INTO FUTURE OF LONDON AREAS.

**THANK YOU**