

An aerial night photograph of the Mexico City skyline, showing a dense cluster of illuminated buildings and streets against a dark sky.

# Clustering Neighborhoods in Mexico City

Uriel Ibarra

February 29th, 2020

# Digitalization has reach the old city

- Mexico City is one of the largest and most important human metropolises on the planet and the oldest city in the Americas. Its history is surely complex, and despite the country's issues, its capital is a beautiful place with many things to offer.
- Today more than ever, access to information is a right, data has become an indispensable tool for the improvement of human activities.
- This capital had been lagging behind and far away of the smartness that involves other metropolis, but recently new boost citizen policies has been implemented as part of a collaborative, multi-stakeholder process.

# Technology helps us to take decisions

- The most common factors stakeholders take into account to select a place to move have been analyzed to **group Mexico City neighborhoods into clusters**, using data since tools.

# Data acquisition and cleaning

Data from Mexico City's new Digital Agency of Public Innovation:

- Neighborhood location
- Neighborhood Human Development Index
- Seismic Risk Atlas
- Public transport stations (Metro, Bus, Trolley)
- Public health center location

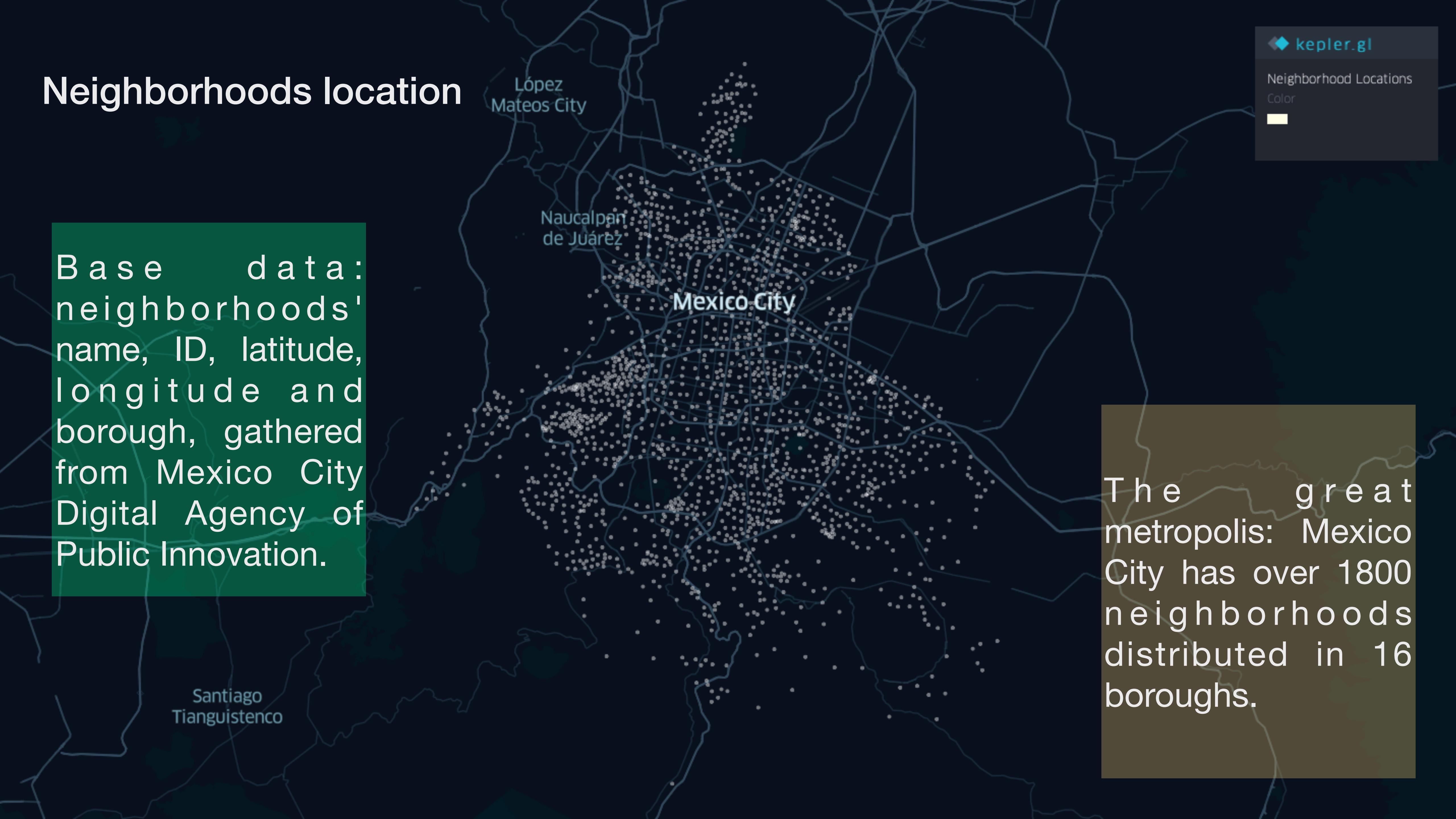
Data from Foursquare API:

- Recreation center location (*Art and entertainment* Foursquare category)
- Food shop location (*Food and drink shop* Foursquare category)



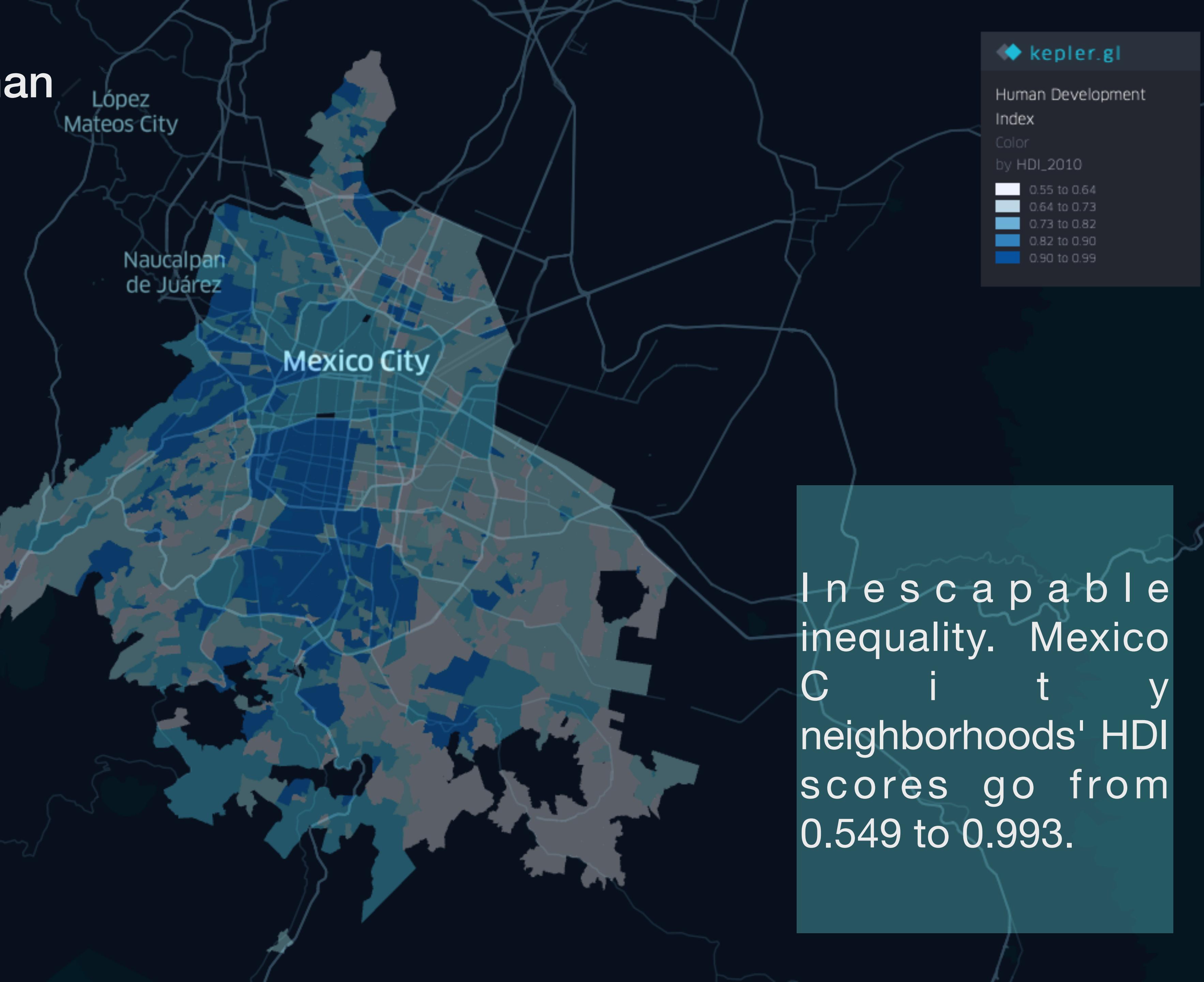
# Neighborhoods location

Base data:  
neighborhoods' name, ID, latitude, longitude and borough, gathered from Mexico City Digital Agency of Public Innovation.



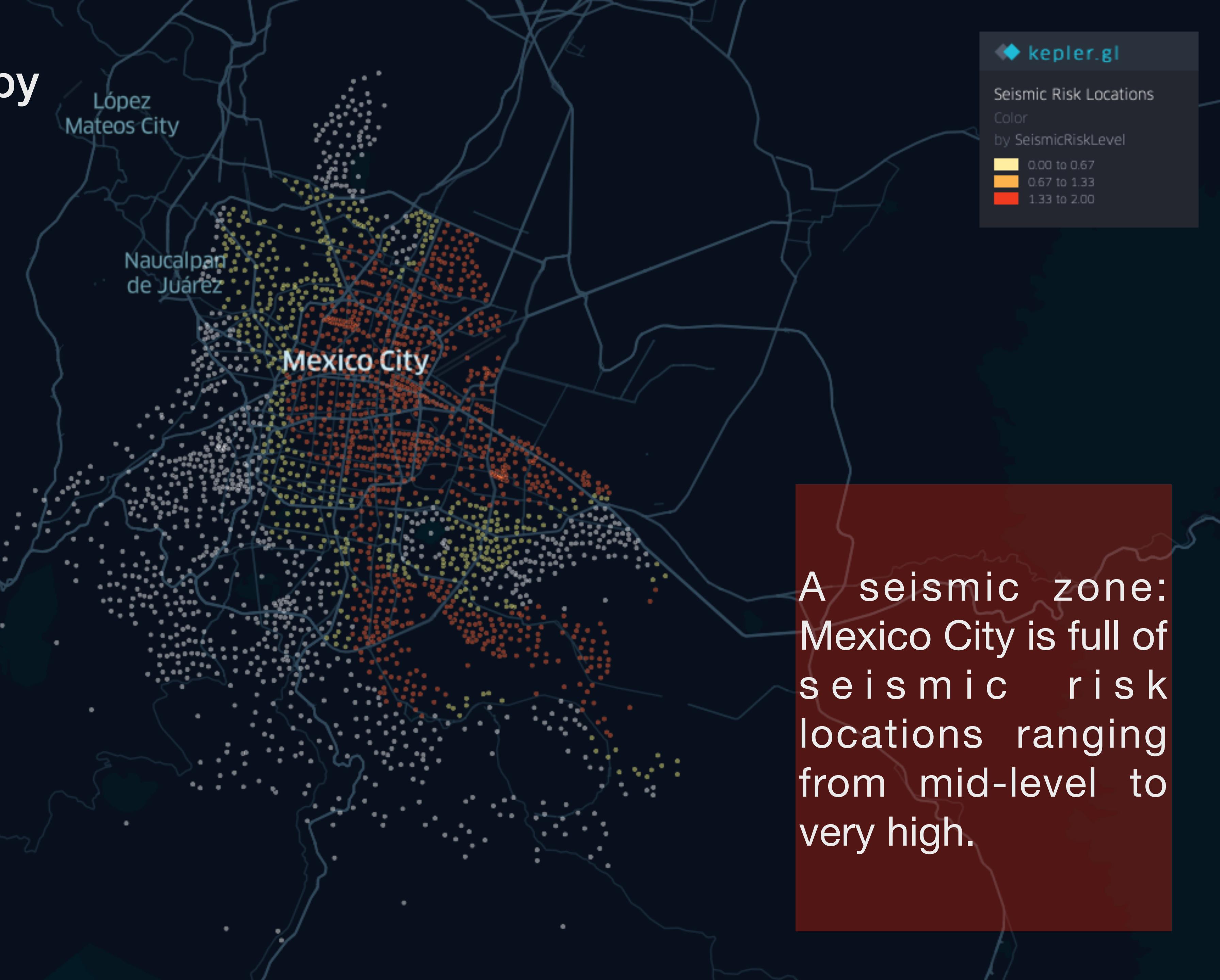
# Neighborhoods by Human Development Index

HDI data:  
neighborhoods' 2010 Human Development Index score, gathered from Mexico City Social Development Evaluation Council.



# Seismic risk locations by level

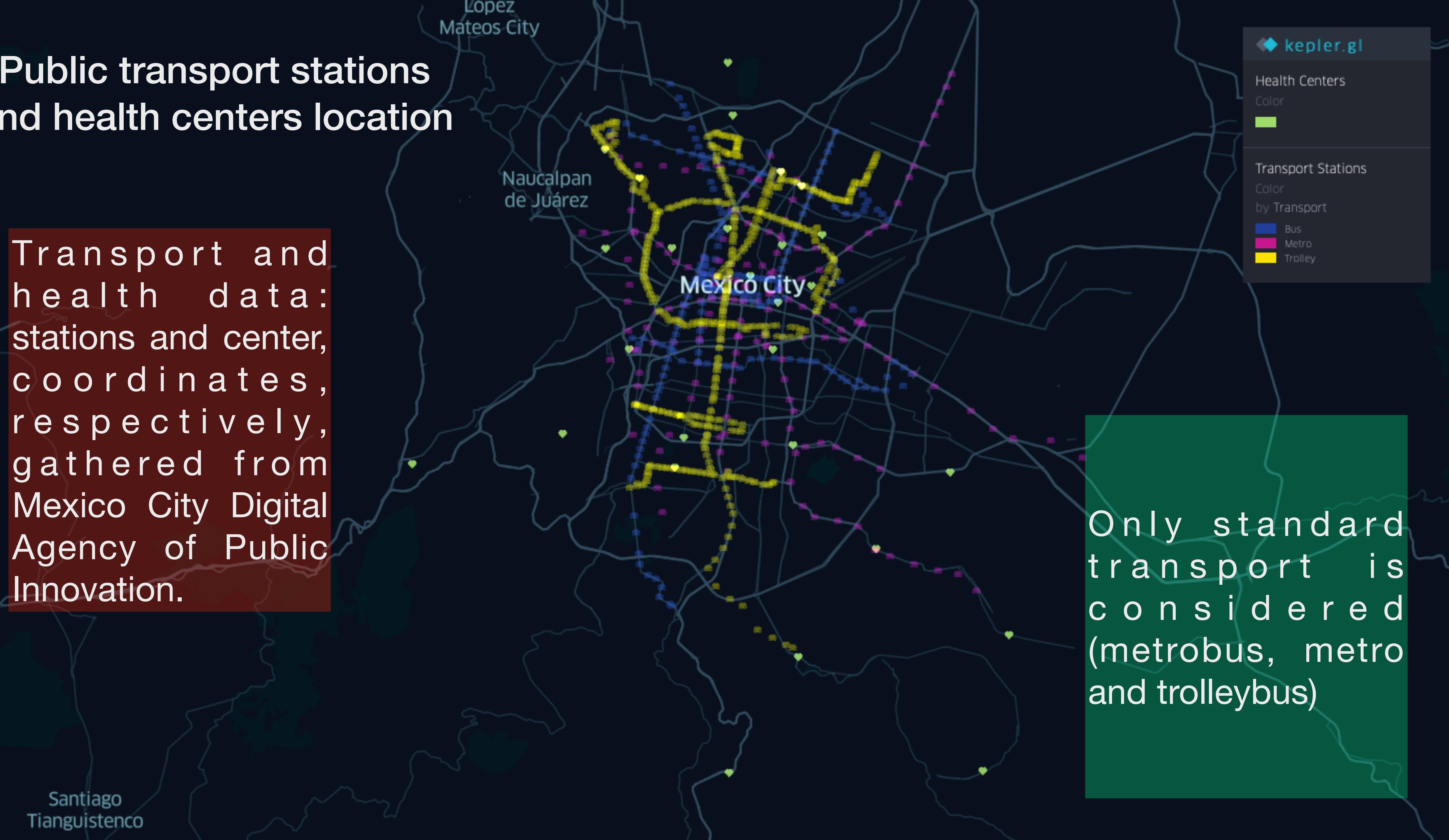
Seismic data: seismic risk locations and level, gathered from Mexico City Digital Agency of Public Innovation.



A seismic zone: Mexico City is full of seismic risk locations ranging from mid-level to very high.

# Public transport stations and health centers location

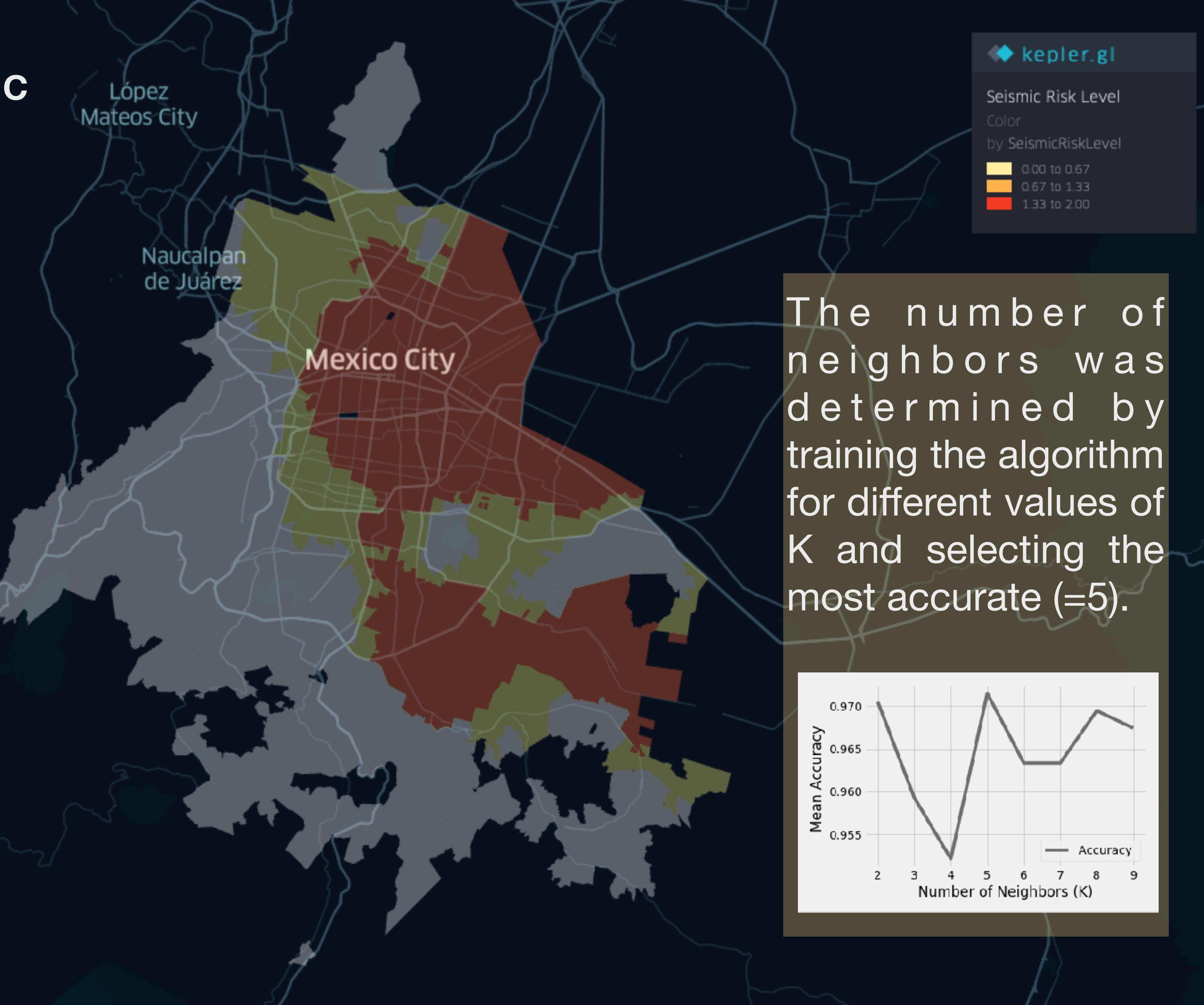
Transport and health data: stations and center, coordinates, respectively, gathered from Mexico City Digital Agency of Public Innovation.



Only standard  
transport is  
considered  
(metrobus, metro  
and trolleybus)

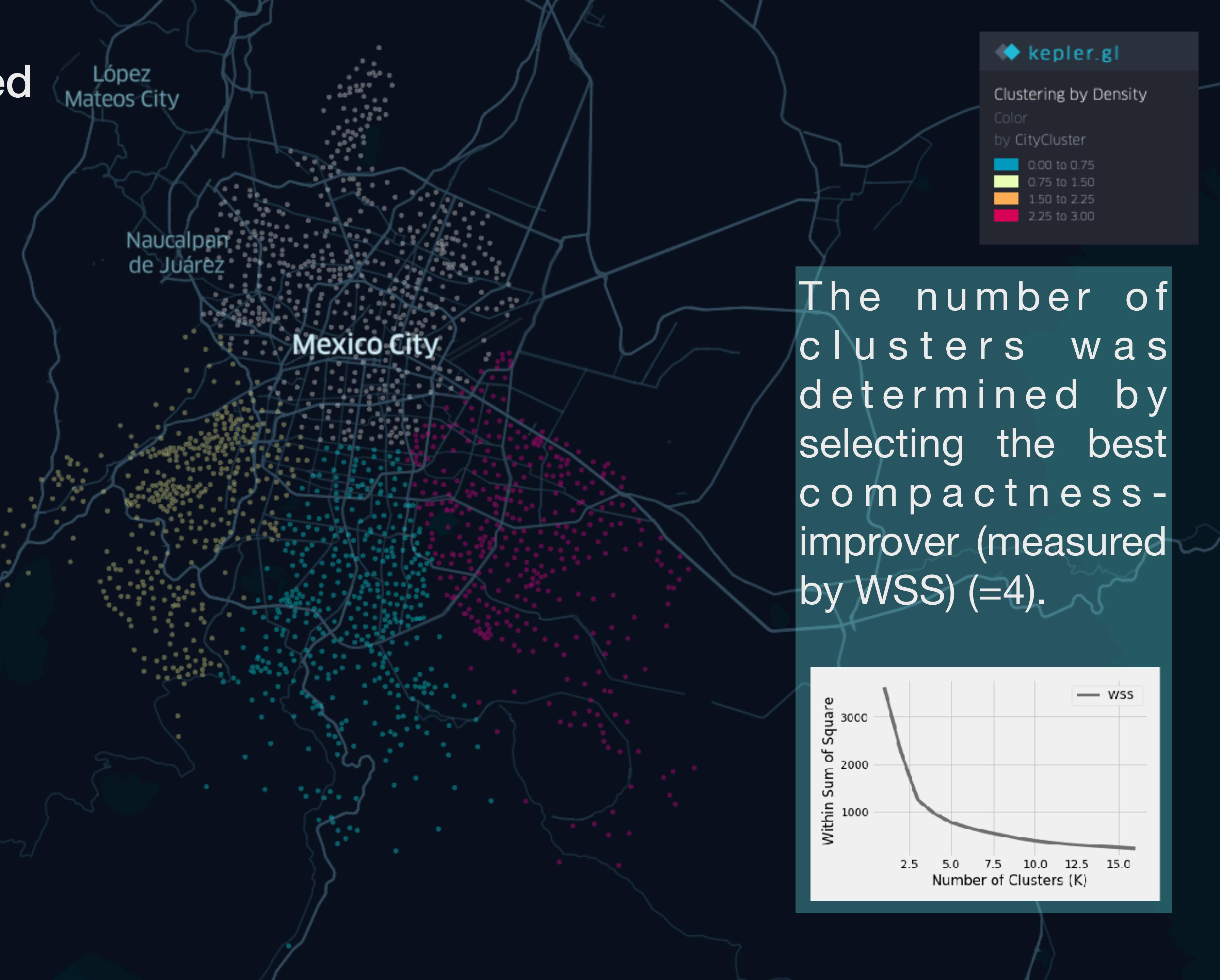
# Neighborhoods by seismic risk level

K-nearest neighbor algorithm was used to classify each Mexico City neighborhood according to its proximity to seismic risk locations.

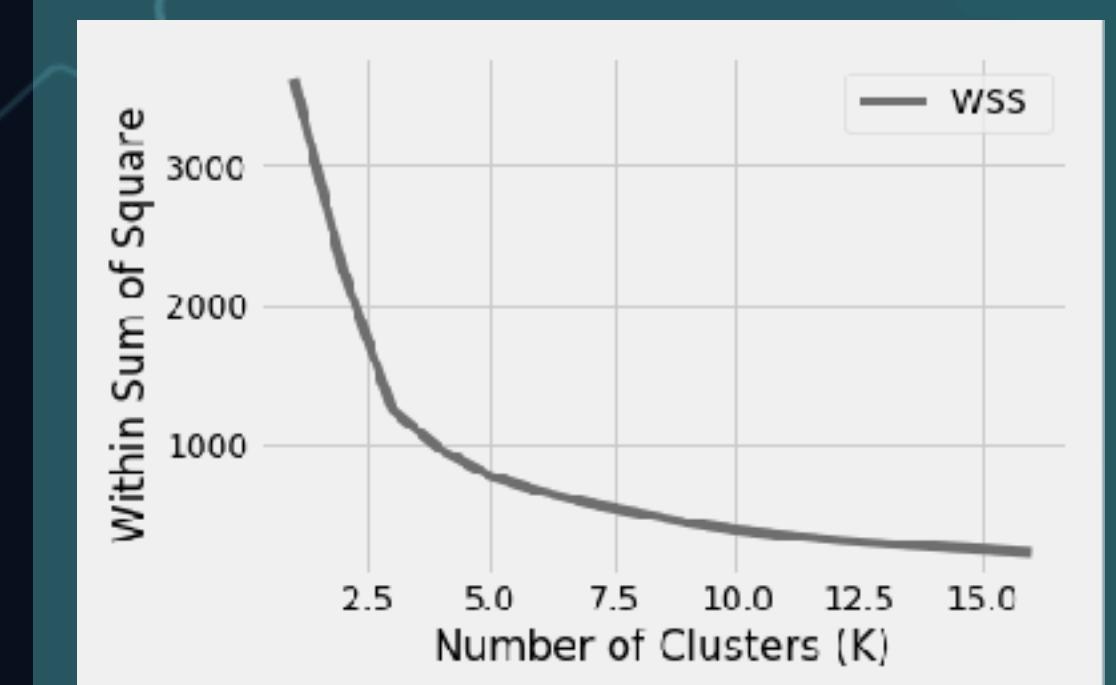


# Neighborhood clustered by density

Mexico City neighborhoods were clustered by density using their latitude and longitude values and the K-means algorithm.

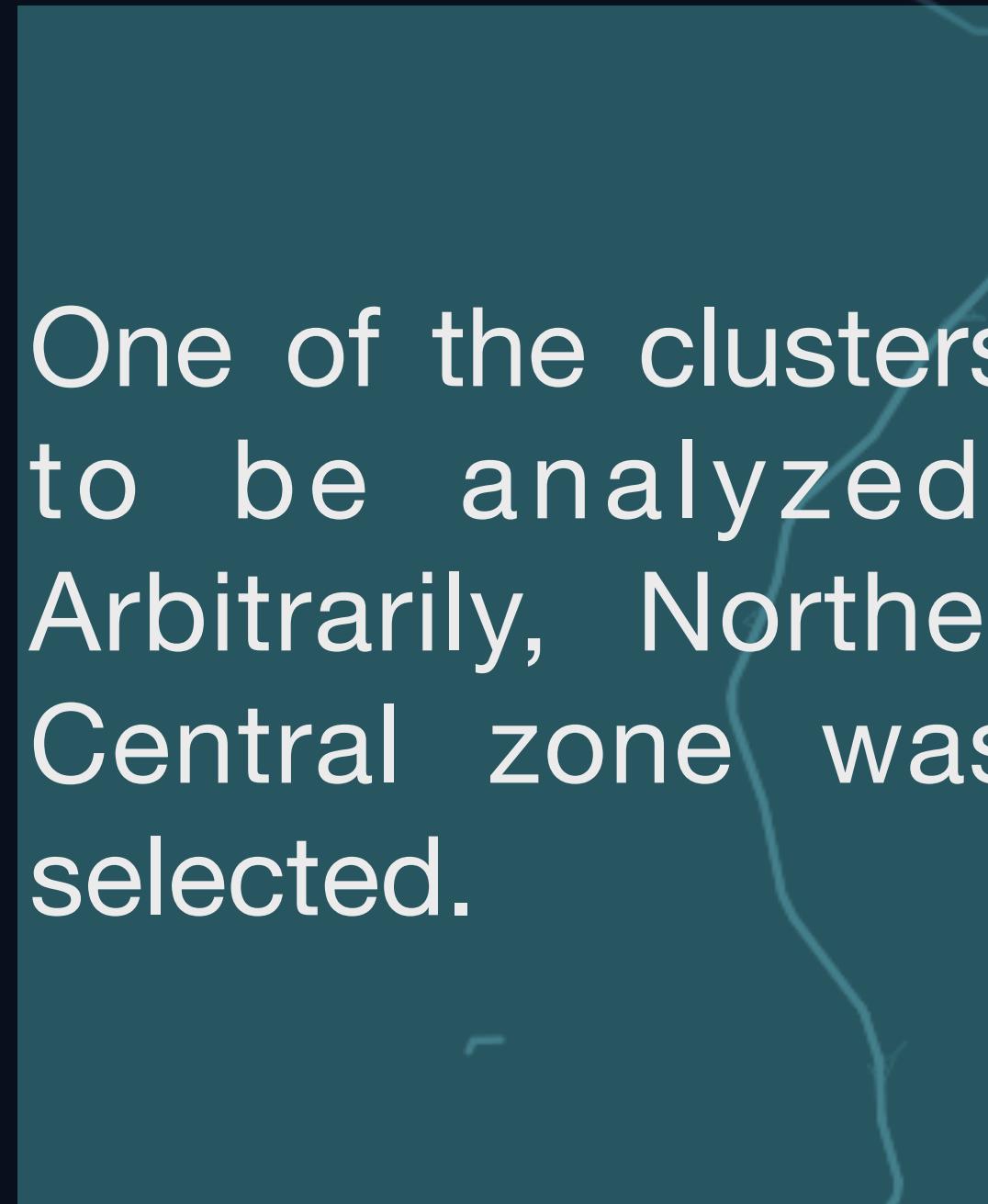


The number of clusters was determined by selecting the best compactness-improver (measured by WSS) (=4).



# Northern Central Mexico City area

One of the clusters to be analyzed. Arbitrarily, Northern Central zone was selected.



Norther Central cluster has 583 neighborhoods, included the well known Mexico City Main Square: The Zocalo (CENTRO IV).

# Venues Data

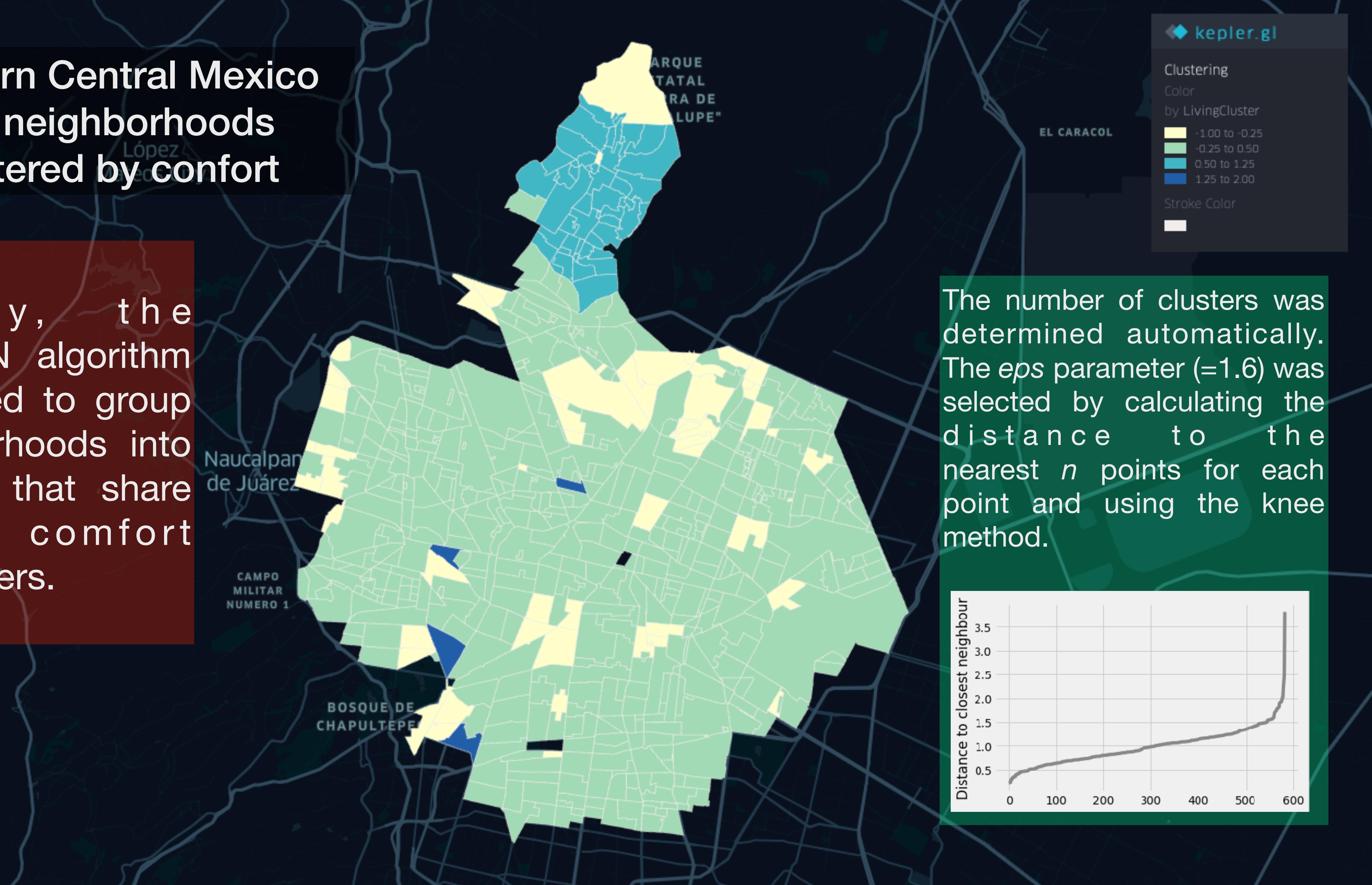
- Foursquare API was used to get info on venues in each neighborhood (latitude, longitude, name and category).
- They were taken into account: venues in the 'Food and drink shop' category, but only those that are proper places to get food supplies (food shops), and also all venues in 'Art and entertainment' category (recreation centers).

# Feature selection

- Using all data presented before, next features were extracted and generated.
  - Human Development Index score.
  - Seismic risk level.
  - Number of transport stations within 1 km around.
  - Distance to the nearest transport station (trolley, bus or metro).
  - Number of health centers within 5 km around.
  - Distance to the nearest public health center.
  - Number of recreation centers within 3 km around.
  - Distance to the nearest recreation center.
  - Number of food shops within 1 km around.
  - Distance to the nearest food shop.

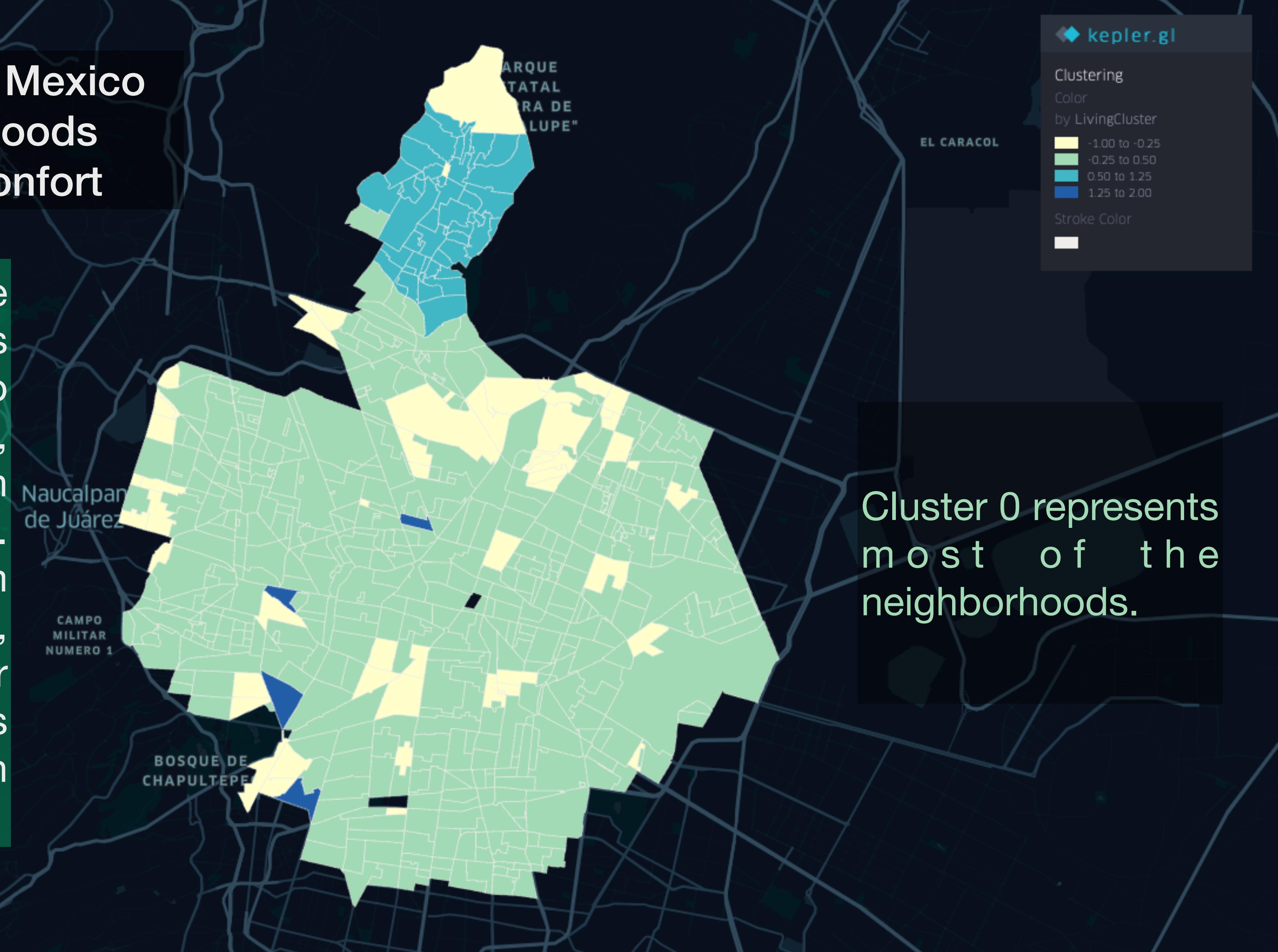
# Northern Central Mexico City neighborhoods clustered by confort

Finally, the DBSCAN algorithm was used to group neighborhoods into clusters that share similar comfort parameters.



# Northern Central Mexico City neighborhoods clustered by confort

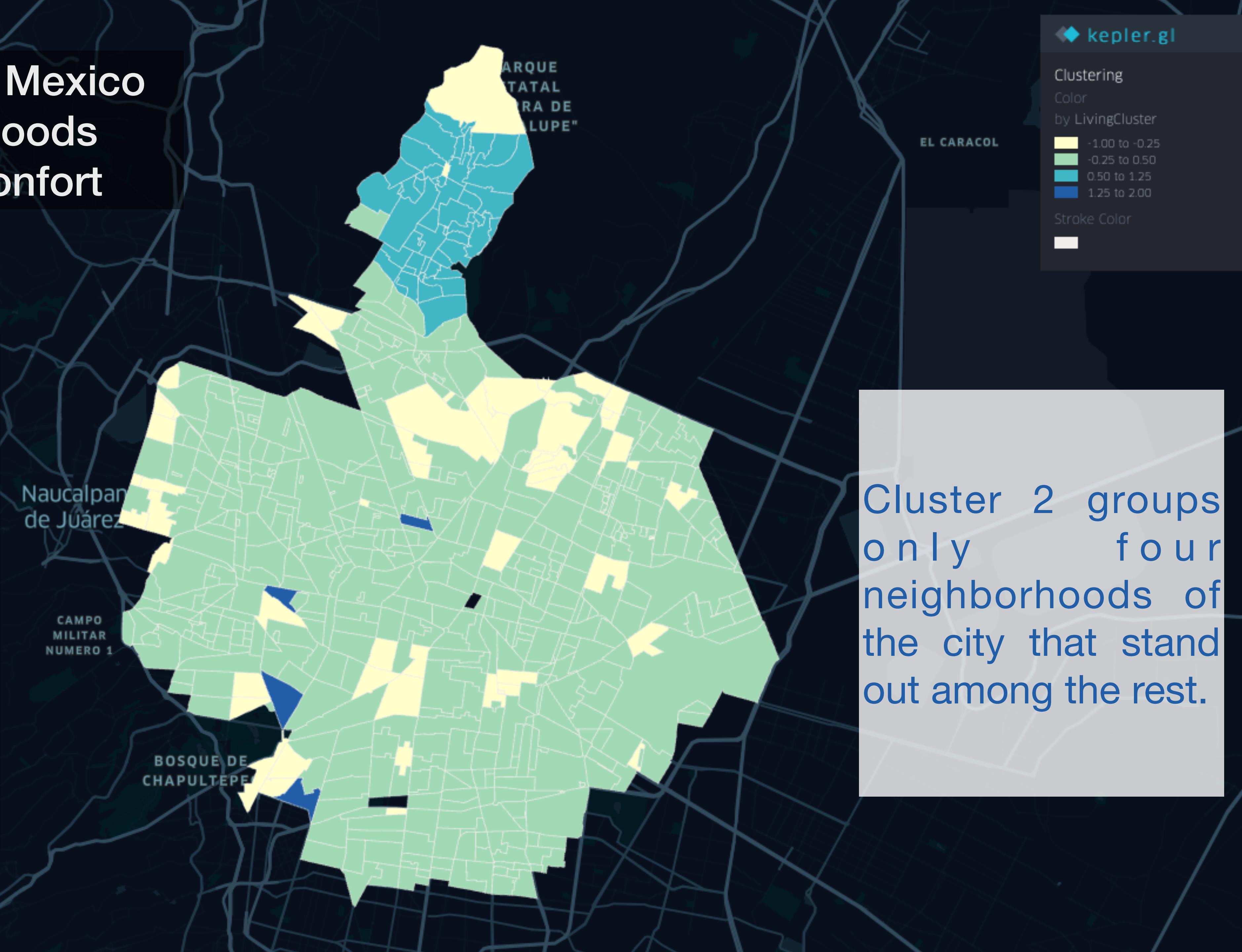
Candidate neighborhoods were grouped into three clusters, labeled with numbers 0, 1 and 2. DBSCAN algorithm is sensitive to noise, thus outlier neighborhoods were labeled with “-1”.



Cluster 0 represents most of the neighborhoods.

# Northern Central Mexico City neighborhoods clustered by confort

Cluster 1 describes a specific part of the city where seismic risk is low compared with the majority and where transport stations, health centers, food shops, and recreation centers are far away.

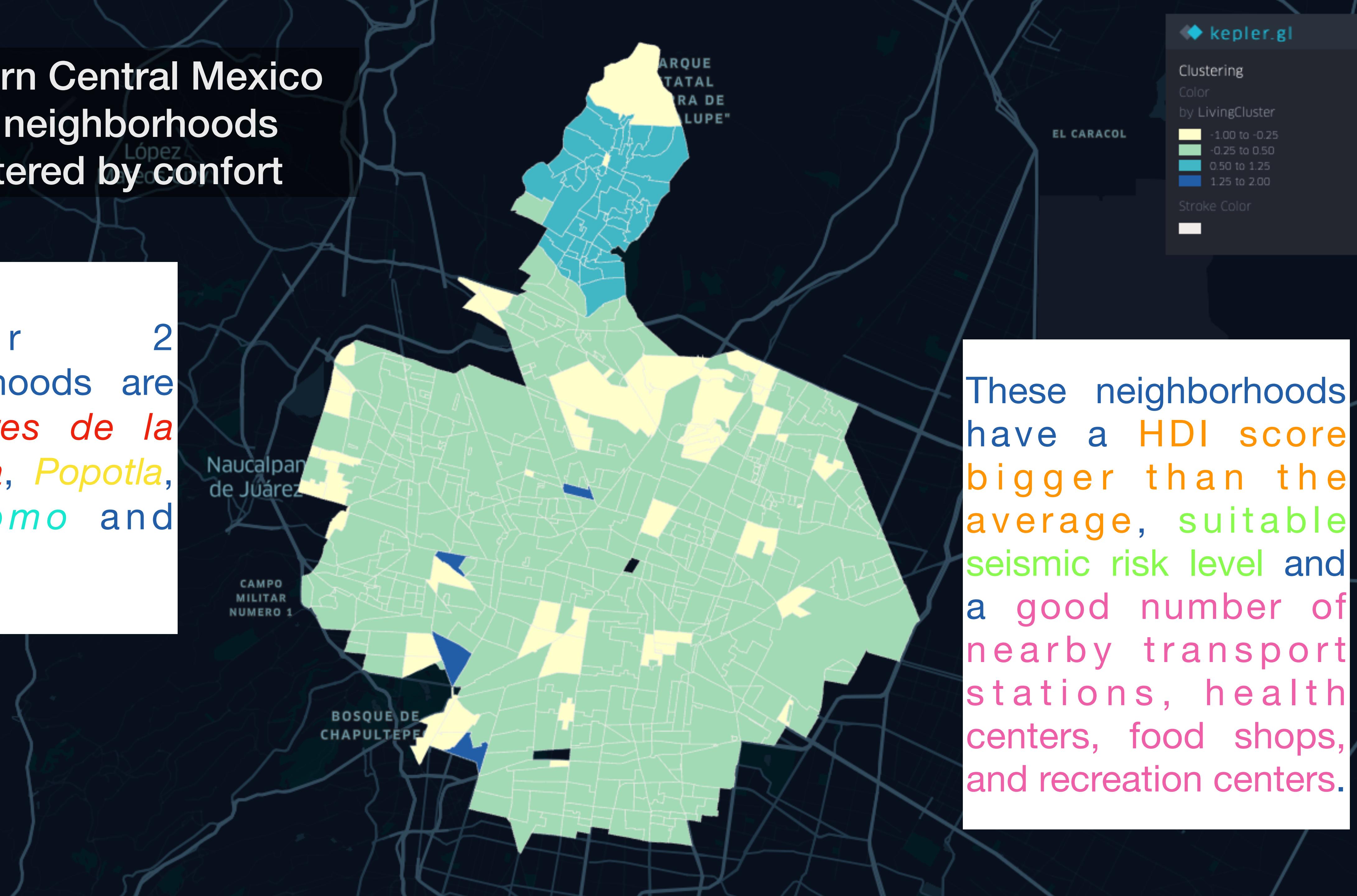


Cluster 2 groups only four neighborhoods of the city that stand out among the rest.

# Northern Central Mexico City neighborhoods clustered by confort

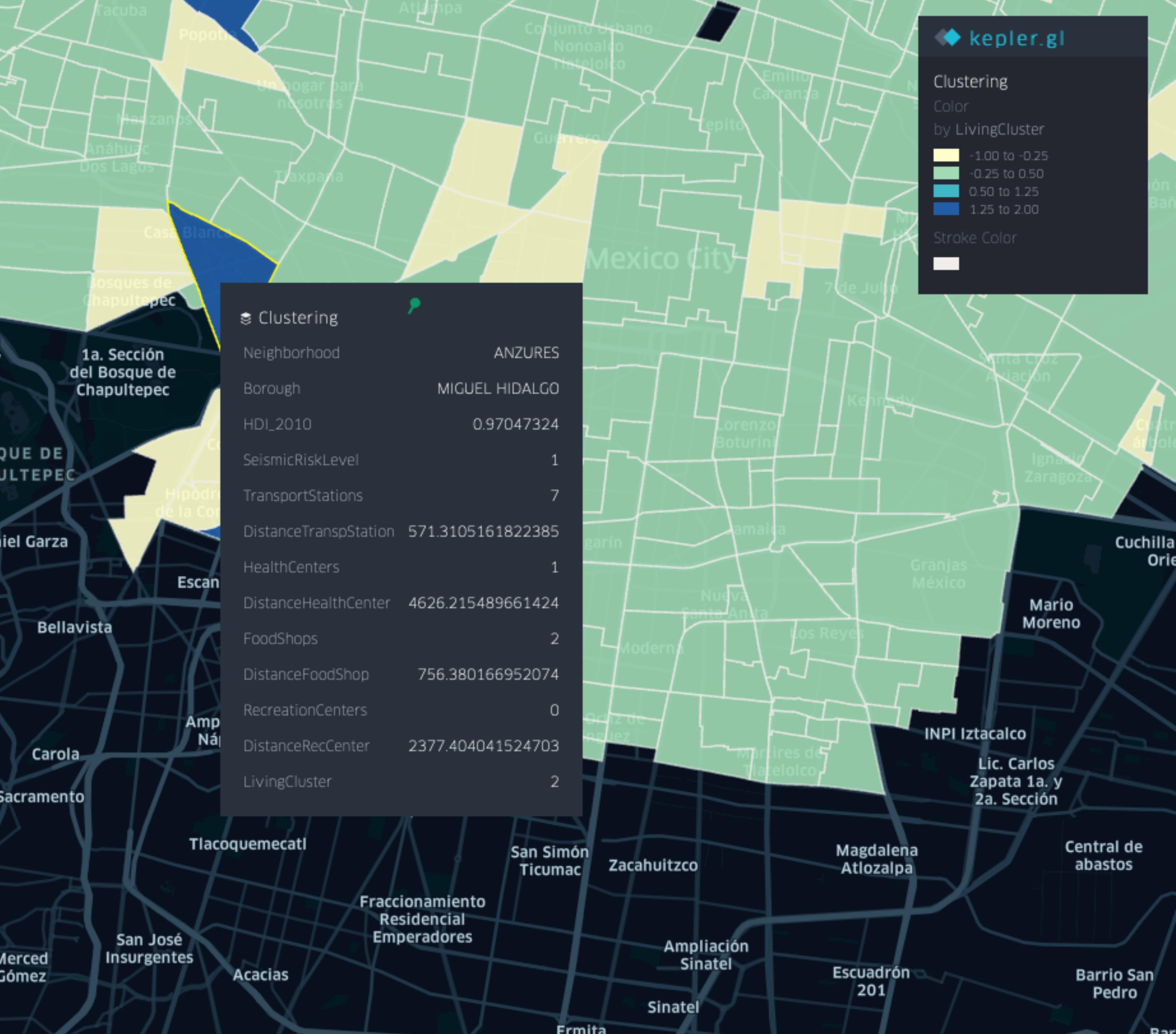
Cluster 2 neighborhoods are *Defensores de la República*, *Popotla*, *Hipódromo* and *Anzures*.

These neighborhoods have a HDI score bigger than the average, suitable seismic risk level and a good number of nearby transport stations, health centers, food shops, and recreation centers.



# Conclusion

The purpose of this project was to identify the optimal places to live in Mexico City according to factors that could influence people to move to a particular area. Clustering of neighborhoods was then performed in order to create a tool to sustain that kind of decisions and insights that were starting points for exploration by stakeholders.



# Future directions

The final decision on optimal house location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood, etc.

