

# IBM Applied Data Science Capstone Project

**Open a new Shopping Mall in Chisinau, Moldova**



## **Introduction**

Chisinau is a fast growing city with high potential. The city is Moldova's main industrial and commercial center, and is located in the middle of the country, on the river Bâc, a tributary of Dniester. According to the results of the 2014 census, the city proper had a population of 532,513, while the population of the Municipality of Chisinau (which includes the city itself and other nearby communities) was 662,836. Chișinău is the most economically prosperous locality in Moldova and its largest transportation hub. Historically, the city was home to fourteen factories in 1919. Chișinău is the financial and business capital of Moldova. Its GDP comprises about 60% of the national economy reached in 2012 the amount of 52 billion lei (US\$4 billion). Thus, the GDP per capita of Chisinau stood at 227% of the Moldova's average. Chisinau has the largest and most developed mass media sector in Moldova, and is home to several related companies ranging from leading television networks and radio stations to major newspapers. All national and international banks (15) have their headquarters located in Chisinau.

## **Business Problem**

In Chisinau there are districts like Codru and Ciocana that have no malls situated nearby. Therefore people prefer to use local markets to go shopping, instead of travelling 5-10 km to the nearest big mall. Moreover, existing malls are not able to satisfy all buyers' needs because they are not big enough and are mostly crowded all the time. Thus, the decision is to analyze the city of Chisinau and to find the best place for a new mall to be built.

# Data

**In order to make an analysis for the specified problem we will need the following data:**

- List of neighborhoods in Chisinau and their latitudes, longitudes and borders accordingly.
- Venues allocated within these neighborhoods obtained using foursquare API. (location, name, type, etc.)
- Chisinau land use (whether a specified area is of commercial use or urban use, etc.)
- Changes in land use 2014-2019 (to understand the city's growth)
- Population density heatmap

## Sources of data and methods to extract them

This Wikipedia page (<https://en.wikipedia.org/wiki/Chișinău>) contains a list of neighbourhoods in Chisinau and general information about the city, with a total of 7 neighbourhoods. Land use, population density heat maps data will be obtained from [https://eo4society.esa.int/ eo\\_clinic/](https://eo4society.esa.int/ eo_clinic/) (maps created by GeoVille and SIRS).

Moreover we will use Foursquare API to get venue data for venues in neighborhoods specified. Data cleaning, data wrangling and data visualization will be done mostly via matplotlib, folium and pandas packages in Jupyter Notebook.

# Methodology

Firstly, we need to get the list of neighbourhoods in the city of Chisinau. Fortunately, the list is available in the Wikipedia page (<https://en.wikipedia.org/wiki/Chi%C8%99in%C4%83u>) or using the data provided by GeoVille (which also includes borders in shapely format).

We will use geopandas and matplotlib to get the data from .shp file. Then we create choropleth map using folium which will show us population distribution over neighborhoods. We also need to get list of malls in these neighborhoods. We will use Foursquare API for this. Send a request to foursquare to explore venues in neighborhood of several "main" points. These "main" points will consist of each neighborhood centroid, most east/west/north/south points. This way we will be able to get more malls' locations. Use KMeans algorithm on data to get some insights. Also use land use and change in land use 2014-2019 data to understand where urban areas are located in the city, therefore we can use appropriate location for a new mall. One of the most important plot will be population density in 2018 in nighttime. This will show us where people are mostly situated in evening-night time (so they are potential clients for a new mall).

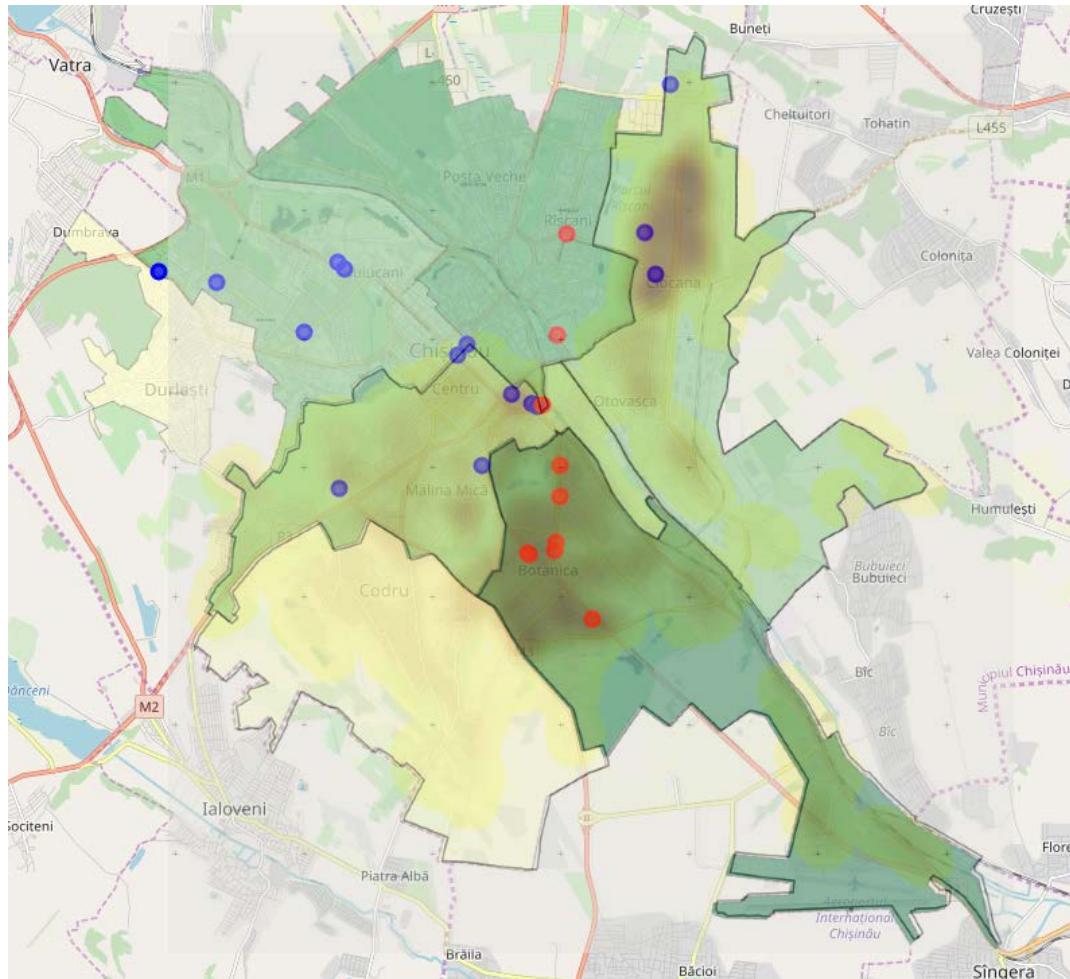
# Results

The results obtained from the KMeans algorithm helps us to divide neighborhood in 3 main clusters:

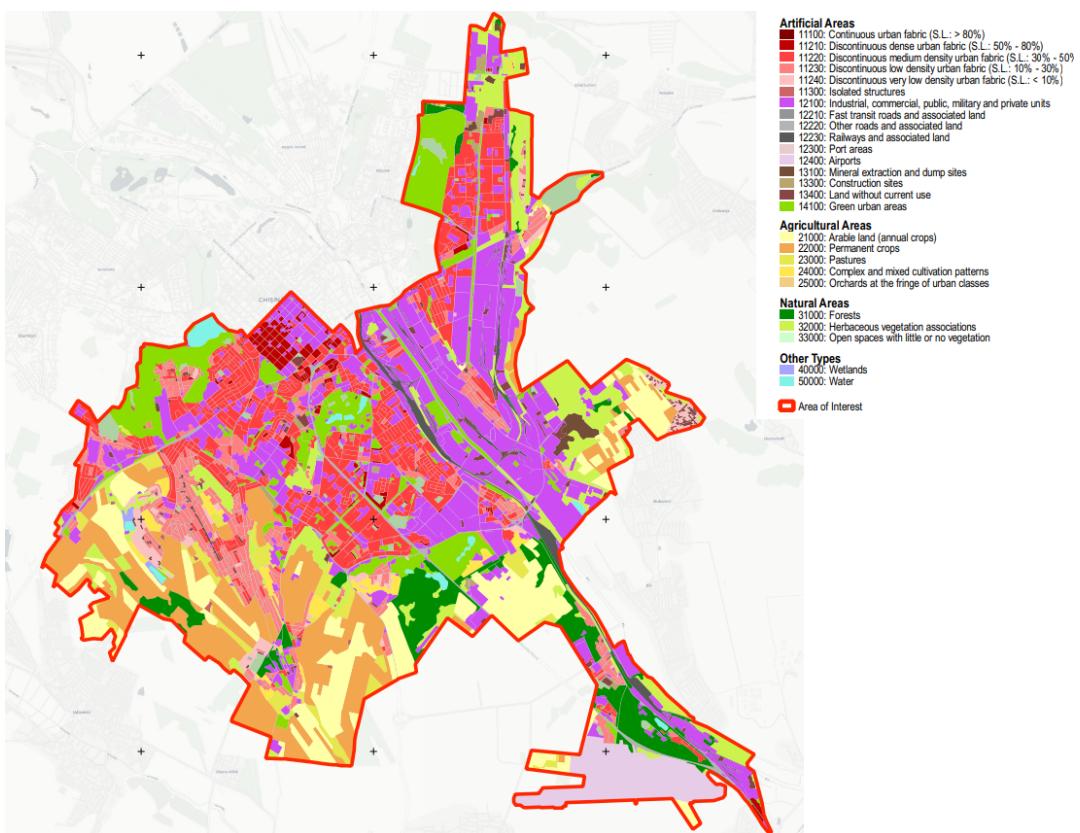
1. Cluster 0 - High population, med-high number of malls
2. Cluster 1 - Low population, no malls
3. Cluster 2 - Med population, med-high number of malls

Let's also apply population density map to our findings.

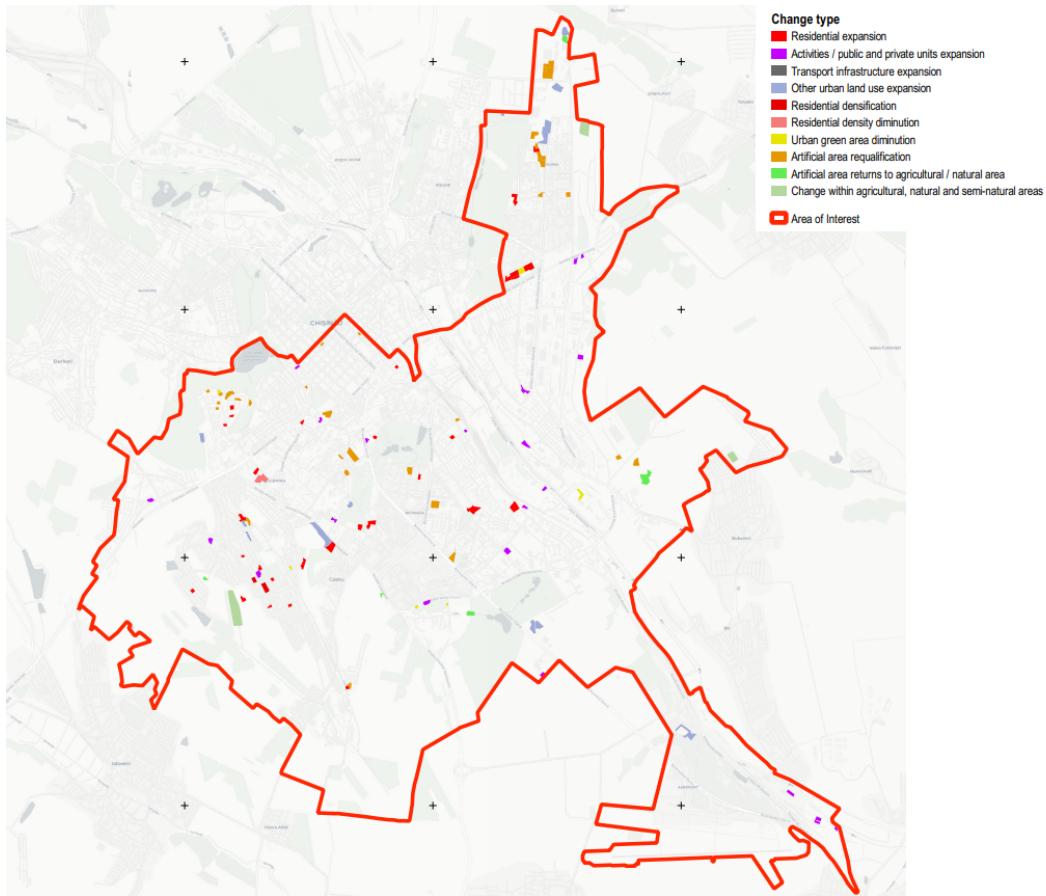
One can clearly see that, although Ciocana (top-left neighborhood) is considered not highly populated, it has very dense location, which is covered by only 3 malls. Also, there is Codru neighborhood which has population ~ 20 000 and has no mall nearby. This will lead people to travel to Botanica in order to go shopping. We can also take a look at land use map of the Chisinau. Note that Codru and Ciocana has very dense urban areas (in red). Moreover we can take a look at changes in land use 2014-2019 to understand how Chisinau is growing. One can notice that there are a lot of new residential expansion occurred in the Codru city. Also some new urban areas were built in Centru and Ciocana locations.



*Malls` location in the city of Chisinau and population density distribution.*



*Land use in Chisinau*



*Changes in land use in Chisinau 2014-2019*

## Discussions

As observations noted from the map in the Results section, most of the shopping malls are concentrated in the central area of Chisinau city, with the highest number in cluster 2 and moderate number in cluster 0. However there are no malls in cluster 0. This represents a great opportunity and high potential areas to open new shopping malls as there is very little to no competition from existing malls. The nearest mall to Codru is within 5-10 km which makes really hard for population to get there. Also, we can see from changes in land use plot that there were a lot of new urban areas built in this neighborhood, which means that city will be expanded in that direction.

## **Conclusion**

In this project, we have analyzed malls distribution in the city of Chisinau. We have analyzed population distribution and malls location to identify the best location for a new mall. In order to do this we have used KMeans algorithm to cluster neighborhoods, also we have analyzed changes in land use 2014-2019 in the city of Chisinau and density population in nighttime which gave us useful insights. During analysis we have concluded that Codru neighborhood may be the perfect choice to build a new shopping mall.