**Cluj-Napoca:**

**Opportunity for Health and Fitness Investments**

October 2019

Teodora Lazar

Data Science Enthusiast

[teolazar@gmail.com](mailto:teolazar@gmail.com)

**INTRODUCTION**

**Background**

Cluj-Napoca, known as Cluj, is a city located in Romania, Eastern Europe that is attrackting more international attention due to its innovation ecosystem and access to employment. It has 393 companies, 8 universities and 22 catalyst organizations, which encourage start-up formation and give plenty of hiring opportunities to the already established businesses. At the same time, it was ranked the second European city in terms of access to employment which attracts youth to establish in this area of the country.

A good indicator of the growth of the city is the real estate market. As study published in 2018, shows that real estate transactions increased by 30% for a total of approximately 600 millions. This sum represents twice the amount of the city budget.

Being on a steady growth, Cluj in adapting to the desires of its population with investments in numerous aspects that increase well being.

**Problem**

The population density of neighbourhoods increases and access to health and fitness is also on the rise. At the moment there is a disproportionate access to gyms depending on the area of the city and with is a business opportunity lies.

**Interest**

The purpose of this paper is to look at the population density of Cluj by neighbourhoods and availability of health and fitness centers. This should help investors make a more informed decision on where to open a new location.

**DATA**

In order to be able to decide which neighbourhoods are the best places for opening a health and fitness location, I will look at the population density, growth of the neighbourhood based on real estate acquisitions and their prices and the availability of gym services in the area.

**Data sources**

The information regarding population density and real estate transactions can be found in publications done by the city hall. Population density per neighbourhoods can be found [here](https://issuu.com/primariamunicipiuluiclujnapoca/docs/brosura_web), while real estate transaction can be found [here](https://www.primariaclujnapoca.ro/userfiles/files/Ghid%20tranzactii%20Cluj-Napoca_2017_final.pdf) and [here](https://files.primariaclujnapoca.ro/2019/07/05/Ghid-tranzactii-Cluj-Napoca_2018_Final_lansare_5iulie2019.pdf). And last but not least the information regarding the venues will be accessed from Foursquare via an API.

**Data formatting and cleaning**

As it can be seen from the above links, the upside is that the data is clear and is not missing information. At the same time, it comes in formats that are not easily transformed into dataframes for analysis. For example, population density per neighbourhood is available in picture format that represents the map of Cluj. Thus, I went through the publications, gathered the data that was relevant to this report and manually aggregated it into a database that can be found [here](http://nopurpose.ca/wp-content/uploads/2019/09/Cluj.csv).

**Feature selection**

After data aggregation, 45 features were available for analysis. Upon analysis, it was clear that redundant or irrelevant features were present. For example, data on how individuals financed their real estate purchases (via mortgage or cash) has no relevance on whether a health and fitness center should be opened in a specific neighbourhood, thus it was excluded.

Also, the number of real estate transactions per year was available as well as percentage change over time. These features convey similar information, thus only the latter was kept.

The dropped features can be found in the Appendix, Table 1.

As a final result, 15 features were identified to be relevant for this report.

Table 2: Kept features

|  |  |  |
| --- | --- | --- |
| Feature Name | Feature Definition | Reason for keeping |
| Neighbourhood | Name of neighbourhood from city | Main feature for the report. |
| Population Density | Population density for each neighbourhood | Feature used in the analysis as predictor of potential market.  Assumption: the more dense the population, the more opportunities for higher demand for health and fitness services. |
| Latitude | Latitute of neighbourhood | Feature needed in venue analysis from Foursquare |
| Longitude | Longitute of neighbourhood | Feature needed in venue analysis from Foursquare |
| Growth\_Land | % change in the number of land properties acquired | Feature represents investment in a particular neighbourhood. Relationship with population density and potential investment to be investigated further. |
| Growth\_House | % change in the number of house properties acquired | Feature represents investment in a particular neighbourhood. Relationship with population density and potential investment to be investigated further. |
| Growth\_Condo | % change in the number of condo properties acquired | Feature represents investment in a particular neighbourhood. Relationship with population density and potential investment to be investigated further. |
| Growth\_Other | % change in the number of other properties acquired | Feature represents investment in a particular neighbourhood. Relationship with population density and potential investment to be investigated further. |
| Growth\_Total | % change in the number of total properties acquired | Feature represents investment in a particular neighbourhood. Relationship with population density and potential investment to be investigated further. |
| Growth\_Price\_Land | % change in the price of land properties acquired | Feature represents investment in a particular neighbourhood. Relationship with population density and potential investment to be investigated further. |
| Growth\_Price\_House | % change in the price of house properties acquired | Feature represents investment in a particular neighbourhood. Relationship with population density and potential investment to be investigated further. |
| Growth\_Price\_Condo | % change in the price of condo properties acquired | Feature represents investment in a particular neighbourhood. Relationship with population density and potential investment to be investigated further. |
| Growth\_Price\_Other | % change in the price of other properties acquired | Feature represents investment in a particular neighbourhood. Relationship with population density and potential investment to be investigated further. |
| Growth\_Price\_Total | % change in the price of total properties acquired | Feature represents investment in a particular neighbourhood. Relationship with population density and potential investment to be investigated further. |
| Growth\_Average\_Price | % change in the average price of total properties acquired | Feature represents investment in a particular neighbourhood. Relationship with population density and potential investment to be investigated further. |

**APPENDIX**

Table1 : Dropped features and the reason for it

|  |  |  |
| --- | --- | --- |
| **Feature Name** | **Feature Definition** | **Reason** |
| Type\_Land\_2017 | Number of land properties purchased in 2017 | Redundant |
| Type\_House\_2017 | Number of house properties purchased in 2017 | Redundant |
| Type\_Condo\_2017 | Number of condo properties purchased in 2017 | Redundant |
| Type\_Others\_2017 | Number of other properties purchased in 2017 | Redundant |
| Type\_Total\_2017 | Number of total properties purchased in 2017 | Redundant |
| Type\_Fin\_Alone\_2017 | Number of properties purchased by cash in 2017 | Irrelevant |
| Type\_Fin\_Bank\_2017 | Number of properties purchased by mortgage in 2017 | Irrelevant |
| Price\_Land\_2017 | EURs spent on land transactions in 2017 | Redundant |
| Price\_House\_2017 | EURs spent on house transactions in 2017 | Redundant |
| Price\_Condo\_2017 | EURs spent on condo transactions in 2017 | Redundant |
| Price\_Others\_2017 | EURs spent on other real estate in 2017 | Redundant |
| Price\_Fin\_Alone\_2017 | Total cash real estate transactions in 2017 | Irrelevant |
| Price\_Fin\_Bank\_2017 | Total mortgage real estate transactions in 2017 | Irrelevant |
| Price\_Fin\_Total\_2017 | EURs spent on real estate properties in 2017 | Redundant |
| Average\_Price\_2017 | Average price of real estate properties in 2018 | Redundant |
| Type\_Land\_2018 | Number of land properties purchased in 2018 | Redundant |
| Type\_House\_2018 | Number of house properties purchased in 2018 | Redundant |
| Type\_Condo\_2018 | Number of condo properties purchased in 2018 | Redundant |
| Type\_Others\_2018 | Number of other properties purchased in 2018 | Redundant |
| Type\_Total\_2018 | Number of total properties purchased in 2018 | Redundant |
| Type\_Fin\_Alone\_2018 | Number of properties purchased by cash in 2018 | Irrelevant |
| Type\_Fin\_Bank\_2018 | Number of properties purchased by mortgage in 2018 | Irrelevant |
| Price\_Land\_2018 | EURs spent on land transactions in 2018 | Redundant |
| Price\_House\_2018 | EURs spent on house transactions in 2018 | Redundant |
| Price\_Condo\_2018 | EURs spent on condo transactions in 2018 | Redundant |
| Price\_Others\_2018 | EURs spent on other real estate transactions in 2018 | Redundant |
| Price\_Fin\_Alone\_2018 | Total cash real estate transactions in 2018 | Irrelevant |
| Price\_Fin\_Bank\_2018 | Total mortgage real estate transactions in 2018 | Irrelevant |
| Price\_Fin\_Total\_2018 | EURs spent on real estate properties in 2018 | Redundant |
| Average\_Price\_2018 | Average price of real estate properties in 2018 | Redundant |