Capstone Project – The Battle of Neighborhoods Area price in Vilnius (Lithuania)

1. Introduction

1.1 Introduction/Business Problem

More and more people want to live in a house and have space and private backyard. This desire grow even further this year then a half of people should work from home. There are many problems with building a house. The bigest problem is price of area. You want an area which is cose enough to the city, but you dont want to pay huge price for that area. The next big question is area size you dont want to small and you dont want to big because price is going to increse.

In this project we will help people who are looking for to buy area for building a house. Our analysis should show how much should you invest in area if you want to live 5km, 10km or 15km from city. How big area for which price client can buy. Should the client to look at areas at south or west from the city center.

1.2 Data

Data about area prcie in Vilnius city was taken from local website www.aroudas.lt. Filter used to filter area just in city, ant the higer price limit was 500000 euros. Also we filter area for house building purposes.

1.3 Methodology

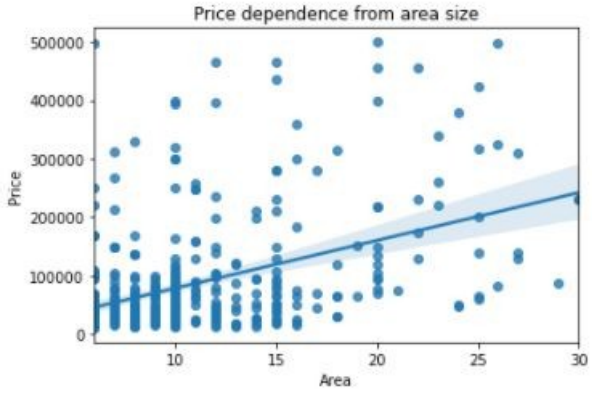
To get data from website we biuld a web scraper. We collected date about: price, address, area. Address contains district and street so we seperate those. After seperating we write a code to filter villages close to city, because they are not in the city, bet often sellers write so to get more popularity.

Geolocator was used to find cordinates of the area and calculate distance to the city center.

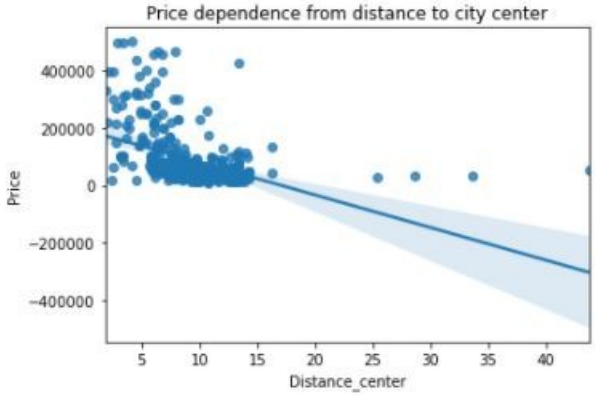
Group by was used to determine all the districts and analyze. Which distrcit is the closest to city center, which has the most selling areas and the price range.

1.4 Results and discusion

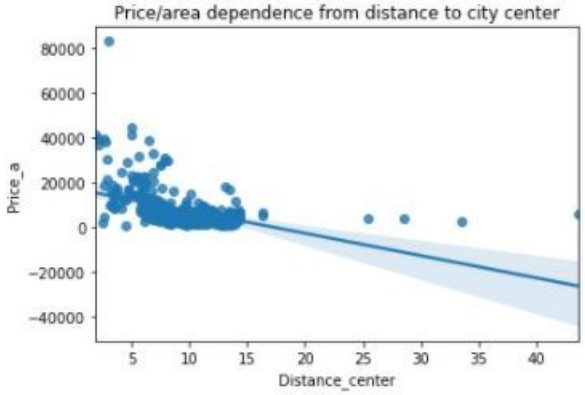
If we look at price dependence from area size we can see that there are no clear relationship, points quite scattered. From linear approximation we can conclued that prices are rising with biger area, but the scattering shows that there are other factors.



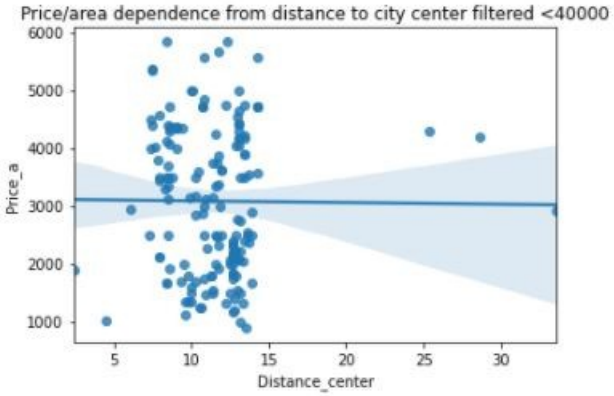
Price dependece from distance to city center could be one of factors. Graph below show quite nonlinear relationship beatween price and distance to the city center. From 6 to 0 km price rises quit quiqly and for 6 to 15 km we can see linear relationship.



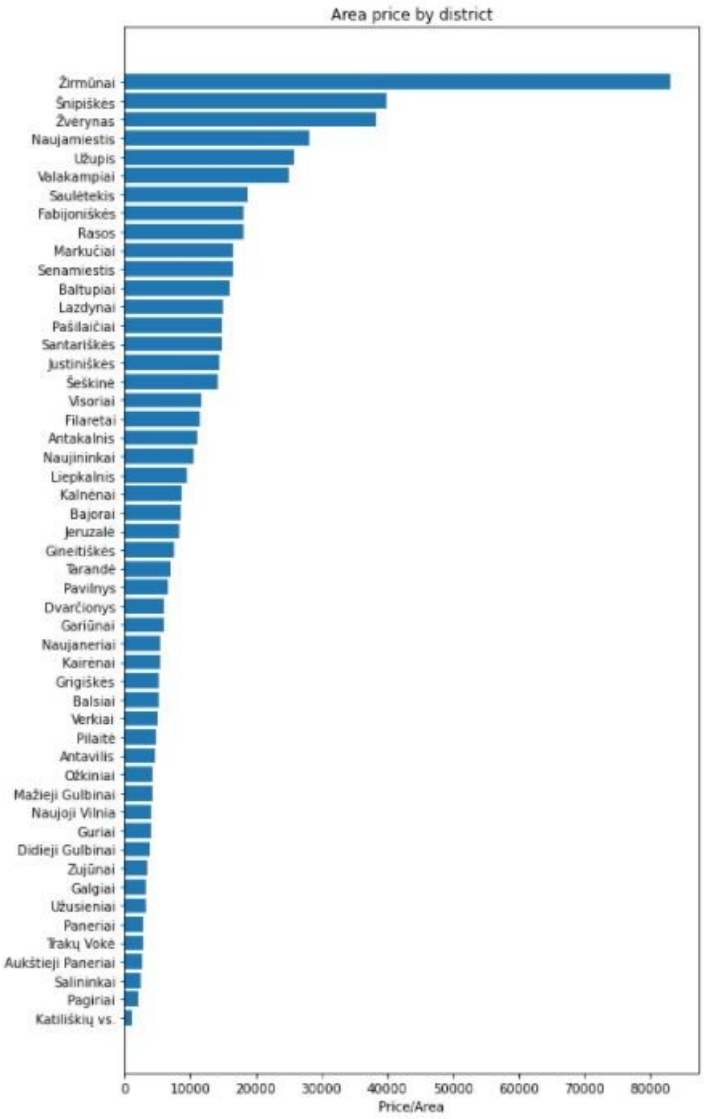
If we dived price by area size we can further analize price dependece from distance to the city center. Dependece is quite similar to the above graph but data is les scatered. In a range of 0 to 6km price going up nonlinear and below taht range price normalizes.

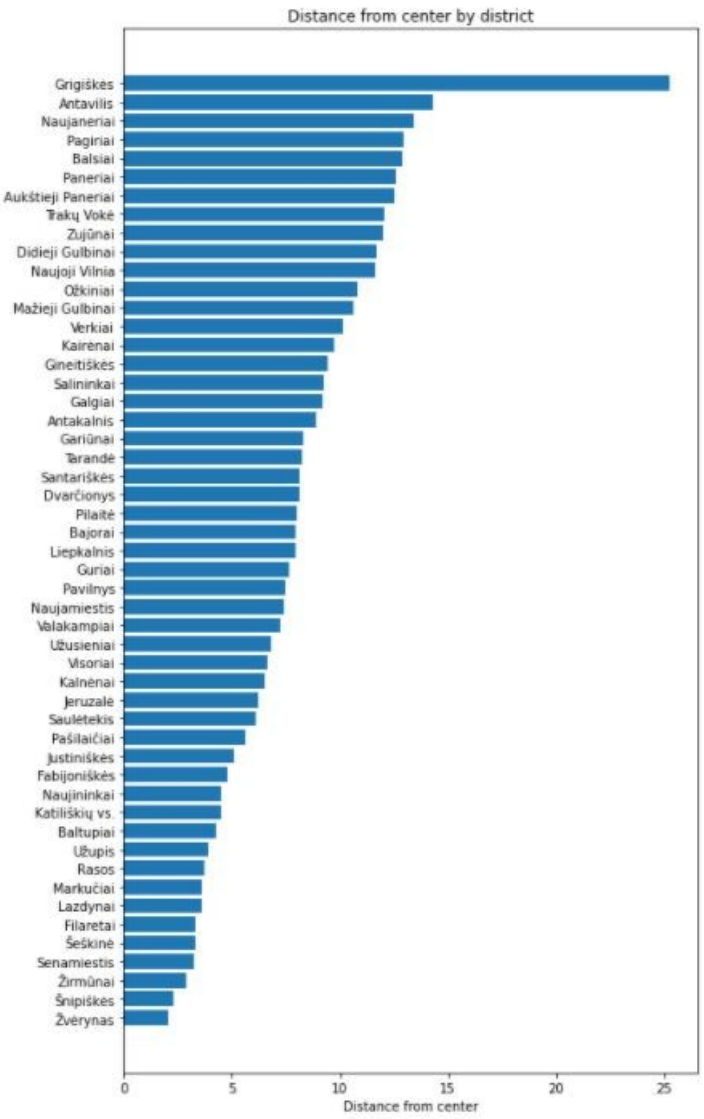


If we filtered price below 40000 for an area, there is almost no areas in a range from 0 to 6km, add then almost all areas are in a range between 6 and 15 km. In this rrange there are no linear relationshipt between price normalized for one acre adn distance to city center. Porbably other factors play a roll such as electrcicity lines, gas lines, city sewer and water system, availability of public transport.

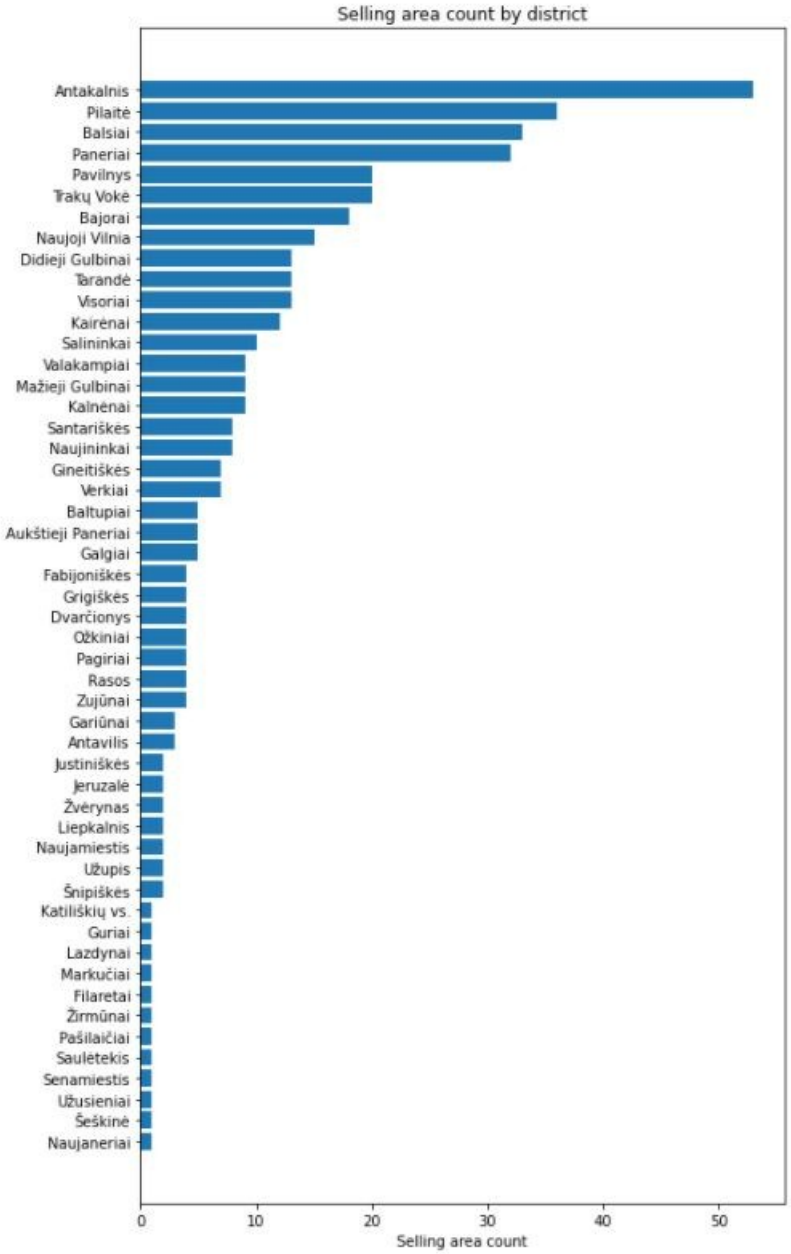


Top three areas by price are also top three areas closes to the city center. If client wants average price areas to buidl a house for living, we should look at distrcits, in range 7,5 to 15km form city center.

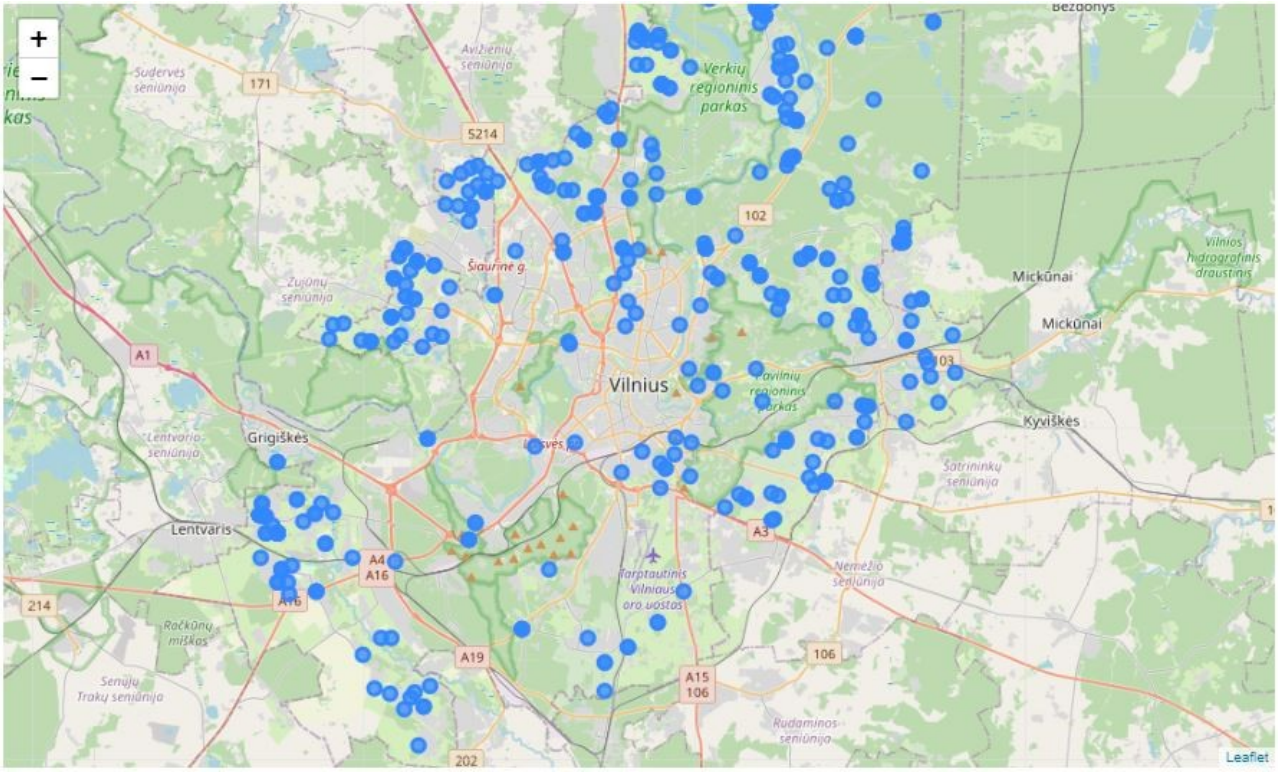




In the grap below we can analyze how many areas are for slae in each district. The client should look at the top 10 districts.



All the areas that are for sale in the map.



1.7 Conculsions

- We can determine area price just by area size or distance to city center there are other parameters.
- Prices of areas that are very close to the city in a range between 0 and 6 km rises quit quikly (may be even exponationaly).
- Top 3 districts in price are also the closest to the city center.
- Pirces in rage 6 to 15 km are quite stady.