Report
"Where I can buy medicaments?"

What we can do when the nearest pharmacy is miles away? Sometimes it is vital to buy pills or drugs as soon as possible for anyone of us. For example, in case of bronchitis or diabetes, it is a question of life or death.

The entrepreneurs are also interested in the answer to the following question: "Where a new pharmacy may be opened?" The answer is a part of a pharmacy`s business plan.

Therefore, the following questions are of high interest:
- Is it enough pharmacies in a district, borough or a town,
- Where a new pharmacy should be opened.

We can define the following groups, which are interested in results of this project:
- Inhabitants,
- Local authorities,
- pharmaceutical companies

In this project I will discover Canadian town Edmonton, Alberta. The town is characterized by large area and population and, subsequently low density of the population.

To conduct the research I need the following information:
- neighborhood: population, area, coordinates;
- coordinates of pharmacies in the town.

Fortunately, the city create Web source "Open data portal". The source contains all the census data of Edmonton I need for the research. The data will be imported from official source https://data.edmonton.ca/.

Pharmacies data will be imported via API from the Web source https://developer.foursquare.com/. The source gives the full details about a venue including location, tips, and categories.

First of all import the neighborhood data.

	Ward	Neighbourhood Number	Neighbourhood Name	0 - 4	5 - 9	10 - 14	15 - 19	20 - 24	25 - 29	30 - 34	35 - 39	40 - 44	45 - 49	50 - 54	55 - 59	60 - 64	65 - 69	70 - 74	75 - 79	80 - 84	85+	No Response
0	WARD 1	3140	CRESTWOOD	68	83	97	99	76	62	59	74	77	110	116	148	98	100	91	80	64	49	800
1	WARD 1	3330	PARKVIEW	126	162	138	165	155	96	128	149	164	164	218	247	197	146	89	90	82	64	705
2	WARD 11	6110	CPR IRVINE	0	1	0	0	18	44	39	4	1	1	6	1	0	1	0	0	2	0	48
3	WARD 9	5350	RHATIGAN RIDGE	73	130	117	112	112	84	56	83	107	103	142	185	205	171	102	74	42	39	1140
4	WARD 5	4140	ELMWOOD	131	121	130	169	180	185	173	135	154	154	161	133	145	140	92	117	60	51	290

Figure 1. The population and area of the neighborhoods

The imported data (fig.1) includes the numbers and the names of Neighborhoods, as well as quantity of population divided by age groups. It is interesting question: is there some correlation between percentage of age groups of local population and income of a local pharmacy, or by another words, who is more often buy medicaments – youth, middle age with children or the elderly? The answer at that question can founded using internal information of a pharmacy retail chain. This question can be decided for a particular drug store as part of its business plan. This is beyond the project.

As we can see from the table there are no any information about the geographical position of the Neighborhoods. This information can be found at the same source from the other file. Import the file to the notebook (fig.2).

|	Neighbourhood Number	Neighbourhood Name	Area Sq Km	Latitude	Longitude	Location	Neighbourhood Boundaries : 2019	Roadway Maintenance Area Polygon	Edmonton Public School Board (EPSB) Ward Boundaries (effective at 12:00 AM on Oct 16, 2017)	Edmonton Catholic School District Ward Boundaries (effective at 12:00 AM on Oct 16, 2017)	City of Edmonton - Ward Boundaries (effective at 12:00 AM on Oct 16, 2017)
0	6790	Wild Rose	1.772290	53.470564	-113.381167	(53.47056409015751, -113.3811672388747)	74	135.0	7.0	7.0	4.0
1	1200	River Valley Victoria	1.298977	53.534988	-113.523686	(53.53498830876779, -113.52368603311965)	344	144.0	4.0	5.0	9.0
2	4330	Mitchell Industrial	0.592507	53.574016	-113.584201	(53.574015777213816, -113.58420113302314)	112	206.0	3.0	1.0	5.0
3	5350	Rhatigan Ridge	1.344078	53.474506	-113.587569	(53.474506145469945, -113.5875691214093)	71	64.0	6.0	6.0	12.0
4	1280	Yellowhead Corridor West	1.164872	53.582841	-113.512729	(53.58284074814297, -113.51272869912015)	257	223.0	3.0	1.0	5.0

Figure 2. The coordinates of the neighborhoods

For processing the data it is necessary to merge this two Dataframes into one and clean from the excessive columns, which I do not need in this research (Neighborhood Boundaries, Roadway Maintenance Area Polygon, etc.). As the result I received the table that contains all the data about Neighborhoods in Edmonton. It includes Neighborhood`s names, inhabitants, coordinates, area. Some cells do not contain coordinates, this cells doe not contain information about population as well. When I saturated myself in a subject, I founded up that such neighborhoods are industrial, subsequently they are useless for the project. I cut them out of the table. After that I calculated sum of inhabitants of the Neighborhoods in Edmonton and place it in new column "Inhabitants" (fig.3).

	Neighbourhood Number	Neighbourhood Name	Area Sq Km	Latitude	Longitude	Inhabitants
0	3140	Crestwood	1.168158	53.535434	-113.569038	2351
1	3330	Parkview	1.546448	53.524060	-113.567914	3285
2	6110	CPR Irvine	0.663610	53.507527	-113.490549	166
3	5350	Rhatigan Ridge	1.344078	53.474506	-113.587569	3077
4	4140	Elmwood	1.025925	53.515738	-113.605993	2721

Figure 3. The neighborhoods data

As the result, I received the DataFrame that contains all necessary information for estimation of distribution of inhabitants over the Edmonton area. The DataFrame has been visualized on the map of the town (fig.4).
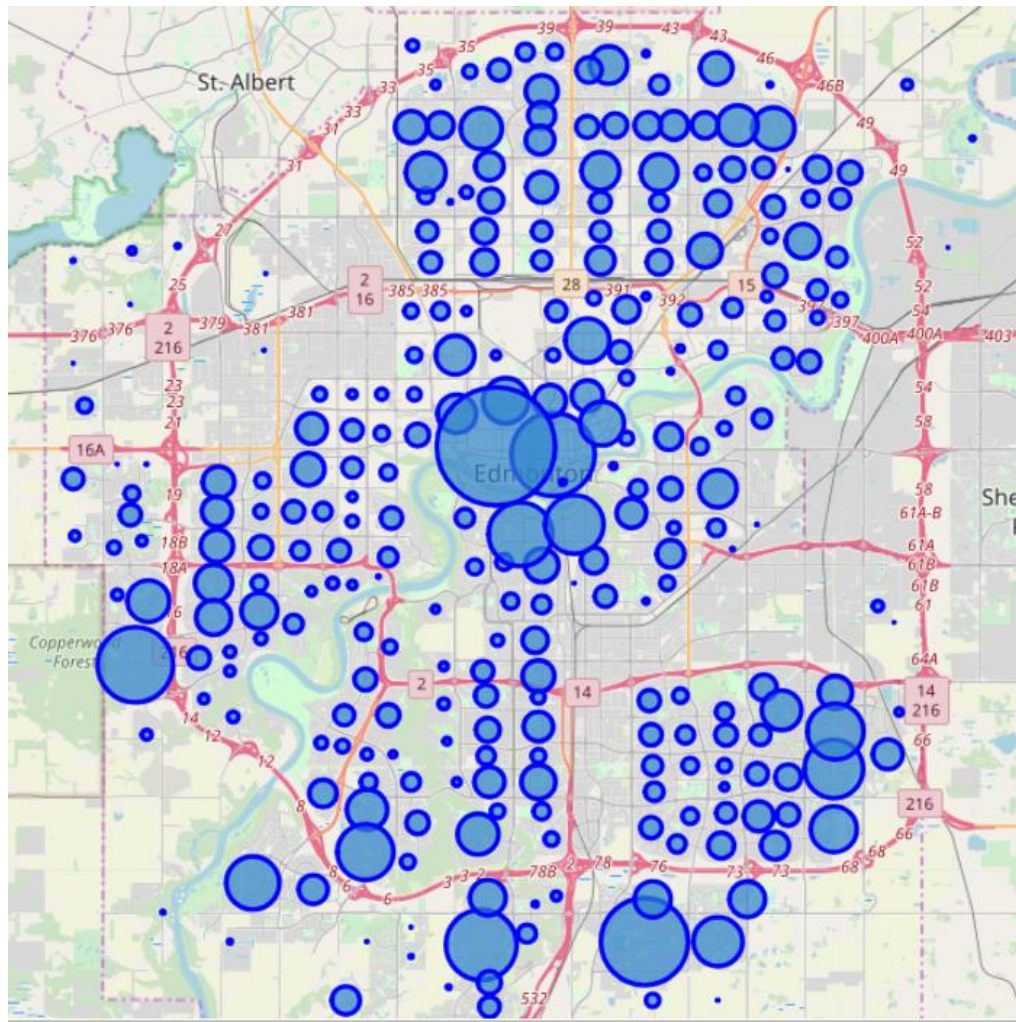
Figure 4. Edmonton neighborhoods

At the map we can see blue bubbles of different size. The bubbles size similar to quantity of habitants in neighborhoods. As we can see, there is extremely high population density in the center of the city. The south, the south-west and the west are also have high level of population density. The north and the north-east inhabited evenly. The east and the north-west are almost deserted.

The next step of the discovery is acquisition and discovering of data about quantity and distribution of pharmacies in the city. For that purpose have been chosen Web-service https://developer.foursquare.com. The service gives to a developer wide range of possibilities to work with API data as well as to receive a venue information of a chosen category. In this research the following categories is used: Pharmacy, Apothecary, Drugstore as well as city molls (all the city molls contains a shop that sales drugs). A request has a limitation of maximum number of results. To overcome constraints it is possible to divide the town into few areas, and make request for every of them. Therefore a table of pharmacies have been received (fig.5)

| ce	location.formattedAddress	location.labeledLatLngs	location.lat	location.lng	location.neighborhood	location.postalCode	location.state	name	referralId
22	[10818 Jasper Ave NW, Edmonton AB T5J 2B3, Can...	[{'label': 'display', 'lat': 53.54098281712579...	53.540983	-113.507116	NaN	T5J 2B3	AB	Rexall	v-1556520357
05	[11835 26 Ave SW (Opening soon), Edmonton AB T...	[{'label': 'display', 'lat': 53.40813625251965...	53.408136	-113.536784	NaN	T6W 0C8	AB	Real Canadian Superstore	v-1556520357
32	[1494, 8882 170 St. NW (West Edmonton Mall), E...	[{'label': 'display', 'lat': 53.52350461852903...	53.523505	-113.626965	West Edmonton	T5T 4M2	AB	London Drugs	v-1556520357
14	[8872 170 St. (West Edmonton Mall), Edmonton A...	[{'label': 'display', 'lat': 53.52232638052472...	53.522326	-113.626023	NaN	T5T 4M2	AB	World Waterpark	v-1556520357
16	[10405 Jasper Ave NW (Jasper Ave), Edmonton AB...	[{'label': 'display', 'lat': 53.54072099103654...	53.540721	-113.499483	NaN	T5J 1G5	AB	Shoppers Drug Mart	v-1556520357

Figure 5. Table of pharmacies

The table contains columns with excessive information for the research. This columns have been deleted (fig.6).

	location.lat	location.lng	name
0	53.540983	-113.507116	Rexall
1	53.523505	-113.626965	London Drugs
2	53.408136	-113.536784	Real Canadian Superstore
3	53.523619	-113.622429	West Edmonton Mall
4	53.522326	-113.626023	World Waterpark

Figure 6. Cleaned table of pharmacies

The pharmacies are visualized on the map of the city, green bubbles at the figure 7.

Figure 7. Pharmacies
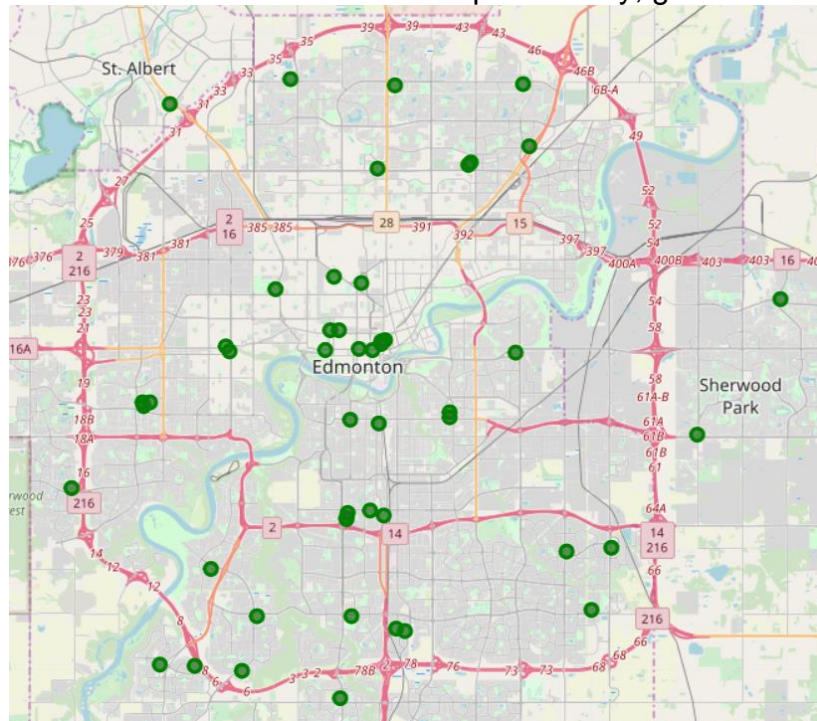
In the next step of the research have been founded the minimum distance between every neighborhood and a nearest pharmacy. The method *geopy.distance* is used for this purpose. As the result the DataFrame that contains distance from every neighborhood to every pharmacy was received. At the next step the minimum distance to the nearest pharmacy was founded and converted to meters. Finally, neighborhoods with distance to a nearest pharmacy less than 2500 meters (half an hour walk distance) were removed from the table. The remaining 74 neighborhoods have distance to a nearest pharmacy ranged from 2 536 to 8 858 meters (fig.7).

	Neighbourhood	distance	Area Sq Km	Latitude	Longitude	Inhabitants
0	Keswick Area	2536	3.870061	53.417600	-113.635797	589
1	Fraser	2557	1.308507	53.610382	-113.374782	3489
2	Homesteader	2563	1.577836	53.584720	-113.406459	3545
3	Potter Greens	2586	1.182586	53.518290	-113.672602	1490
4	Maple Ridge Industrial	2594	3.745517	53.498009	-113.356365	77

Figure 7. Table of neighborhoods remote from pharmacies

The remote neighborhoods are visualized on the map of the city (fig.8). The neighborhoods distributed in the north, the south-east and the west of Edmonton mostly.
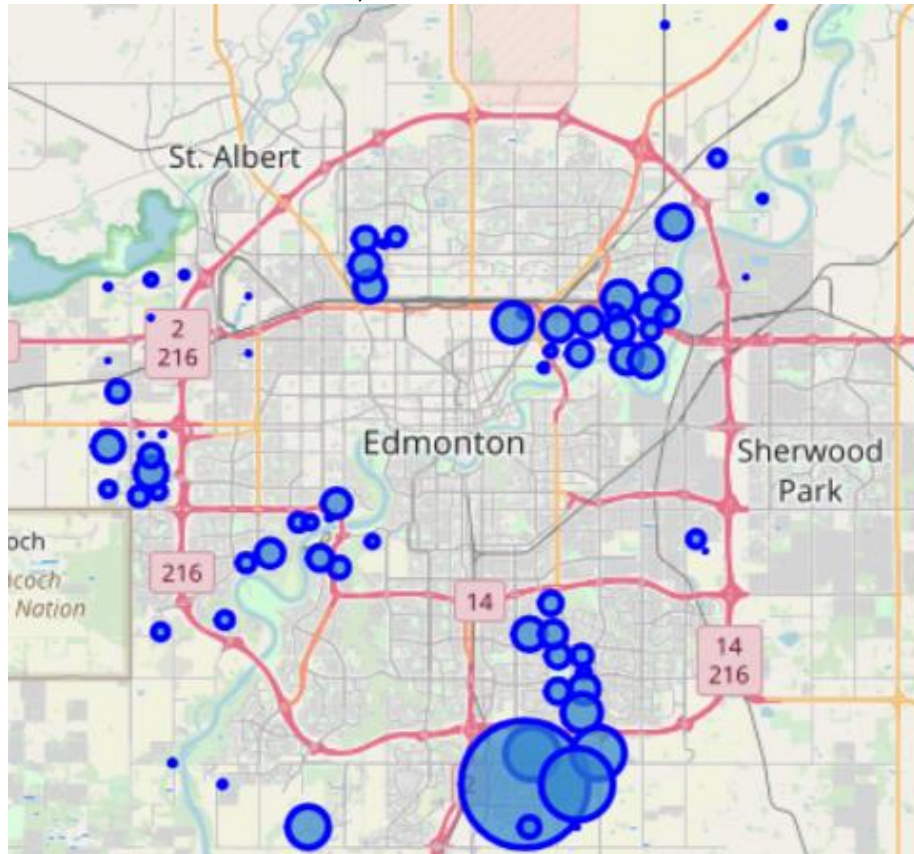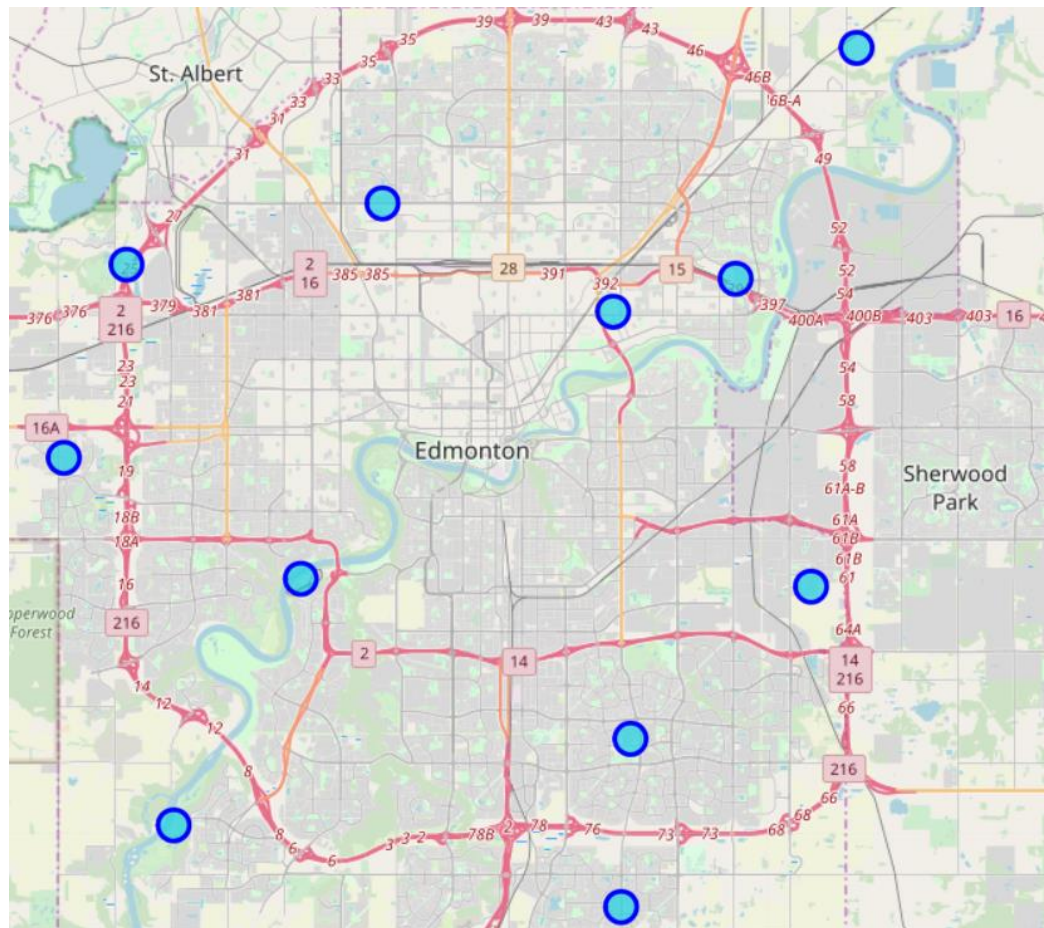

Figure 8. Neighborhoods remote from pharmacies

The next question is the determination of quantity of needed pharmacies. It is obvious, that the quantity most has a proportion similar to existing quantity pharmacies serving nearest neighborhoods. This number of needed pharmacies is calculated and is 10.875. As we can see, it should be opened 11 (10.875 rounded to nearest) new pharmacies.

The next problem is to find coordinates of the new pharmacies. The good idea is to use K-Means Clustering for that purpose. The initialization method of the centroids k-means++ was used. The number of time the k-means algorithm run with different centroid seeds was 10, that is enough for this project.

Finally, the table with coordinates of new pharmacies was received. The coordinates are the approximate positions where new pharmacies would be accessible for the population of remote neighborhoods. The new pharmacies are visualized on the map of the city (fig.9).

Figure 9. New pharmacies

Note: if founded coordinates are match with an existing building, another closest place should be reviewed.

To sum up the research it is important to mention the followings:
- The aim of the research is archived;
- The free census data can be found, downloaded and used in a project;
- The service https://developer.foursquare.com has limitations and venue information has to be examined;
- The Jupiter Notebook gives to a user wide verity of tools for wide range of tasks.

Reference index

- International Pharmaceutical Federation – FIP (2017). Pharmacy at a glance – 2015-2017. The Hague, The Netherlands: International Pharmaceutical Federation
- https://data.edmonton.ca/
- https://developer.foursquare.com/
- https://github.com/
- https://pandas.pydata.org/
- https://www.ibm.com/