

Рекомендация по открытию ресторана по населению районов города Алматы

Джантыков Айдар

20.05.2019

1. Вступление: Бизнес проблема

Для открытия успешного ресторанного бизнеса, нам требуется узнать где этот бизнес будет пользоваться спросом. В этом проекте мы постараемся найти оптимальное место для ресторана. В частности, этот отчет будет адресован заинтересованным сторонам, заинтересованным в открытии ресторана в городе Алматы, Казахстан. В данном проекте мы постараемся выяснить оптимальное местонахождение ресторана по соотношению плотности населения и количеству ресторанов по близости.

2. Сбор и очистка данных

Нам потребуются данные о количестве проживающих в разных районах города. Мы будем производить сбор всех необходимых данных с открытых источников, таких как [1],[2] и др. Также нам потребуется использование Foursquare API, для определения ресторанов по близости и их классификации [5].

Для того, чтобы мы могли использовать данные предоставляемые [1], [2], [5], нам нужно произвести очистку данных. На сайте [1] были предоставлены данные о населении всего Казахстана, включая районы. Нам нужно убрать лишние значения не интересующих нас городов и районов. В итоге у нас осталось только население всего города Алматы и его районы. С ним мы и будем дальше работать.

	Neighborhood	Count	lat	long
0	Almaly	215768	43.255220	76.927133
1	Alatau	260441	43.298628	76.831061
2	Auezov	295543	43.223376	76.847334
3	Bostandyq	343541	43.169788	76.929244
4	Zhetysu	166001	43.323875	76.929575
5	Medeu	209836	43.144485	77.028917
6	Naurzbyay	128169	43.189284	76.802733
7	Turksib	235357	43.347832	77.012400

Для визуализации данных нам также потребуется обработать данные для правильной визуализации. Убрать ненужные для нас столбцы и обозначить названия районов как заголовки строк.

Count	
Neighborhood	
Almaly	215768
Alatau	260441
Auezov	295543
Bostandyq	343541
Zhetysu	166001
Medeu	209836
Nauryzbay	128169
Turksib	235357

3. Методология

В этом проекте мы направим наши усилия на выявление районов Алматы с высокой плотностью населения и определение количества ресторанов по близости.

На первом этапе мы собрали все необходимые данные по населению районов города Алматы на 1 апреля 2019 года (согласно Комитету по статистике Министерства национальной экономики Республики Казахстан [1]).

Вторым этапом была обработка данных. Исключение неиспользуемых значений и городов РК, построение новой таблицы данных для последующей работы с ней.

Третьим этапом нам требовалось визуализировать все имеющиеся данные, для построения более точной рекомендации.

Четвертым этапом потребовалось найти GEOjson файл, в котором состояли координаты всех районов города Алматы, которые требовались для классификации по плотности населения[3].

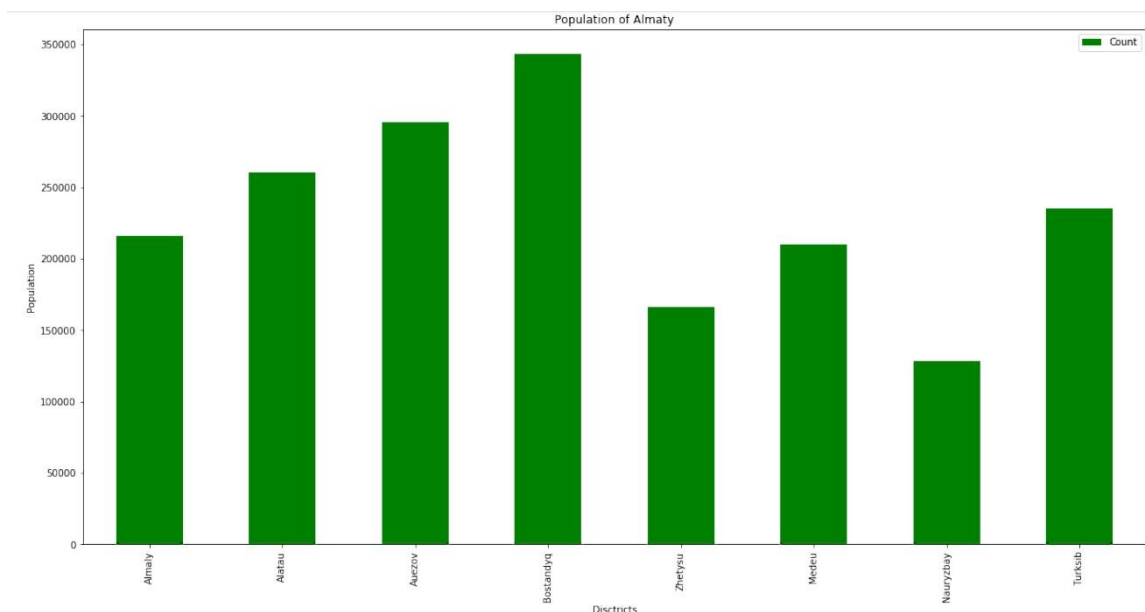
Пятым этапом было построение карты плотности населения и последующий поиск ресторанов в городе Алматы.

Шестой и заключительный этап состоял из вывода результатов и построения рекомендации.

4. Анализ и визуализация

Приступаем к анализу собранных данных и их визуализации.

Данная диаграмма показывает нам базовое соотношение населения по районам города Алматы:



Здесь видно, что Бостандыкский и Ауэзовский районы занимают лидирующие позиции по количеству населения.

Для того, чтобы мы могли получше разглядеть разницу населения районов, нам потребуется построить еще несколько диаграмм.

Диаграмма «пирог»:

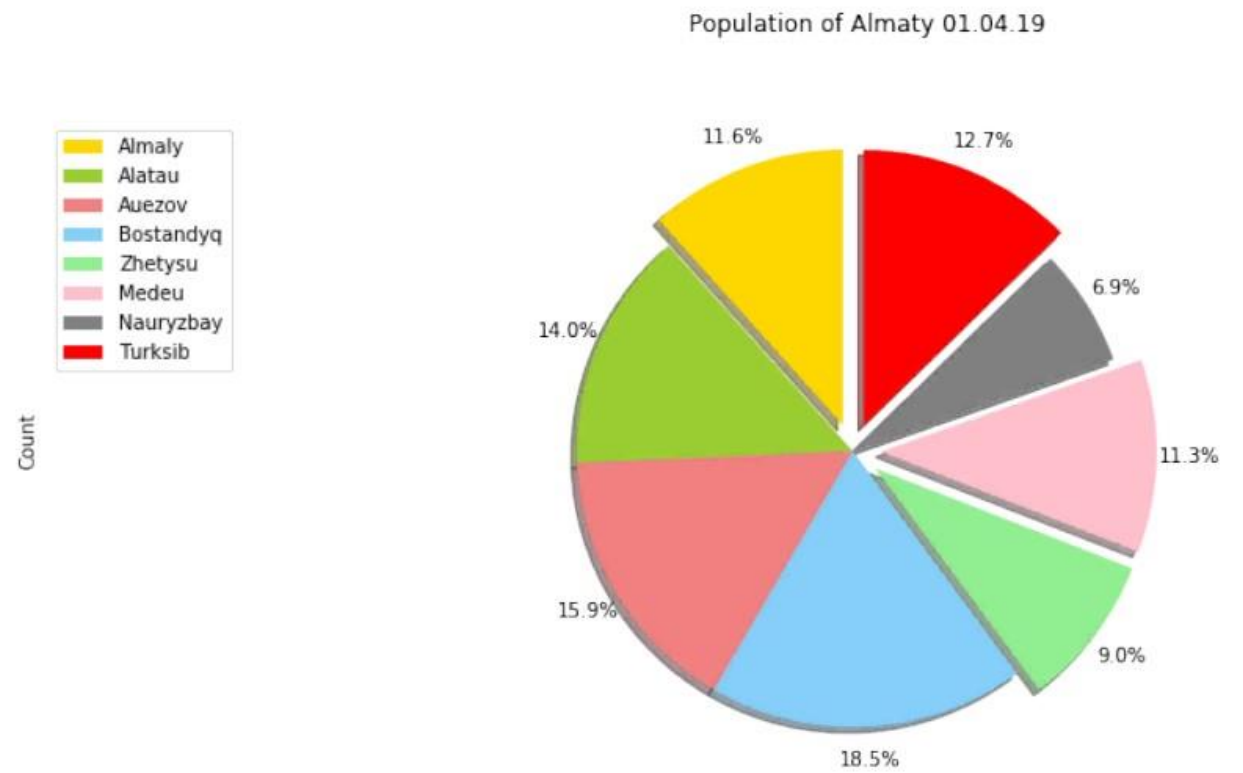
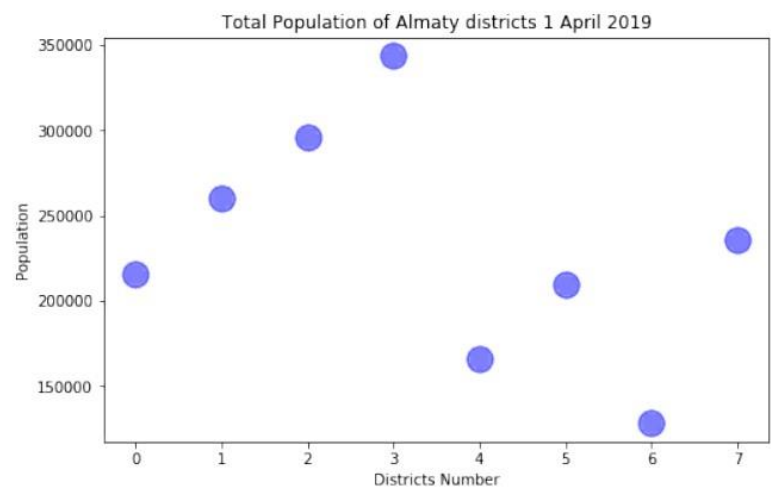


Диаграмма «Waffle»:

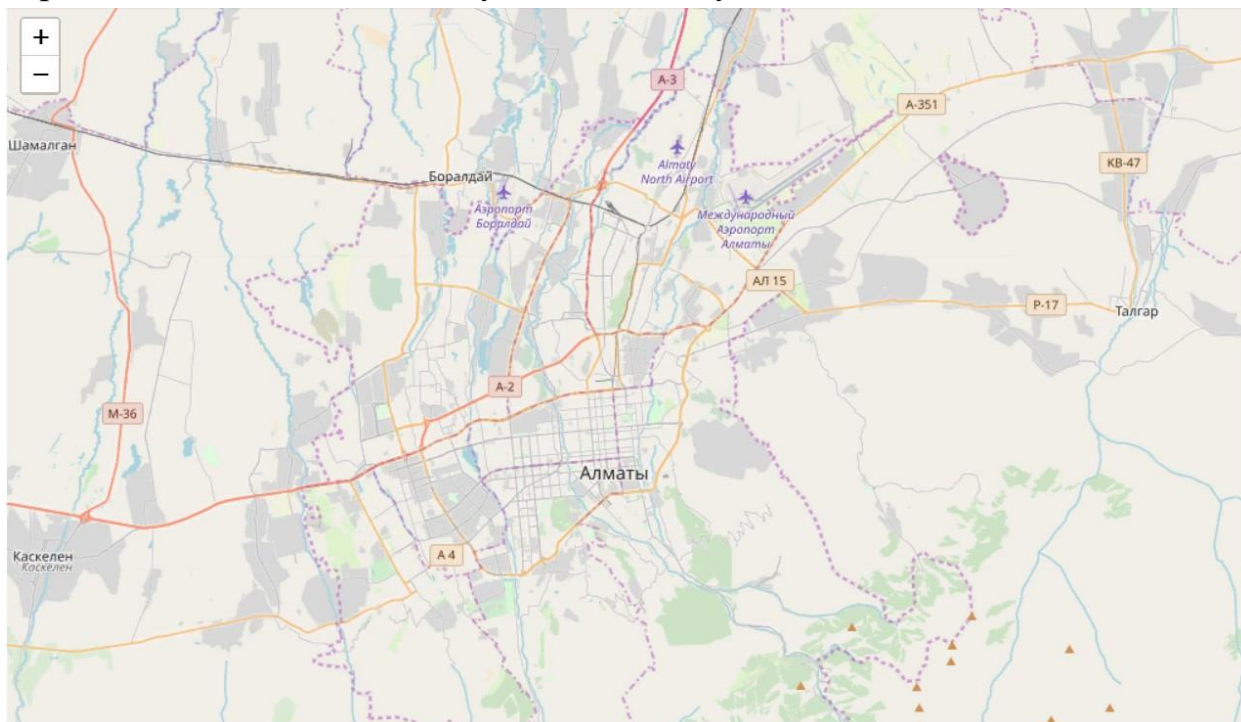


И также «Scatter plot»:

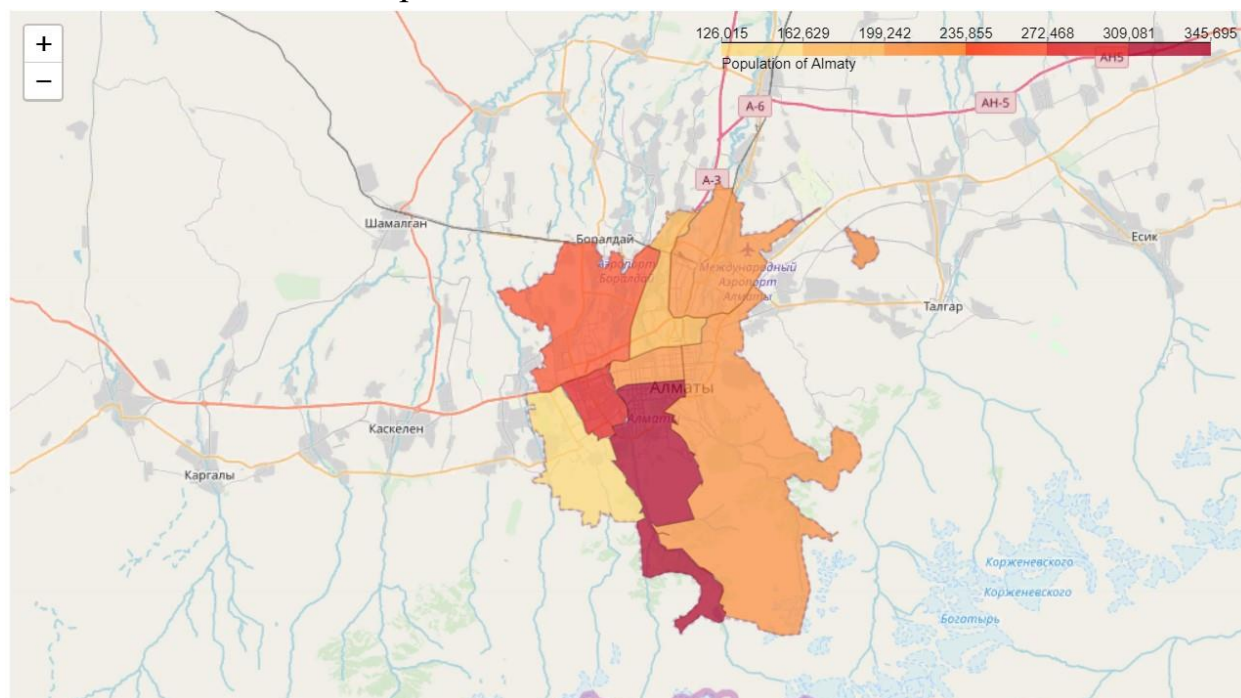


Эти диаграммы отлично показывают нам соотношение различных районов города Алматы.

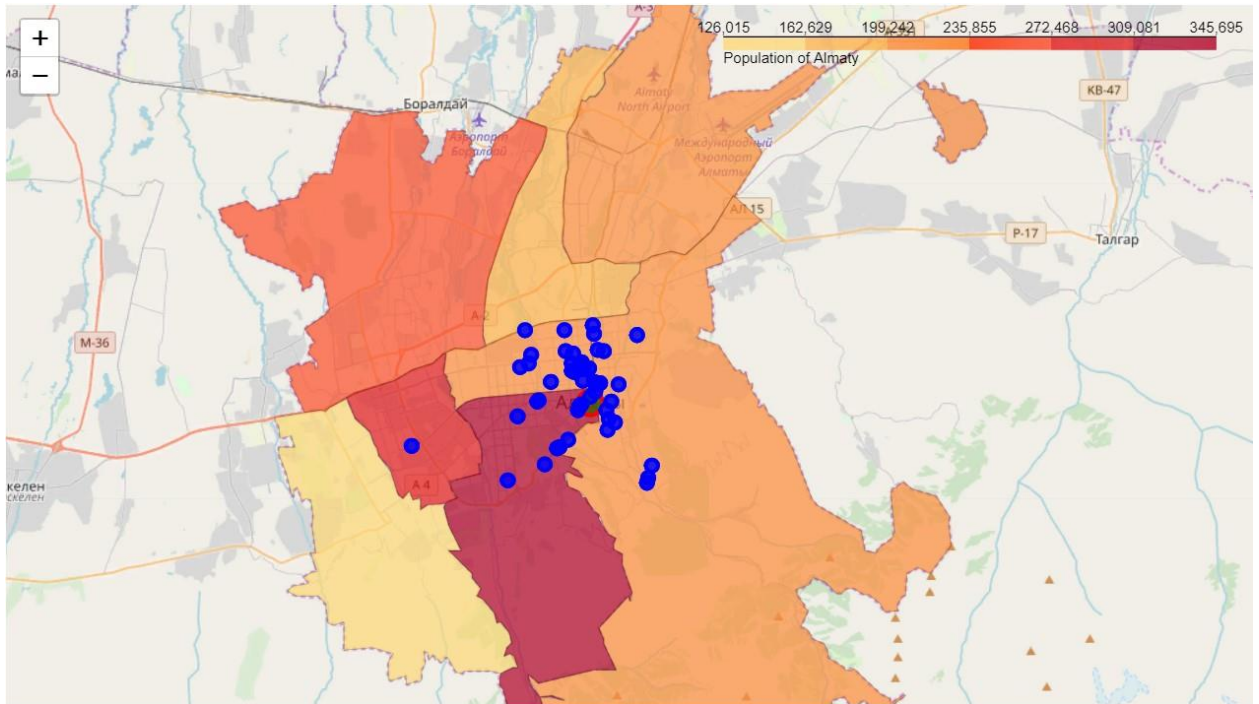
Далее нам следует обозначить плотность населения на самой карте города. Для этого мы используем библиотеку под названием «folium»:



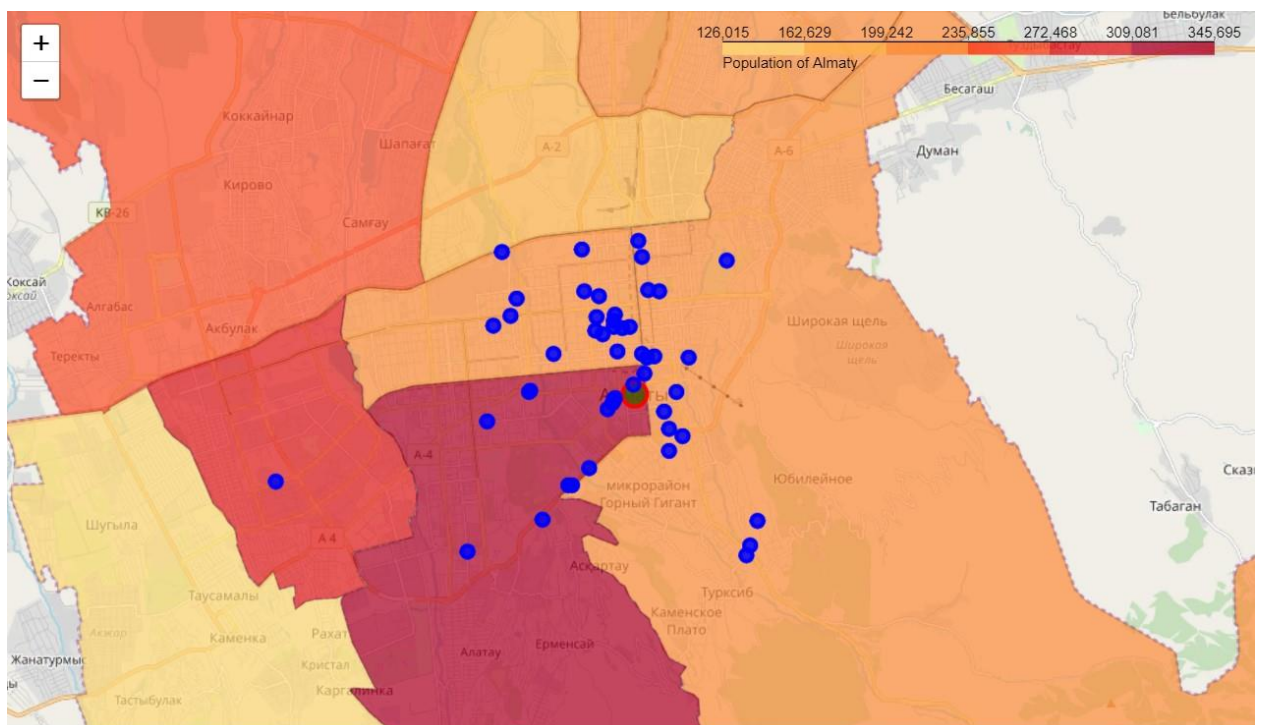
Теперь разделим карту города по районам используя GeoJSON файл, предоставленный [3]. Также используя функцию «choropleth» построим карту плотности населения по районам:



После построения карты города и классификации по плотности населения и различным районам, нам следует обратиться к сервису Foursquare для получения информации о ближайших ресторанах в городе. Зарегистрировавшись бесплатно на данном сервисе, мы получили доступ к Foursquare API. Далее давайте воспользуемся ими и добавим метки ресторанов на уже имеющуюся карту города Алматы:



Так как масштаб немного некорректно отображает информацию, нам следует немного приблизить картинку:



Теперь нам ясно видно, что большая часть ресторанов расположена около центра города, в частности в Медеуском районе и в самом населенном районе, количество ресторанов намного меньше, а в Ауэзовском районе, который стоит после Бостандыкского по плотности населения, находится только лишь один ресторан.

5. Результаты и обсуждения

После визуализации всех необходимых данных, пора приступать к результатам исследования.

Наш анализ показывает, что, хотя в Алматы существует большое количество ресторанов (около 50 в нашей первоначальной области интереса), довольно близко к центру города (немного западнее) есть места с низкой плотностью ресторанов и высокой плотностью населения. Наибольшая концентрация ресторанов была обнаружена в самом центре города и севернее от центра на расстоянии 5 км. Наше исследование было сосредоточено на выявление плотности населения и ресторанов по районам города Алматы. Из всех районов города, наиболее интересным нам показался Бостандыкский район. Так как он имеет максимальную плотность населения (около 350 000 человек) и минимальное количество ресторанов в округе. Наша рекомендация состоит в том, чтобы открыть ресторан именно в этом районе, так как он с наиболее высокой вероятностью будет пользоваться спросом. Либо посмотреть в сторону Ауэзовского района, так как плотность населения в нем почти так же высока, как и в Бостандыкском, а конкуренция в разы меньше.

Результатом всего этого является определение зон, содержащих наибольшее количество потенциальных новых ресторанов, в зависимости от количества и расстояния до существующих ресторанов в целом. Это, конечно, не означает, что эти зоны - действительно оптимальные места для нового ресторана! Цель этого анализа состояла в том, чтобы предоставить информацию о районах, их плотности населения и расположенных недалеко ресторанов, но не переполненных существующими ресторанами - вполне возможно, что есть очень веская причина для небольшого количества ресторанов в любом из этих районов, причины, которые сделает их непригодными для нового ресторана, несмотря на отсутствие конкуренции в этом районе. Поэтому рекомендуемые зоны следует рассматривать только как отправную точку для более подробного анализа, что в конечном итоге может привести к локациям, в которой не только не будет соревнования

поблизости, но также будут учтены другие факторы и соблюдены все другие соответствующие условия.

6. Заключение

Целью этого проекта было выявление районов Алматы, расположенных близко к центру, с небольшим количеством ресторанов, чтобы помочь заинтересованным сторонам сузить поиск оптимального местоположения для нового итальянского ресторана.

Рассчитав распределение плотности населения по данным Комитета по статистике Министерства национальной экономики Республики Казахстан и плотности ресторанов по данным Foursquare, мы построили рекомендательную систему по открытию нового ресторана в городе Алматы, которая с большей вероятностью будет пользоваться спросом.

Окончательное решение об оптимальном расположении ресторана будет приниматься заинтересованными сторонами на основе конкретных характеристик окрестностей и мест в каждой рекомендуемой зоне с учетом дополнительных факторов, таких как привлекательность каждого места (близость к парку или воде), уровень шума / близость к основным дорогам, наличие недвижимости, цены, социальная и экономическая динамика каждого района и т. д.

7. Источники

1. <http://stat.gov.kz>
2. <https://data.egov.kz>
3. <https://github.com/akilbekov/almaty.geo.json>
4. <https://2gis.kz>
5. <https://foursquare.com>
6. <https://www.coursera.org/learn/applied-data-science-capstone/>