Тема: **Анализ территории конкурентной среды торговой точки «Булочная 83» в ЖК «Морской Фасад»**

Выполнил: **Зиновьев Евгений (менеджер тт)**

Используемый язык программирования: Python (Anaconda)

Используемые карты: OpenStreetMap (openstreetmap.org)

Используемые библиотеки Python: osmnx – доступ к операциям над картами OpenStreetMap, pandas – для анализа данных, random – для случайного выбора между параметрами, matplotlib – для визуализации данных (! для реализации данного алгоритма необходимо, чтобы все библиотеки были установлены в активной среде Python!)

Задача: извлечь и отформатировать объекты-клиенты, объекты-конкуренты из OpenStreetMap, отобразить их месторасположения, показать зоны охвата торговых точек, а также их вероятности посещения со стороны объектов-клиентов и высчитать за счет этих вероятностей потенциальный клиентопоток за день.

План:

1. Извлечение ГеоДатаФреймов (ГДФ) из OpenStreetMap

2. Преобразование данных в более удобный для анализа вид

3. Нанесение на график фигуры нижнего слоя из ГДФ объектов-клиентов и верхнего слоя из ГДФ объектов-конкурентов

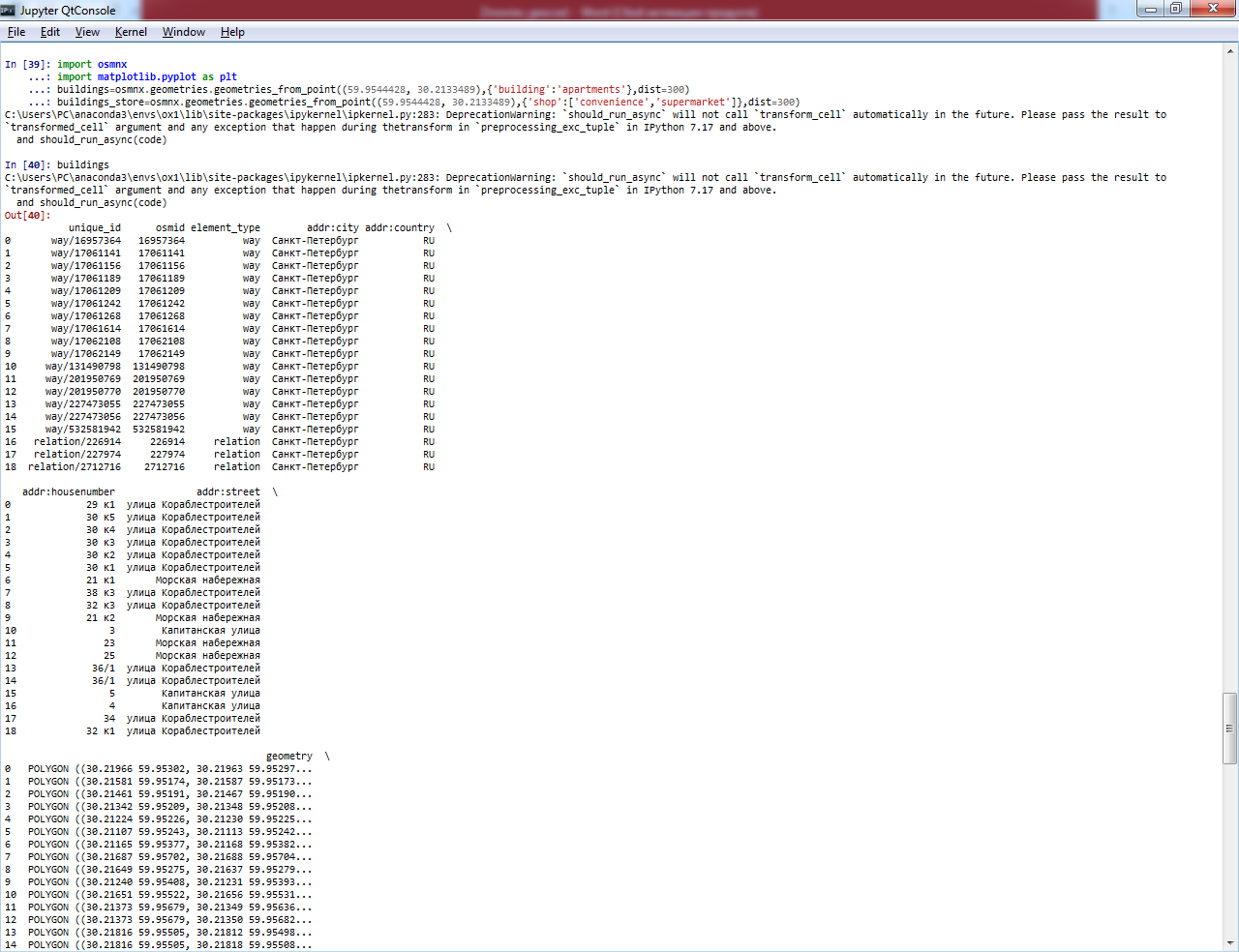
4. Применение инструментов редактирования графиков matplotlib

5. Статистический расчет

6. Визуализация статистического расчета

1. Извлечение ГеоДатаФреймов (ГДФ) из OpenStreetMap

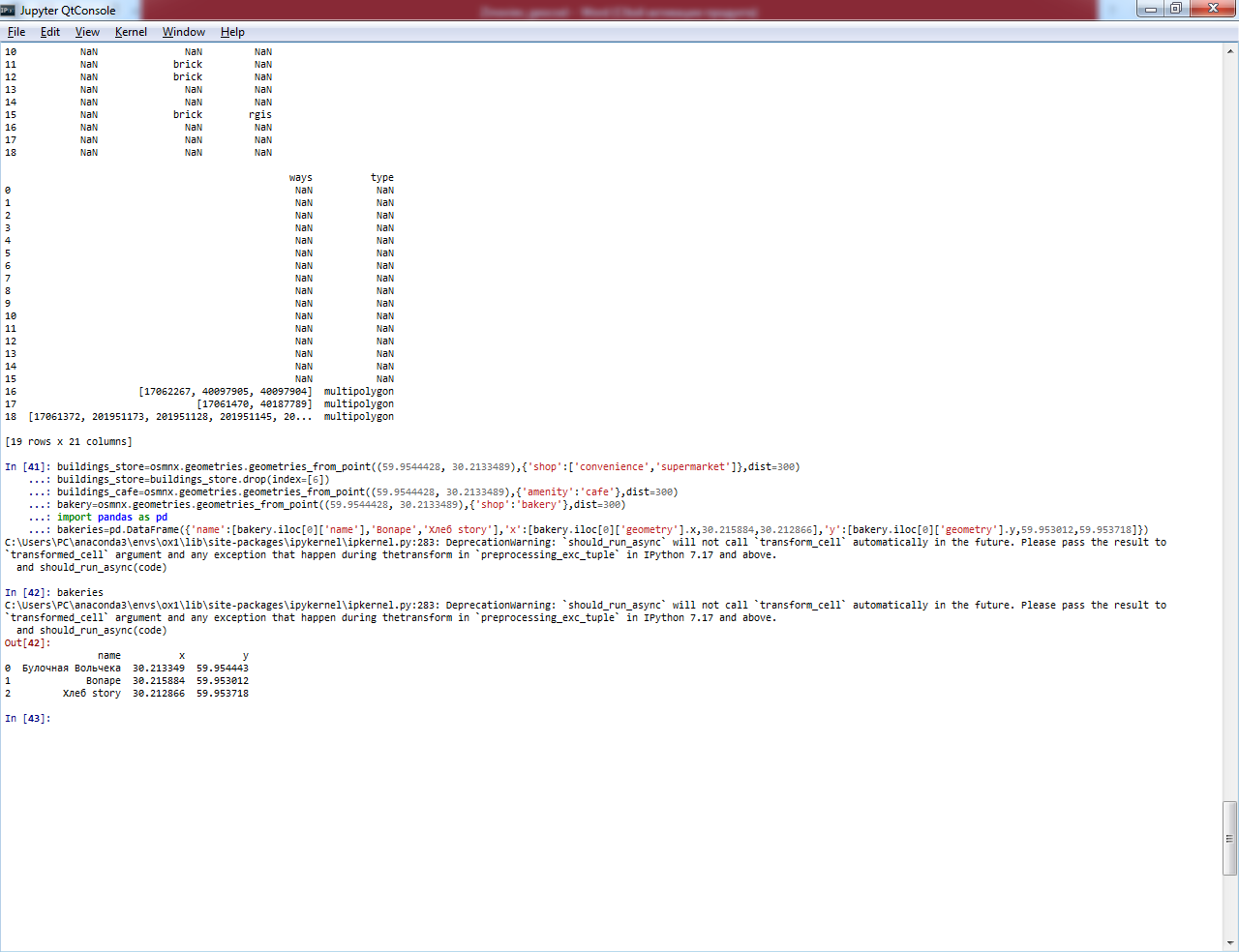
Для извлечения ГДФ у библиотеки osmnx есть метод geometries.geometries\_from\_point, который возвращает таблицу данных об указанном участке карт (параметры: координаты центра с радиусом - dist). Эту таблицу можно отфильтровать по тэгам: ‘building’:’apartments’ – например, значит, показать все жилые дома:



Таким образом, получаем такого вида таблицы данных, которые необходимо далее редактировать.

2. Преобразование данных в более удобный для анализа вид

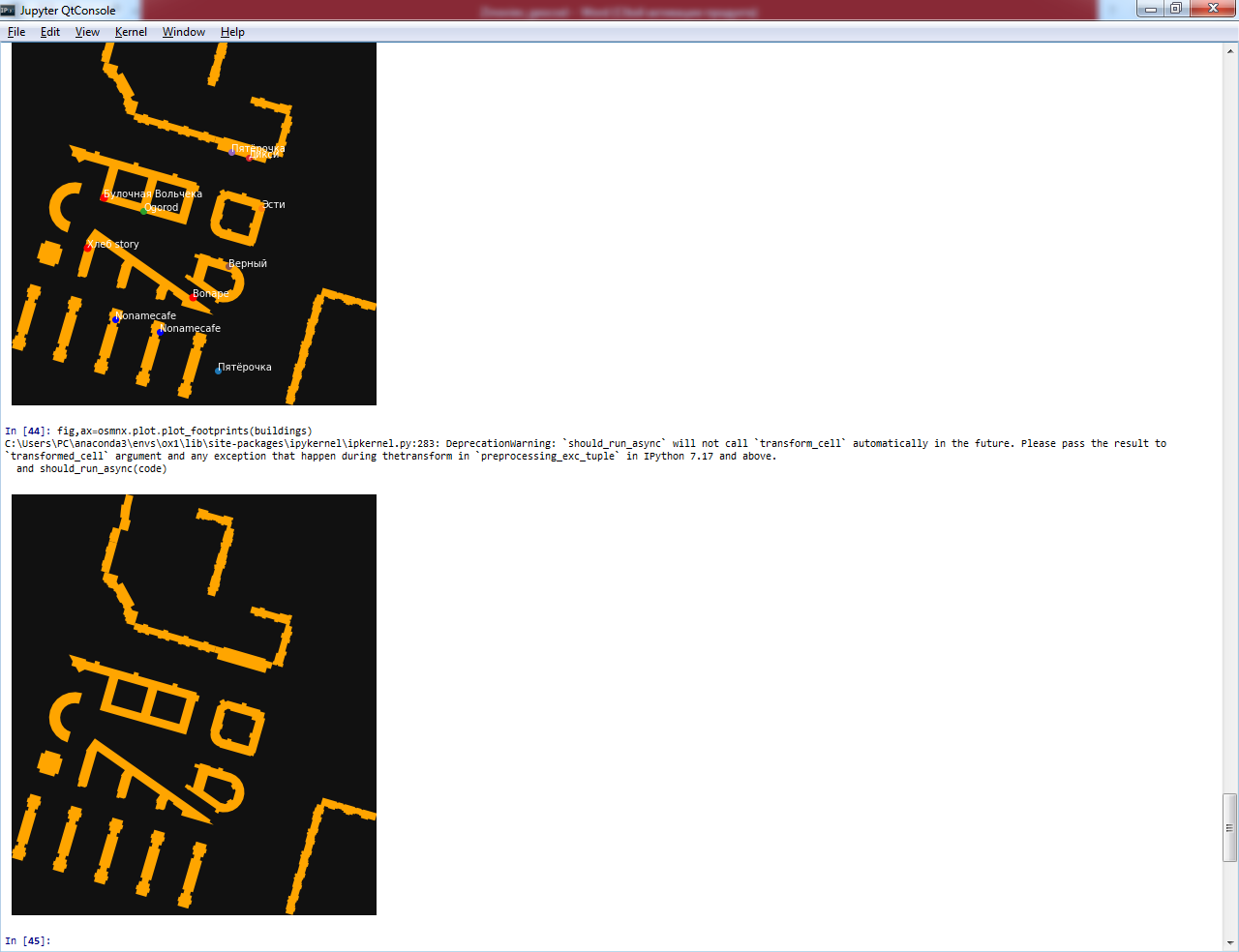
Необходимо удалить некоторые пустые данные, создать из имеющихся ГДФ свой собственный ДатаФрейм, с которым будет удобно работать.



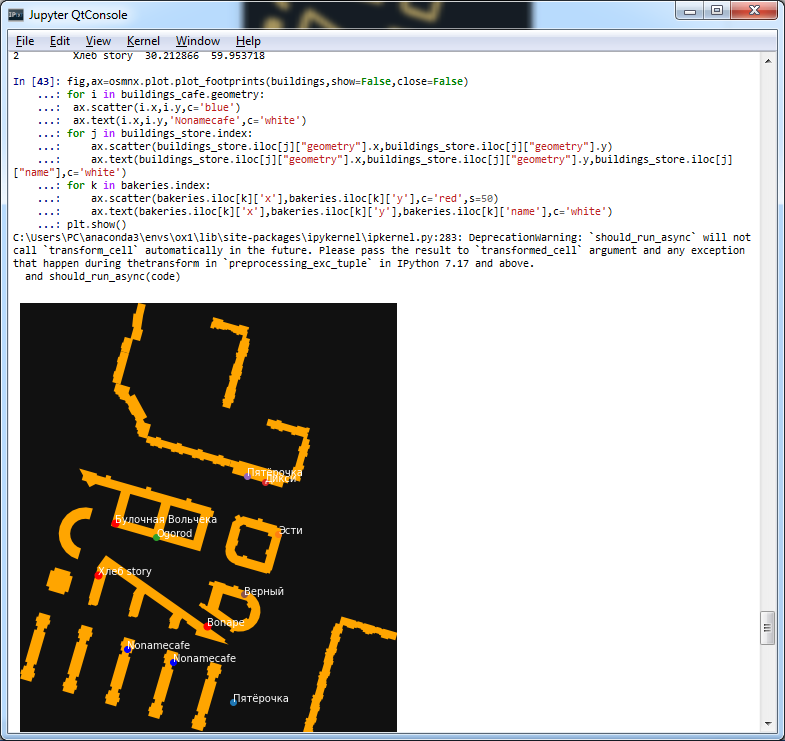
Получили массив bakeries – объекты-конкуренты

3. Нанесение на график фигуры нижнего слоя из ГДФ объектов-клиентов и верхнего слоя из ГДФ объектов-конкурентов

Вывод нижнего слоя из ГДФ объектов-клиентов:



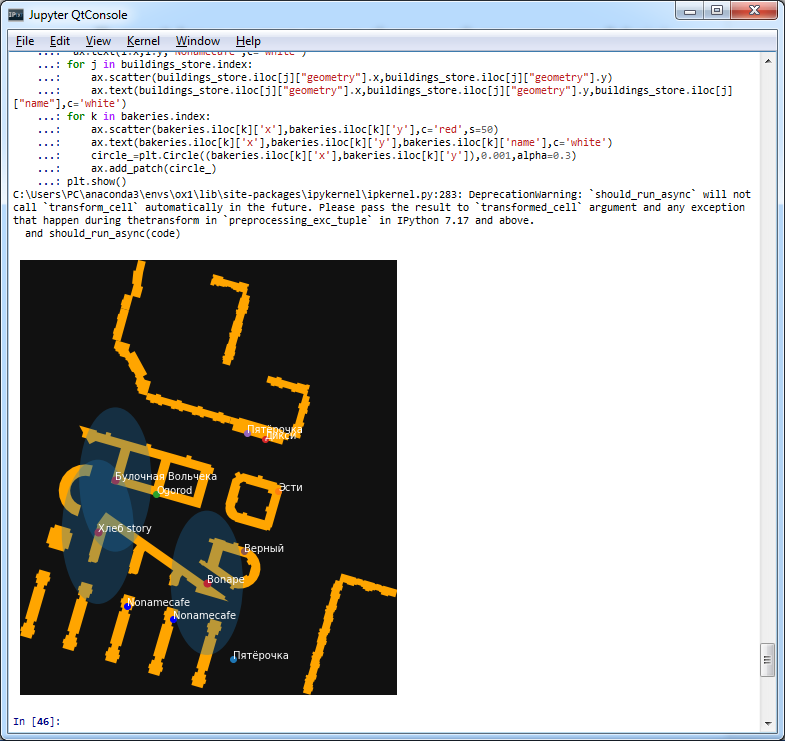
Нанесение на нижний слой точек верхнего слоя объектов-конкурентов:



На данной фигуре красными точками обозначены объекты-конкуренты (bakeries), другими цветами – магазины другого формата.

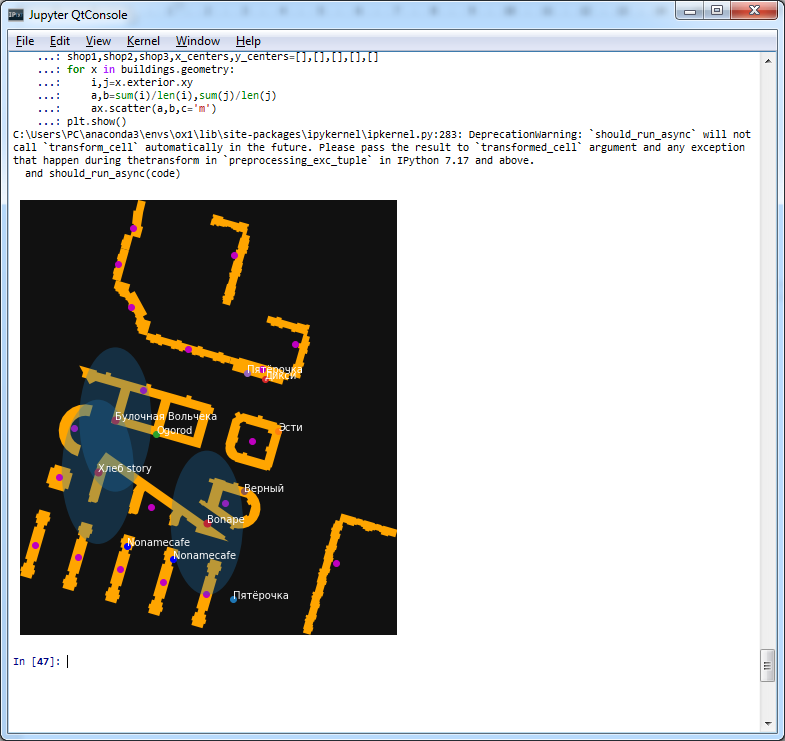
4. Применение инструментов редактирования графиков matplotlib

Нарисуем 100-метровые зоны охвата объектов-конкурентов:

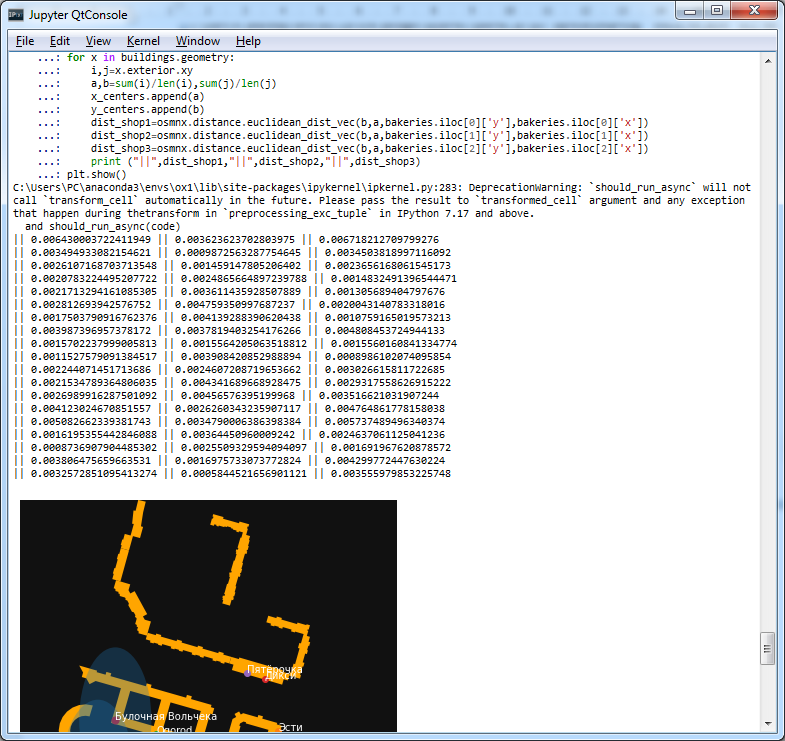


5. Статистический расчет

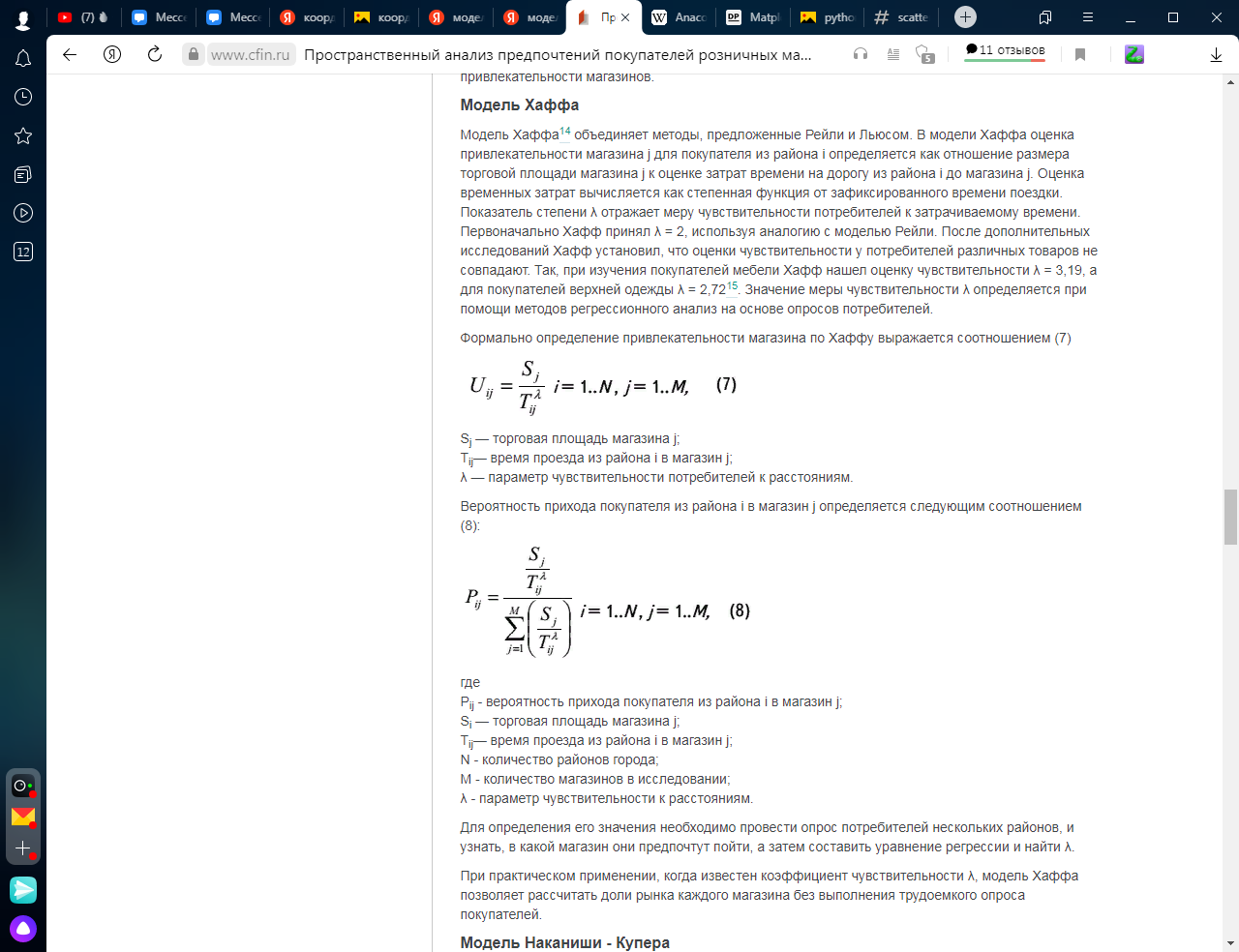
Найдем входы из жилых зданий. Каждое жилое здание представляет полигон точек (самый первый рисунок в данной работе – POLYGON), то есть набором точек, соединенных между собой, но надо как-то извлечь всего одну из них точку. Предположим, что вход в жилое здание – центр пересекающихся диагоналей многоугольника и просто извлечем все х и у координаты, как кортеж, в 2 отдельных списка с помощью метода класса polygon - exterior.xy и найдем среднее арифметическое значение (пурпурные неподписанные точки):



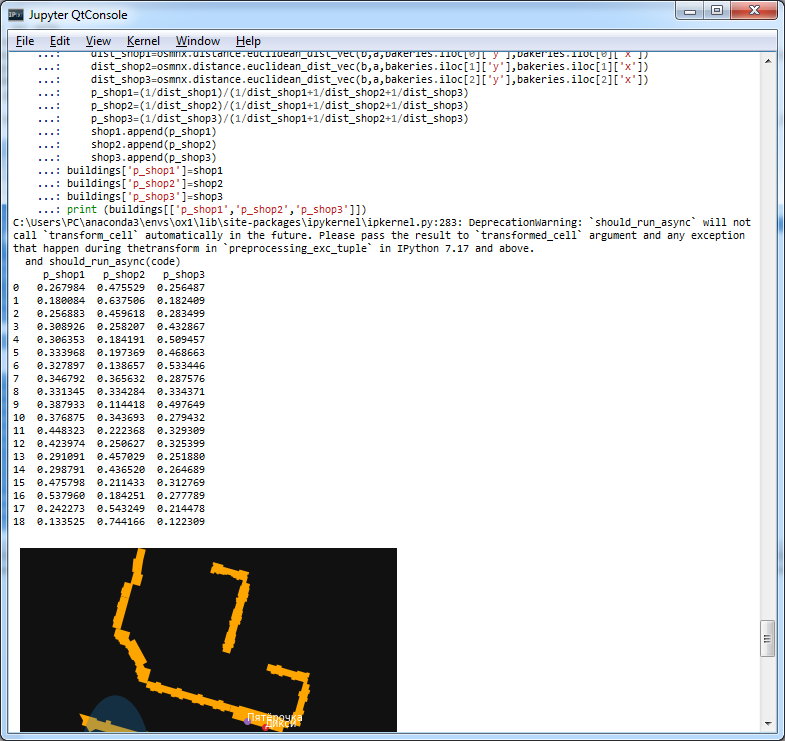
Далее вычислим расстояния от каждого объекта-конкурента до пурпурной точки (значения координат представлены в десятичном виде параметров широты и долготы):



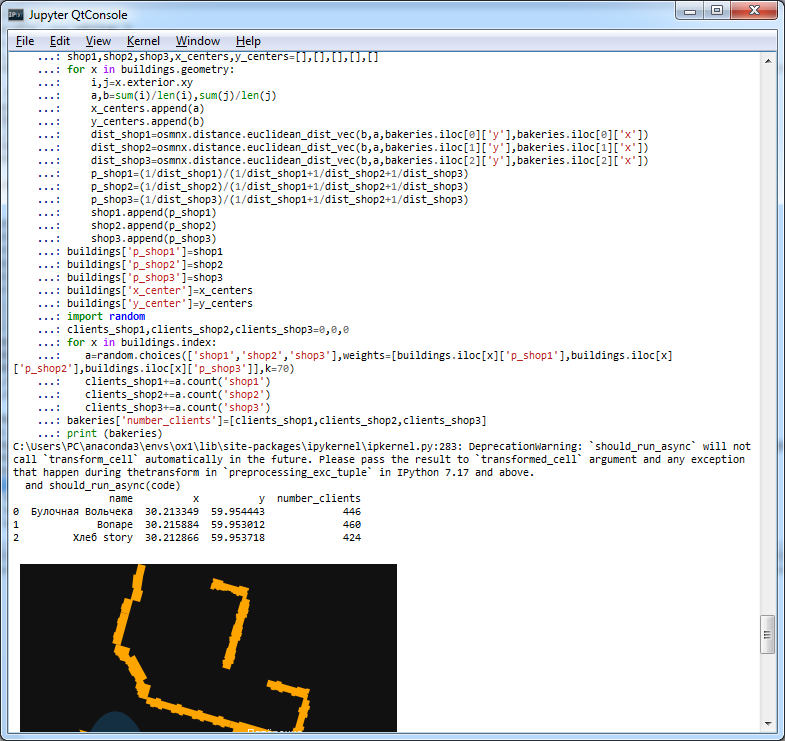
Далее вычислим вероятность прихода клиента из i-ой точки объекта-клиента (жилого здания) в j-й объект-конкурент по модели Хаффа без учета качественного параметра, в связи с отсутствующей информацией:



(и запишем значение вероятности в соответствующие столбцы массива:



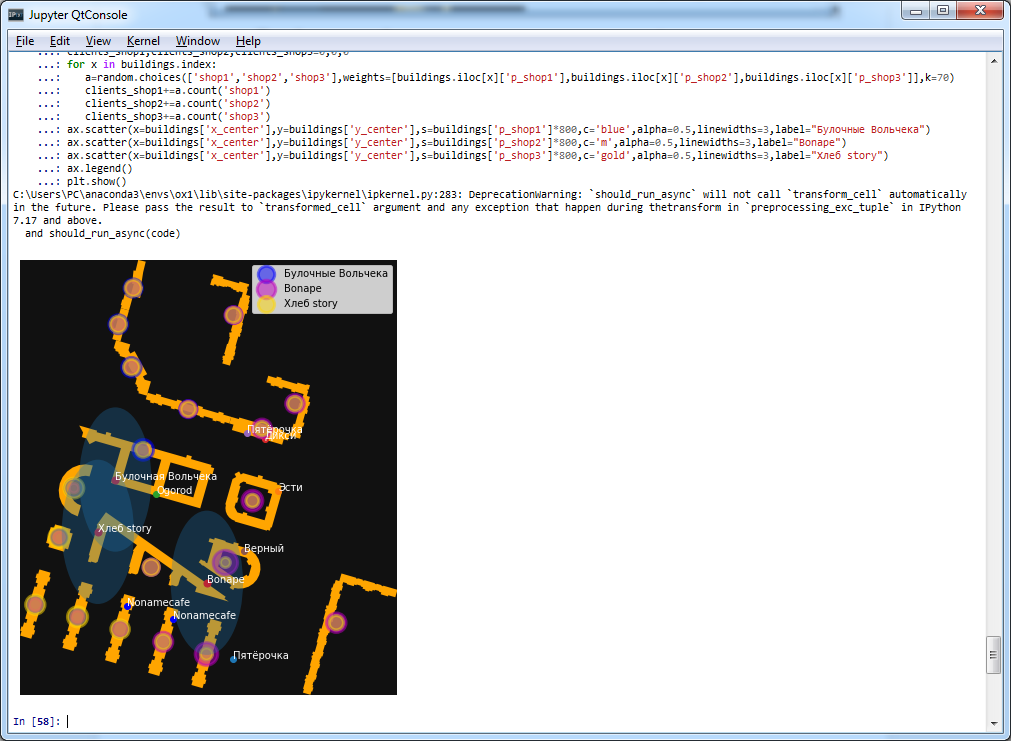
Далее вычислим примерное количество клиентов из одного объекта-клиента. По данным сайта <https://www.primeestate.ru/buildings/morskoy-fasad> : в исследуемом ЖК находится 1909 квартир на 19 домов (количество элементов в индексе таблицы на рисунке выше), то есть расчетное кол-во квартир – 100 шт. Среднее число заселенности квартиры – 3 чел и доля булочных на рынке ~ 15-25%. Перемножив эти значения, получим, что в каждом доме живет примерно 70 потенциальных клиентов булочных (в нашем случае – 3 шт). Теперь смоделируем выбор каждого из этих 70 человек для каждого объекта-клиента в соответствии с вычисленными весами в предыдущем пункте работы и посчитаем голоса, отданные в пользу каждого объекта-конкурента:



При бесконечно большом выполнении верхнеописанной программы в среднем случае у «Булочной Вольчека» потенциальных клиентов меньше, чем у «Bonape», но больше, чем у «Хлеб story». Высчитанное количество человек за день в торговой точке «Булочная 83» примерно соответствует реальному количеству человек, хотя на самом деле занижено на процентов 15 в связи с неучтенным качественным фактором бренда, так как он у Булочных Вольчека выше, чем у представленных конкурентов.

6. Визуализация статистического расчета

В конечном итоге, покажем распределения вероятности посещения клиента из каждого дома каждый объект-конкурент. Для этого изменим размер каждой точки входа в дом в соответствии с вероятностью (чем больше вероятность посещения у определенной булочной, тем ее цветовой круг больше в соответствии с легендой в верхнем правом углу):



Заключение: зоны охвата объектов-конкурентов в пределах 300 –метровой области распределены равномерно (с 5 %-м отклонением), расчетный клиентопоток составил 400-500 человек в день. Полная версия кода представлена в закрепленном блокноте.