**Навигация по файлам (каждый файл запускается сам в себе):**

Видео по программе (с голосовыми комментариями): https://drive.google.com/file/d/1fRyeMLP70qOgl5VZhP1F3z-b5AtoNp\_e/view?usp=sharing

[описание видео.mp4](описание%20видео.mp4)

дополнительное видео после добавление информации по домам:

<https://drive.google.com/file/d/1MceXNvV3UfXCB7IJWkr5cHYAG0sO_dvF/view?usp=sharing>

* **Папка utils:** - основные функции в которых все вычисления (файлы запускать не требуется) а также добавление координат через google api.
* **Папка parser** – парсеры сайтов с данными по домам, а также parser yandex-справочника
* **Папка files** – файлы для проекта полученные после запуска файлов и проекта
* **Папка Flask** – обертки для Flask

**Корень проекта:**

**а) 1.Get\_info\_for\_many\_new\_coords -** получения списка в котором будут как координаты и информация по действующим **торговым точкам, так и вновь полученные координаты точек и информация по ним, которые потом передаются в tableue**\_desctop. Для работы необходимы в наличии только файлы func, csvj, ostanovk.

**- ссылки** на отчеты из табло (интерактивные карты по новым и действующим бизнесам, карты можно увеличивать и при наведении на точку по ней отображается основная информация) и сам отчет можно скачать или открыть с того компьютера на котором он составлялся:

<https://public.tableau.com/views/Book16_16143677148020/Sheet1?:language=en&:display_count=y&publish=yes&:origin=viz_share_link>

<https://public.tableau.com/views/Book16_16143677148020/Satellite?:language=en&:display_count=y&publish=yes&:origin=viz_share_link>

**- эксель** файл со всей информацией как для новых так и для действующих бизнесов (ниже ссылка):

[files\ITOGI\_FOR\_MANY\_POINS.xlsx](files/ITOGI_FOR_MANY_POINS.xlsx)

**1.1.add\_rating** – скрипт который добавляет рейтинг к новым точкам.

**2.Get\_inf\_for\_1\_done** – в файле на вход приходят координаты адреса, где пользователь хочет открыть новую торговую точку(координаты получаются с яндекс карт), далее данная точка анализируется относительно всех действующих точек которые есть в базе. На выходе получаем список с информацией по новой точке и по всем действующим точкам, которые потом передаются в tableue\_desctop. Для работы данного файла (2.Get\_inf\_for\_1\_done) должен быть в наличии файл(“FOR\_CHEK\_SPISOK\_TEK\_KOORD\_WITH\_ALL\_INFO) который получается после запуска файла 3.ZAPIS\_VSEH\_COORD, а также файлы func, csvj, ostanovk.

На выходе получаем:

**- ссылки** на отчеты из табло:

<https://public.tableau.com/views/Project_test_16131009080200/NEWPOINT?:language=en&:display_count=y&publish=yes&:origin=viz_share_link>

- файлы с информацией по всем точкам (ниже ссылка) и новой точке(она со status “new”):

[files\ALL\_POINT\_INFO.xlsx](files/ALL_POINT_INFO.xlsx)

- и отдельно файл с информацией по запрошенной точке(ниже ссылка):

[files\one\_POINT\_INFO.xlsx](files/one_POINT_INFO.xlsx)

**в) 3.ZAPIS\_VSEH\_COORD** – вспомогательный файл, который заранее готовит файл с дополнительной информацией по действующим торговым точкам. Для работы необходимы в наличии только файлы func и csvj.Для работы программы данный файл запускать не нужно, так как необходимый файл (ниже) уже записан.

[**files\FOR\_CHEK\_SPISOK\_TEK\_KOORD\_WITH\_ALL\_INFO.json**](files/FOR_CHEK_SPISOK_TEK_KOORD_WITH_ALL_INFO.json)

**Дополнительные комментарии по работе:** для начала работы каждого файла нужен файл csvjs. Все вновь полученные json файлы из программ а-д) автоматически обновляются в отчетах в tableue. То есть изначально формируется формат отчета в tableue desktop и далее в отчетах данные обновляются автоматически при запуске программ, но при автоматическом обновлении ссылку на tableue online может получить только зарегистрированный на сайте пользователь. На данном этапе, предполагается, что заранее будет формироваться отчет по отраслям, и далее «по запросу» будет направляться на него ссылка и файлы эксель. Проверить какой-то 1 адрес можно файлом 2.Get\_inf\_for\_1\_done, но отрисовать карту по нему в автоматическом режиме на данном этапе нельзя. Программы под пп. а-в) работают не зависимо друг от друга.

**При запуске любой программы просит ввести категорию объекта:** cheap, midle, expensive. При этом можно ничего не вводить и просто нажать ENTER – в данном случае показатели пригодности точки будут рассчитаны без учета «классовости объекта».

**В итоговых файлах для новых точек (status = new) указаны оценочные значения основных метрик:**

Расстояние до конкурента (distance\_to\_closer\_conquer)

Расстояние до метро (distance\_to\_closer\_metro)

Численность населения в радиусе до ближайшего конкурента (chisl)

С учетом того что были спарсены дынные по численности населения в каждом доме Москвы, также дополнительно приведен расчет населения в радиусе до конкурента по численности домов которые находятся в радиусе.

Добавлена информация по количеству конкурентов в радиусе 2 км, численности населения в радиусе 2 км и соответственно населения на 1 конкурента в радиусе 2 км. Численно методов домов ниже чем численность методом плотности ввиду того что на сайте gosjkh дана информация по домам на общую численность 6,5 млн человек, то есть ~ 50% от официальной численности в Москве.

В итоговый файл (для ускорения скорости работы) попадают только точки, у которых все 3 вышеописанных компонента = “good”, или может быть ‘bad’ только расстояние до метро и до остановок общественного транспорта, и далее данные точки уже фильтруются в tablue. То есть суть работы программы найти точки с хорошими показателями локации с учетом их классовости.

Все новые точки получают rating, который рассчитывается как сумма логарифмов основных метрик (расстояние до метро, численность населения в радиусе до конкурента и в радиусе 2 км, количество конкурентов в радиусе 2 км, население на 1 конкурента). Далее на карте точки с лучшим рейтингом подсвечиваются цветом.

**Все файлы лежат в папке files.**

**В папке Flask скрипты с обертками для фласк.**

**Основное описание ПРОЕКТА:**

Приложение: электронный консалтинг (в начале все будет на Базе г Москвы).

Место внедрения приложения в продакшн: консалтинговый сервис в интернете(продажа отчетов в tablue).

Кому интересно приложение:

- малый и средний бизнес (сфера услуг)

- тем кто хочет открыть новый бизнес (сфера услуг)

Принцип работы:

На основании расположения торговой точки и срока ее работы на 1 месте, оценивается влияние внешних факторов на успешность такой точки:

* расстояние до ближайшего конкурента,
* близость метро,
* плотность населения в районе и в радиусе до ближайшего конкурента (определяется средняя по городу плотность населения в радиусе до ближайшего конкурента)
* сегмент в котором планирует работать клиент. Модель работает как с указанием сегмента и без, если сегмент указан, то анализ показателей новой торговой точки ведется на основании показателей действующих бизнесов в соответствующем сегменте.
* Выручке (в Проектt выручка поставлена синтетически) далее на базе юр лица возможна доработка проекта по получению ИНН и выручки из открытых источников)
* Срок работы (начала статистика по данной компании) – при отсутствии выручки срок работы это основной критерий прибыльности бизнеса, если > 1,5 лет то вероятно бизнес прибыльный.

Визуализация потенциальных новых торговых точек на карте в Tableue desktop

**Описание работы приложения:**

1. **В начале:**
   1. аккумулируются данные по действующим бизнесам из открытых источников. Есть парсер Яндекса, но в бесплатной версии только 10 компаний. Поэтому использована база 2гис.
   2. Генерация новых точек с заданным шагом (расстояние между точками)
   3. Из массива сгенерированных точек определяются те, что подходят под критерии(далее критериев может быть больше):

- расстояние до метро <= (среднего по действующим точкам)/1,2

- расстояние до ближайшего конкурента >= (среднего по действующим точкам)\*1,2

- численность населения в радиусе до ближайшего конкурента >=>= (среднего по действующим точкам)\*1,2

Если выбран сегмент, то анализ по действующим точкам проводится в рамках соответствующего сегмента.

**Что улучшить:**

Исключить из координат координаты рек, самих шоссе крупных и дорог. Добавить фичи при наличии информации по клиентам, близости проходных дорог, центров и так далее чтобы более детально и реалистично определять привлекательность точки.