Исследование рынка заведений общественного питания Москвы

Описание проекта

Заказчик: Инвесторы из фонда «Shut Up and Take My Money»

Исходные данные: датасет с заведениями общественного питания Москвы, составленный на основе данных сервисов Яндекс Карты и Яндекс Бизнес на лето 2022 года.

Цель исследования: подготовить исследование рынка Москвы, найти интересные особенности и презентовать полученные результаты, которые в будущем помогут в выборе подходящего инвесторам места.

Шаг 1. Загрузка данных и изучиение общей информации

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import plotly.express as px
from plotly import graph_objects as go
from io import BytesIO
import requests
import numpy as np

df = pd.read_csv('moscow_places.csv')
df.head()
```

6]:	ı	name	category	address	district	hours	lat	Ing	rating	price	avg_bill	middle_avg_bill	middle_coffee_cup	chain	seats
	0 WoV	Vфли	кафе	Москва, улица Дыбенко, 7/1	Северный административный округ	ежедневно, 10:00—22:00	55.878494	37.478860	5.0	NaN	NaN	NaN	NaN	0	NaN
	1	етыре инаты	ресторан	Москва, улица Дыбенко, 36, корп. 1	Северный административный округ	ежедневно, 10:00-22:00	55.875801	37.484479	4.5	выше среднего	Средний счёт:1500—1600 ₽	1550.0	NaN	0	4.0
	2 >	Хазри	кафе	Москва, Клязьминская улица, 15	Северный административный округ	пн-чт 11:00– 02:00; пт,с6 11:00–05:00; вс 11:00	55.889146	37.525901	4.6	средние	Средний счёт:от 1000₽	1000.0	NaN	0	45.0
	3 0	nouse Coffee Shop	кофейня	Москва, улица Маршала Федоренко, 12	Северный административный округ	ежедневно, 09:00—22:00	55.881608	37.488860	5.0	NaN	Цена чашки капучино:155– 185₽	NaN	170.0	0	NaN
	4	Иль 1арко	пиццерия	Москва, Правобережная улица, 1Б	Северный административный округ	ежедневно, 10:00—22:00	55.881166	37.449357	5.0	средние	Средний счёт:400—600₽	500.0	NaN	1	148.0

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8406 entries, 0 to 8405
 Data columns (total 14 columns):
                           Non-Null Count Dtype
      Column
                                         8406 non-null object
        name
        name
category
address
district
hours
                                        8406 non-null object
8406 non-null object
                                        8406 non-null
                                                                  object
                                        7870 non-null
8406 non-null
        lat
lng
rating
                                                                  float64
                                        8406 non-null
8406 non-null
3315 non-null
        price
                                                                  object
  ### Sprice 3315 non-null
### say_bill 3816 non-null
### niddle_avg_bill 3149 non-null
### niddle_coffee_cup 535 non-null
### chain 8406 non-null
### sets 4795 non-null
                                        3816 non-null
3149 non-null
                                                                 object
float64
                                                                    float64
                                         8406 non-null int64
4795 non-null float64
         seats
dtypes: float64(6), int64(1), object(7)
memory usage: 919.5+ KB
По сырым данным имеем данные о 8406 заведениях.
```

Шаг 2. Предобработка данных

```
In [5]: # проверми маличие полишх дубликатов-строк в таблице
df.duplicated().sum()

Out[5]: 0

CAEAJAEM AON. Проверку На Дубликаты по выбранным столбцам:

In [6]: w1 =df[['address','district','lat','lng','seats']].duplicated().sum()

Out[6]: 31

In [7]: w2 =df[['category','address','district','lat','lng','seats']].duplicated().sum()

Out[7]: 15

In [8]: w3 =df[['category','address','district','lat','lng','seats','hours']].duplicated().sum()

Out[8]: 5
```

По результату проверки на дубликаты, можно сделать предположение, что при записи данных в базу заведения были записаны несколько раз под разными категориями, часами работы и т.д. Но их одинаковое местоположение и одинаковое кол-во мест говорит о том, что это одни и те же заведения.

Количество пропусков в столбцах таблицы:

```
name 0
category 0
address 0
district 0
hours 536
lat 0
lng 0
rating 0
price 5091
avg_bill 4590
middle_avg_bill 5257
middle_coffee_cup 7871
chain 0
seats 3611
```

Пропуски имеем в столбцах hours, price, avg_bill, middle_avg_bill, middle_coffee_cup.seats. Для дальнейшей работы с данными, сделаем замену пропусков в столбцах: hours, price, avg_bill, а в остальных оставим пропуски как есть.

```
In [10]:

df['hours']= df['hours'].fillna('не указано')

df['price']= df['price'].fillna('не указано')

df['avg_bill']= df['avg_bill'].fillna('не указано')
```

Столбец с информацией об улице

```
In [11]:

# cosganne cronfig c naspannew ynum
words1 = ['ynuma','yn','mepeynon','mocce','npocnent','nnomanb','npoesn',
'ceno','annes','fynbap','hafepexhas','rynum','nuhus']

str_pat = r".*,\s*\b([^,]*?(?:())\b[^,]*)[,$]+".format("|".join(words1))

df["street"] = df["address"].str.extract(str_pat)
df.head(2)
```

Out[11]:		name	category	address	district	hours	lat	Ing	rating	price	avg_bill	middle_avg_bill	middle_coffee_cup	chain	seats	street
	0	WoWфли	кафе	Москва, улица Дыбенко, 7/1	Северный административный округ	ежедневно, 10:00-22:00	55.878494	37.478860	5.0	не указано	не указано	NaN	NaN	0	NaN	улица Дыбенко
	1	Четыре комнаты	ресторан	Москва, улица Дыбенко, 36, корп. 1	Северный административный округ	ежедневно, 10:00-22:00	55.875801	37.484479	4.5	выше среднего	Средний счёт:1500- 1600₽	1550.0	NaN	0	4.0	улица Дыбенко

Столбец с информацией о круглосуточной работе

```
In [12]: # Напишем функцию, с помощью которой в зависимости от наличия круглосуточного режима работы присвоим значение 'True' или 'False'

def alert_hours(x):
    if x == 'exequebno, круглосуточно':
        return 'True'
    return 'False'
#print( alert_hours('exequebno, круглосуточно')) #выполним проверку работы функции
```

Теперь создадим столбец 'is_24/7' и проверим полученный результат:

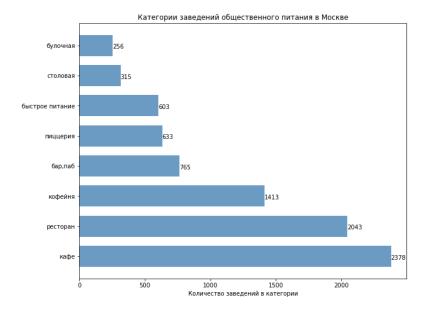
]:	nam	e category	address	district	hours	lat	Ing	rating	price	avg_bill	middle_avg_bill	middle_coffee_cup	chain	seats
	0 WoWφ <i>i</i>	іи кафе	Москва, улица Дыбенко, 7/1	Северный административный округ	ежедневно, 10:00-22:00	55.878494	37.478860	5.0	не указано	не указано	NaN	NaN	0	NaN
	1 Четыр комнат	ое ресторан ы	Москва, улица Дыбенко, 3б, корп. 1	Северный административный округ	ежедневно, 10:00-22:00	55.875801	37.484479	4.5	выше среднего	Средний счёт:1500– 1600₽	1550.0	NaN	0	4.
	2 Хазр	и кафе	Москва, Клязьминская улица, 15	Северный административный округ	пн-чт 11:00— 02:00; пт,с6 11:00—05:00; вс 11:00	55.889146	37.525901	4.6	средние	Средний счёт:от 1000₽	1000.0	NaN	0	45
	Dormou 3 Coffe Sho	е кофейня	Москва, улица Маршала Федоренко, 12	Северный административный округ	ежедневно, 09:00-22:00	55.881608	37.488860	5.0	не указано	Цена чашки капучино:155– 185₽	NaN	170.0	0	Na
	4 Иль Маря	о пиццерия	Москва, Правобережная улица, 1Б	Северный административный округ	ежедневно, 10:00-22:00	55.881166	37.449357	5.0	средние	Средний счёт:400—600 ₽	500.0	NaN	1	148
840	1 Суц Маня		Москва, Профсоюзная улица, 56	административный	ежедневно, 09:00-02:00	55.670021	37.552480	4.4	не указано	не указано	NaN	NaN	0	86
840	2 Миславні	ес кафе	Москва, Пролетарский проспект, 19, корп. 1		ежедневно, 08:00—22:00	55.640875	37.656553	4.8	не указано	не указано	NaN	NaN	0	150
840	З Самова	р кафе	Москва, Люблинская улица, 112А, стр. 1	Юго-Восточный административный округ	ежедневно, круглосуточно	55.648859	37.743219	3.9	не указано	Средний счёт:от 150₽	150.0	NaN	0	150
840	4 Чайхан Sa		Москва, Люблинская улица, 112А, стр. 1	Юго-Восточный административный округ	ежедневно, круглосуточно	55,648849	37.743222	4.2	не указано	не указано	NaN	NaN	1	15

Maces Lowers &

Шаг 3. Анализ данных

Распределение заведений по категориям.

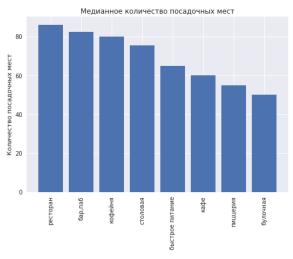
[14]:		category	number
	0	кафе	2378
	1	ресторан	2043
	2	кофейня	1413
	3	бар,паб	765
	4	пиццерия	633
	5	быстрое питание	603
	6	столовая	315
	7	булочная	256



Промежуточные выводы: наиболее распространённой категорией заведений общественного питания являются кафе. На втором месте - рестораны, на третьем-кофейни.

Медианное количество посадочных мест

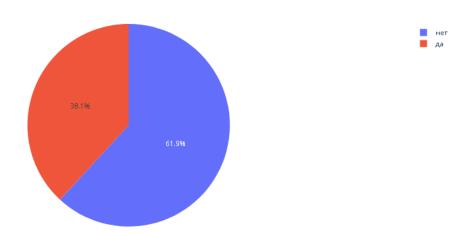
Out[17]:		category	category_cnt	seat_median
	0	ресторан	2043	86.0
	1	бар,паб	765	82.5
	2	кофейня	1413	80.0
	3	столовая	315	75.5
	4	быстрое питание	603	65.0
	5	кафе	2378	60.0
	6	пиццерия	633	55.0
	7	булочная	256	50.0



Промежуточные выводы: по количеству посадочных мест лидируют бары/пабы, рестораны и кофейни (чаще всего от 80-90) В остальных категориях в среднем кол-во посадочных мест находится в диапазоне 50-75.

Соотношение сетевых и несетевых заведений по количеству

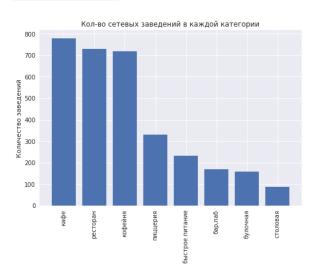
chain number 0 0 5201 **1** 1 3205



Промежуточные выводы: несетевый заведений примерно в 1,5 раза больше, чем сетевых.

Какие категории заведений чаще являются сетевыми

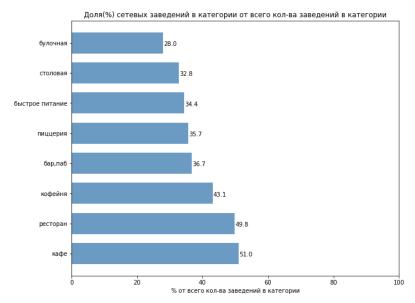
	category	number
0	кафе	779
1	ресторан	730
2	кофейня	720
3	пиццерия	330
4	быстрое питание	232
5	бар,паб	169
6	булочная	157
7	столовая	88



Рассчитаем долю сетевых заведений (выраженную в %) в каждой категории относительно всего количества, для этого разделим кол-во сетевых заведений в категории на всё кол-во заведений в категории, которое мы рассчитали ранее:

Charles	FOD I
Out	201

	category	number	part %
2	кофейня	720	51.0
6	булочная	157	49.8
3	пиццерия	330	43.1
4	быстрое питание	232	36.7
1	ресторан	730	35.7
7	столовая	88	34.4
0	кафе	779	32.8
5	бар,паб	169	28.0

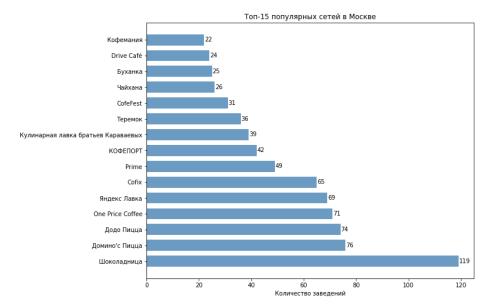


Промежуточные выводы: из графиков по абсолюным и % показателям видим,что в Топ-3 сетевых заведений входят типы объектов общественного питаний: кафе, рестораны и кофейни.

Топ-15 популярных сетей в Москве

Out[26]:

	name	category	number
0	Шоколадница	кофейня	119
1	Домино'с Пицца	пиццерия	76
2	Додо Пицца	пиццерия	74
3	One Price Coffee	кофейня	71
4	Яндекс Лавка	ресторан	69
5	Cofix	кофейня	65
6	Prime	ресторан	49
7	КОФЕПОРТ	кофейня	42
8	Кулинарная лавка братьев Караваевых	кафе	39
9	Теремок	ресторан	36
10	CofeFest	кофейня	31
11	Чайхана	кафе	26
12	Буханка	булочная	25
13	Drive Café	кафе	24
14	Кофемания	кофейня	22



Промежуточные выводы: большая часть самых популярных сетевых заведений относится к категории "кофейня"

Какие административные районы Москвы присутствуют в датасете

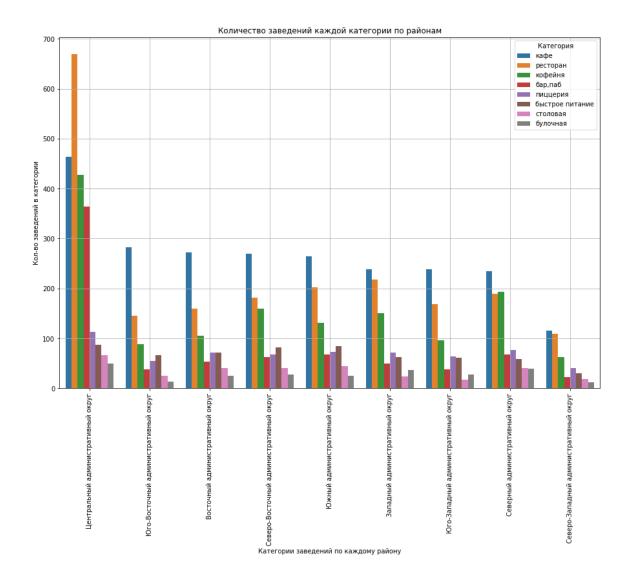
Посмотрим, какие районы присутсвует в имеющихся данных, а также какое кол-во заведений в каждом из них:

ut[29]:		district	number
	0	Центральный административный округ	2242
	1	Северный административный округ	900
	2	Южный административный округ	892
	3	Северо-Восточный административный округ	891
	4	Западный административный округ	851
	5	Восточный административный округ	798
	6	Юго-Восточный административный округ	714
	7	Юго-Западный административный округ	709
	8	Северо-Западный административный округ	409

Количество заведений каждой категории по районам

3]:	district		rest	cof_hause	pab	pizza	fast	stol	cake
0	Центральный административный округ	464	670	428	364	113	87	66	50
1	Юго-Восточный административный округ	282	145	89	38	55	67	25	13
2	2 Восточный административный округ		160	105	53	72	71	40	25
3	3 Северо-Восточный административный округ		182	159	63	68	82	40	28
4	Южный административный округ	264	202	131	68	73	85	44	25
5	Западный административный округ	239	218	150	50	71	62	24	37
6	 Юго-Западный административный округ 		168	96	38	64	61	17	27
7	7 Северный административный округ		189	193	68	77	58	41	39
8	8 Северо-Западный административный округ		109	62	23	40	30	18	12

Отразим данные по количеству завдений по каждой категории по каждому району на графике:

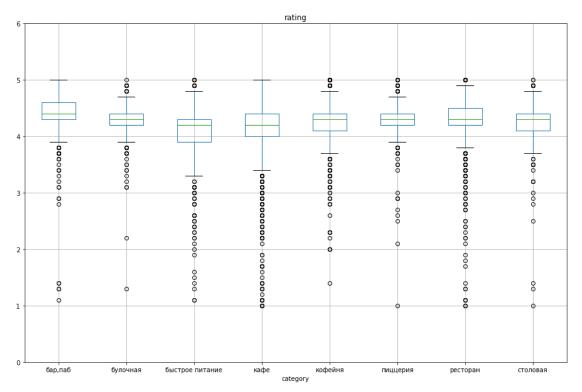


Промежуточные выводы: по результатам рассчётов и визуализации видим, что по количеству заведений лидирует Центральный адм.округ, меньше всего заведений в Северо-Западном адм. округе.По всем районам прослежвается тенденция, характерная для всего датасета - лидируют по кол-ву заведений рестораны, кафе и кофейни. Также видим, что в отличии от других районов, на первом месте в Центральном районе находятся рестораны, а потом кафе.

Распределение средних рейтингов по категориям заведений

[36]:		category	rating_mean
	0	бар,паб	4.387712
	1	пиццерия	4.301264
	2	ресторан	4.290357
	3	кофейня	4.277282
	4	булочная	4.268359
	5	столовая	4.211429
	6	кафе	4.123886
	7	быстрое питание	4.050249

Построим диаграмму размаха ("Ящик с усами") по рейтингам каждой категории заведений:



Промежуточные выводы: по сделанному расчёту среднего рейтинга, можем сделать вывод, что между разными типами заведений есть различия в рейтингах, но они являютя не значительными (в десятых и сотых).

Промежуточные выводы: по сделанному расчёту среднего рейтинга, можем сделать вывод, что между разными типами заведений есть различия в рейтингах, но они являютя не значительными (в десятых и сотых).

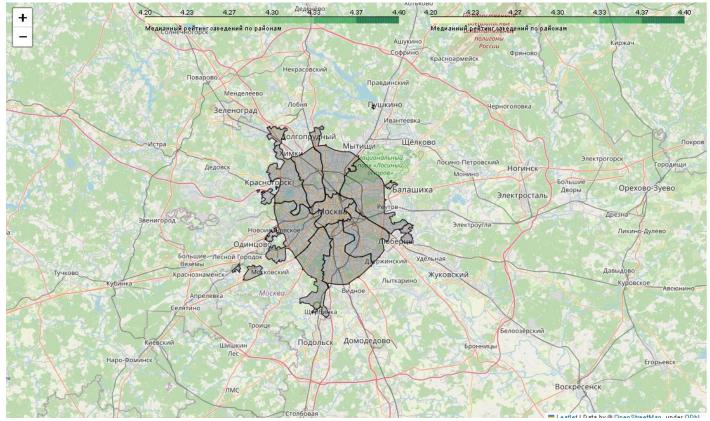
По диаграмме размаха ("Ящик с усами") видим, что нормальным распределением значений рейтинга для каждой категории являются следующие значения:

- бар,паб 3,9 5,0
- булочная 3,9 4,7
- пиццерия 3,9 4,7
- ресторан 3,8 4,9
- кофейня 3,7 4,8
- столовая 3,7 4,8
- кафе 3,5 5,0
- быстрое питание 3,2 4,8 Значения, расположенные до или после, являются выбросами.

Фоновая картограмма (хороплет) со средним рейтингом заведений каждого района.

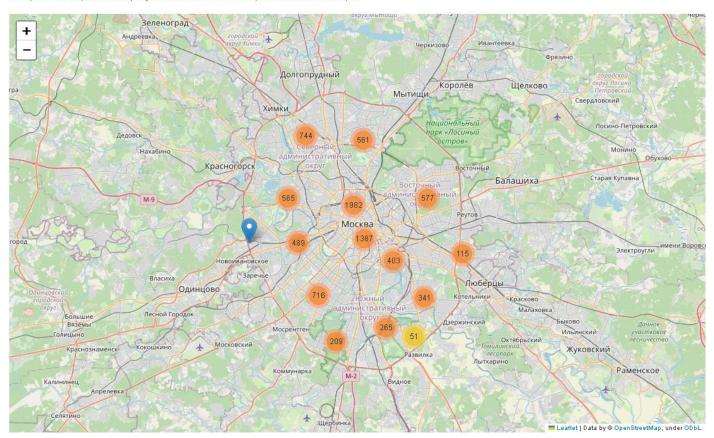
В датасете представлено 8 районов Москвы. Для каждого района посчитаем медианный рейтинг торговых центров, которые находятся на его территории:

Out[39]:		district	rating
	0	Восточный административный округ	4.3
	1	Западный административный округ	4.3
	2	Северный административный округ	4.3
	3	Северо-Восточный административный округ	4.2
	4	Северо-Западный административный округ	4.3
	5	Центральный административный округ	4.4
	6	Юго-Восточный административный округ	4.2
	7	Юго-Западный административный округ	4.3
	8	Южный административный округ	4.3



Кластеры заведений

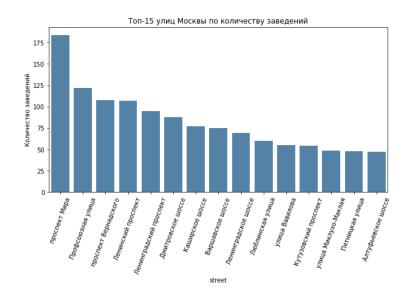
Отобразим все торговые центры, указанные в датасете на карте с помощью кластеров:



Промежуточные выводы: на графике, отображающем распределение заведений по кластерам также видим, что самые крупные кластеры находятся в центральной части Москвы, далее следуют Сервный, Западный и Восточные районы. Меньшая плотность наблюдается в Южной части.

Топ-15 улиц по количеству заведений

[45]:		street	number_object
	0	проспект Мира	184
	1	Профсоюзная улица	122
	2	проспект Вернадского	108
	3	Ленинский проспект	107
	4	Ленинградский проспект	95
	5	Дмитровское шоссе	88
	6	Каширское шоссе	77
	7	Варшавское шоссе	75
	8	Ленинградское шоссе	69
	9	Люблинская улица	60
	10	улица Вавилова	55
	11	Кутузовский проспект	54
	12	улица Миклухо-Маклая	49
	13	Пятницкая улица	48
	14	Алтуфьевское шоссе	47



Количество заведений каждой категории по улицам Тор-15

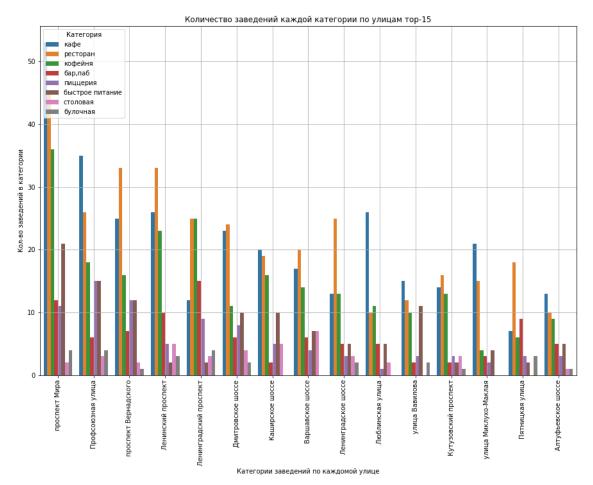
Out[47]:		street	cafe	number_object
	0	проспект Мира	53	184
	1	Профсоюзная улица	35	122
	2	проспект Вернадского	25	108
	3	Ленинский проспект	26	107
	4	Ленинградский проспект	12	95
	5	Дмитровское шоссе	23	88
	6	Каширское шоссе	20	77
	7	Варшавское шоссе	17	75
	8	Ленинградское шоссе	13	69
	9	Люблинская улица	26	60
	10	улица Вавилова	15	55
	11	Кутузовский проспект	14	54
	12	улица Миклухо-Маклая	21	49
13		Пятницкая улица	7	48
	14	Алтуфьевское шоссе	13	47

Control	DAOI:
Ou L	49 3

	street	cafe	number_object	rest	cof_hause	pab	pizza	fast	stol	cake
0	проспект Мира	53	184	45	36	12	11	21	2.0	4.0
1	Профсоюзная улица	35	122	26	18	6	15	15	3.0	4.0
2	проспект Вернадского	25	108	33	16	7	12	12	2.0	1.0
3	Ленинский проспект	26	107	33	23	10	5	2	5.0	3.0
4	Ленинградский проспект	12	95	25	25	15	9	2	3.0	4.0
5	Дмитровское шоссе	23	88	24	11	6	8	10	4.0	2.0
6	Каширское шоссе	20	77	19	16	2	5	10	5.0	NaN
7	Варшавское шоссе	17	75	20	14	6	4	7	7.0	NaN
8	Ленинградское шоссе	13	69	25	13	5	3	5	3.0	2.0
9	Люблинская улица	26	60	10	11	5	1	5	2.0	NaN
10	улица Вавилова	15	55	12	10	2	3	11	NaN	2.0
11	Кутузовский проспект	14	54	16	13	2	3	2	3.0	1.0
12	улица Миклухо-Маклая	21	49	15	4	3	2	4	NaN	NaN
13	Пятницкая улица	7	48	18	6	9	3	2	NaN	3.0
14	Алтуфьевское шоссе	13	47	10	9	5	3	5	1.0	1.0

На основе сводной таблицы построим график

На основе сводной таблицы построим график



Промежуточные выводы: по результатам рассчётов и визуалици видим, что но самых популярных улицах, где расположены заведения, сохранятеся тенденция, характерная для всего датасета - самыми часто встерчающимися типами заведений являются кафе или рестораны, и в тройке лидеров в большенстве случаев- кофейни.

Улицы, на которых находится только один объект общепита

Найдём число улиц с одним объектом общественного питания:

Число улиц с одним объектом общественного питания: 414

Out[52]:		street	number_object	category	seats
	0	улица Усиевича	1	пиццерия	48.0
	1	улица Уткина	1	кофейня	NaN
	2	2-й Хорошёвский проезд	1	столовая	NaN
	3	2-й Щемиловский переулок	1	кафе	30.0
	4	улица Мусоргского	1	столовая	NaN
	409	Чонгарский бульвар	1	кофейня	55.0
	410	Черёмушкинский проезд	1	кафе	NaN
	411	Чермянский проезд	1	ресторан	NaN
	412	Челюскинская улица	1	ресторан	NaN
	413	1-й Автозаводский проезд	1	кафе	20.0

414 rows × 4 columns

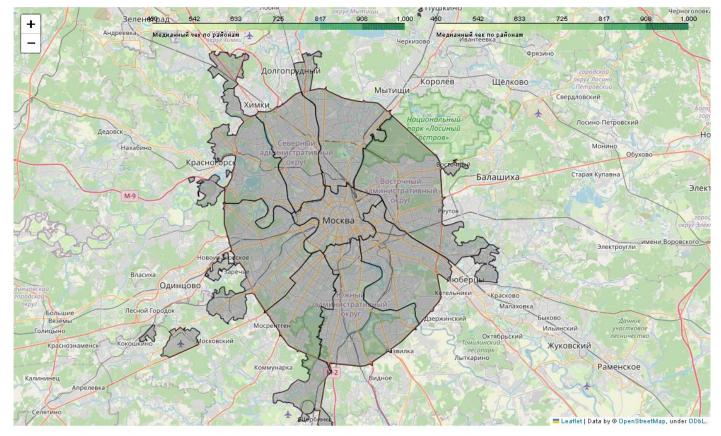
rt[53]:		category	category_cnt	seat_mean
	0	столовая	36	83.285714
	1	ресторан	88	71.428571
	2	бар,паб	40	66.000000
	3	пиццерия	14	59.888889
	4	булочная	9	59.600000
	5	кофейня	73	47.166667
	6	быстрое питание	17	47.000000
	7	кафе	137	43,234043

Промежуточные выводы: заведения, которые являются единственными на своей улице - предствлены всеми категориями заведение, при этом, ср. кол-во посадочных мест в таких заведениях ниже (в диапазоне 43-83, тогда как по всем данным этот диапазоне 90-120 мест.

Медиана средних чеков по каждому району

В датасете представлено 8 районов Москвы. Для каждого района посчитаем медианный средний чек:

Out[54]:		district	median_avg_biil
	0	Западный административный округ	1000.0
	1	Центральный административный округ	1000.0
	2	Северо-Западный административный округ	700.0
	3	Северный административный округ	650.0
	4	Юго-Западный административный округ	600.0
	5	Восточный административный округ	575.0
	6	Северо-Восточный административный округ	500.0
	7	Южный административный округ	500.0
	8	Юго-Восточный административный округ	450.0



Промежуточные выводы: по данным рассчётов и графика видим, что есть расположение влияет на средний чек заведения:

- самый высокий средний чек, равный 1000 руб, в Западном и Центральном адм. округах,
- средний чек в диапазоне 700-600 руб в Северо-Западной, Северном, Юго-Западном адм. округах *средний чек в диапазоне 575-450 руб. в Восточный, Северо-Восточный,Южный, Юго-Восточный адм. округе

Обобщённый вывод по анализу данных

Какие административные районы Москвы присутствуют в датасете Центральный,Северный,Южный,Северо-Восточный,Западный,Восточный,Юго-Восточный,Юго-Западный, Северо-Западный

Топ-15 улиц по количеству заведений Проспект Мира.Проф союзная улица Проспект Вернадского,Ленинский проспект, Ленинградский проспект,Дмитровское шоссе,Каширское шоссе, Варшавское шоссе,Ленинградское шоссе,Люблинская улица,улица Вавилова,Кутузовский проспект,улица Миклухо-Маклая,Пятницкая улица,Алтуфьевское шоссе

улица,Алтуфьевское шоссе Распределение заведений по категориям: наиболее распространённой категорией заведений общественного питания являются кафе. На втором месте - рестораны, на

Соотношение сетевых и несетевых заведений по количеству несетевый заведений примерно в 1,5 раза больше, чем сетевых.

Категории заведений, которые чаще являются сетевыми Топ-3 сетевых заведений входят типы объектов общественного питаний: кафе, рестораны и кофейни

Ton-15 популярных сетей в Москве Шоколадница, Домино'с Пицца, Додо Пицца, One Price Coffee кофейня,Яндекс Лавка,Соfіх,Prime, КОФЕПОРТ, Кулинарная лавка братьев Караваевых,Теремок,СofeFest,Чайхана,Буханка, Drive Café,Кофемания.

Медианное количество посадочных мест по количеству посадочных мест лидируют бары/пабы, рестораны и кофейни (чаще всего от 80-90) В остальных категориях в среднем кол-во посадочных мест находится в диапазоне 50-75.

Распределение средних рейтингов по категориям заведений: между разными типами заведений есть различия в рейтингах от 4,05 до 4,38, но они являютя не значительными (в десятых и сотых).

Медиана средних чеков по районам: колеблется от 450 до 1000 руб.

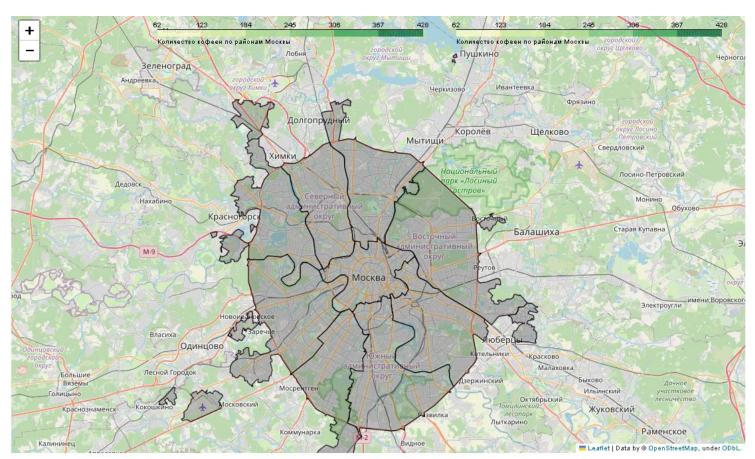
Шаг 4. Детализация исследования: открытие кофейни

Сколько всего кофеен в датасете

```
# рассчитаем кол-во кофеен, записанных в дата сете
print(len(df_3))
```

1413

]:		district	number
	0	Центральный административный округ	428
	1	Северный административный округ	193
	2	Северо-Восточный административный округ	159
	3	Западный административный округ	150
	4	Южный административный округ	131
	5	Восточный административный округ	105
	6	Юго-Западный административный округ	96
	7	Юго-Восточный административный округ	89
	8	Северо-Западный административный округ	62

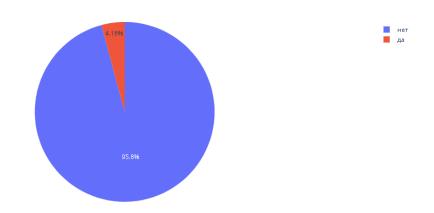


Промежуточные выводы: по результатам расчётов и визуализации видим, что:

- наибольшее кол-во кофеен расположено в Центральном адм. округе 428
- в диапазоне от 193 до 105 кофеен расположено в Северо-Восточном, Западном, Южном,Восточном адм. округах.
- в диапазоне 96-62 кофейни расположено в Юго-Западном, Юго-Восточном и Северо-Западном адм. округах.

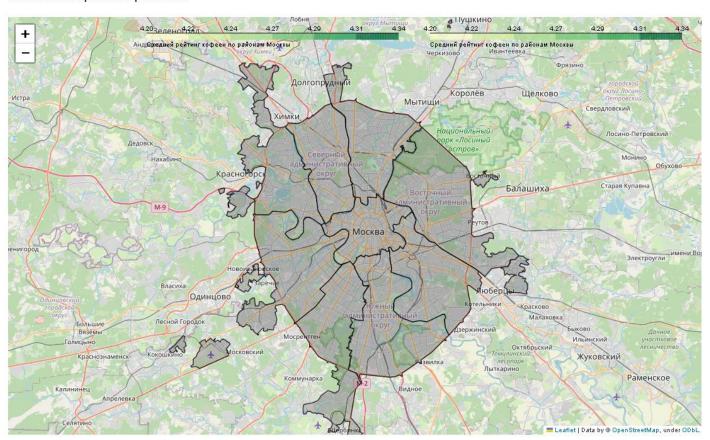
Доля круглосуточных кофеен

	is_24/7	number
0	False	1354
1	True	59



На графике видим, что только 4,2% кофеен являются круглосуточными.

Рейтинги кофеен по районам



Промежуточные выводы: по результатам расчётов и визуализации видим, что диапазон рейтиргов по районом колеблится в диапазоне 4.1- 4.3:

- на 1-м месте по рейтингу кофеени в Центральном и Северо-Западном адм.округе.
- на 2-м месте кофейни Северного и Юго-Западный адм.округов.
- на 3-м месте кофейни Южного и Юго-Восточного адм.округов.
- на 4-м месте кофейни Северо-Восточного и Западного адм. округов.

Средняя стоимость чашки кофе

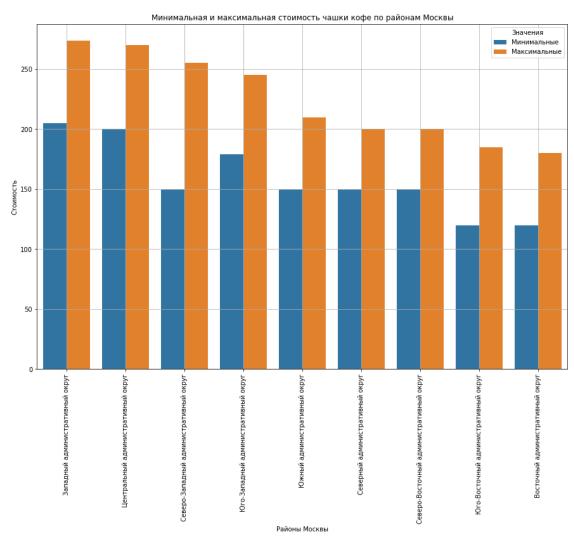
Для того, чтобы рассчитать рекомендуемую стоимость чашки кофе, сначала выделим из столбца "avg_bill" указанный диапазон значений в отдельные столбцы:

	name	category	address	district	hours	lat	Ing	rating	price	avg_bill	middle_avg_bill	middle_coffee_cup	chain	seats	st
3	Dormouse Coffee Shop	кофейня	Москва, улица Маршала Федоренко, 12	Северный административный округ	ежедневно, 09:00-22:00	55.881608	37.488860	5.0	не указано	Цена чашки капучино:155– 185 ₽	NaN	170.0	0	ИвИ	у/ Мари Федорі
58	Coffeekaldi's	кофейня	Москва, Угличская улица, 13, стр. 8	Северо-Восточный административный округ	ежедневно, 09:00-22:00	55.900316	37.570558	4.1	средние	Средний счёт:500—800₽	650.0	NaN	1	NaN	Углич ул
72	Лакрица	кофейня		Северо-Восточный административный округ	ежедневно, 08:00-21:00	55.891196	37.585653	4.5	средние	Цена чашки капучино:150– 190₽	NaN	170.0	0	12.0	Алтуфьев ш
$+\parallel$															-
di	# измения тип данных у столбцов min_bill и max_bill df_3_filtr['min_bill'] = df_3_filtr['min_bill'].astype('int') df_3_filtr['max_bill'] = df_3_filtr['max_bill'].astype('int')														
pı	# рассчитаем средние значения границ диапазона винимальной и максимальной стоимости чашки кофе по всем данным: print('Среднее минимальное значение стоимости чашки кофе, руб:', round(df_3_filtr['min_bill'].mean())) print('Среднее максимальное значение стоимости чашки кофе, руб:', round(df_3_filtr['max_bill'].mean()))														

Среднее минимальное значение стоимости чашки кофе, руб: 245 Среднее максимальное значение стоимости чашки кофе, руб: 349

Рассчитаем минимальное и максимальное значение стоимости чашки кофе для каждого района Москвы, в которых есть кофейни, по которым мы имеем данные о дипазоне стоимости чашки кофе:

	district	max_bill	min_bill
(Эападный административный округ	274	205
1	I Центральный административный округ	270	200
2	 Северо-Западный административный округ 	255	150
3	В Юго-Западный административный округ	245	179
2	4 Южный административный округ	210	150
	Б Северный административный округ	200	150
6	 Северо-Восточный административный округ 	200	150
7	 Юго-Восточный административный округ 	185	120
8	Восточный административный округ	180	120



По резульатам рассчётов по медиане, видим, что по районам:

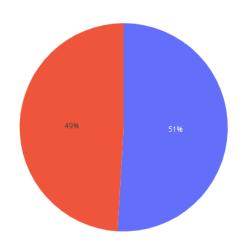
- минимальная стоимость чашки находится в диапазоне 120-205 руб,
- максимальная стоимость чашки кофе колеблется в перделах 180-274 руб,

что значительно отличается в меньшую сторону от усреднённых данных по всему датасету:

- среднее минимальное значение чашки кофе, руб: 245
- среднее максимальное значение чашки кофе, руб: 349

Возможность развития сети

L]:		chain	number
	0	1	720
	1	0	693



■ да ■ нет Из рассчётов видим, что чуть больше половины кофеен являются сетевыми заведениями.

Рекомендации по открытию кофейни

Т.к. основателям фонда «Shut Up and Take My Money» не даёт покоя успех сериала «Друзья», то для того, чтобы соблюсти дух сериала, рекомендуется открыть кофейню в Центральный административный округ, где находится значительная часть Московских кофеен, что говорит о том, что в этом районе этот тип заведений пользуется большой популярностью.

Кафейня может быть круглосуточной, учитывая её тематику и планируемый район размещения - в любое время суток центральная часть Москвы оживлена.

Минимальная стоимость чашки кофе рекомендована в дот 200 руб, исходня из данных по Центральному адм. округу.

Шаг 5. Презентация

Ссылка на презентацию: https://cloud.mail.ru/public/QGuX/zzMGDsgfJ