

# Рынок заведений общественного питания Москвы

## Описание проекта

Вы решили открыть небольшое кафе в Москве. Оно оригинальное — гостей должны обслуживать роботы. Проект многообещающий, но дорогой. Вместе с партнёрами вы решились обратиться к инвесторам. Их интересует текущее положение дел на рынке — сможете ли вы снискать популярность на долгое время, когда все зеваки посмотрят на роботов-официантов? Вы — гуру аналитики, и партнёры просят вас подготовить исследование рынка. У вас есть открытые данные о заведениях общественного питания в Москве.

## Инструкция по выполнению проекта

Загрузите данные и подготовьте их к анализу

- 1. [Загрузите данные о заведениях общественного питания Москвы. Убедитесь, что тип данных в каждой колонке — правильный, а также отсутствуют пропущенные значения и дубликаты. При необходимости обработайте их.](#)

Анализ данных

- 1. [Исследуйте соотношение видов объектов общественного питания по количеству. Постройте график.](#)
- 2. [Исследуйте соотношение сетевых и не сетевых заведений по количеству. Постройте график.](#)
- 3. [Для какого вида объекта общественного питания характерно сетевое распространение?](#)
- 4. [Что характерно для сетевых заведений: много заведений с небольшим числом посадочных мест в каждом или мало заведений с большим количеством посадочных мест?](#)
- 5. [Для каждого вида объекта общественного питания опишите среднее количество посадочных мест. Какой вид предоставляет в среднем самое большое количество посадочных мест? Постройте графики.](#)
- 6. [Выделите в отдельный столбец информацию об улице из столбца address.](#)
- 7. [Постройте график топ-10 улиц по количеству объектов общественного питания. Воспользуйтесь внешней информацией и ответьте на вопрос — в каких районах Москвы находятся эти улицы?](#)
- 8. [Найдите число улиц с одним объектом общественного питания. Воспользуйтесь внешней информацией и ответьте на вопрос — в каких районах Москвы находятся эти улицы?](#)
- 9. [Посмотрите на распределение количества посадочных мест для улиц с большим количеством объектов общественного питания. Какие закономерности можно выявить?](#)

Сделайте общий вывод и дайте рекомендации о виде заведения, количестве посадочных мест, а также районе расположения. Прокомментируйте возможность развития сети.

- 1. Подготовка презентации Подготовьте презентацию исследования для инвесторов. Для создания презентации используйте любой удобный инструмент, но отправить презентацию нужно обязательно в формате pdf. Приложите ссылку на презентацию в markdown-ячейке в формате:

Презентация: <ссылка на облачное хранилище с презентацией>

Следуйте принципам оформления из темы «Подготовка презентации». Оформление: Задание выполните в Jupyter Notebook. Программный код заполните в ячейках типа code, текстовые пояснения — в ячейках типа markdown. Примените форматирование и заголовки.

Описание данных

- Таблица rest\_data:
- id — идентификатор объекта;
- object\_name — название объекта общественного питания;
- chain — сетевой ресторан;
- object\_type — тип объекта общественного питания;
- address — адрес;
- number — количество посадочных мест.

Комментарий ревьюера v. 1: Молодец, что кратко описал проект и добавил содержание со ссылками.

**Загрузите данные о заведениях общественного питания Москвы. Убедитесь, что тип данных в каждой колонке — правильный, а также отсутствуют пропущенные значения и дубликаты. При необходимости обработайте их.**

</div>

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from scipy import stats
from numpy import median
import seaborn as sb
from plotly import graph_objects as go
import plotly.express as px
import plotly.io as pio
pio.renderers.default='notebook'
```

In [2]:

```
data = pd.read_csv('/datasets/rest_data.csv')
```

In [3]:

```
data.info()
```

<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 15366 entries, 0 to 15365  
Data columns (total 6 columns):  
id 15366 non-null int64  
object\_name 15366 non-null object  
chain 15366 non-null object  
object\_type 15366 non-null object  
address 15366 non-null object  
number 15366 non-null int64  
dtypes: int64(2), object(4)  
memory usage: 720.4+ KB

In [4]:

```
data.describe()
```

Out[4]:

	id	number
count	15366.000000	15366.000000
mean	119720.066901	59.547182
std	73036.130732	74.736833
min	838.000000	0.000000
25%	28524.000000	12.000000
50%	144974.500000	40.000000
75%	184262.250000	80.000000
max	223439.000000	1700.000000

In [5]:

```
data.query('number==1700')
```

Out[5]:

	id	object_name	chain	object_type	address	number
12723	199696	Arena by Soho Family	нет	бар	город Москва, Кутузовский проспект, дом 36, ст...	1700

<https://sova.ponominalu.ru/industry/sovety-i-pravila/arena-by-soho-family/> сайтутверждает что вместимость 6-6,5 к зрителей, так что сделаем вид что все ок

In [6]:

```
data.duplicated().sum()
```

Out[6]:

0

In [7]:

```
columns=['object_name', 'chain', 'object_type', 'address']
for column in columns:
    data[column] = data[column].str.lower()
```

In [8]:

```
data.head()
```

Out[8]:

	id	object_name	chain	object_type	address	number
0	151635	сметана	нет	кафе	город москва, улица егора абакумова, дом 9	48
1	77874	родник	нет	кафе	город москва, улица талалихина, дом 2/1, корпус 1	35
2	24309	кафе «академия»	нет	кафе	город москва, абельмановская улица, дом 6	95
3	21894	пиццетория	да	кафе	город москва, абрамцевская улица, дом 1	40
4	119365	кафе «вишневая метель»	нет	кафе	город москва, абрамцевская улица, дом 9, корпус 1	50

In [9]:

```
def logical(x):
    if x=='да':
        return True
    elif x == 'нет':
        return False
data['chain'] = data['chain'].apply(logical)
```

In [10]:

```
data.head()
```

Out[10]:

	id	object_name	chain	object_type	address	number
0	151635	сметана	False	кафе	город москва, улица егора абакумова, дом 9	48
1	77874	родник	False	кафе	город москва, улица талалихина, дом 2/1, корпус 1	35
2	24309	кафе «академия»	False	кафе	город москва, абельмановская улица, дом 6	95
3	21894	пиццетория	True	кафе	город москва, абрамцевская улица, дом 1	40
4	119365	кафе «вишневая метель»	False	кафе	город москва, абрамцевская улица, дом 9, корпус 1	50

посмотрели на данные (не обнаружили пропусков и других отклонений), заменили chain на более привычный вид

Исследуйте соотношение видов объектов общественного питания по количеству. Постройте график.

</div>

In [11]:

```
fig = go.Figure(data=[go.Pie(labels=data['object_type'].value_counts().reset_index()['index'], values=data['object_type'].value_counts())])
fig.update_layout(
    title={
        'text':'Соотношение видов объектов общественного питания по количеству',
        'y':0.9,
        'x':0.38})
fig.show()
```

In [12]:

```
object_types = data.groupby('object_type', as_index=False).agg({'object_name': 'count'}).sort_values(by='object_name', ascending=False)
object_types['percent'] = (object_types['object_name']/object_types['object_name'].sum()*100).round(1).astype('str') + '%'
fig = px.bar(object_types, y='object_name', x='object_type', text='object_name', color='object_type', width=1000, height=500)
fig.update_xaxes(tickangle=45)
fig.update_layout(
    title='Соотношение видов объектов общественного питания по количеству',
    xaxis_title="Тип объекта",
    yaxis_title="Количество")
for trace, percent in zip(fig.data, object_types['percent']) :
    trace.name = trace.name.split('=')[1] + ' (' + percent + '%)'
fig.show()
```

Как мы видим лидерером являются кафе, затем столовые, потом рестораны и предприятия быстрого обслуживания, замыкают список закусочные и магазины с отделом кулинарии

Комментарий ревьюера v. 1: Отличные графики :) Подумай, пожалуйста, с чем может быть связано преобладание кафе?

думаю что все очень просто - открытие кафе сделать более просто например по сравнению с рестораном (т.к. у ресторан должен

обеспечить полную термическую обработку блюда с начала готовки до конечного результата, т.е. недопускаются полуфабрикаты, а также нельзя ворваться "с одним шашлыком в меню" в ресторанный бизнес), а по примеру со столовой - то у населения более распространено кафе

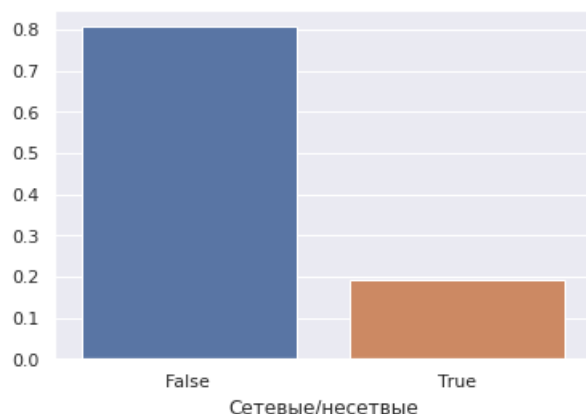
Комментарий ревьюера v. 2: Да, всё верно :)

**Исследуйте соотношение сетевых и несетевых заведений по количеству. Постройте график.**

</div>

In [13]:

```
sb.set()
chains = data.groupby('chain')['id'].count().reset_index().rename(columns={'id':'counts'})
chains['ratio'] = chains['counts'] / chains['counts'].sum()
sb.barplot(x='chain',y='ratio', data=chains)
plt.xlabel('Сетевые/несетевые')
plt.ylabel("")
plt.show()
```



достаточно много несетевых заведений - 80,7%

Комментарий ревьюера v. 1: Как Ты думаешь почему преобладают несетевые заведения?

думаю что не у всех хватает ресурсов содержать сеть - денег, времени и т.д.

Комментарий ревьюера v. 2: Ты верно ответил, что возможно не хватает ресурсов - для открытия сети их требуется намного больше.

**Для какого вида объекта общественного питания характерно сетевое распространение?**

</div>

In [14]:

```
object_chain = (data
    .groupby(['object_type','chain'])['id'].count()
    .reset_index()
    .rename(columns={'id':'counts'})
    .merge(data.groupby('object_type').chain.count(), on='object_type',suffixes=('_', '_total'))
)
```

In [15]:

```
object_chain['ratio_percent'] = round((object_chain.counts / object_chain.chain_total) * 100)
object_chain
```

Out[15]:

object_type	chain	counts	chain_total	ratio_percent
-------------	-------	--------	-------------	---------------

	object_type	chain	counts	chain_total	ratio_percent
0	бар	False	819	856	96.0
1	бар	True	37	856	4.0
2	буфет	False	574	585	98.0
3	буфет	True	11	585	2.0
4	закусочная	False	304	360	84.0
5	закусочная	True	56	360	16.0
6	кафе	False	4703	6099	77.0
7	кафе	True	1396	6099	23.0
8	кафетерий	False	346	398	87.0
9	кафетерий	True	52	398	13.0
10	магазин (отдел кулинарии)	False	195	273	71.0
11	магазин (отдел кулинарии)	True	78	273	29.0
12	предприятие быстрого обслуживания	False	1132	1923	59.0
13	предприятие быстрого обслуживания	True	791	1923	41.0
14	ресторан	False	1741	2285	76.0
15	ресторан	True	544	2285	24.0
16	столовая	False	2584	2587	100.0
17	столовая	True	3	2587	0.0

In [16]:

```
fig = px.bar(object_chain, x='object_type', y='counts', color='object_type')
fig.update_xaxes(tickangle=45)
fig.show()
```

Видно что предприятия быстрого обслуживания - лидеры, кафе на втором месте, вообще нет сетевого обслуживания у столовых

Комментарий ревьюера v. 1: Отличный анализ, а как Ты думаешь с чем связано отсутствие сетевых столовых?

Ахах) ответа на этот вопрос у меня нет - но не думаю что какая-то сложная причина в этом

Комментарий ревьюера v. 2: На самом деле причина в том, что столовые как правило существуют при организациях (заводах, театрах, университетах и т. д.) поэтому сетевыми быть не могут.

Что характерно для сетевых заведений: много заведений с небольшим числом посадочных мест в каждом или мало заведений с большим количеством посадочных мест?

</div>

In [17]:

```
chain = data.groupby('object_type').agg({'number':'mean','id':'count'}).reset_index().rename(columns={'id':'counts'})
chain['ratio'] = chain.number / chain.counts
chain = chain.sort_values(by='ratio', ascending=False)
fig = px.bar(chain.query('object_type in ["кафе","ресторан","предприятие быстрого обслуживания", "магазин (отдел кулинарии)", "столовая", "бар"]'), x='object_type', y='ratio',color='object_type', title='Отношение среднего числа мест к количеству заведений')
fig.update_xaxes(title="")
fig.update_layout(
    autosize=False,
    width=700,
    height=400)
fig.show()
```

In [18]:

```
chain.query('object_type in ["кафе","ресторан","предприятие быстрого обслуживания", "магазин (отдел кулинарии)", "столовая", "бар"]')
```

Out[18]:

	object_type	number	counts	ratio
0	бар	43.494159	856	0.050811
8	столовая	130.246618	2587	0.050347
7	ресторан	96.834136	2285	0.042378
5	магазин (отдел кулинарии)	5.589744	273	0.020475
6	предприятие быстрого обслуживания	20.552262	1923	0.010688
3	кафе	39.716019	6099	0.006512

бары и столовые с ресторанами берут большим количеством мест, а остальные заведения берут количеством точек

Для каждого вида объекта общественного питания опишите среднее количество посадочных мест. Какой вид предоставляет в среднем самое большое количество посадочных мест? Постройте графики.

</div>

In [19]:

```
fig = px.bar(chain.sort_values(by='number'), x='object_type', y='number', color='object_type', title='Среднее число посадочных мест по типам заведений')
fig.update_yaxes(title='среднее число мест')
fig.show()
```

столовые и рестораны рассчитаны на большее количество мест

Комментарий ревьюера v. 1: Проанализируй, пожалуйста, среднее число посадочных мест более детально.

**Выделите в отдельный столбец информацию об улице из столбца address.**

</div>

In [20]:

```
streets = data['address'].to_list()
```

In [21]:

```
streets_only = []
for street in streets:
    start = street.find('город москва')
    end = street.find('дом')
    streets_only.append(street[start+14:end])
```

In [22]:

```
data['street_name'] = streets_only
```

In [23]:

```
data
```

Out[23]:



id		object_name	chain	object_type	address	number	street_name
0	151635	сметана	False	кафе	город москва, улица егора абакумова, дом 9	48	улица егора абакумова,
1	77874	родник	False	кафе	город москва, улица талалихина, дом 2/1, корпус 1	35	улица талалихина,
2	24309	кафе «академия»	False	кафе	город москва, абельмановская улица, дом 6	95	абельмановская улица,
3	21894	пиццетория	True	кафе	город москва, абрамцевская улица, дом 1	40	абрамцевская улица,
4	119365	кафе «вишневая метель»	False	кафе	город москва, абрамцевская улица, дом 9, корпус 1	50	абрамцевская улица,
...	...	...	...	...	...	...	...
15361	208537	шоколадница	True	кафе	город москва, 3-й крутицкий переулок, дом 18	50	3-й крутицкий переулок,
15362	209264	шоколадница	True	кафе	город москва, улица земляной вал, дом 33	10	улица земляной вал,
15363	209186	шоколадница	True	кафе	город москва, улица земляной вал, дом 33	20	улица земляной вал,
15364	221900	шоколадница	True	кафе	город москва, поселение московский, киевское ш...	36	поселение московский, киевское шоссе, 22-й кил...
15365	222535	шоколадница	True	кафе	город москва, ходынский бульвар, дом 4	10	ходынский бульвар,

15366 rows × 7 columns

Постройте график топ-10 улиц по количеству объектов общественного питания. Воспользуйтесь внешней информацией и ответьте на вопрос — в каких районах Москвы находятся эти улицы?

</div>

In [24]:

```
import os
os.listdir()
cwd = os.getcwd()
cwd
```

Out[24]:

'/home/user-0-818729446/work'

In [25]:

```
pip install -U pip setuptools
```

Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: pip in /home/jovyan/.local/lib/python3.7/site-packages (21.0.1)
Requirement already satisfied: setuptools in /home/jovyan/.local/lib/python3.7/site-packages (53.0.0)
Note: you may need to restart the kernel to use updated packages.

ссылку на файл <https://yadi.sk/i/fw0h1mS4Wdl28w>

Комментарий ревьюера v. 3: Ок :)

In [26]:

```
file = 'moscow.xlsx'
xl = pd.ExcelFile(file)
print(xl.sheet_names)
df1 = xl.parse('Sheet1')
```

[Sheet1]

In [27]:

```
#df1 = pd.read_excel('moscow.xlsx')
```

In [28]:

```
In [29]: df1 = pd.read_csv('/home/user-0-818729446/work/moscow.csv')
```

таблица взята с сайта data.mos.ru

In [29]:

df1

Out[29]:

Unnamed: 0		address	adm_area	district	street_name
0	0	город москва, варшавское шоссе, дом 150, строе...	южный административный округ	муниципальный округ чертаново южное	варшавское шоссе
1	1	город москва, косинская улица, дом 26а	восточный административный округ	муниципальный округ вешняки	косинская улица
2	2	город москва, гороховский переулок, дом 21	центральный административный округ	муниципальный округ басманный	гороховский переулок
3	3	город москва, тагильская улица, дом 4, строени...	восточный административный округ	муниципальный округ метрогородок	тагильская улица
4	4	город москва, ленинградский проспект, дом 73а,...	северный административный округ	муниципальный округ сокол	ленинградский проспект
...	...	...	...	...	...
370773	370773	город москва, поселение новофедоровское, дерев...	троицкий административный округ	поселение новофедоровское	изумрудная улица
370774	370774	город москва, поселение щаповское, 1-й дунайск...	троицкий административный округ	поселение щаповское	1-й дунайский переулок
370775	370775	город москва, город троицк, индустриальная ули...	троицкий административный округ	городской округ троицк	индустриальная улица
370776	370776	город москва, город троицк, индустриальная ули...	троицкий административный округ	городской округ троицк	индустриальная улица
370777	370777	город москва, город троицк, песчаная улица, до...	троицкий административный округ	городской округ троицк	песчаная улица

370778 rows × 5 columns

удалим что не надо

In [30]:

```
df1 = df1.drop(['Unnamed: 0', 'street_name'], axis=1)
df1
```

Out[30]:

	address	adm_area	district
0	город москва, варшавское шоссе, дом 150, строе...	южный административный округ	муниципальный округ чертаново южное
1	город москва, косинская улица, дом 26а	восточный административный округ	муниципальный округ вешняки
2	город москва, гороховский переулок, дом 21	центральный административный округ	муниципальный округ басманный
3	город москва, тагильская улица, дом 4, строени...	восточный административный округ	муниципальный округ метрогородок
4	город москва, ленинградский проспект, дом 73а,...	северный административный округ	муниципальный округ сокол
...	...	...	...
370773	город москва, поселение новофедоровское, дерев...	троицкий административный округ	поселение новофедоровское
370774	город москва, поселение щаповское, 1-й дунайск...	троицкий административный округ	поселение щаповское
370775	город москва, город троицк, индустриальная ули...	троицкий административный округ	городской округ троицк
370776	город москва, город троицк, индустриальная ули...	троицкий административный округ	городской округ троицк
370777	город москва, город троицк, песчаная улица, до...	троицкий административный округ	городской округ троицк

370778 rows × 3 columns

In [31]:

data

Out[31]:

	id	object_name	chain	object_type	address	number	street_name
0	151635	сметана	False	кафе	город москва, улица егора абакумова, дом 9	48	улица егора абакумова,
1	77874	родник	False	кафе	город москва, улица талалихина, дом 2/1, корпус 1	35	улица талалихина,
2	24309	кафе «академия»	False	кафе	город москва, абельмановская улица, дом 6	95	абельмановская улица,
3	21894	пиццетория	True	кафе	город москва, абрамцевская улица, дом 1	40	абрамцевская улица,
4	119365	кафе «вишневая метель»	False	кафе	город москва, абрамцевская улица, дом 9, корпус 1	50	абрамцевская улица,
...	...	...	...	...	...	...	...
15361	208537	шоколадница	True	кафе	город москва, 3-й крутицкий переулок, дом 18	50	3-й крутицкий переулок,
15362	209264	шоколадница	True	кафе	город москва, улица земляной вал, дом 33	10	улица земляной вал,
15363	209186	шоколадница	True	кафе	город москва, улица земляной вал, дом 33	20	улица земляной вал,
15364	221900	шоколадница	True	кафе	город москва, поселение московский, киевское ш...	36	поселение московский, киевское шоссе, 22-й кил...
15365	222535	шоколадница	True	кафе	город москва, ходынский бульвар, дом 4	10	ходынский бульвар,

15366 rows × 7 columns

In [32]:

```
data = data.merge(df1, on='address', how='inner')
data
```

Out[32]:

	id	object_name	chain	object_type	address	number	street_name	adm_area	district
0	151635	сметана	False	кафе	город москва, улица егора абакумова, дом 9	48	улица егора абакумова,	северо-восточный административный округ	муниципальный округ ярославский
1	77874	родник	False	кафе	город москва, улица талалихина, дом 2/1, корпус 1	35	улица талалихина,	центральный административный округ	муниципальный округ таганский
2	24309	кафе «академия»	False	кафе	город москва, абельмановская улица, дом 6	95	абельмановская улица,	центральный административный округ	муниципальный округ таганский
3	21894	пиццетория	True	кафе	город москва, абрамцевская улица, дом 1	40	абрамцевская улица,	северо-восточный административный округ	муниципальный округ лианозово
4	119365	кафе «вишневая метель»	False	кафе	город москва, абрамцевская улица, дом 9, корпус 1	50	абрамцевская улица,	северо-восточный административный округ	муниципальный округ лианозово
...	...	...	...	...	...	...	...	...	...
15269	222491	кальянная «мята lounge»	True	кафе	город москва, профсоюзная улица, дом 142, корп...	40	профсоюзная улица,	юго-западный административный округ	муниципальный округ теплый стан
15270	212216	мята lounge	True	кафе	город москва, привольная улица, дом 11	56	привольная улица,	юго-восточный административный округ	муниципальный округ выхино-жулебино
15271	206341	мята lounge	True	кафе	город москва, салтыковская улица, дом 7г	100	салтыковская улица,	восточный административный округ	муниципальный округ новокосино
15272	213772	тануки	True	ресторан	город москва, осенний бульвар, дом 9	98	осенний бульвар,	западный административный округ	муниципальный округ крылатское
15273	74972	шоколадница	True	кафе	город москва, улица новый арбат, дом 13	30	улица новый арбат,	центральный административный округ	муниципальный округ арбат

15274 rows × 9 columns

теперь перейдем непосредственно к заданию, т.к. у нас появились районы в таблице

In [33]:

```
top_objects_count = (data
```

```

top_objects_count = data
    .groupby('street_name', as_index=False)
    .agg({'id': 'count', 'district': 'max'})
    .sort_values(by='id', ascending=False)
    .rename(columns={'id': 'id_count_top'})
)

```

In [34]:

```

temp = data.groupby('street_name').count().sort_values(by='object_type', ascending=False).head(10).reset_index()
top_10 = list(temp['street_name'])

```

In [35]:

```

ax = sns.barplot(x='object_type', y='street_name', data=temp, palette="Set3")
ax.set_xlabel('Количество объектов')
ax.set_ylabel('Название улицы')
ax.set_title('Топ-10 улиц по количеству объектов общественного питания')
plt.show()

```



In [36]:

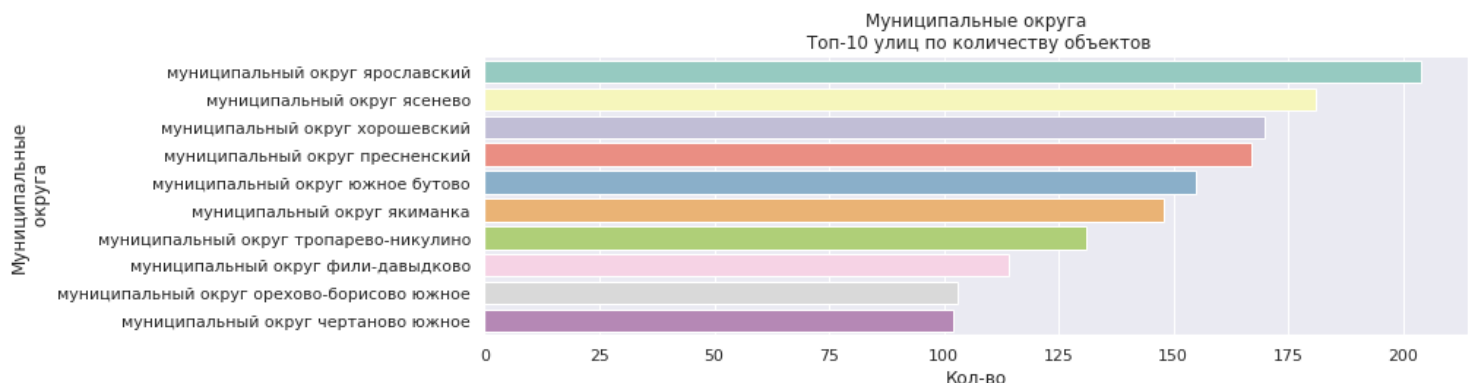
```

ax=sns.catplot(x='id_count_top', y='district',
    kind='bar', data=top_objects_count.head(10),
    palette='Set3',
    height=3.5, aspect=4)
(ax
    .set_xticklabels(rotation=0)
    .set(title = 'Муниципальные округа\n Топ-10 улиц по количеству объектов')
    .set(ylabel = 'Муниципальные\n округа')
    .set(xlabel = 'Кол-во')
)

```

Out[36]:

<seaborn.axisgrid.FacetGrid at 0x7f833f322350>



южное орехово-борисово и южное чертаново имеет самую низкую конкуренцию в общепите, нужно дополнительное исследование чем это обусловлено, можно сделать 2 гипотезы:

- H0 - районы более привлекательны "для освоения общепита"
- H1 - районы имеют более низкую платежеспособность, не выгодно с точки зрения комерциализации, др. проблемы

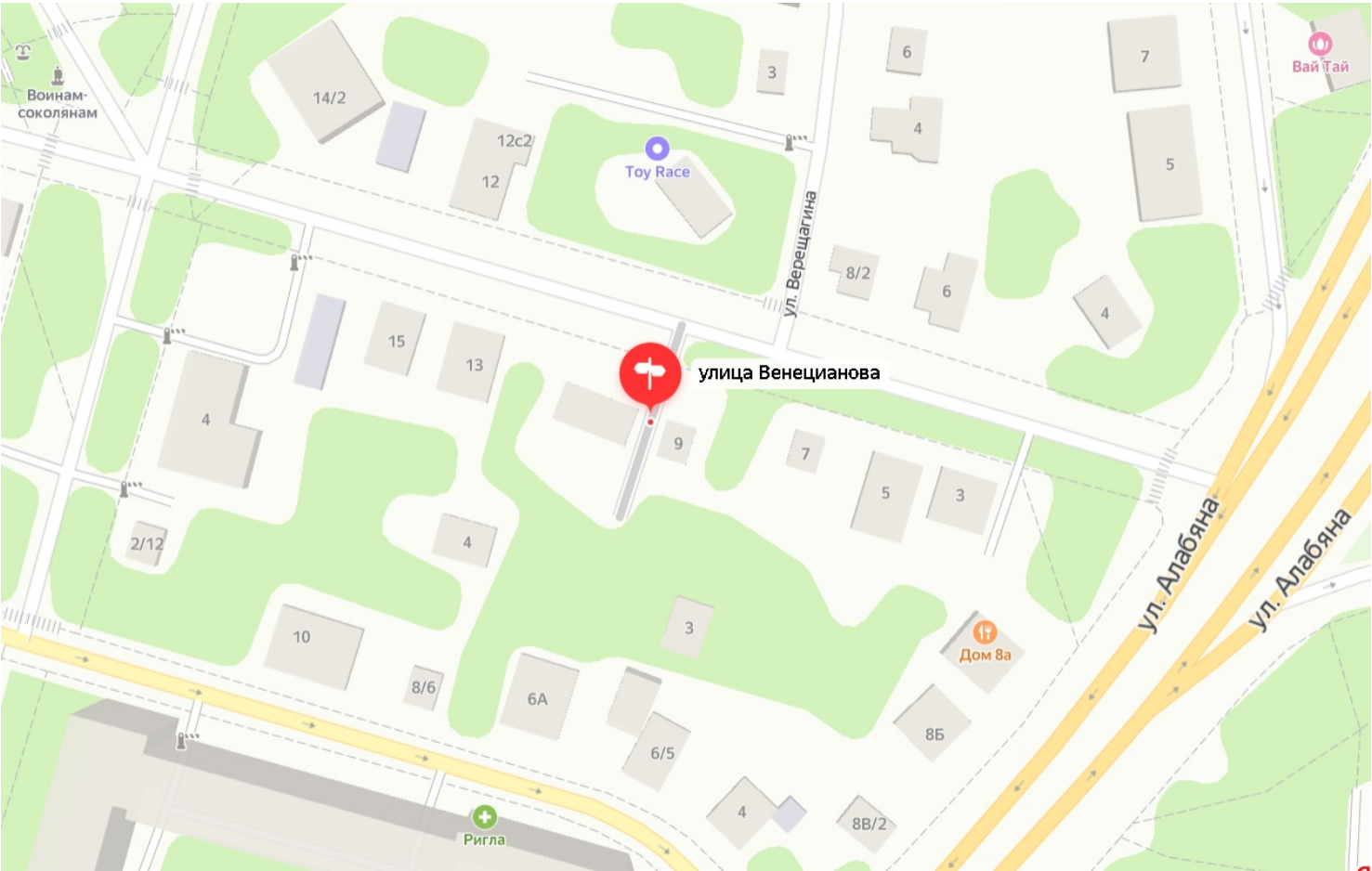
```
one_object = top_objects_count.query('id_count_top==1').count()[1]
one_object
```

802

802 улицы с 1 объектом

Комментарий ревьюера v. 1: Как Ты думаешь с чем связано наличие только одного заведения на улице?

возможно улица слишком "маленькая"



что-то в этом духе, протяженность, если верить открытым источникам, 48 метров

Комментарий ревьюера v. 2: Да, Ты прав, в большинстве случаев это связано с длиной этих улиц. Кстати, если Тебе интересно, то вот самые короткие улицы в мире: <https://www.smithsonianmag.com/travel/top-ten-small-streets-worth-visiting-180954589/>

```
one_object = top_objects_count.query('id_count_top==1')
```

one\_object

Out[39]:

	street_name	id_count_top	district
1732	улица брусилова,	1	муниципальный округ южное бутово
2044	улица рословка,	1	муниципальный округ митино
17	1-й кожевнический переулок,	1	муниципальный округ даниловский
18	1-й кожуховский проезд,	1	муниципальный округ даниловский
218	6-й лучевой просек,	1	муниципальный округ сокольники
...	...	...	...
1394	производственная улица,	1	муниципальный округ солнцево
1395	прокатная улица,	1	муниципальный округ ивановское
469	воротниковский переулок,	1	муниципальный округ тверской
1398	пролетарский проспект, владение 26	1	муниципальный округ царицыно
2252	яхромская улица,	1	муниципальный округ дмитровский

802 rows × 3 columns

In [40]:

```
not_top_streets = one_object.groupby('street_name').agg({'id_count_top':'count', 'district':'max'}).reset_index().sort_values(by='id_count_top')
not_top_streets = not_top_streets[(not_top_streets['id_count_top']==1) & (not_top_streets['district'].isna()==False)]
not_top_streets
```

Out[40]:

	street_name	id_count_top	district
0	1-й балтийский переулок,	1	муниципальный округ аэропорт
528	поселение щаповское, деревня овечкино, владение	1	поселение щаповское
529	поселение щаповское, деревня шаганино,	1	поселение щаповское
530	поселение щаповское, посёлок курилово, школьна...	1	поселение щаповское
531	поселение щаповское, посёлок щапово, улица сад...	1	поселение щаповское
...	...	...	...
271	город щербинка, улица авиаторов,	1	городской округ щербинка
272	город щербинка, улица барышевская роща, владен...	1	городской округ щербинка
273	городская улица,	1	муниципальный округ даниловский
263	город троицк, пионерская улица,	1	городской округ троицк
801	яхромская улица,	1	муниципальный округ дмитровский

802 rows × 3 columns

In [41]:

```
districts_one_place = not_top_streets.groupby('district')['id_count_top'].count().sort_values(ascending=False)
districts_one_place.head(15)
```

Out[41]:

district	
муниципальный округ таганский	29
муниципальный округ хамовники	27
муниципальный округ басманный	24
муниципальный округ пресненский	21
муниципальный округ тверской	20
муниципальный округ марьяна роща	18
муниципальный округ сокольники	15
муниципальный округ крюково	14
муниципальный округ мещанский	14
муниципальное образование савелки	13
муниципальный округ матушкино	13
муниципальный округ южное бутово	13
муниципальный округ замоскворечье	13
муниципальный округ даниловский	12
муниципальный округ красносельский	12
Name: id_count_top, dtype: int64	

в Таганском районе больше всего улиц всего с 1м заведением общепита, Хамовники и Басманный район на 2ом и Зьем месте соответственно, возможно это связано с большим наличием "маленьких" улиц

Посмотрите на распределение количества посадочных мест для улиц с большим количеством объектов общественного питания. Какие закономерности можно выявить?

</div>

In [42]:

```
top_10_objects = top_objects_count.head(10)
top_10_objects = top_10_objects.drop(['district', 'id_count_top'], axis=1)
```

In [43]:

```
top_10_objects
```

Out[43]:

	street_name
1412	проспект мира,
1416	профсоюзная улица,
879	ленинградский проспект,
1366	пресненская набережная,
432	варшавское шоссе,
883	ленинский проспект,
1408	проспект вернадского,
870	кутузовский проспект,
787	каширское шоссе,
797	кировоградская улица,

In [44]:

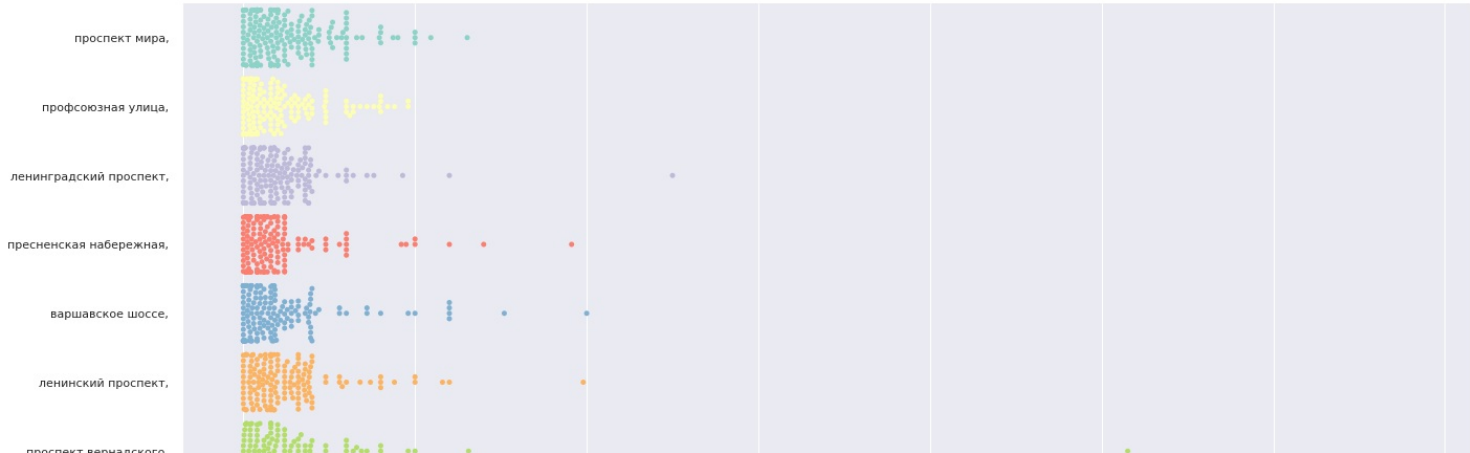
```
top_10_streets = top_10_objects.merge(data, on='street_name', how='left')
```

In [45]:

```
ax=sns.catplot(y='street_name', x='number',
               kind='swarm', data=top_10_streets,
               palette='Set3',
               height=10, aspect=2)
(ax
 .set_xticklabels(rotation=45)
 .set(title = "")
 .set(ylabel = "")
 .set(xlabel = "")
 )
```

Out[45]:

<seaborn.axisgrid.FacetGrid at 0x7f831af883d0>





In [46]:

```
top_10_streets.query('number >=300').count().sum()
```

Out[46]:

216

216 объектов с посадочными местами более 300

In [47]:

```
top_10_streets.query('number >=200').count().sum()
```

Out[47]:

720

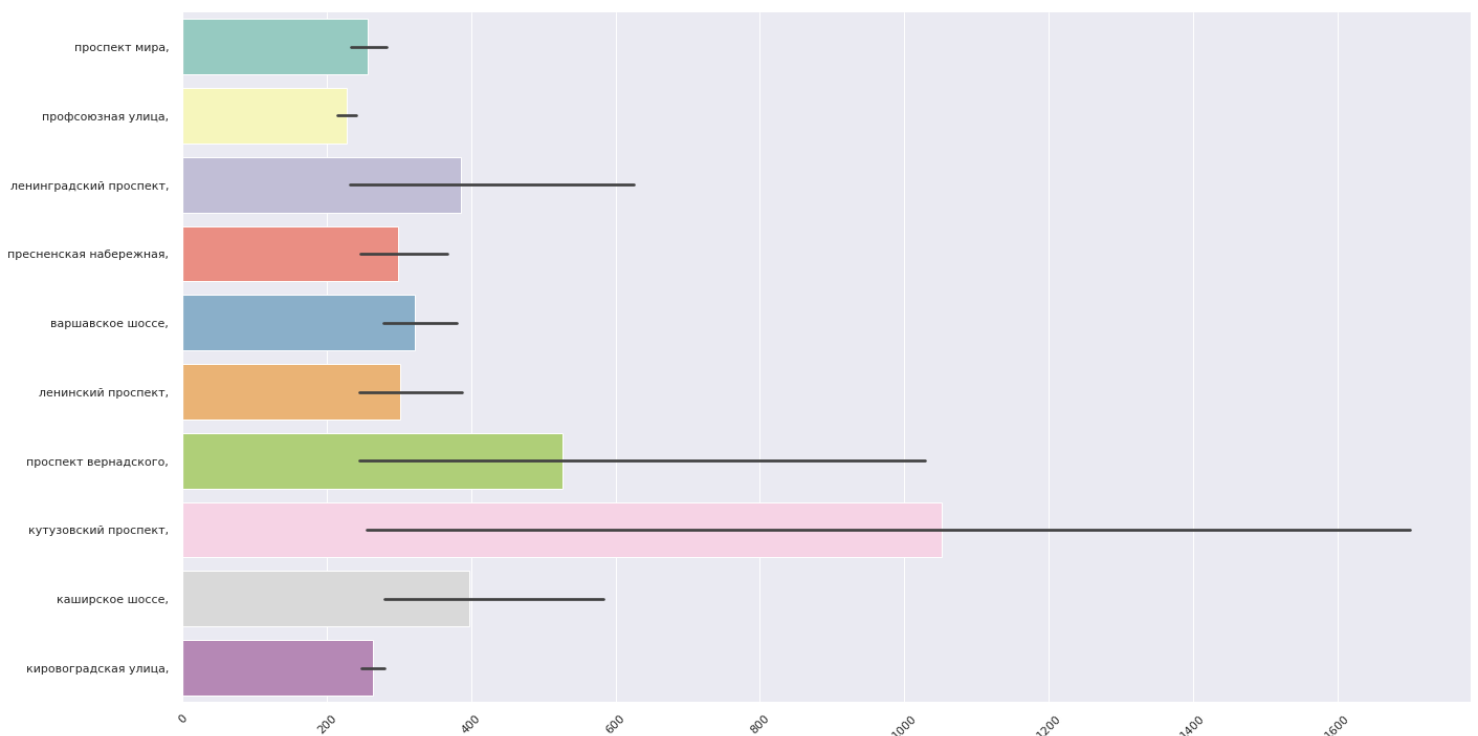
720 объектов с посадочными местами более 200

In [48]:

```
ax=sns.catplot(y='street_name', x='number',
               kind='bar', data=top_10_streets.query('number>200'),
               palette='Set3',
               height=10, aspect=2)
(ax
 .set_xticklabels(rotation=45)
 .set(title = "")
 .set(ylabel = "")
 .set(xlabel = "")
 )
```

Out[48]:

<seaborn.axisgrid.FacetGrid at 0x7f8339d49950>





- кутузовский проспект и проспект вернадского находятся около кремля, следовательно имеют огромный пешеходный трафик, тоже самое можно сказать о ленинградском проспекте
- каширское шоссе находится не так близко к кремлю, но по протяженности больше чем вышеперечисленные проспекты, возможно это выезд из города, поэтому более привлекателен для владельцев общепита

**Сделайте общий вывод и дайте рекомендации о виде заведения, количестве посадочных мест, а также районе расположения. Прокомментируйте возможность развития сети.**

Выводы:

- среди всех заведений общественного питания 40% кафе, 17% столовые, 15% рестораны, 12.5% предприятия быстрого обслуживания, 6% бары
- из всех заведений 80% являются несетевыми
- южное орехово-борисово и южное чертаново имеет самую низкую конкуренцию в общепите
- в Таганском районе, районе Хамовники и Басманном районе больше всего улиц всего с 1м заведением общепита

Рекомендации для бизнеса:

- если открывать точки общепита в привлекательных районах города надо учитывать большую конкуренцию
- провести дополнительное исследование гипотез для "малозаполненных" районов Москвы:
  - Н0 - районы более привлекательны "для освоения общепита"
  - Н1 - районы имеют более низкую платежеспособность, не выгодно с точки зрения коммерциализации, др. проблемы

Комментарий ревьюера v. 1: Отличные рекомендации :)

<https://docs.google.com/presentation/d/1WpUvPori5cE71pJ1SpUCIte5klBTnuLekEsv3X3TPHg/edit?usp=sharing> ссылка на презентацию

Комментарий ревьюера v. 2: Отличная презентация, но нужно добавить слайд с постановкой задачи/проблемы и перенести выводы в конец слайда. Кроме того, нужно добавить последний слайд "Спасибо за внимание!" или что-то похожее.

добавил

Комментарий ревьюера v. 2: Отлично :)

In []: