

## Практическая работа – «Облачные сервисы для анализа данных»

### Сценарий 1. Создание и публикация диаграммы из CSV-файла с картой Москвы

С помощью сервиса Yandex DataLens можно предоставлять публичный доступ к созданным чартам, диаграммам и дашбордам.

#### Необходимо помнить...

Доступ к опубликованным данным не ограничен!

Перед публикацией собственных данных убедитесь, что данные на чартах и дашбордах не являются персональными данными или коммерческой тайной!

В этом задании с помощью DataLens будем визуализировать показатели населения на карте районов Москвы и опубликуем их в DataLens Public.

В качестве исходных данных будет использоваться файл [Moscow Population 2018.csv](#) с информацией о населении и площади Москвы на 2018 год.

**Шаг 0.** Для визуализации и исследования данных подготовьте DataLens к работе и выполните следующие действия:

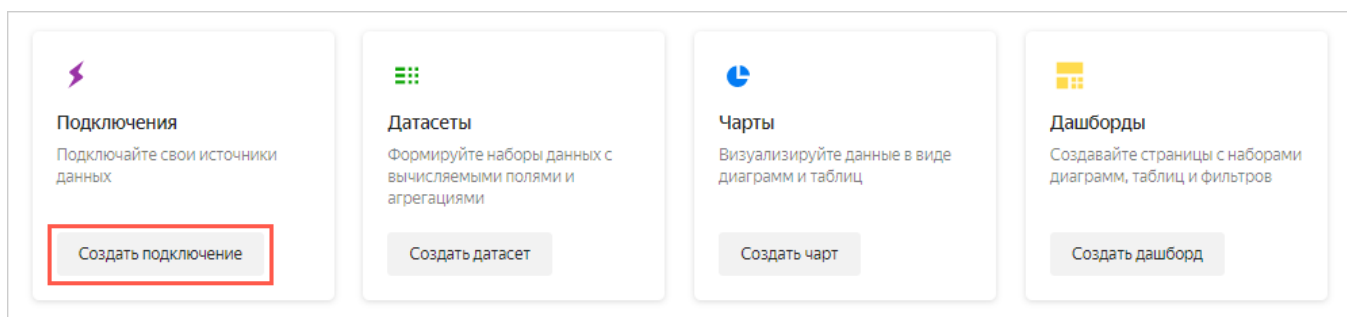
Чтобы начать работать с сервисом Яндекс.DataLens:

1. Войдите в ваш аккаунт на Яндексе. Если у вас еще нет аккаунта, создайте его.
2. Откройте главную страницу DataLens= <https://datalens.yandex.ru/>
3. Нажмите **Открыть DataLens** и затем нажмите **Войти**.

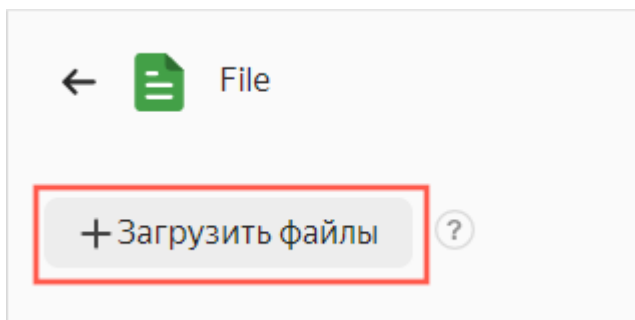
#### Шаг 1. Создайте подключение и датасет

Создайте датасет на основе подключения типа **File**.

1. Перейдите в интерфейс [DataLens](#).
2. Нажмите кнопку **Создать подключение**.



3. Выберите тип подключения **File**.
4. Нажмите кнопку **Загрузить файлы**, затем выберите необходимый файл. В данном примере **Moscow Population 2018.csv**. Загрузка может занять до нескольких минут в зависимости от скорости подключения к интернету.



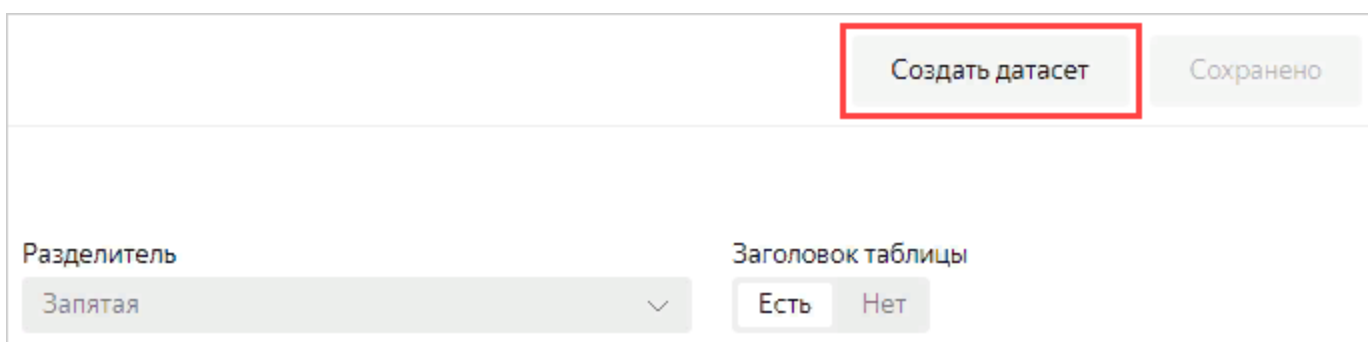
После успешной загрузки вы увидите начальный фрагмент данных из файла. Для просмотра можно использовать полосы прокрутки справа и внизу.

5. В правом верхнем углу нажмите кнопку **Создать подключение**.
6. Введите имя подключения «Moscow Population 2018» и нажмите кнопку **Создать**. Подключение к файлу создано. Данные будут представлены в том же виде, что и в файле.

## Шаг 2. Создайте датасет (набор данных для аналитики)

1. Для дальнейшей обработки данных нужно создать «датасет».

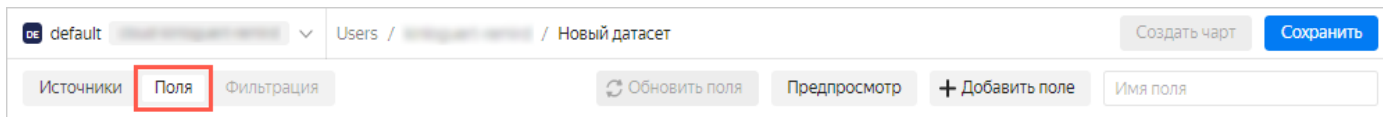
Нажмите кнопку **Создать датасет**.



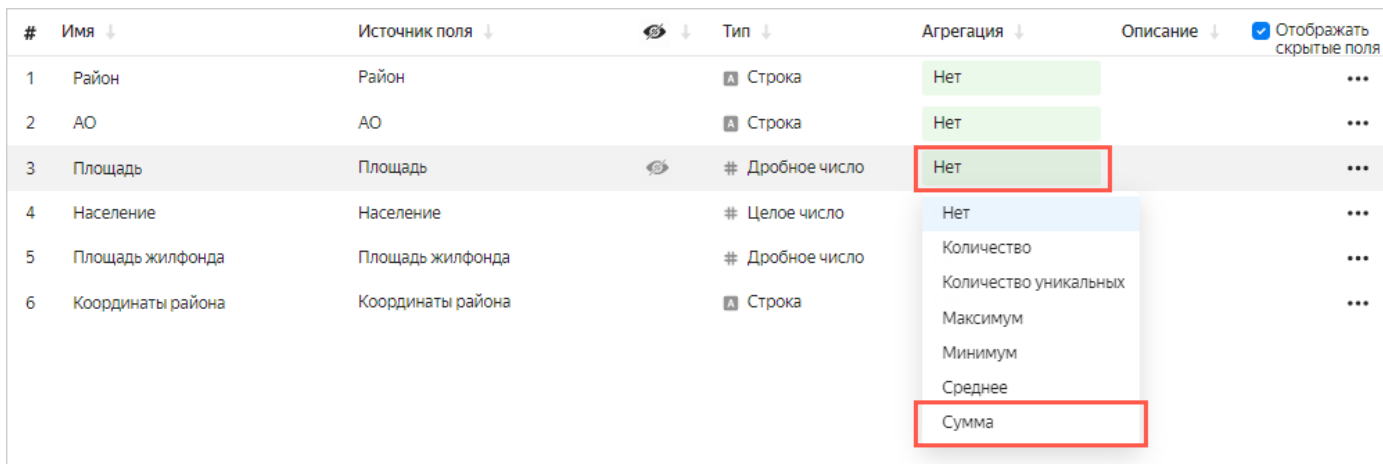
2. Если рабочая область пустая, то перетащите таблицу **Moscow Population 2018** из панели выбора на рабочую область.

#	Район	АО	Площадь	Население	Площадь жилфонда	Координаты района
1	Академический	ЮЗАО	5.83	109387	2467.0	[[[55.68403, 37.55226], [55.68408, 37.55232], [55.68428, 37.55256], [55.684
2	Алексеевский	СВАО	5.29	80534	1607.9	[[[55.79714, 37.63499], [55.79965, 37.63552], [55.80134, 37.63592], [55.804
3	Алтуфьевский	СВАО	3.25	57596	839.3	[[[55.8873, 37.5674], [55.88724, 37.57245], [55.88716, 37.57635], [55.88725

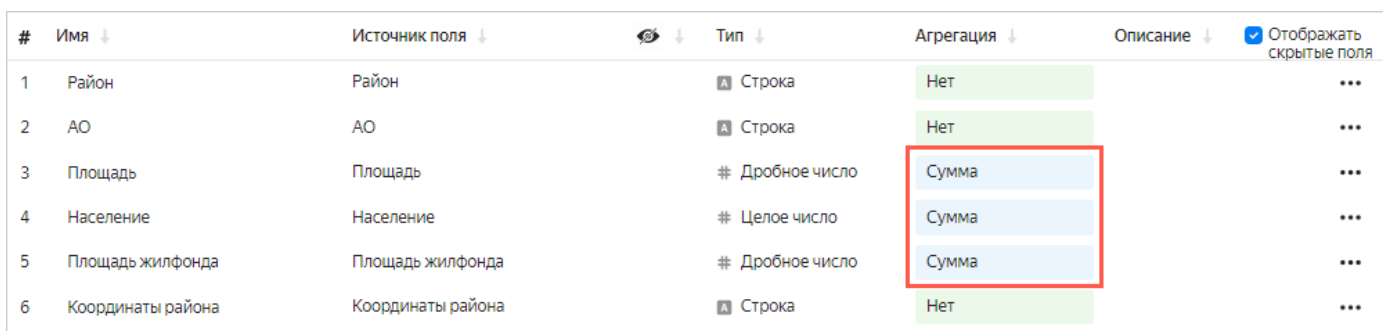
3. Перейдите на вкладку **Поля**.



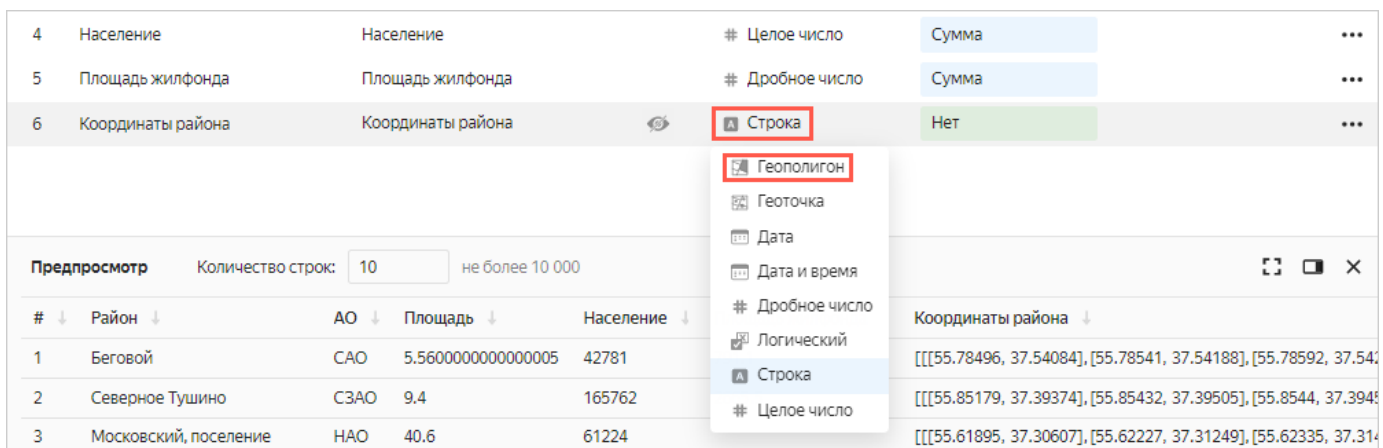
4. Для полей **Площадь**, **Население** и **Площадь жилфонда** укажите тип агрегации **Сумма**.



Заметьте, что поля с настроенными функциями агрегации становятся показателями и меняют свой цвет фона на голубой.



5. Для поля **Координаты района** измените тип данных на **ГеополYGON**.



6. В правом верхнем углу нажмите кнопку **Сохранить**.

default Users / / Новый датасет Создать чарт Сохранить

Источники Поля Фильтрация Обновить поля Предпросмотр + Добавить поле

#	Имя ↓	Источник поля ↓	Тип ↓	Агрегация ↓	Описание ↓	<input checked="" type="checkbox"/> Отображать скрытые поля
1	Район	Район	Строка	Нет		...
2	АО	АО	Строка	Нет		...
3	Площадь	Площадь	Дробное число	Сумма		...
4	Население	Население	Целое число	Сумма		...
5	Площадь жилфонда	Площадь жилфонда	Дробное число	Сумма		...
6	Координаты района	Координаты района	ГеополYGON	Нет		...

7. В открывшемся окне введите имя датасета «**Население Москвы**» и нажмите кнопку **Создать**.

8. После сохранения датасета нажмите **Создать чарт**.

default Users / / Население Москвь ☆ Создать чарт Сохранить

Источники Поля Фильтрация Обновить поля Предпросмотр + Добавить поле

#	Имя ↓	Источник поля ↓	Тип ↓	Агрегация ↓	Описание ↓	<input checked="" type="checkbox"/> Отображать скрытые поля
1	Район	Район	Строка	Нет		...
2	АО	АО	Строка	Нет		...

### Шаг 3. Создайте чарт с картой районов Москвы

1. Выберите тип визуализации **Карта**.

Датасет: Население Мо... Столбчатая диаграмма

Поиск  +

**Измерения**

- АО
- Координаты района
- Район

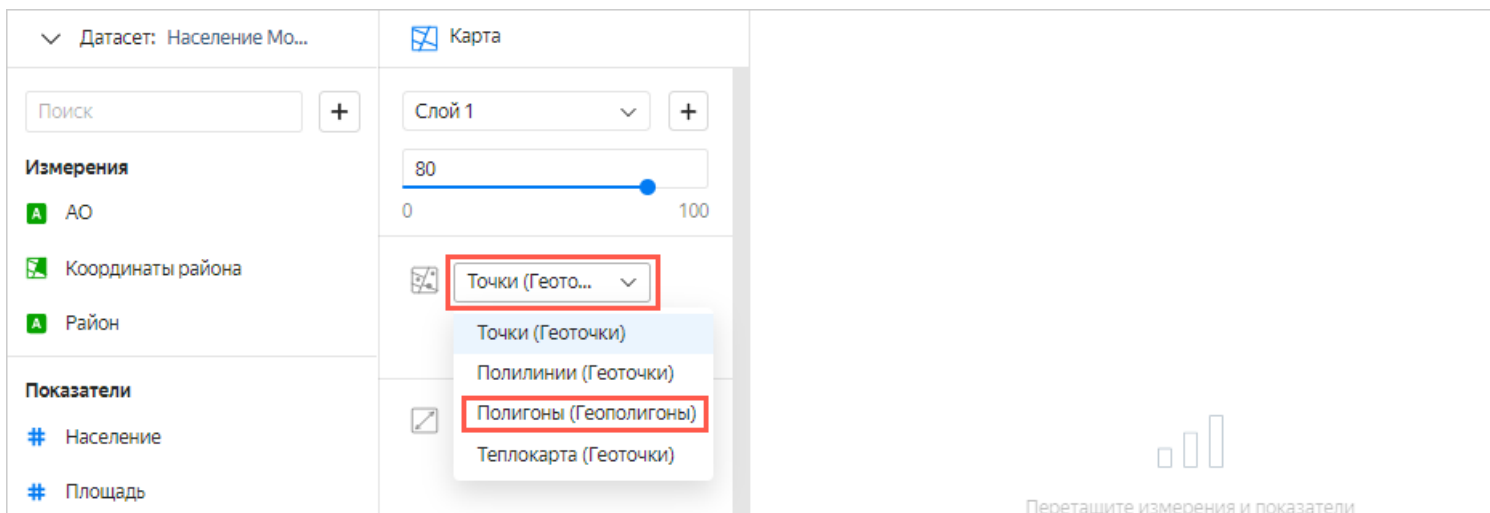
**Показатели**

- Население
- Площадь
- Площадь жилфонда

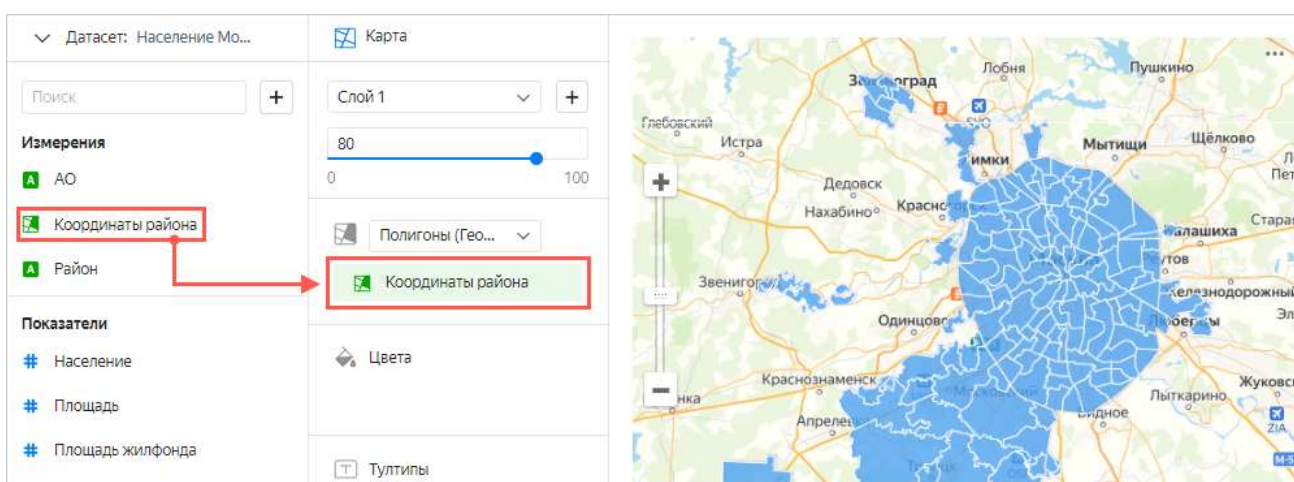
- Линейная диаграмма
- Диаграмма с областями
- Нормированная диаграмма с областями
- Столбчатая диаграмма**
- Нормированная столбчатая диаграмма
- Линейчатая диаграмма
- Нормированная линейчатая диаграмма
- Точечная диаграмма
- Круговая диаграмма
- Кольцевая диаграмма
- Индикатор
- Древовидная диаграмма
- Таблица
- Сводная таблица
- Карта**

Перетащите измерения и показатели в соответствующие секции, чтобы построить визуализацию

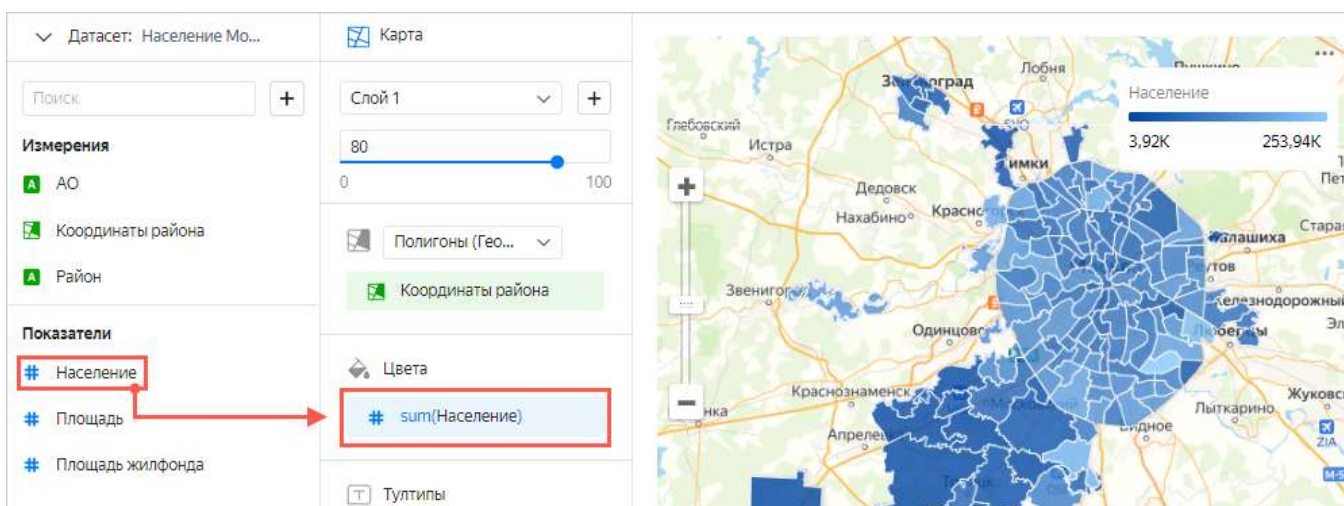
2. Выберите тип слоя **Полигоны (ГеополYGONы)**.



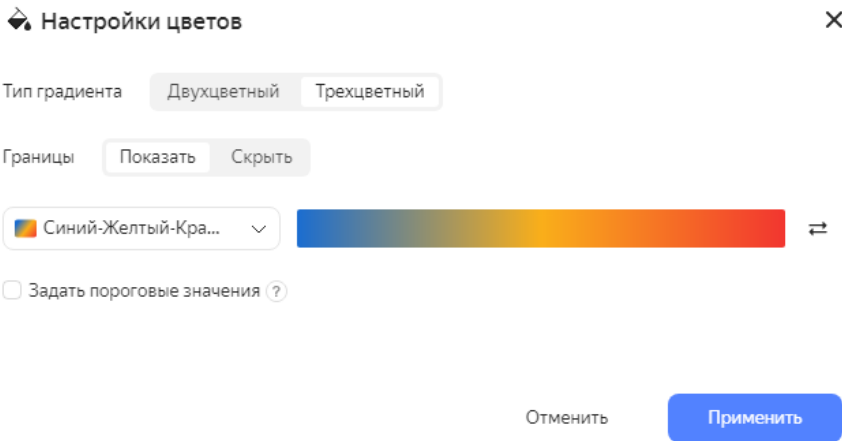
3. Добавьте на карту полигоны районов. Для этого из раздела **Измерения** перетащите поле **Координаты** в секцию **Полигоны (Геополигоны)**.



4. Измените цвета районов относительно показателя населения. Для этого из раздела **Показатели** перетащите поле **Население** в секцию **Цвета**.



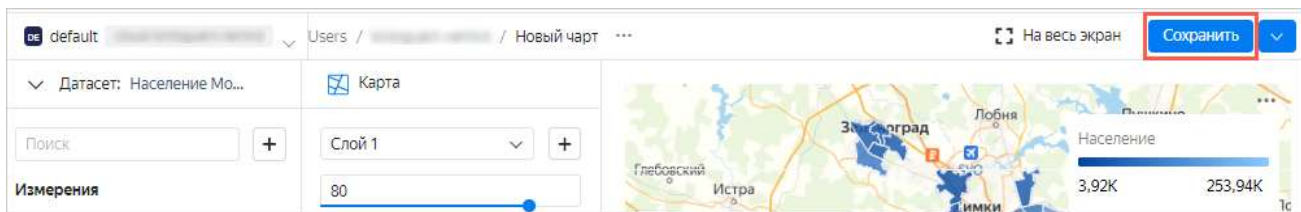
Выполните настройку цветов полигонов карты. Нажмите шестеренку, в панели настройки включите тип градиента – трехцветный, цветовой вариант – синий-желтый-красный. Нажмите – Применить.



Добавьте дополнительную информацию для полигонов. Для этого из раздела **Измерения** перетащите поле **Район** в секцию **Тултипы**.

Измените прозрачность слоя 1 на значение 60.

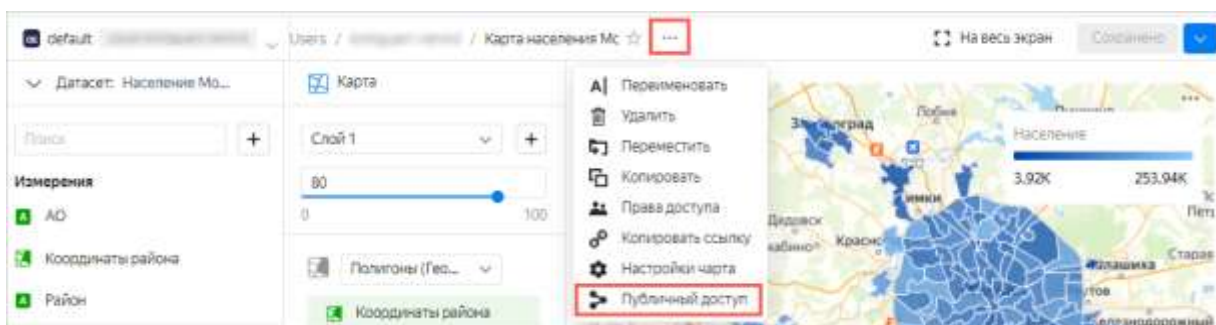
5. В правом верхнем углу нажмите кнопку **Сохранить**.



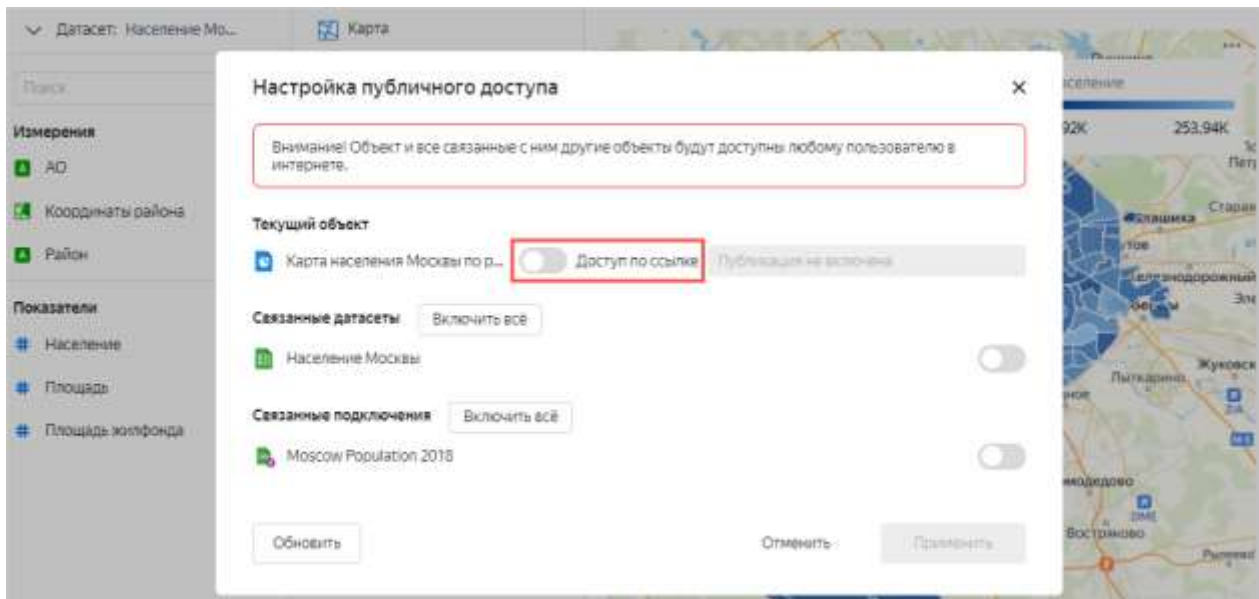
6. В открывшемся окне введите название чарта «**Карта населения Москвы по районам**» и нажмите кнопку **Сохранить**.
7. Проверьте работу фильтров слоя. Для этого нажмите в секции **Фильтры слоя** кнопку +. В открывшейся панели выберите какой-нибудь округ, например, ЦАО. Нажмите – Применить фильтр. Посмотрите и запомните результат в отчете. Измените фильтр на другой округ. Также запомните результат в отчете. Удалите фильтры слоя.
8. Проверьте работу фильтров данных. Для этого нажмите в секции **Фильтры** кнопку +. В списке опций выберите Площадь жилфонда. На панели настройки фильтра выберите операцию **Больше**, в поле **Значение** напишите 2000. Нажмите – Применить фильтр.

## Шаг 4. Опубликуйте чарт в DataLens Public

1. Нажмите значок  и в открывшемся списке выберите **Публичный доступ**.

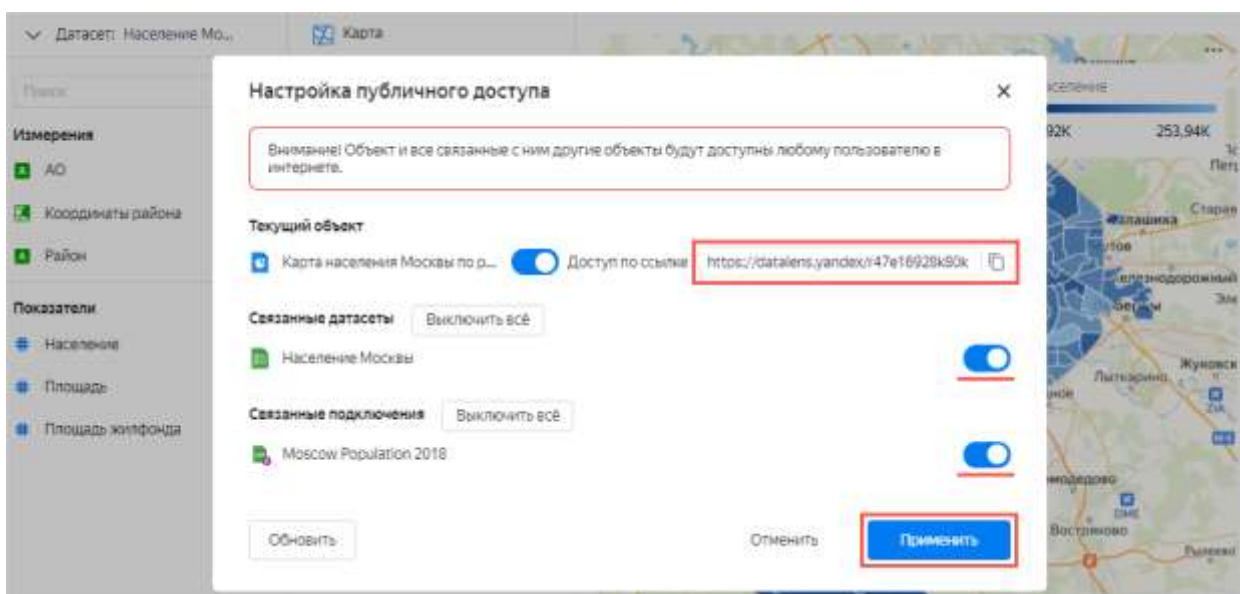


2. В открывшемся окне **Настройка публичного доступа** включите **Доступ по ссылке**.

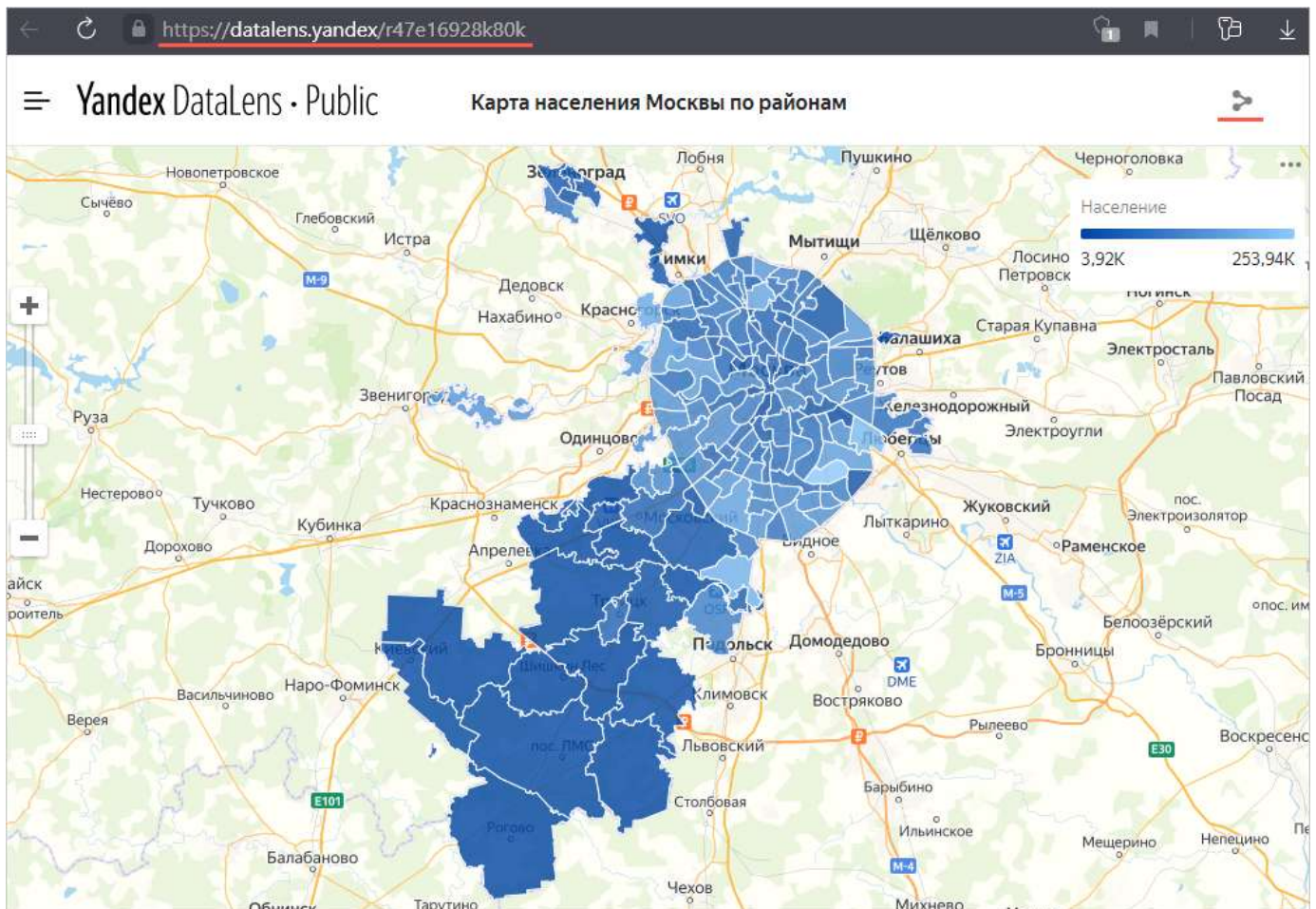


3. Подтвердите публикацию чарта и связанных с ним данных.

4. Скопируйте уникальный URL чарта. Нажмите **Применить**.



5. Откройте новую вкладку в браузере и перейдите по скопированной ссылке.



Этой ссылкой можно поделиться с коллегами, партнерами или друзьями. DataLens Public не требует авторизации и аутентификации при переходе по ссылке чарта. Отправьте ссылку 2-3 своим одноклассникам. Получите от них подтверждение доступности вашего чарта.

## Шаг 5. Столбчатая диаграмма и дашборд

Для визуализации распределения жителей по районам и населению создайте чарт — столбчатую диаграмму.

1. Добавьте на график районы. Для этого из раздела **Измерения** перетащите поле **Район** в секцию **X**.
2. Добавьте на график население. Для этого из раздела **Показатели** перетащите поле **Население** в секцию **Y**.
3. Отсортируйте график по населению. Из первой колонки из раздела **Показатели** перетащите поле **Население** в секцию **Сортировка**.
4. Оставьте на графике только те районы, где население больше 100 000.
5. Из первой колонки из раздела **Показатели** перетащите поле **Население** в секцию **Фильтры**.
6. В открывшемся окне укажите операцию **Больше** и значение 100 000.



7. Нажмите кнопку **Применить фильтр**.

Создайте дашборд с картой, столбчатой диаграммой и селекторами.

Для создания дашборда выполните следующее:

1. В левой колонке виджетов сервиса Yandex DataLens нажмите **Создать дашборд**.
2. В открывшемся окне введите название дашборда. Дашборд появится в списке на странице навигации.

Добавьте на дашборд полученные ранее чарт с картой районов и чарт Столбчатая диаграмма.

Чтобы добавить селектор на дашборд:

1. Откройте дашборд.
2. В верхней части страницы нажмите кнопку **Редактировать**.
3. Нажмите кнопку **Добавить** и выберите **Селектор**.
4. Выберите тип селектора: **На основе датасета**

Укажите параметры селектора:

- В блоке **Общие настройки**:
  - **Датасет**. Определяет датасет с данными для селектора.
  - **Поле**. Определяет поле датасета со значениями селектора. Может быть как измерением, так и показателем.
  - **Тип селектора**. Определяет тип селектора: выпадающий список, поле ввода или календарь. Тип селектора **Календарь** доступен только для полей датасета типа Дата или Дата и время. Если в **Поле** выбран показатель, то доступен только тип селектора **Поле ввода**.
  - **Операция**. Определяет операцию сравнения, по которой селектор фильтрует значения чарта (например, **Равно**, **Больше** или **Меньше**). Если оставить поле пустым, селектор по умолчанию будет фильтровать по операции **Равно**. Список доступных операций зависит от типа поля. Не указывайте операцию, если селектор будет фильтровать QL-чарт.
  - **Множественный выбор**. Определяет возможность выбора нескольких значений. Опция доступна только для селекторов типа **Список**.
  - **Диапазон**. Определяет возможность выбора временного промежутка. Опция доступна только для селекторов типа **Календарь**.
  - **Значение по умолчанию**. Отображается изначально при открытии дашборда.

В блоке **Внешний вид**:

- **Название**. Определяет название селектора. Используется для выбора селектора при установлении связи с другими виджетами. Опция позволяет управлять отображением названия на дашборде.
- **Внутренний заголовок**. Текст, который отображается в селекторе для обозначения операции сравнения. Вы можете изменить значение по умолчанию на свое. Например, для

операции **Равно** можно указать значение = или равно. Параметр доступен только для селекторов типа **Список**.

5. Нажмите кнопку **Добавить**. Виджет отобразится на дашборде.

Ограничения:

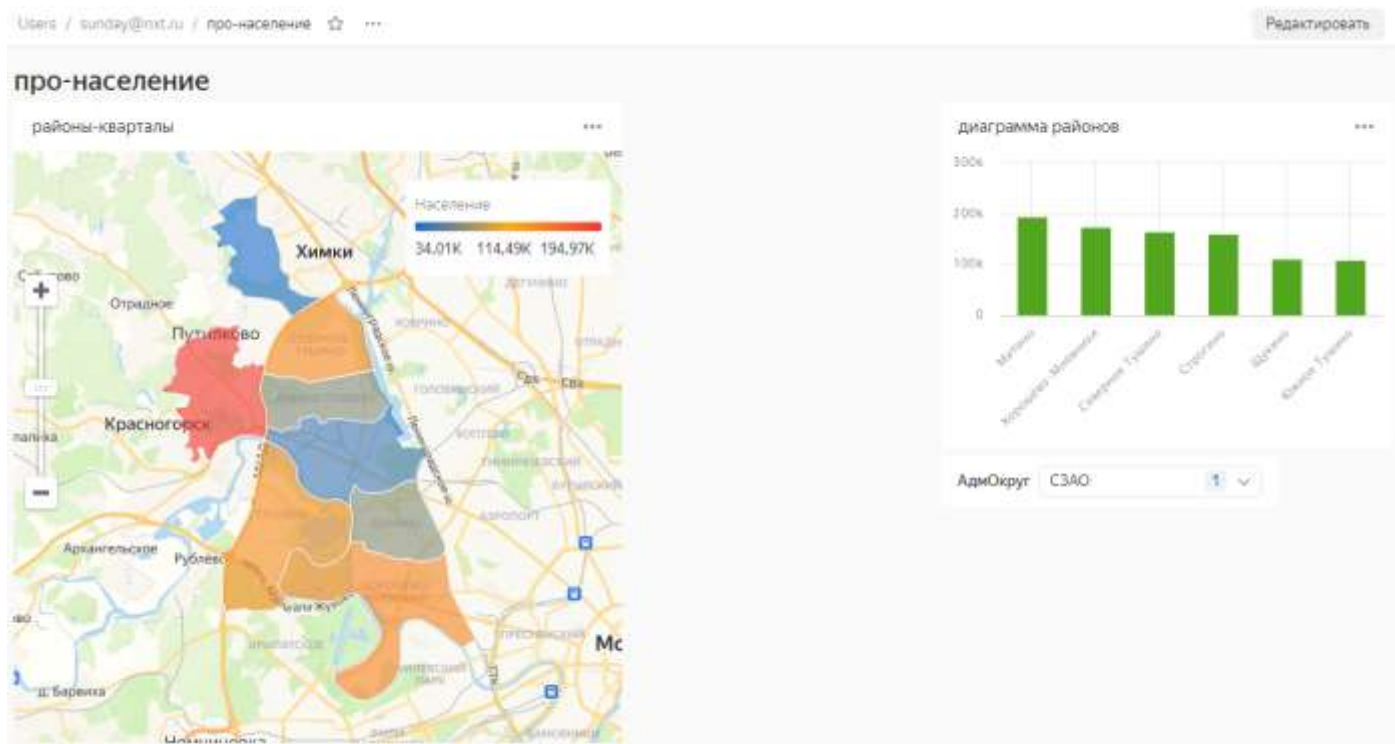
- Для селекторов по показателям доступен только один тип селектора — **Поле ввода**.
- Селекторы по показателям рекомендуется делать только независимыми от других селекторов (необходимо для них задать тип **связи Игнор** с другими селекторами в разделе **Связи** при редактировании дашборда).

Добавьте селектор, чтобы иметь возможность фильтровать по полю АО (округ).

1. Нажмите кнопку **Добавить**.
2. Выберите **Селектор**.
3. Выберите датасет **ваш датасет**.
4. Выберите поле АО.

После этого автоматически заполнится поле **Название**.

5. Нажмите галочку **Показывать** напротив заголовка селектора.
6. Включите опцию **Множественный выбор**.
7. Нажмите кнопку **Добавить**.



Настройте отображение виджетов и приступите к исследованию зависимостей

1. Перетащите селекторы наверх страницы рядом друг с другом.
2. Растяните чарты по ширине и расположите рядом друг с другом.
3. В правом верхнем углу нажмите кнопку **Сохранить**.

Примените различные фильтры и исследуйте возможные зависимости данных.

В отчете напишите:

«В датасете, который был исследован, прослеживаются следующие зависимости:....»

Отправьте отчет на почту преподавателю.