1. Пояснение к дашборду:

Сделала модель через справочник, так как соединить (значение поля «№ распоряжения/поручения» задачи содержится в значении поля «№ проверки» мероприятия) так и не получилось. Что я сделала и это НЕ увенчалось успехом:

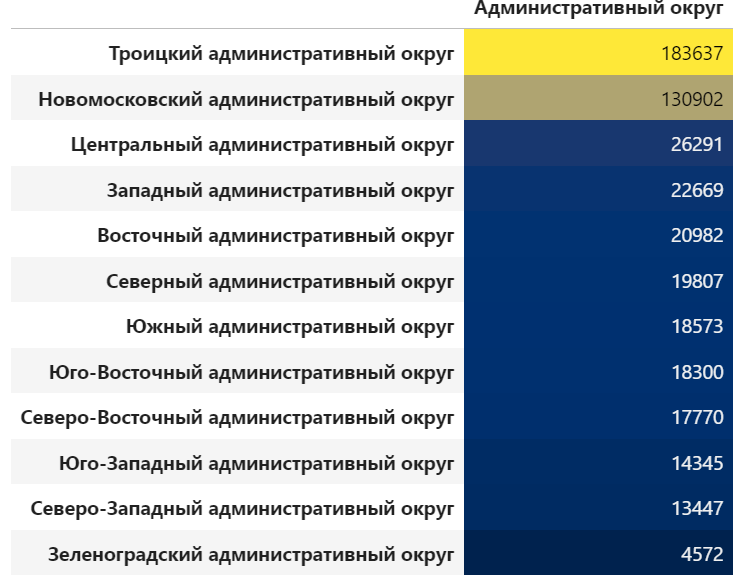
1. Меняла тип данных;
2. Удаляла версионность «№ проверки» (так как по заданию связь 1 ко многим)

В итоге решила через создание справочника, так как нужно было посмотреть по округам.

1. Общее заключение по качеству данных:
   1. Из всех предоставленных датасетов хуже всего качество данных в файле в формате .xls для тестового задания на аналитика BI (более всего нуждались в предобработке и местами противоречат друг другу).

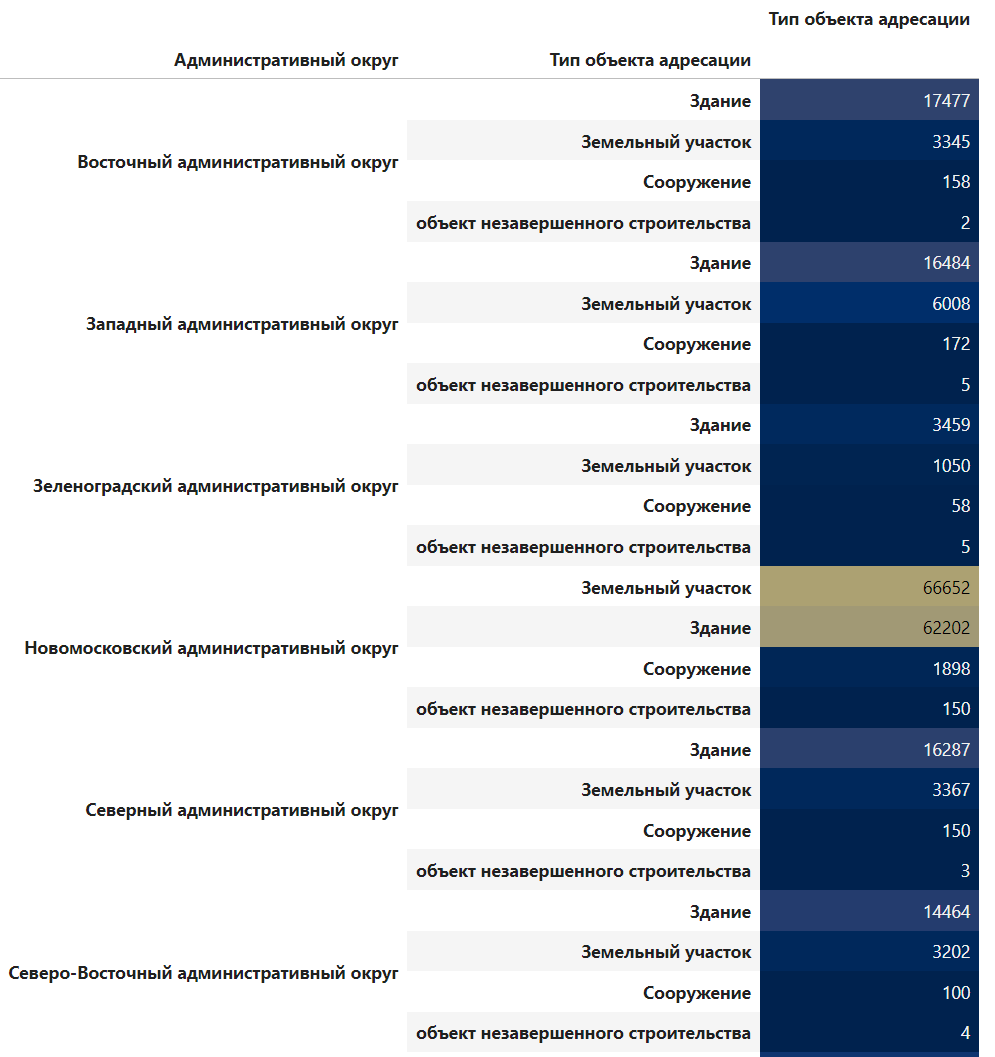
Выводы по итогу анализа данных:

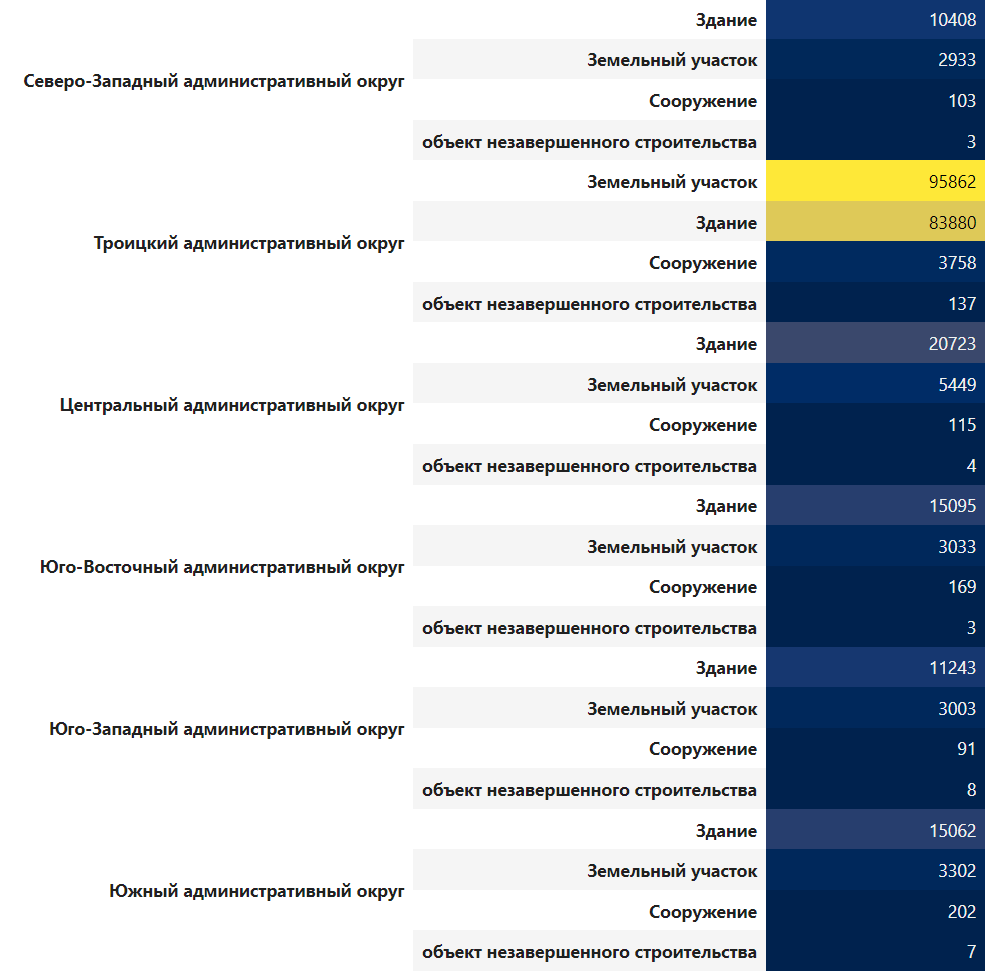
1. Общее количество объектов в разрезе округов:



1. Способы классификации (типизации) объектов – в датасете всего 4ре типа:
   1. Здание
   2. Сооружение
   3. Земельный участок
   4. объект незавершенного строительства

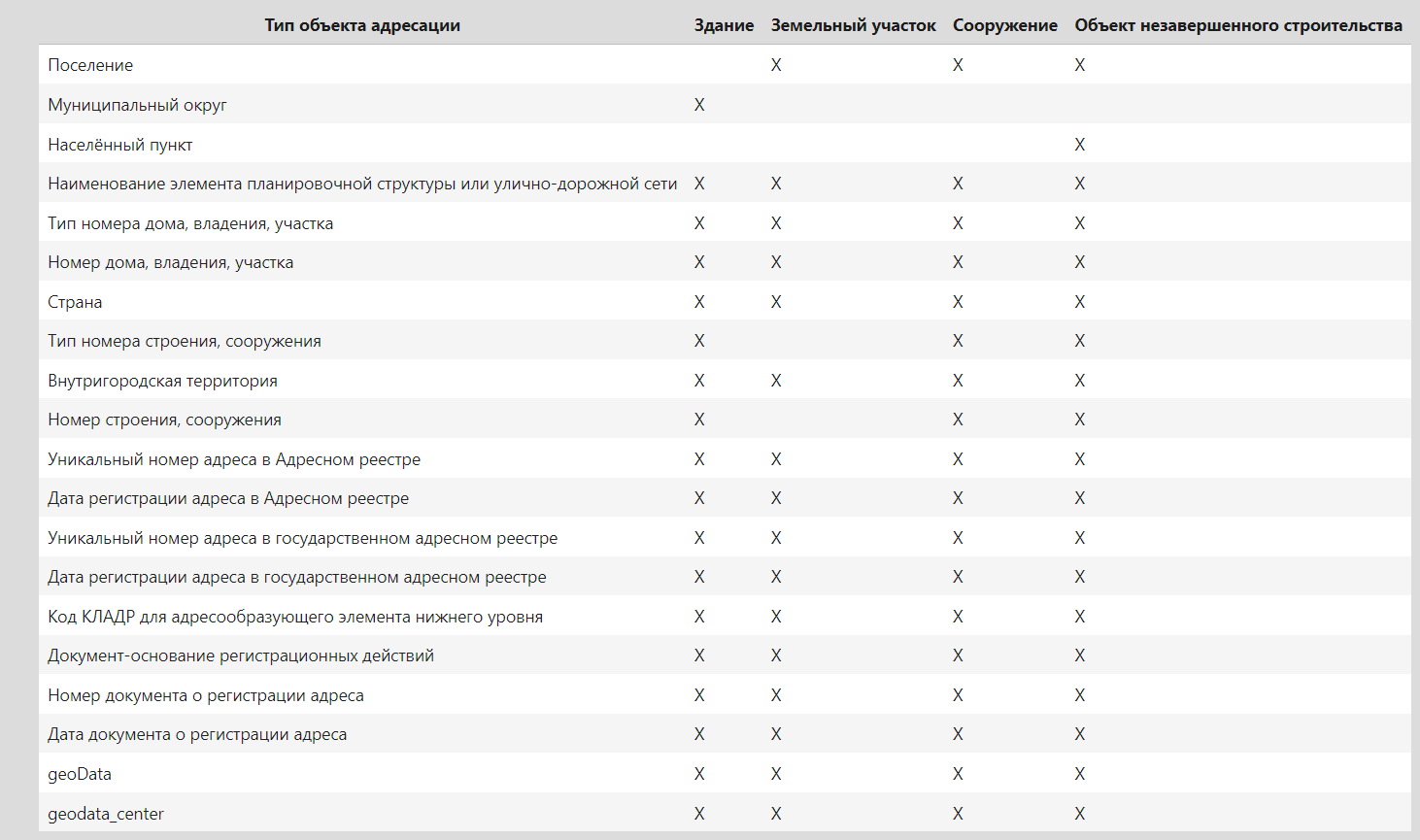
В разрезе административных округов они распределены следующим образом:





* 3. Общие характеристики объектов. Для всех типов объектов полностью указаны следующие характеристики:
* global\_id - доля пропусков (0%)
* Тип объекта адресации - доля пропусков (0%)
* На территории Москвы - доля пропусков (0%)
* Полное юридическое написание адреса или описание местоположения - доля пропусков (0%)
* Учётный номер объекта адресации в БД БТИ (кроме помещения и комнаты) - доля пропусков (0%)
* Субъект РФ - доля пропусков (0%)
* Упрощённое написание адреса или описание местоположения - доля пропусков (0%)
* Административный округ - доля пропусков (0%)
* Муниципальный округ, поселение - доля пропусков (0%)
* Кадастровый номер объекта недвижимости (кроме земельного участка) - доля пропусков (0%)
* Кадастровый номер земельного участка (для ОКС) - доля пропусков (0%)
* Тип адреса - доля пропусков (0%)
* Вид адреса - доля пропусков (0%)
* Состояние адреса - доля пропусков (0%)
* Статус адреса - доля пропусков (0%)

Если смотреть на признаки с пропусками, то наиболее полно представлены след. признаки для каждого типа:



4. Локализацию объектов – см тетрадку в .ipynb

5. Количество и распределение проверок по объектам – см дашбод, вкладка «Распределение проверок». Районы, которые подвергаются проверкам чаще всего:



Существуют ли районы с повышенным количеством нарушений?

Можно сказать, что кол-во нарушений в Басманном районе и районе Печатники существенно отличается от кол-ва нарушений в др. районах. Лидер по кол-ву нарушений – Басманный район.



Больше всего выявлено «нецелевое использование ЗУ».

6. Распределение нагрузки между проверками - – см дашбод,

7. Частота проведения проверок и количество выявленных нарушений.