Так как записи без цены не дадут нам возможности обучить модель или проверить ее качество, удалим их.

Для предсказания цены нам необходим только год и месяц. Выделим их из данных о дате публикации объявления.

В адресе оставим только улицу, для упрощения обработки категориального признака.

В списке улиц присутствуют неявные дубликаты. Приведем названия улиц к единому написанию.

Добавим столбец с информацией о том, является ли этаж квартиры последним.

Преобразуем типы данных у числовых столбцов.