В таблице по недвижимости России имеются пропущенные ячейки (NaN) – это говорит нам о том что либо значение не было записано (было пропущено) в таблицу во время ее создания. Например высота потолка или квадратура.

Также какой-либо параметр может не соответствовать (не подоходить) текущему виду недвижимости. Например: поле аэропорт может быть пустым если в городе его вовсе нет, то же можно отнести и к паркам; в квартире может отсутсвувовать балкон, соответственно заполнять ячейку нет смысла;

Ячейки могут содержать в себе ложное значение, т.е. значение которое какому-либо параметру несоотвествует. Другими словами, значение которое выходит из диапазона большинства значений. Такие параметры могут повлиять на рассчеты, скажем, среднего значения или медианы. Такие значения называют **Выбросами**. Для того чтобы определить диапазон нормальных значений удобнее всего использовать библиотеки для визуализации, *seaborn* например. Таким образом мы можем визуально определить выбросы и не включать их в расчеты.

Также в датасете могут присутсвовать дубликаты – повторяющиеся записи. В таблице по недвижимости России их нет. Но при наличии дубликатов, от них можно избавиться с помощью метода *.drop\_duplicates().*