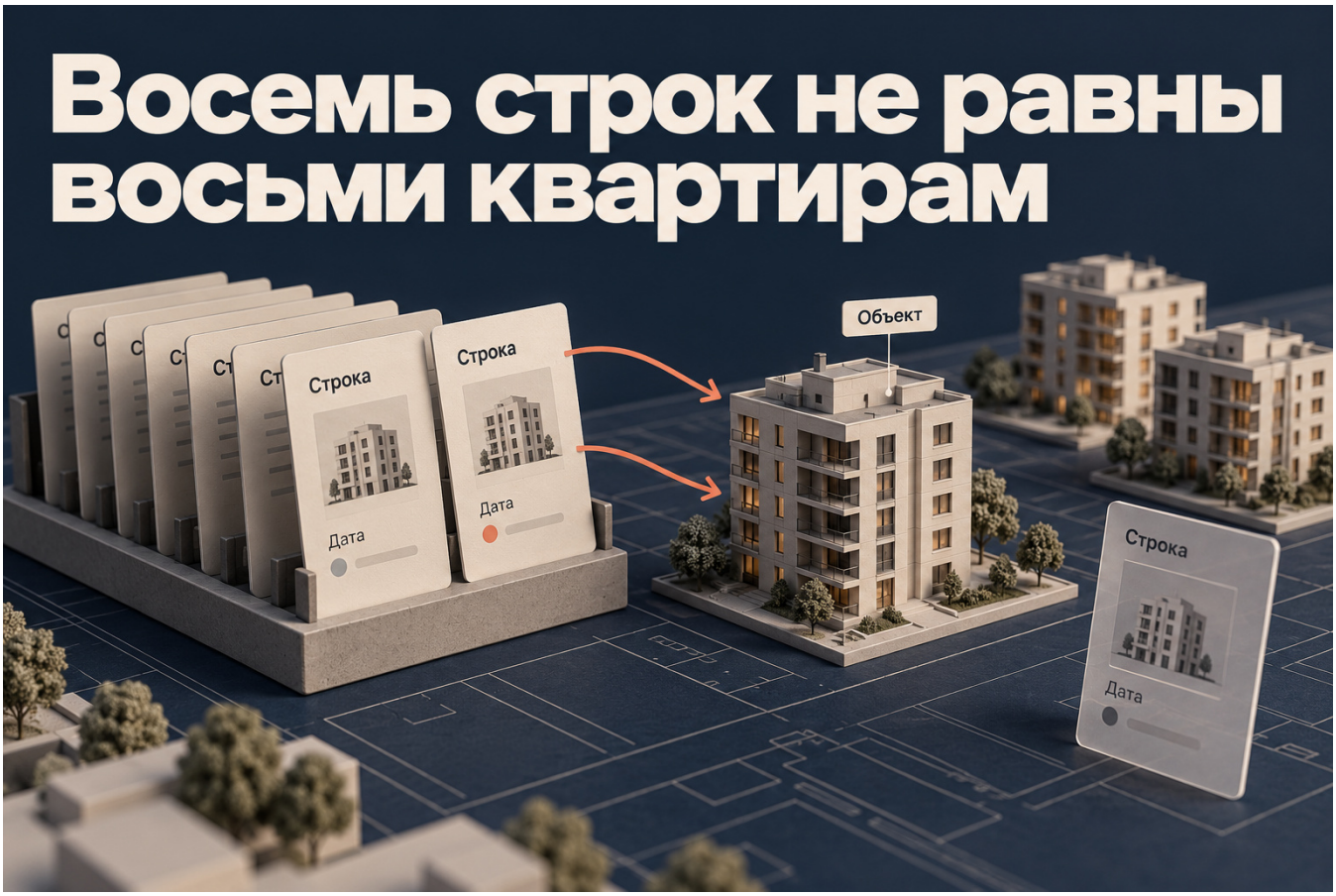
Д781: паспорт данных и проверяемая обработка

Восемь учебных строк, словарь, журнал обработки и честный срез без скрытых подстановок.



Практический гайд с заполненным учебным примером и чистыми рабочими листами. Подготовлено редакцией Нейросистемы риелтора.

Учебные цифры не являются данными о текущем рынке. Обложка представляет концептуальную иллюстрацию.

Д781: паспорт набора

Самостоятельный комплект «Нейросистемы риелтора». Все восемь строк и источник У781 вымышлены. Даты учебные, 2026 год. Набор не описывает рынок города и не содержит сведений реальных клиентов.

Поле	Значение
Единица	Состояние объекта на дату в источнике
Ключ	Код объекта + дата + источник
Назначение	Срез активных записей на 08.10.2026
Исходный объём	8 строк, 5 объектов
После исключения дубля	7 наблюдений на двух датах

Мой вопрос / единица	Источник / охват / право использования

Восемь исходных строк

Строка	Объект	Дата	Цена: исходник	м²	Статус
1	O1	01.10	6 400 000 ₽	44	Активно
2	O1	08.10	6 200 000 ₽	44	Активно
3	O1	08.10	6 200 000 ₽	44	Активно
4	O2	01.10	7 100 000 ₽	50	Активно
5	O2	08.10	7 100 000 ₽	Пусто	Активно
6	O3	08.10	7 200 000 ₽	0	Активно
7	O4	08.10	8 500 000 ₽	55	Снято
8	O5	08.10	7,8 млн ₽	60	Активно

Пояснение
Все даты 2026 года, источник У781. Строка 3 точно повторяет 2. Соответствие кодов объектам задано условиями, а не выведено из сходства цен.

Сохраняйте этот исходник неизменным. Все преобразования относятся к расчётной копии. Строки O1 и O2 на разных датах не являются техническими дублями.

Словарь полей и допустимых значений

Поле	Смысл / правило
ID строки	Постоянный указатель исходной записи
Объект	Учебный код; не номер случайного объявления
Дата	Дата наблюдения с годом
Источник	У781; сохраняется в ключе
Цена исходная	Строка с единицей как получена
Цена в рублях	Число после обоснованного преобразования
Площадь	м²; для деления нужна положительная
Статус	Активно или снято; не статус сделки
Причина пропуска	Отсутствует либо некорректный ноль

Моё поле / тип / единица	Пропуск / допустимые значения

Журнал обработки Д781

Строка	Действие	Основание
3	Не включать в расчётную копию	Точная копия 2, тот же ключ и поля
8	7,8 млн ₽ → 7 800 000 ₽	Явная единица миллионов
5	Площадь остаётся пропуском	В текущей строке отсутствует
6	Расчётная площадь пустая; исходный 0 сохранён	Нулевой знаменатель непригоден
7	Сохранить статус снято	Продажа не подтверждена

Без переноса из прошлого
50 м² из строки 4 не подставляется в строку 5. В Д781 перенос запрещён. В другой методике допустимая подстановка должна иметь источник и отдельный признак.

Строка / поле / было	Правило / стало / причина

Срез на 8 октября: активные записи

Объект	Цена, ₽	Расчётная площадь	Результат ₽/м²
O1	6 200 000	44	140 909,09
O2	7 100 000	Неизвестна	Не рассчитывать
O3	7 200 000	Некорректный 0	Не рассчитывать
O5	7 800 000	60	130 000

O4 исключён только из этого среза
Он остаётся в наборе со статусом снято. Это не означает, что объект продан.

Четыре активных записи; показатель вычислим для двух. 2 / 4 = 50% вычислимости. Это не точность модели и не оценка полноты всего рынка.

Мой фильтр даты / статуса	Число записей / пригодных для показателя

Разные количества, разные вопросы

Что считаем	Ответ Д781
Исходные строки	8
Технические копии	1
Уникальные наблюдения после очистки	7
Разные объекты во всём наборе	5
Активные записи на 08.10	4
Пригодные для цены за м² в срезе	2

Не подменяйте единицу
Семь наблюдений не являются семью квартирами. Две пригодные строки не представляют автоматически все четыре активных объекта.

Мой показатель / единица	Фильтр / знаменатель / ограничения

Запрос к нейросети для проверки

Промпт
Проверь набор по словарю. Не изменяй исходник. Для замечания укажи ID строки, поле, исходное значение, правило и предлагаемое действие. Не заполняй неизвестную площадь, не считай снятое проданным, не объединяй разные даты. Выдай журнал исключений и ограничения расчёта.

Приёмка предложения	Как проверить
Ключ совпал	Сопоставить дату, объект и источник
Единица нормализована	Найти явную единицу в исходнике
Пропуск сохранён	Не появился придуманный факт
Фильтр объяснён	Каждая исключённая строка имеет причину

Предложение модели	Решение человека / основание

Чистый журнал преобразований

Версия исходника / дата	Версия расчётной копии / исполнитель

Строка / поле / исходное значение	Правило / новое значение / основание

Исключение / причина	Что осталось неизвестным / источник уточнения

Воспроизводимость
Коллега должен повторить результат из исходника по сохранённым правилам. Нельзя заменять историю обработки только готовой красивой таблицей.

Передача набора и его ограничений

Назначение / единица / ключ	Период / охват / источник
Разрешённое использование	Пропуски / исключения / ограничения
Версия словаря / преобразования	Принял / дата / следующий пересмотр

Проверено 19.10.2026. Д781 предназначен для освоения структуры данных. Фактические разрешения на использование источников и свойства реального рынка в этом упражнении не устанавливаются.

Источники и проверка

W3C: модель табличных данных и метаданных