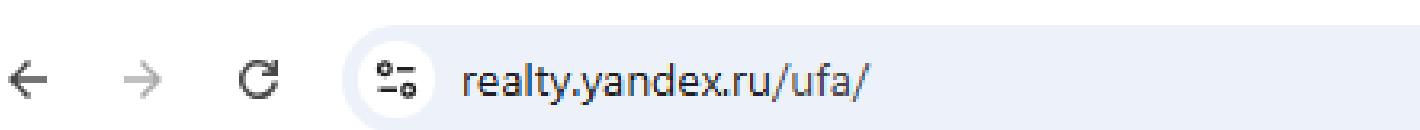


Недвижимость

Регрессия

Источники данных

- Площадки объявлений
- Яндекс Недвижимость
 - «Вторичка»
 - Цена – Площадь – Комнат – Источник



Ввод данных и контроль

The screenshot shows a Microsoft Excel spreadsheet titled "2026-02-26 ПИ432 ЛР1 Подгруппа 1". The menu bar includes File, Edit, View, Insert, Format, Data, Tools, Add-ins, Help, and a Cloud icon. The ribbon toolbar below the menu bar includes search, navigation, print, zoom (100%), orientation (Portrait), cell style (p.), percentage (%), decimal (.0), currency (.00), number (123), date/time (По ум...), and font size (10). The active cell is J13, with a formula bar showing "fx". The table has columns labeled A through F. Column A contains row numbers from 1 to 13. Columns B, C, D, and E contain data entries. Column F contains hyperlinks for each row. Row 13 is highlighted in blue.

| | A | B | C | D | E | F |
|----|------------|----------------------|--------|----------|----------|---|
| 1 | Цена | Площадь | Комнат | Оператор | Контроль | Источник |
| 2 | 18 403 500 | 112,7 м ² | 4 | | | Купить квартиру 112.7 м² |
| 3 | 7 700 000 | 84 м ² | 4 | | | Купить квартиру 84 м², 6 |
| 4 | 38 224 940 | 174,6 м ² | 4 | | | Купить квартиру 174.56 м² |
| 5 | 17 885 400 | 108,6 м ² | 4 | | | Купить квартиру 108.66 м² |
| 6 | 13 500 000 | 145 м ² | 4 | | | Купить квартиру 145 м², 1 |
| 7 | | | | | | |
| 8 | 4640000 | 46,3 м ² | 2 | | | https://realty.yandex.ru/offers/1127000 |
| 9 | 4750000 | 50 м ² | 2 | | | https://realty.yandex.ru/offers/1128000 |
| 10 | 8200000 | 65 м ² | 2 | | | https://realty.yandex.ru/offers/1129000 |
| 11 | 4699999 | 39,4 м ² | 2 | | | https://realty.yandex.ru/offers/1130000 |
| 12 | 6232000 | 41 м ² | 2 | | | https://realty.yandex.ru/offers/1131000 |
| 13 | | | | | | |

Google Colab

- <https://colab.research.google.com/>
- Создать блокнот
- Название: Фамилия, группа, ЛР1

Загрузка данных

- pandas
 - read_csv
 - Ссылка на таблицу
- Выбрать столбцы
 - Цена
 - Площадь
 - Комнат

DataFrame

- df
- df.dtypes

Очистка данных

- Тип данных по столбцам
 - int
 - float
 - object – string
- NaN – Not a Number
 - Пропуски
 - Пустые строки

Описательная статистика

- df.describe()
- Смысл полученных характеристик
- Гистограмма
- Диаграмма разброса / рассеяния
- «Матрица» графиков

Линейная регрессия

- Уравнение связи
- Линия регрессии на диаграмме разброса
 - Цена = f (Площадь)
 - Цена = f (Площадь, Комнат)
 - Фиктивные переменные для числа комнат
 - Dummy Variables
 - One Hot Encoding
- Диаграмма: Реальная цена vs. Прогноз
- Метрики качества модели
 - Сравнить модели

Оформление отчета

- Текстовые ячейки и комментарии
- Заголовки – структура
- Пояснения основных действий
- Доступ на чтение
- Ссылка на блокнот
- Форма на гитхабе