



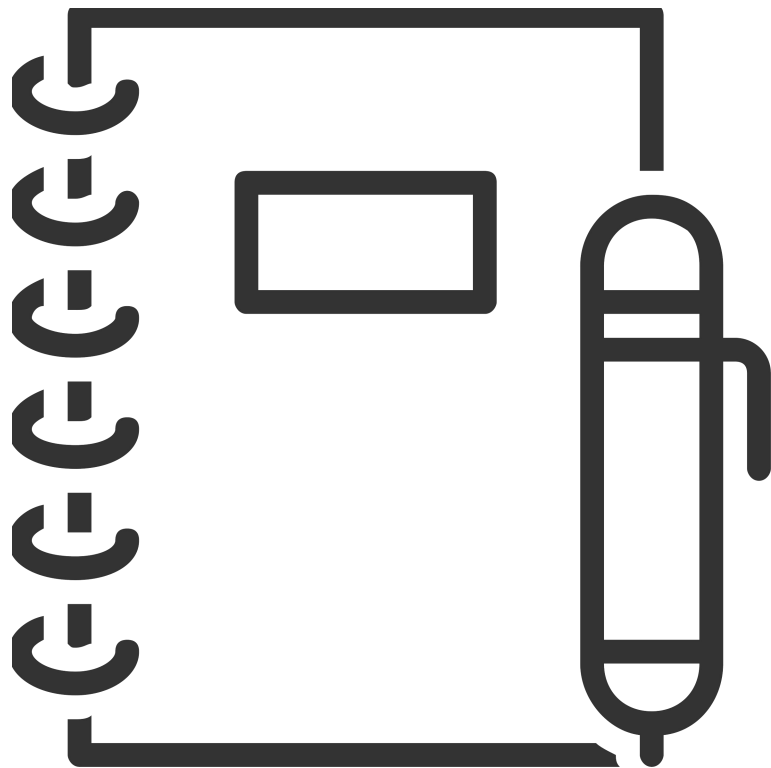
Тестовое задание на должность Аналитик данных

Исследование провел: Аникин Александр
тел: +7 (999) 562-02-35
Дата исследования: 24.04.23

<https://public.tableau.com/app/profile/aleksanr.anikin/viz/anikinTest/Dashboard1?publish=yes>



ссылка дашборда на сайте Tableau Public



План презентации:

1. Оглавление
2. Ссылка дашборда и план
- 3-4. Сводные таблицы
5. Функция ВПР
6. Функции
- 7-8. Предложение по ценам
7. Вывод

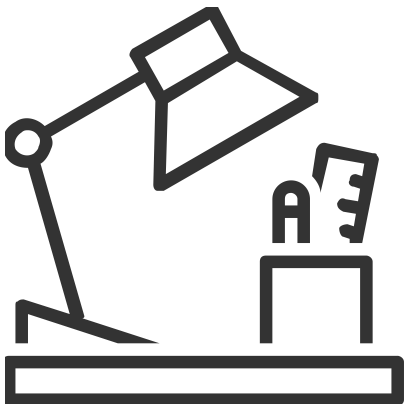


Сводные таблицы:

Go критерии составления таблицы были построены формулы среднего значения и количества - AVERAGEIFS, COUNTIFS- ссылающиеся на алазируемую базу, из-за пропусков данных в таблице произошли ошибки, они заменены на 0, не во всех проектах присутствуют количества комнат 4 и тд. В Формулах, статичные движения закреплены для удобство в применении, при смене значений, происходит автоматический расчет, данные проверены на Питоне, тетрадка Юпитера во вложении. Так же сделано построение методом PivotTable и Tableau

Проект	Космонавтов 11			
Корпус	2Б			
Сроки строительства (старт/дата ввода)	июля 2021		июля 2021	
Количество комнат (Данные объектива)	Ср.площ., м² (ПД)	Ср. цена за м², ₺	Ср.расчетный бюджет, ₺	Кол-во, шт.
студия	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	88	100,800	8,850,240	1
4	0	0	0	0
Итого	18	20,160	1,770,048	0

Проект	Космос	Y2		
Корпус	1	Y2		
Планируемая дата ввода 12/31/2022 Старт продаж Values Статус Y2				
Ср. Площадь, м² (ПД) Ср. Цена за м², ₺ Ср. Расчетный бюджет, ₺ Кол-во комнат шт				
Количество комнат (Данные Объектива) Y2				
2	68	118000	7820643	4
3	83	114000	9304680	1
студия	31	153333.3333	4382733	6
1	51	127000	6155160	6
Итого	58	128,083	6,915,804	4



Сводные таблицы:

Используя библиотеку pandas посчитаны показатели таблицы для проверки, построен живой дашборд, ссылка на 2ом слайде

средняя площадь студии = nan
средняя площадь 1к = 50.52
средняя площадь 2к = 68.14
средняя площадь 3к = 83.49
средняя площадь 4+к = nan
Ср. цена за м² студии = nan
Ср. цена за м² 1к = 127000.0
Ср. цена за м² 2к = 118000.0
Ср. цена за м² 3к = 114000.0
Ср. цена за м² 4+к = nan
Ср. расчетный бюджет = 4382733.33
Ср. расчетный бюджет = 6155160.0
Ср. расчетный бюджет = 7820642.5
Ср. расчетный бюджет = 9304680.0
Ср. расчетный бюджет = nan
Количество штук = 6
Количество штук = 6
Количество штук = 4
Количество штук = 1
Количество штук = 0

Планируемый Год Ввода
☒ (All)
☒ 2020
☒ 2021
☒ 2022
☒ 2023

Проект
☐ (All)
☐ Космонавтов 11
☒ Космос
☐ Северное сияние

Корпус
(All)
Статус
В продаже

Год Старта Продаж
☒ (All)
☒ Null
☒ 2018
☒ 2019
☒ 2020
☐ 2021

Дашборд составлен на основе данных из Тех задания с добавлением фильтров

Сводная таблица по средней площади, средней цене за метр и среднему расчетному бюджету

Количество Комнат	Проект / Корпус / Планируемая Дата Ввода / Старт Продаж							
	1				3			
	2022 2021				2023 2021			
	Средний бюджет	Средняя цена	Средняя площадь	Количество	Средний бюджет	Средняя цена	Средняя площадь	Количество
3	9,304,680	114,000	83	1	8,404,800	102,000	84	15
2	7,656,515	115,500	68	4	6,997,800	109,000	67	9
1	6,150,430			6	5,929,800	119,000	52	20
студия	4,291,500			6	4,069,800	151,000	29	32
Grand Total	6,150,430	128,000	53	17	5,866,700	121,000	51	76

Количество Комнат: 2
Корпус: 1
Проект: Космос
Year of Планируемая Дата Ввода: 2022
Year of Старт Продаж: 2021
Средняя цена: 115,500



ВПР

Построена функция по заданиям в формате эксель, для этого создан доп столбец и по нему найдены нужны показатели

```
=VLOOKUP(A6&B6&C6,'расчеты VLOOKUP'!A:E,5,0)
```

Функцией merge так же построен нудный алгоритм

```
data_merge = data.merge(df, on=['Проект', 'Корпус'], how='left')
```

```
data_merge.head()
```

	Проект	Корпус	Номер лота_у	Площадь, м² (ПД)	Расчетный бюджет лота, ₽	Цена за м², ₽
0	Космонавтов 11	3.1	1	53.0	4101120.0	76800.0
1	Космонавтов 11	3.1	10	52.0	3818100.0	71500.0
2	Космонавтов 11	3.1	2	25.0	NaN	NaN
3	Космонавтов 11	3.1	3	65.0	4746560.0	72800.0
4	Космонавтов 11	3.1	4	37.0	3749760.0	99200.0



Формулы

Составлена формула по заданию в рамках
ексель и библиотеки пандас, построена сводная
таблица, так же двумя способами

x = AVERAGEIF(\$C\$16:\$C\$45,G10,\$B\$16:\$B\$45)	
G	H
3. Собрать из полученной таблицы сводную вида:	
Интервалы площадей	Ср. площадь, м2
22	#DIV/0!
27	26
32	#DIV/0!
37	38
42	40
47	47
52	53
57	59
62	64
67	68
72	71

```
df_intel = pd.merge(datal, df_inte, how='left')
new = df_intel.groupby(['Интервалы площадей']).agg({'Площадь, м² (ПД)': ['mean']}).reset_index()
new['Площадь, м² (ПД)'] = round(new['Площадь, м² (ПД)'])
new
```

Интервалы площадей	Площадь, м² (ПД)	
	mean	
0	27	26.0
1	37	38.0
2	42	40.0
3	47	47.0
4	52	53.0
5	57	59.0
6	62	64.0
7	67	68.0
8	72	71.0

IFS

Logical_test1	AND(B16>=20,B16<25)	= FALSE
Value_if_true1	22	= 22
Logical_test2	AND(B16>=25,B16<30)	= FALSE
Value_if_true2	27	= 27
Logical_test3	AND(B16>=30,B16<35)	= FALSE

- 52

#define function for classifying players based on points

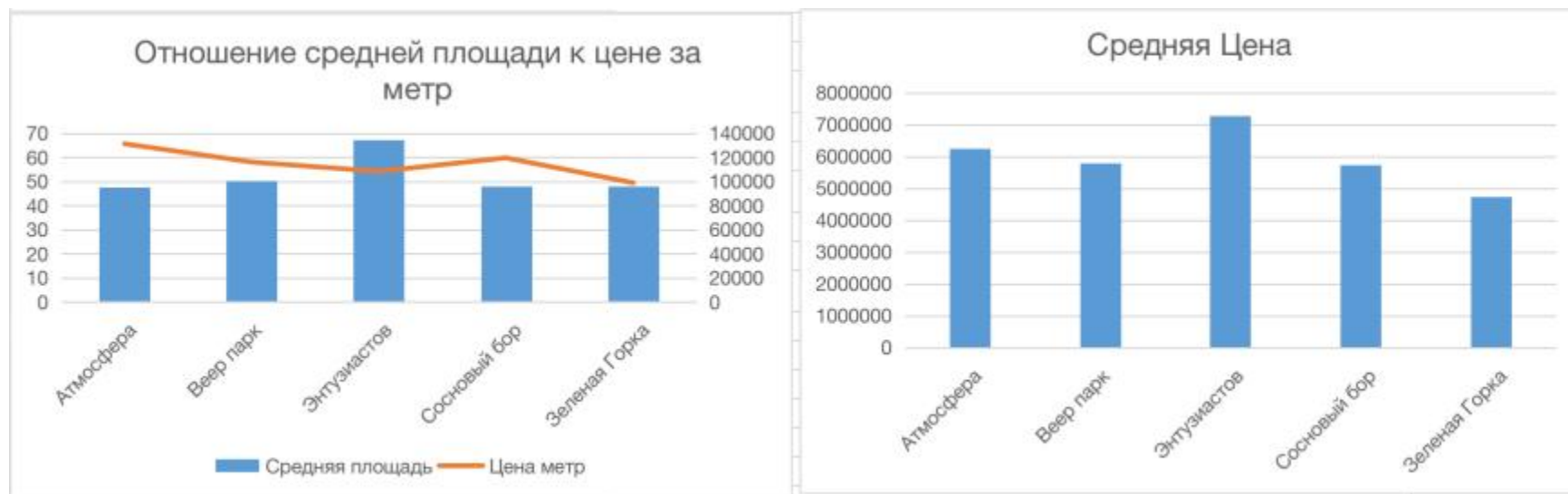
```
def f(row):
    if row['Площадь, м² (ПД)'] >= 20 and row['Площадь, м² (ПД)'] < 25:
        val = '22'
    elif row['Площадь, м² (ПД)'] >= 25 and row['Площадь, м² (ПД)'] < 30:
        val = '27'
    elif row['Площадь, м² (ПД)'] >= 30 and row['Площадь, м² (ПД)'] < 35:
        val = '32'
    elif row['Площадь, м² (ПД)'] >= 35 and row['Площадь, м² (ПД)'] < 40:
        val = '37'
    elif row['Площадь, м² (ПД)'] >= 40 and row['Площадь, м² (ПД)'] < 45:
        val = '42'
    elif row['Площадь, м² (ПД)'] >= 45 and row['Площадь, м² (ПД)'] < 50:
        val = '47'
    elif row['Площадь, м² (ПД)'] >= 50 and row['Площадь, м² (ПД)'] < 55:
        val = '52'
    elif row['Площадь, м² (ПД)'] >= 55 and row['Площадь, м² (ПД)'] < 60:
        val = '57'
    elif row['Площадь, м² (ПД)'] >= 60 and row['Площадь, м² (ПД)'] < 65:
        val = '62'
    elif row['Площадь, м² (ПД)'] >= 65 and row['Площадь, м² (ПД)'] < 70:
        val = '67'
    else:
        val = '72'
    return val
```

#create new column 'Good' using the function above
datal['Интервалы площадей'] = datal.apply(f, axis=1)

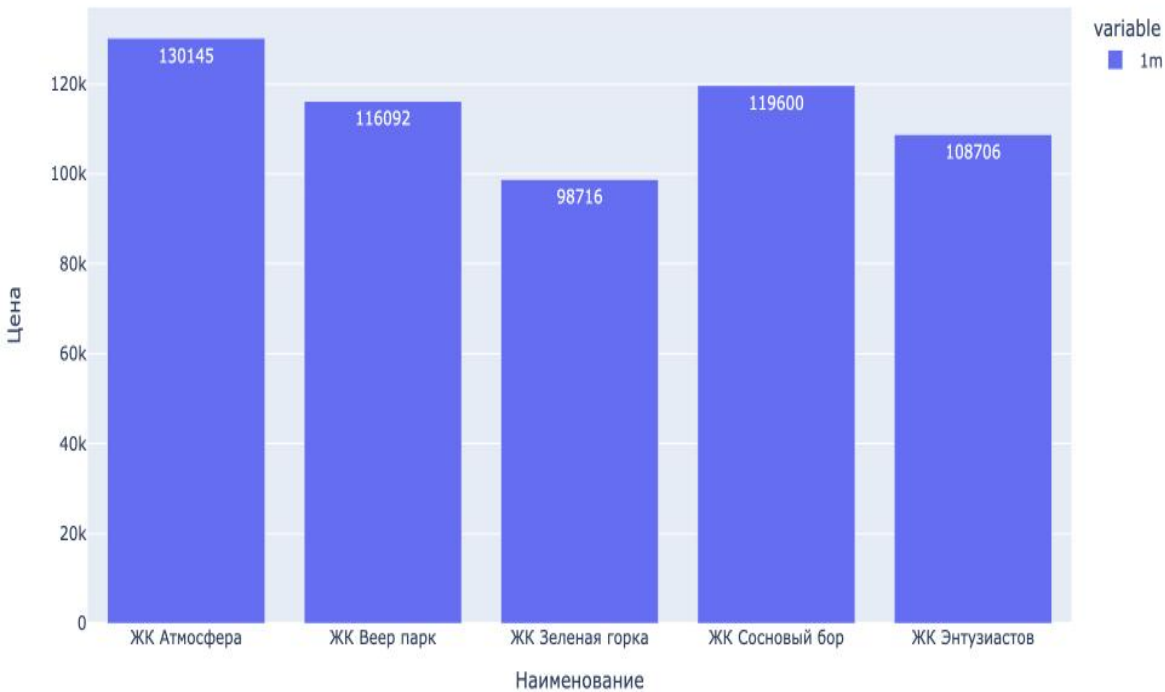


Предложение по ценам

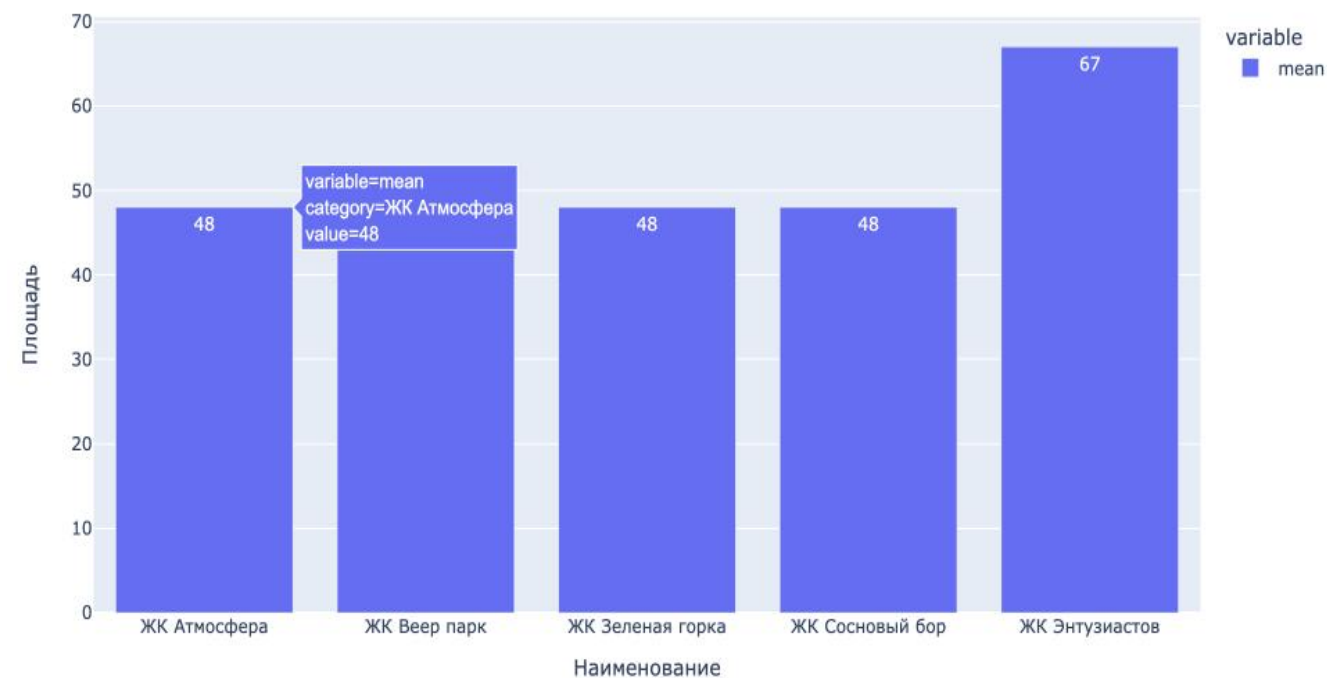
- Средняя цена по квадратному метру выше чем у конкурентов, средняя цена на втором месте после Энтузиастов, квартир в продаже у конкурентов, которые сдают объекты в 1ом квартале значительно меньше, что связано с более высокой ценой за метр
- Цена за объект складывается не только из цены у конкурентов, тут не учтены классы домов, целевая аудитория, инфраструктура и тд
- Исходя только из цен, площади и количества объектов в продаже выходит вывод, что цена завышена по сравнению с локальными конкурентами
- Решения о снижении цены принимаются на большем количестве факторов, чем 1 цена



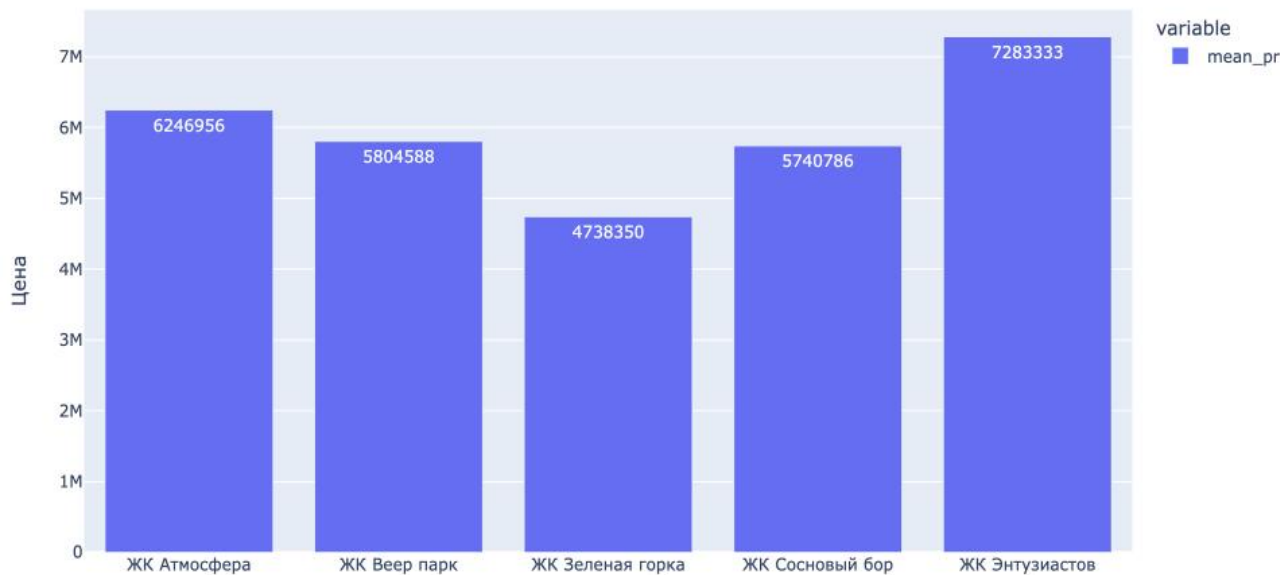
Средняя цена за 1м



Средняя площадь



Средняя цена



Так же посчитано библиотекой plotly