

1. Разработайте Google-форму, позволяющую планировать отпуск. Форма должна содержать следующие поля: страна, город, дата начала, дата окончания, планируемый бюджет.

**Информация об отпуске**

Описание

**Страна \***

Краткий ответ

**Город \***

Краткий ответ

**Дата начала \***

День, месяц, год

**Дата окончания \***

Дата

Обязательный вопрос

**Планируемый бюджет \***

Краткий ответ

[https://docs.google.com/forms/d/1mRiTclFgWvKKeXWM0pVpJlrbR\\_VwgRcPxAx2b5Za6OY/edit](https://docs.google.com/forms/d/1mRiTclFgWvKKeXWM0pVpJlrbR_VwgRcPxAx2b5Za6OY/edit)

Результаты должны сохраняться в Google-таблицу:

Регистрация на мероприятие (Ответы) ☆ 📁 ☁

Файл Правка Вид Вставка Формат Данные Инструменты Расширения Справка Последнее изменение: 4 мин...

100% 123 По умолча... 10 В I S A

	A	B	C	D	E	F
1	Отметка времени	Страна	Город	Дата начала	Дата окончания	Планируемый бюджет
2	23.12.2021 0:31:27	Россия	Москва	31.12.2021	07.01.2022	60000
3	23.12.2021 0:32:03	Россия	Сочи	03.01.2022	06.01.2022	55000
4	23.12.2021 0:32:43	Германия	Берлин	07.01.2022	09.01.2022	70000
5	23.12.2021 0:33:46	Греция	Афины	01.01.2022	16.01.2022	160000
6	23.12.2021 0:34:53	Италия	Рим	21.02.2022	25.02.2022	64000
7	23.12.2021 0:35:32	Россия	Казань	18.04.2022	24.04.2022	30000
8	23.12.2021 0:36:45	Франция	Париж	11.07.2022	24.08.2022	250000
9	23.12.2021 0:37:39	Беларусь	Минск	24.12.2021	26.12.2021	24000
10	23.12.2021 0:56:39	Австрия	Вена	11.05.2020	15.05.2020	42000
11	23.12.2021 0:57:30	Ирландия	Дублин	19.09.2019	22.09.2019	77000
12	23.12.2021 0:58:18	Минск	Брест	01.05.2019	12.05.2019	40000
13						

2,3. Добавьте в Google-таблицы столбцы для заполнения информации по фактической информации по отпуску. Столбцы должны содержать следующую информацию: стоимость билетов, стоимость отеля, траты на отдыхе, общая оценка отпуска по шкале от 1 до 5, столбец с итоговой стоимостью отпуска (сумма трех столбцов выше), отклонение итоговой стоимости от плана, стоимость одного дня отпуска.

Последние 3 столбца должны вычисляться с помощью формул:

Регистрация на мероприятие (Ответы) ☆ 📁 ☁

Файл Правка Вид Вставка Формат Данные Инструменты Расширения Справка Последнее изменение: 6 мин...

75% 123 По умолча... 10 В I S A

	A	B	C	D	E	F	G
1	Стоимость билетов	Стоимость отеля	Траты на отдыхе	Общая оценка отпуска 1-5	Итоговая стоимость отпуска	Отклонение итог. стоимости от плана %	Стоимость одного дня отпуска
2	20000	30000	10000	3	60000	0	4285.714286
3	18000	19000	9000	4	46000	-9.09090909	13000
4	45000	19000	7800	3	58000	-1.714285714	34400
5	27000	72000	14000	5	113000	-25.375	7533.333333
6	30000	28000	11000	3	69000	7.6125	17260
7	10000	18000	9000	2	37000	20	9000
8	38000	180000	90000	5	308000	20	9512.121212
9	9800	4000	4800	4	17000	-29.16666667	8500
10	27300	14380	9670	5	51350	23.36065038	12695
11	32065	22800	19600	5	74465	-3.27692377	24905
12	10700	14500	18000	5	38200	-4.5	3472.727273

4. Разработайте отчет (в Google-таблицах или Google Data-Studio) со следующими показателями: вывести самый дорогой отпуск (суммарно), среднее отклонение фактической цены отпуска от планируемой, рейтинга (таблица) стран по стоимости отеля за день, рейтинг (таблица) стран по стоимости трат на месте за день.

Отчет должен содержать фильтр по датам, а также по оценке:

Отчет без названия

Файл Изменить Просмотреть Вставить Страница Упорядочить Ресурс Справка Сбросить Предоставить доступ

Page 1 of 2

Добавить данные Добавить диаграмму Добавить элемент управления

Страна	Город	Дата начала	Дата окончания	Планируемая цена	Фактическая цена	Отклонение	Рейтинг	Стоимость отеля за день	Стоимость трат за день
Россия	Москва	31 дек. 2021 г.	7 янв. 2022 г.	20000	30000	10000	5	30000	6533.33
Россия	Казань	18 апр. 2022 г.	24 апр. 2022 г.	10000	18000	8000	2	36000	3900
Россия	Сочи	3 янв. 2022 г.	6 янв. 2022 г.	15000	15000	9000	4	39000	2750
Минск	Брест	1 мая 2019 г.	12 мая 2019 г.	10700	14500	13000	5	38200	2417.5
Италия	Рим	21 февр. 2022 г.	25 февр. 2022 г.	30000	28000	11000	3	69000	2250
Ирландия	Дублин	19 сент. 2019 г.	22 сент. 2019 г.	32000	22800	19600	5	74400	2113.64
Греция	Афины	1 янв. 2022 г.	16 янв. 2022 г.	27000	72000	14000	5	113000	1920.63
Германия	Берлин	7 янв. 2022 г.	9 янв. 2022 г.	45000	16000	7800	5	68800	1181.82
Беларусь	Минск	24 дек. 2021 г.	26 дек. 2021 г.	8500	4000	4500	4	17000	
Австрия	Вена	11 мая 2020 г.	15 мая 2020 г.	27300	14850	9670	5	51820	

1 - 11 / 11

Самый дорогой отпуск обошёлся в: 300 000,00 €

Среднее отклонение фактической цены от планируемой: -2,37

Дата начала: Выберите диапазон дат

Дата окончания: Выберите диапазон дат

Общая оценка отпуска 1-5

Страна	Стоимость отеля за день	Страна	Стоимость трат за день
1. Германия	8 000	1. Ирландия	6 533.33
2. Ирландия	7 600	2. Германия	3 900
3. Италия	7 000	3. Италия	2 750
4. Греция	4 800	4. Австрия	2 417.5
5. Россия	4 095.24	5. Беларусь	2 250
6. Франция	3 818.18	6. Франция	2 113.64
7. Австрия	3 712.5	7. Россия	1 920.63
8. Беларусь	2 000	8. Минск	1 181.82

1 - 9 / 9

Расход бюджета: какую долю в среднем занимают авиабилеты, проживание и траты на месте:

Отчет без названия

Файл Изменить Просмотреть Вставить Страница Упорядочить Ресурс Справка Сбросить Предоставить доступ

Page 2 of 2

Добавить данные Добавить диаграмму Добавить элемент управления

	Страна	Город	Стоимос...	Стоимос...	Траты на о...	Итоговая стои...	Доля стоимости билет...	Доля стоимости отеля	Доля трат
1.	Франция	Париж	39000	168000	93000	300000	13	56	31
2.	Россия	Москва	20000	30000	10000	60000	33.33333333333336	50	16.66666666666668
3.	Россия	Казань	10000	18000	8000	36000	27.77777777777778	50	22.22222222222222
4.	Россия	Сочи	15000	15000	9000	39000	38.46153846153846	38.46153846153846	23.076923076923077
5.	Минск	Брест	10700	14500	13000	38200	28.010471204188484	37.95811518324608	34.031413612565444
6.	Италия	Рим	30000	28000	11000	69000	43.47826086956522	40.57971014492754	15.942028985507246
7.	Ирландия	Дублин	32000	22800	19600	74400	43.01075268817205	30.645161290322584	26.34408602150538
8.	Греция	Афины	27000	72000	14000	113000	23.893805309734514	63.716814159292035	12.389380530973451

1 - 11 / 11

Средняя доля стоимости авиабилетов: 38,1

Средняя доля стоимости отеля: 40,25

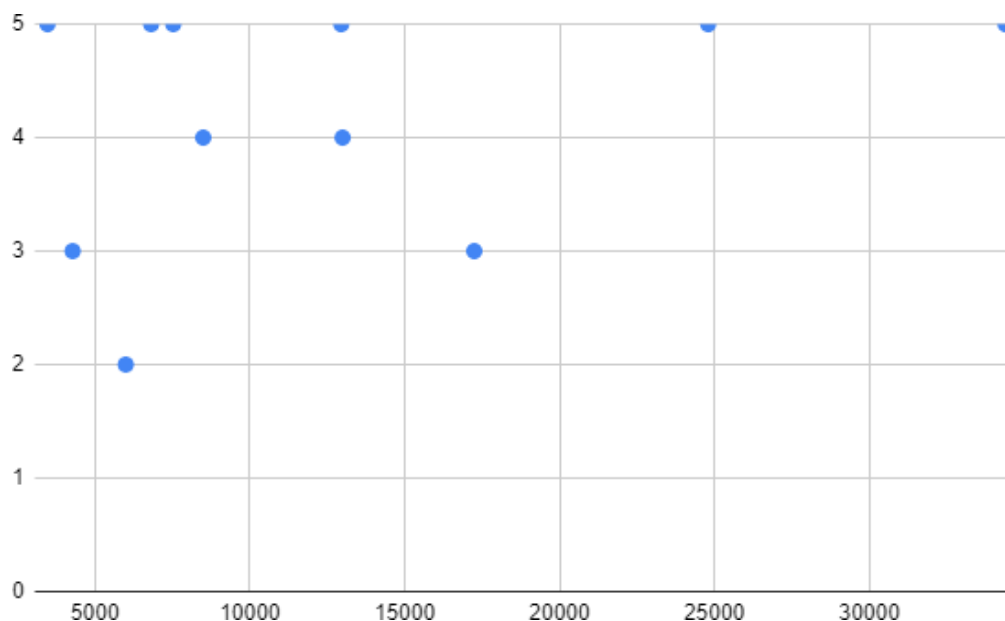
Средняя доля трат: 21,65

5. Рассчитайте корреляцию между ценой отпуска за день и оценкой за отзыв.

=КОРРЕЛ(M2:M12; J2:J12)

M	N
ного дня отпуска	Стоимость отеля за день
	Корреляция
	0,3015844751

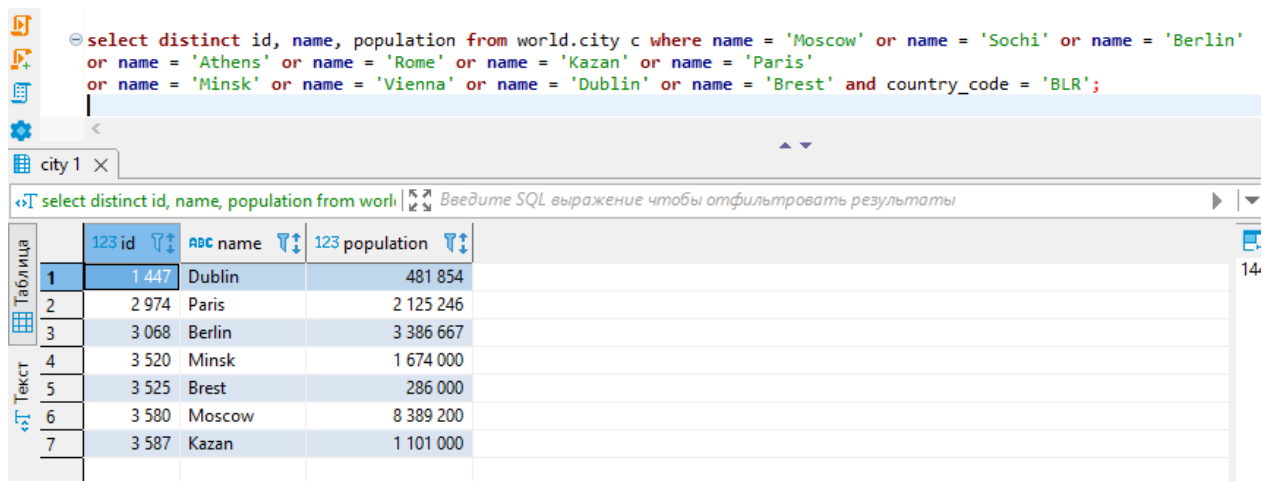
6. Постройте точечный график с этими показателями



7. Проверьте статистическую гипотезу, что вы хорошо планируете отпуска: среднее отклонение планируемых трат от реальных равно 0

	Статистическая гипотеза	
	"Я хорошо планирую отпуск"	
Дисперсия=	394,0563913	
Стандарт.отклон=	19,85085367	
Т-статика=	-0,3953658729	
Критическое значение=	2,228138852	
Гипотезу отклонить нельзя		

8. При помощи Python и SQL найдите, как называются ваши города в таблице City базы данных World-db



The screenshot shows a SQL query interface. The query is:

```
select distinct id, name, population from world.city c where name = 'Moscow' or name = 'Sochi' or name = 'Berlin' or name = 'Athens' or name = 'Rome' or name = 'Kazan' or name = 'Paris' or name = 'Minsk' or name = 'Vienna' or name = 'Dublin' or name = 'Brest' and country_code = 'BLR';
```

The results are displayed in a table with the following data:

	123 id	abc name	123 population
1	1 447	Dublin	481 854
2	2 974	Paris	2 125 246
3	3 068	Berlin	3 386 667
4	3 520	Minsk	1 674 000
5	3 525	Brest	286 000
6	3 580	Moscow	8 389 200
7	3 587	Kazan	1 101 000

9. Создайте соответствующий справочник на отдельном листе Google-таблиц

1	Город	Sity
2	Москва	Moscow
3	Сочи	Sochi
4	Берлин	Berlin
5	Афины	Athens
6	Рим	Rome
7	Казань	Kazan
8	Париж	Paris
9	Минск	Minsk
10	Вена	Vienna
11	Дублин	Dublin
12	Брест	Brest

## 10. Добавьте колонку с английским названием города при помощи функции VLOOKUP (ВПР)

=ВПР(C2; 'Справочник'!A2:B12; 2; ЛОЖЬ)

	В	С	Д	
ени	Страна	Город	Sity	Д
0:31:27	Россия	Москва	Moscow	
0:32:03	Россия	Сочи	Sochi	
0:32:43	Германия	Берлин	Berlin	
0:33:46	Греция	Афины	Athens	
0:34:53	Италия	Рим	Rome	
0:35:32	Россия	Казань	Kazan	
0:36:45	Франция	Париж	Paris	
0:37:39	Беларусь	Минск	Minsk	
0:56:39	Австрия	Вена	Vienna	
0:57:30	Ирландия	Дублин	Dublin	
0:58:18	Минск	Брест	Brest	

## 11. Скачайте данные в CSV и откройте их в Python

```
In [ ]: import pandas as pd
df=pd.read_csv("Регистрация на мероприятие (Ответы) - Ответы на форму (1) (1).csv")
dg=pd.read_csv("city_202112241519.csv")
```

```
In [ ]: df.head(6)
```

Out[ ]:

	Отметка времени	Страна	Город	Sity	Дата начала	Дата окончания	Планируемый бюджет	Стоимость билетов	Стоимость отеля	Траты на отдыхе	Общая оценка отпусков 1-5	Итоговая стоимость отпусков	Отклонение итог. стоимости от плана %	Стоимость одного дня отпусков	Стоимость отеля за день	Стоимость трат за день	Доля стоимости билетов	Доля стоимости отеля	Доля трат	Корреляция
0	23.12.2021 0:31:27	Россия	Москва	Moscow	31.12.2021	07.01.2022	60000.0	20000.0	30000.0	10000.0	3.0	60000.0	0	4285,714286	4285,714286	1428,571429	33,33333333	50	16,66666667	0,301584475
1	23.12.2021 0:32:03	Россия	Сочи	Sochi	03.01.2022	06.01.2022	55000.0	15000.0	15000.0	9000.0	4.0	39000.0	-29,09090909	13000	5000	3000	38,46153846	38,46153846	23,07692308	NaN
2	23.12.2021 0:32:43	Германия	Берлин	Berlin	07.01.2022	09.01.2022	70000.0	45000.0	16000.0	7800.0	5.0	68800.0	-1,714285714	34400	8000	3900	65,40697674	23,25581395	11,3372093	NaN
3	23.12.2021 0:33:46	Греция	Афины	Athens	01.01.2022	16.01.2022	160000.0	27000.0	72000.0	14000.0	5.0	113000.0	-29,375	7533,333333	4800	933,3333333	23,89380531	63,71681416	12,38938053	NaN
4	23.12.2021 0:34:53	Италия	Рим	Rome	21.02.2022	25.02.2022	64000.0	30000.0	28000.0	11000.0	3.0	69000.0	7,8125	17250	7000	2750	43,47826087	40,57971014	15,94202899	NaN
5	23.12.2021 0:35:32	Россия	Казань	Kazan	18.04.2022	24.04.2022	30000.0	10000.0	18000.0	8000.0	2.0	36000.0	20	6000	3000	1333,333333	27,77777778	50	22,22222222	NaN



12. При помощи SQL и Python получите датафрейм с названием города и его населением из таблицы City

```
In [ ]: dg.head(8)
```

```
Out[ ]:
```

	id	name	population
0	1447	Dublin	481854
1	2974	Paris	2125246
2	3068	Berlin	3386667
3	3520	Minsk	1674000
4	3525	Brest	286000
5	3580	Moscow	8389200
6	3587	Kazan	1101000

13. При помощи Python соедините данные из выгруженного CSV файла и таблицы с населением гор

df.merge(dg, left_on='Sity', right_on='name', how='inner')										
Отметка времени	Страна	Город	Sity	Дата начала	Дата окончания	Планируемый бюджет	Стоимость билетов	Стоимость отеля	Траты на отдыхе	Общая оценка отпуска 1-5
0 23.12.2021 0:31:27	Россия	Москва	Moscow	31.12.2021	07.01.2022	60000.0	20000.0	30000.0	10000.0	3.0
1 23.12.2021 0:32:43	Германия	Берлин	Berlin	07.01.2022	09.01.2022	70000.0	45000.0	16000.0	7800.0	5.0
2 23.12.2021 0:35:32	Россия	Казань	Kazan	18.04.2022	24.04.2022	30000.0	10000.0	18000.0	8000.0	2.0
3 23.12.2021 0:36:45	Франция	Париж	Paris	11.07.2022	24.08.2022	250000.0	39000.0	168000.0	93000.0	5.0
4 23.12.2021 0:37:39	Беларусь	Минск	Minsk	24.12.2021	26.12.2021	24000.0	8500.0	4000.0	4500.0	4.0
5 23.12.2021 0:57:30	Ирландия	Дублин	Dublin	19.09.2019	22.09.2019	77000.0	32000.0	22800.0	19600.0	5.0
6 23.12.2021 0:58:18	Минск	Брест	Brest	01.05.2019	12.05.2019	40000.0	10700.0	14500.0	13000.0	5.0

df.merge(dg, left_on='Sity', right_on='name', how='inner')										
Отклонение итог. стоимости от плана %	Стоимость одного дня отпуска	Стоимость отеля за день	Стоимость трат за день	Доля стоимости билетов	Доля стоимости отеля	Доля трат	Корреляция	Id	name	population
0	4285.714286	4285.714286	1428.571429	33.33333333	50	16.66666667	0.3015844751	3580	Moscow	8389200
-1.714285714	34400	8000	3900	65.40697674	23.25581395	11.3372093	NaN	3068	Berlin	3386667
20	6000	3000	1333.333333	27.77777778	50	22.22222222	NaN	3587	Kazan	1101000
20	6818.181818	3618.181818	2113.636364	13	56	31	NaN	2974	Paris	2125246
-29.16666667	8500	2000	2250	50	23.52941176	26.47058824	NaN	3520	Minsk	1674000
-3.376623377	24800	7600	6533.333333	43.01075269	30.64516129	26.34408602	NaN	1447	Dublin	481854
-4.5	3472.727273	1318.181818	1181.818182	28.0104712	37.95811518	34.03141361	NaN	3525	Brest	286000



14. Сгруппируйте итоговый датафрейм по странам и рассчитайте среднюю численность населения в городах, в которых вы отдыхали

```
In [ ]: ddf = df.merge(dg, left_on='City', right_on='name', how='inner')
ddf[['Страна', 'population']].groupby(['Страна']).mean()
```

Out[ ]: **population**

Страна	
Беларусь	1674000
Германия	3386667
Ирландия	481854
Минск	286000
Россия	4745100
Франция	2125246

---

Ссылка на среду:

[https://github.com/JustRuslashka/-  
Just\\_Ruslashka/blob/main/%22Untitled\\_ipynb%22.ipynb](https://github.com/JustRuslashka/-Just_Ruslashka/blob/main/%22Untitled_ipynb%22.ipynb)