

Задание

Подготовить пример предметной области (описание сущностей и мер + проект в PowerBI):

- Сделать загрузку из файлов по меньшей мере двух форматов
- Добавить по меньшей мере три преобразования на уровне загрузки данных
- Подготовить по меньшей мере 4 меры
- Использовать по меньшей мере 5 DAX различных функций
- Подготовить Дашборд используя не менее 4-х различных визуальных элементов
- На оценку хорошо, одна таблица фактов, на оценку отлично 1 таблицы фактов и схема снежинка
- Результат в виде power bi отчёта + файл с описанием, предметной области мер и используемых функций и преобразований требуемых в задании

Выполнила: Калмыкова М.А.

Группа: МИВТ-24-3-1

Предметная область

Тема: "Управление бронированием отеля"

- Общие показатели бизнеса (доход, количество бронирований, процент отмен)
- Анализ стоимости номеров/количества бронирований во времени
- Анализ каналов привлечения клиентов
- Анализ состава гостей (в разрезе взрослые/дети)
- Анализ взаимосвязи между отказом от бронирования и временем, оставшимся до даты заезда
- Анализ влияния типа питания на цены/продолжительность пребывания в отеле

Условности

- Цены в датасете даны в евро, но для интереса попробуем перевести их согласно соответствующему курсу 2017-2018 годы.
- Данные взяты с сайт ЦБ РФ (Таблица 1 соответствующего файла). Данные были предобработаны, т.к. имелись пропуски.

Ссылка на датасет: <https://www.kaggle.com/datasets/ahmedwaelnasef/hotel-booking/data>

Файлы

Hotel.csv

Основной файл с данными бронирования (номер бронирования, количество взрослых и детей, количество ночей, цена за ночь и так далее)

meal.txt

Дополнительные данные по типу питания

RC_F01_01_2017_T31_12_2018 (1).xlsx

Данные курса евро к рублю за 2017-2018 годы по курсу ЦБ РФ

Преобразование данных

Для файла meal.txt:

- Использовать первую строку в качестве заголовков
- Удалить лишние пробелы в столбце name
- Пустое значение изменено на новое

A ^B _C Column1	A ^B _C Column2
plan	name
Meal Plan 1	Без питания
Meal Plan 2	Завтрак
Meal Plan 3	Полупансион
Not Selected	



A ^B _C plan	A ^B _C name
Meal Plan 1	Без питания
Meal Plan 2	Завтрак
Meal Plan 3	Полупансион
Not Selected	Не выбрано

Преобразование данных

Для файла CourseEuro.xlsx:

- Заполнить null вниз в столбце curs

1.2 curs	
	63,1125
	null
	null
	null
	null
	null
	null
	null
	63,1125
	63,5445

[illegible]

Преобразование данных

Для файла Hotel.csv:

- Объединить столбцы день, месяц, год в единую дату
- Установить тип нового столбца - дата
- Т.к. ночи в выходные и будние дни считаются в отдельных столбцах -> добавить столбец общего количества ночей
- Переупорядочить столбцы для удобства
- Отфильтровать записи где 0 ночей
- Изменить тип на логический в столбцах car_parking_space и repeated_guest
- Удалить ошибки в датах
- Заменить нулевые цены на null (это потери данных)
- Заменить точки на запятые и преобразовать столбец price к десятичным числам

Преобразование данных

Пример для файла Hotel.csv

1^2_3 year	1^2_3 month	1^2_3 date
2017	10	2
2018	11	6
2018	2	28
2018	5	20
2018	4	11
2018	9	13
2017	10	15

[illegible]

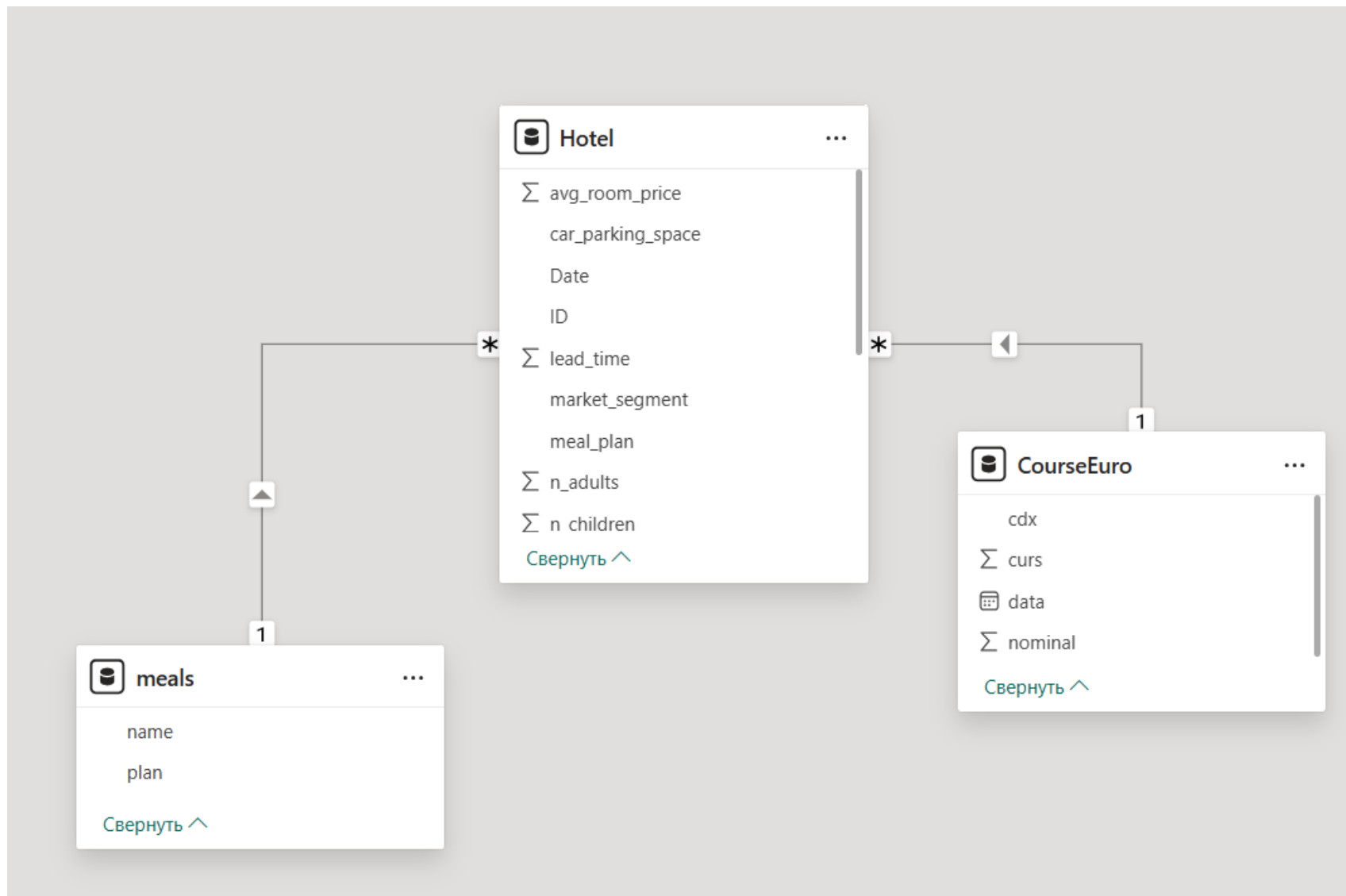
Представление таблиц

- Переименование численных столбцов в валютные там где это необходимо

avg_room_price ▾
76,50 €
76,50 €
68,85 €
76,50 €
76,50 €
55,96 €
80,00 €
80,00 €
85,50 €
76,50 €
90,00 €

nominal ▾	data ▾	curs ▾	cdx ▾
1,00 €	1 января 2017 г.	63,11 ₪	Евро
1,00 €	2 января 2017 г.	63,11 ₪	Евро
1,00 €	3 января 2017 г.	63,11 ₪	Евро
1,00 €	4 января 2017 г.	63,11 ₪	Евро
1,00 €	5 января 2017 г.	63,11 ₪	Евро
1,00 €	6 января 2017 г.	63,11 ₪	Евро
1,00 €	7 января 2017 г.	63,11 ₪	Евро

Представление модели



Вычисляемые столбцы

Пересчитаем цену за номер в рублях, используя таблицу курса евро от ЦБ РФ по датам за 2017-2018 года.

Цена =

```
VAR exchangeRate =  
    LOOKUPVALUE(  
        CourseEuro[curs],  
        CourseEuro[data], 'Hotel'[Date]  
    )  
RETURN  
    'Hotel'[avg_room_price] * exchangeRate
```

avg_room_price ▾	special_requests ▾	status ▾	Цена ▾
76,50 €	1	Canceled	5 187,25 ₽
76,50 €	1	Canceled	5 274,45 ₽
68,85 €	1	Canceled	4 747,00 ₽
76,50 €	1	Canceled	5 229,52 ₽
76,50 €	1	Canceled	5 252,67 ₽

Вычисляемые столбцы

Распределим бронирования по группам, в зависимости от разницы между днем бронирования и датой приезда.

Дней до =

```
SWITCH(  
    TRUE(),  
    Hotel[lead_time] <= 7, "0-7 дней",  
    Hotel[lead_time] <= 14, "8-14 дней",  
    Hotel[lead_time] <= 30, "15-30 дней",  
    "30+ дней"  
)
```

Дней до ▾
30+ дней
30+ дней
30+ дней
30+ дней
30+ дней
30+ дней
0-7 дней
0-7 дней
30+ дней
30+ дней
0-7 дней
0-7 дней
0-7 дней

Меры

Общее количество бронирований = `COUNT(Hotel[ID])`

Общий доход =

```
SUMX(  
    'Hotel',  
    'Hotel'[avg_room_price] * 'Hotel'[nights]  
)
```

Процент отмен =

```
DIVIDE(  
    CALCULATE(COUNT(Hotel[ID]), Hotel[status] = "Canceled"),  
    COUNT(Hotel[ID])  
)
```

Среднее количество ночей = `AVERAGE(Hotel[nights])`

Год

☒ 2017

☒ 2018

11,34 МЛН

Общий доход

36 ТЫС.

Общее количество бронирований

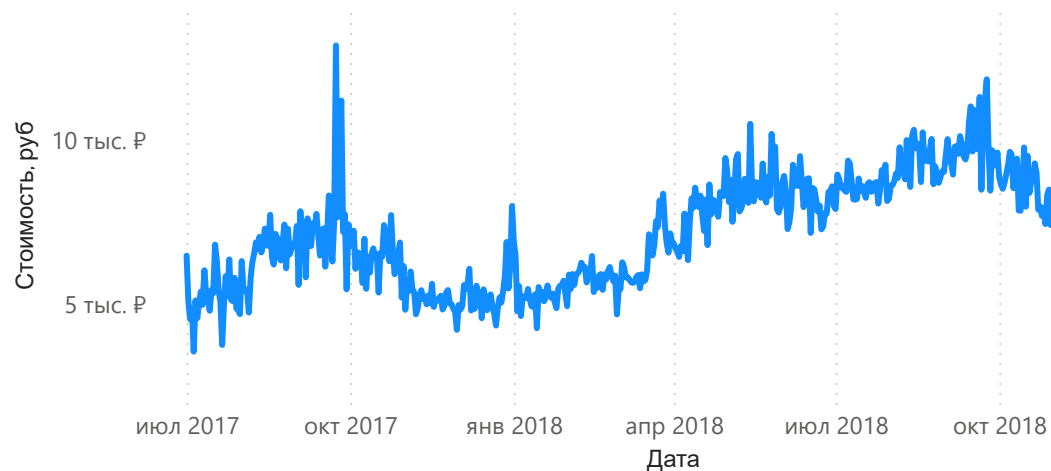
32,84%

Процент отмен

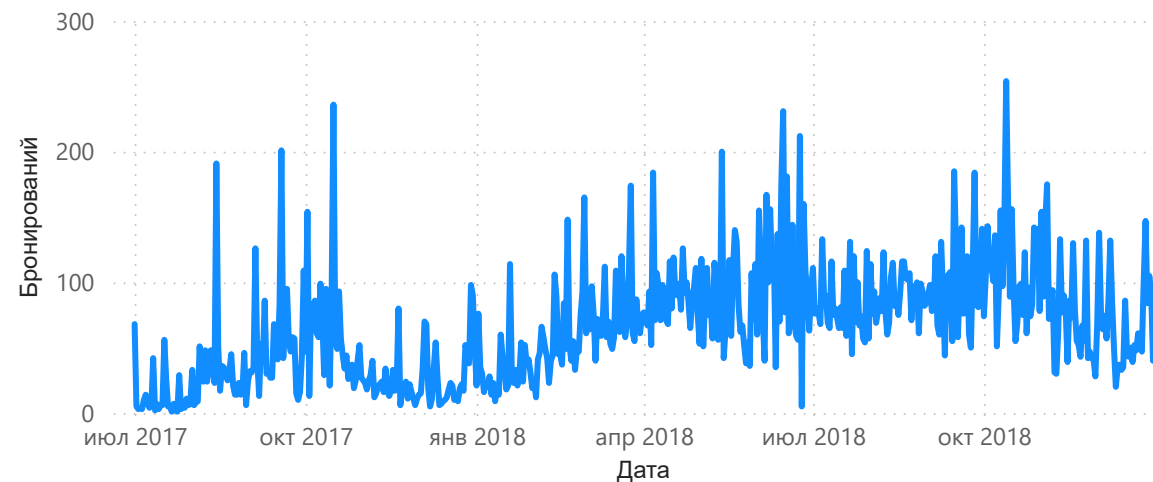
3

Среднее количество ночей

Средняя цена за ночь



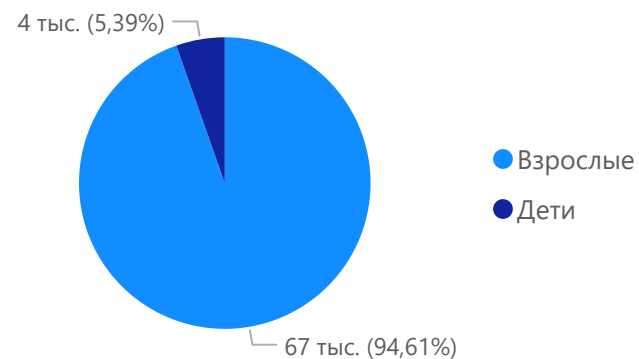
Количество бронирований



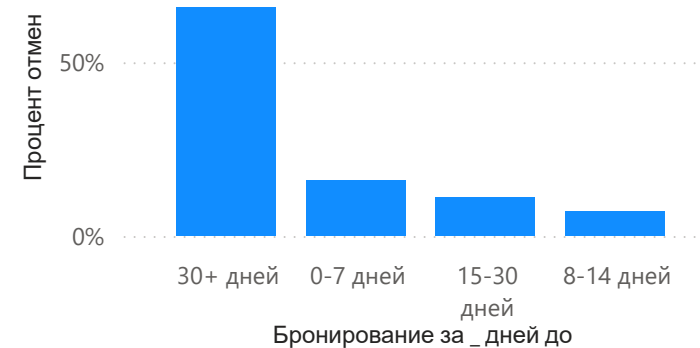
Каналы привлечения клиентов

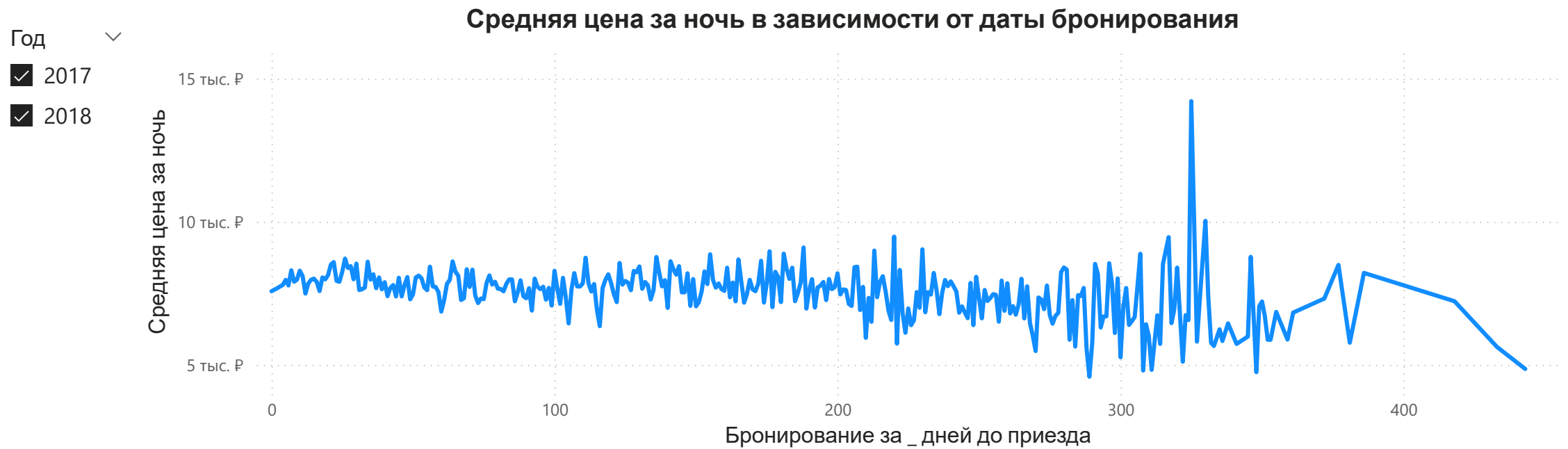


Состав гостей

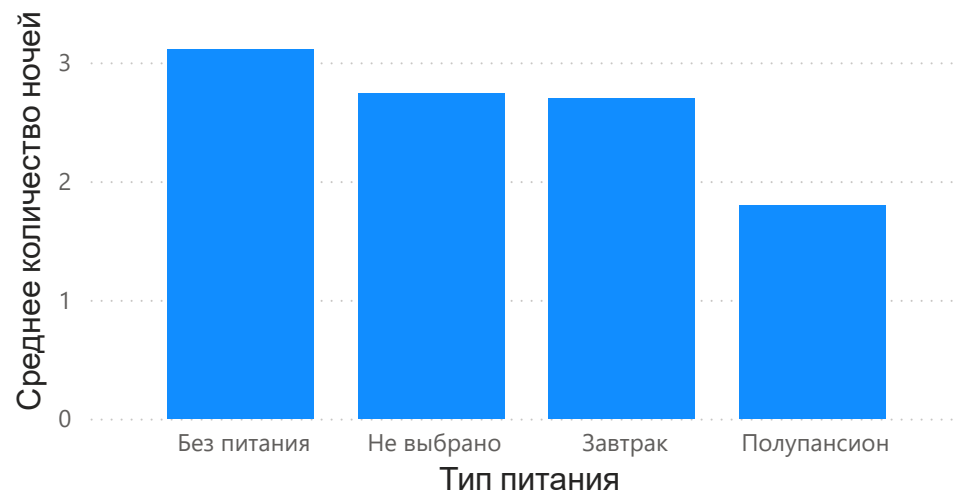


Процент отказов





Продолжительность визита в зависимости от типа питания



Средняя цена за ночь в зависимости от типа питания

