

Финальная работа по модулю “SQL и получение данных”

Мария Рахман

Описание бизнес области

Датасет принадлежит московской сети антикафе Time Club. Антикафе работает 24/7, целевая аудитория - молодежь, студенты и школьники.

В Time Club действует система клубных карт - они нужны для идентификации посетителя клуба и расчета бонусов. Клиенты могут получить карту на одно посещение (стать гостем), либо пройти регистрацию и получить именную карту (стать членом клуба). По карте рассчитывается длительность посещения. На входе и выходе посетитель проводит картой через сканер - так фиксируется время начала и конца посещения. У именных клубных карт есть несколько уровней, которые дают владельцу бонусы.

Уровни клубных карт				
Гостевая	Бронзовая	Серебряная	Золотая	Платиновая

Описание базы данных

Начало датасета - 01.06.2018, конец - 23.11.2018. Датасет состоит из 5 таблиц:

1. Booking - бронирования
2. Client - клиенты
3. Expense - расходы клиента после посещения
4. Shift - смены сотрудников
5. Visit - посещения клиентов

Структура и связи базы данных

booking
booking_id
client_id
place_id
creation_utc
employee_id

client
client_id
card_level
creator_employee_id
creation_utc
creation_place_id

expense
expense_id
place_id
client_id
amount
creator_employee_id
creation_utc

shift
shift_id
place_id
employee_id
start_utc
finish_utc

visit
visit_id
client_id
place_id
start_utc
finish_utc

Таблицы связаны по 3 полям:

client_id - уникальный id клиента

employee_id - уникальный id клуба

place_id - уникальный id сотрудника

Структура таблиц

Таблица Booking - Бронирования

связана с таблицей клиентов, клубов и сотрудников

booking_id	client_id	place_id	creation_utc	employee_id
уникальный id брони	уникальный id клиента	уникальный id клуба	время создания	уникальный id сотрудника

Таблица Client - Клиенты

связана с таблицей клиентов, клубов и сотрудников

client_id	card_level	creator_employee_id	creation_utc	creation_place_id
уникальный id клиента	уровень клубной карты*	уникальный id сотрудника	время создания	уникальный id клуба**

* Уровни клубной карты	
id	name
1	Гостевая
2	Бронзовая
3	Серебряная
4	Золотая
5	Платиновая

** Клубы сети	
id	name
1	Гнездо
2	Реформа (закрылся, нет в выборке)
3	Убежище

Структура таблиц

Таблица Expense - Расходы

связана с таблицей клиентов, клубов и сотрудников

expense_id	place_id	client_id	amount	creator_employee_id	creation_utc
уникальный id расхода	уникальный id клуба	уникальный id клиента	сумма расхода	уникальный id сотрудника	время создания

Таблица Shift - Смены сотрудников

связана с таблицей клиентов, клубов и сотрудников

shift_id	place_id	employee_id	start_utc	finish_utc
уникальный id смены	уникальный id клуба	уникальный id сотрудника	время начала	время окончания

Таблица Visit - Посещения клиентов

связана с таблицей клиентов, клубов и сотрудников

visit_id	client_id	place_id	start_utc	finish_utc
уникальный id посещения	уникальный id клиента	уникальный id клуба	время начала	время окончания

Запросы к базе данных

1. Показать топ 10 клиентов, которые потратили больше всего денег в антикафе [expense]
2. Показать число посещений и выручку за ноябрь 2018 по каждому антикафе [expense + visit]
3. Показать, какую выручку в % приносят клиенты с каждым типом карты (card_level) [expense + client]
4. Кто из сотрудников зарегистрировал больше всего клиентов, топ 10 [client]
5. Выбрать всех клиентов, которых зарегали первые 3 сотрудника (из предыдущего запроса). Посмотреть количество посещений по каждому сотруднику отдельно у топ 40 клиентов [client + visit]
6. Показать долю посещений клиентов с гостевой картой от общего числа посещений по каждому антикафе [client + visit]
7. Показать, клиенты с каким уровнем карты чаще всего бронируют комнаты, по убыванию [booking + client]
8. Топ 10 количества новых клиентов за смену [client + shift]
9. Показать среднее количество посещений за последние 3 месяца в выходной день и в рабочий [visit]
10. Среднее количество гостей за утреннюю, дневную, вечернюю смену за последний месяц по каждому кафе [visit]