

Данные задания предназначены для оценки технических и аналитических навыков кандидата.

Задания №1 и №3 предполагают описание аналитических подходов к решению задач, задание №2 – хорошо читаемый SQL запрос.

1. Требуется спрогнозировать суммарный объем начисленных миль всем участникам программы лояльности в 2023 году.
  - Приведите структуру факторов, которые необходимо учесть при прогнозировании.
  - Опишите подход / алгоритм прогнозирования объема начисленных миль (будет плюсом, если опишете, каким образом будет оценено влияние независимых факторов на целевую переменную). Ответ поясните.
2. Используя диалект Oracle SQL, сформировать запрос к хранилищу данных, позволяющий получить номера участников в программе лояльности, фактически потративших на авиабилеты более 50 000 миль за 2022 год и имевших до совершения первой полетной транзакции (Product\_nm = "Ticket") в 2022 году более 40 000 миль на счету. Требуется вывести номер участника в программе, кол-во транзакций списания на билеты и среднее количество миль, списанных на билеты, в 2022 году.  
Структура необходимых таблиц в хранилище приведена ниже.

Таблица с участниками:

Client id	Client number
Client id 1	Client number 1
Client id 2	Client number 2
Client id 3	Client number 3
...	...

Таблица с транзакциями:

Client id	Txn ID	Txn_date	Txn_type	TXN_sts	Product id	Miles_sum	Client mile balance_after txn
Client id 1	Txn ID 1	01.01.2021	Accrual	Cancelled	Product id 1	Miles_sum 1	Balance_after txn 1
Client id 2	Txn ID 2	27.03.2021	Redemption	Processed	Product id 2	Miles_sum 1	Balance_after txn 2
Client id 3	Txn ID 3	12.11.2020	Redemption	Cancelled	Product id 3	Miles_sum 1	Balance_after txn 3
...	...	...	...	...	...	...	...

Txn\_date – дата транзакции,

Txn\_type – тип транзакции (redemption – списание, accrual – начисление),

Txn\_sts – статус транзакции (Processed – обработанная, cancelled – отменённая),

Product\_id – идентификатор продукта,

Miles\_sum – количество миль, начисленное / списанное за данную транзакцию,

Client\_mile\_balance\_after\_txn – кол-во миль на счету участника после совершения транзакции.

Таблица с продуктами:

Product id	Product_nm
Product id 1	Ticket
Product id 2	Bank card
Product id 3	Hotel
...	...

Product\_nm – название продукта (Ticket – авиабилеты. За них начисляются мили, на них можно списывать мили; Bank card – партнёрская банковская карта, мили начисляются за покупки по этой карте; Hotel – партнёры-отели, мили начисляются за оплату номеров в этом отеле, а также мили можно списывать на оплату отелей.

3. Требуется разработать сегмент для рассылки по стимулированию совершения полетов на Дальний Восток (Владивосток, Хабаровск, Южно-Сахалинск, Петропавловск-Камчатский) в период июнь-август путем предоставления промокода со скидкой на покупку билета.
  - Опишите алгоритм формирования выборки участников программы в сегмент так, чтобы коммуникация была получена наиболее потенциально заинтересованными в предложении, а также приведите алгоритм оценки эффективности коммуникации.

Решение:

1)

Структура факторов для прогнозирования суммарного объема начисленных миль в программе лояльности может включать следующее:

1. Количество участников программы лояльности
2. Уровень удовлетворенности клиентов от программы лояльности
3. Уровень конкуренции на рынке авиаперевозок
4. Экономические показатели страны и мировой экономики (например, ВВП, инфляция, курс валют)
5. Тенденции изменения цен на авиабилеты и услуги авиакомпании
6. Количество рекламных и маркетинговых мероприятий компании

При прогнозировании объема начисленных миль можно использовать различные методы, в том числе:

1. Аналитическое моделирование - разработка математической модели на основе исторических данных о начисленных милях и соответствующих факторов для прогнозирования будущих результатов.
2. Регрессионный анализ - оценка связи между объемом начисленных миль и независимыми факторами с помощью одной или нескольких регрессионных моделей.
3. Методы машинного обучения - использование алгоритмов машинного обучения, таких как случайный лес или нейронные сети, для построения модели прогнозирования.

Для оценки влияния независимых факторов на целевую переменную, в данном случае, объем начисленных миль, можно использовать коэффициенты регрессии или важность признаков, вычисленные с помощью алгоритмов машинного обучения. При аналитическом моделировании можно провести анализ дисперсии, чтобы определить наиболее значимые факторы. В целом, использование различных подходов к прогнозированию позволяет получить более точный прогноз и учесть влияние различных факторов.

2)

```
SELECT t1.client_id, COUNT(t1.txn_id) AS number_of_transactions
AVG(t1.miles_sum) AS avg_miles_spent
FROM transactions t1
```

```

INNER JOIN (
    SELECT client_id, SUM(miles_sum) AS total_miles
    FROM transactions
    WHERE product_id = 'Product id 1' AND txn_date >= '01-JAN-22'
    GROUP BY client_id
    HAVING SUM(miles_sum) > 50000
)
t2 ON t1.client_id = t2.client_id
INNER JOIN (
    SELECT client_id, MIN(txn_date) AS first_ticket_txn_date
    FROM transactions
    WHERE product_id = 'Product id 1' AND txn_date >= '01-JAN-22'
    GROUP BY client_id)
t3 ON t1.client_id = t3.client_id
WHERE t1.product_id = 'Product id 1' AND t1.txn_date >= t3.first_ticket_txn_date AND t1.txn_type = 'Redemption'
AND t1.txn_sts = 'Processed'
AND t1.client_mile_balance_after_txn >= 40000
GROUP BY t1.client_id;

```

---

3)

Алгоритм формирования выборки участников программы в сегмент может включать следующие шаги:

1. Определение критериев сегментации: для формирования выборки участников программы следует определить критерии сегментации, включая местоположение (наличие регулярных рейсов на Дальний Восток), предпочтения путешествий и другие соответствующие параметры.
2. Сбор необходимых данных: на основе критериев сегментации необходимо собрать данные о каждом участнике программы лояльности.
3. Фильтрация и отбор кандидатов: на основании данных, полученных на предыдущем шаге, необходимо отфильтровать кандидатов, которые соответствуют критериям сегментации.
4. Сегментация участников: на основании данных, собранных на предыдущих шагах, необходимо произвести сегментацию участников программы лояльности в соответствии с критериями сегментации.
5. Выборка кандидатов для рассылки: на основании сегментации участников программы лояльности необходимо выбрать подходящих кандидатов для рассылки коммуникации.

Алгоритм оценки эффективности коммуникации может включать следующие шаги:

1. Определение целевого показателя (например, количество проданных билетов на рейсы на Дальний Восток).
2. Сбор данных до рассылки: сбор данных о показателе до проведения рассылки коммуникации.
3. Рассылка коммуникации: проведение коммуникации (например, рассылка промокода со скидкой на покупку билетов).
4. Сбор данных после рассылки: сбор данных о показателе после проведения рассылки коммуникации.
5. Оценка эффективности: сравнение показателей до и после рассылки коммуникации для определения эффективности коммуникации. В качестве дополнительного показателя можно использовать отклик на рассылку (например, количество использованных промокодов со скидкой на покупку билетов).

