12.07.2024

Executive Summary for Traveltide.

1. Introduction

Elena Tarrant, the new Head of Marketing at TravelTide, is tasked with designing and executing a personalized rewards program to enhance customer retention. The aim is to leverage customer insights to tailor rewards that encourage repeat business and build customer loyalty. This summary provides an overview of the project objectives, methodologies, findings, and recommendations for the rewards program.

2. Project Objectives

The primary objectives of the project are:

- To identify customer segments with distinct behaviors and preferences.
- To design personalized rewards that resonate with each customer segment.
- To implement a data-driven approach to enhance the effectiveness of the rewards program.

3. Methodology

To achieve these objectives, the following steps were undertaken:

1. **Exploratory Data Analysis (EDA)**: Understanding the business context and available data, and cleaning and preparing the data for analysis.
2. **Feature Engineering**: Creating new attributes and metrics to better capture customer characteristics.
3. **Customer Segmentation**: Grouping customers based on their behaviors and preferences using clustering methods and predefined criteria.
4. **Rewards Assignment**: Matching each customer segment with the most suitable reward.

4. Data Overview

The analysis was performed on a dataset containing customer interactions, including:

- Session details: duration, clicks, discounts utilized.
- Trip details: flights booked, hotel stays, cancellations.
- Demographic details: age, marital status, children.
- Financial details: money spent on hotels and flights.

5. Key Metrics and Features

Several key metrics were derived from the data to facilitate customer segmentation:

- **Session Metrics**: Number of sessions, average session duration, number of clicks.
- **Trip Metrics**: Number of trips, trip length, distance traveled, money spent.
- **Demographic Metrics**: Age, marital status, presence of children.

6. Customer Segmentation

Customers were segmented into distinct groups based on their behaviors and preferences:

- **Solo Travelers**: Typically single individuals with low spend and fewer trips.
- **Couples**: Married or partnered individuals with moderate spend and trips.
- **Families**: Households with children, characterized by higher spend and longer trips.
- **Business Travelers**: Individuals with frequent short trips and higher spend on flights.
- **Groups**: Larger groups traveling together, often with high spend on accommodations.

7. Rewards Design

Each customer segment was matched with personalized rewards to maximize engagement:

- **Solo Travelers**: No cancellation fees.
- **Couples**: Exclusive discounts and adventure tours.
- **Families**: No cancellation fees and family adventure tours.
- **Business Travelers**: Free checked bag.
- **Groups**: One night free hotel stay with meals.

8. Findings

The analysis revealed significant differences in customer behavior across segments, justifying the need for personalized rewards. Key findings include:

- **High engagement**: Families and groups showed the highest engagement and spend.
- **Value sensitivity**: Solo travelers and business travelers were more sensitive to specific perks like cancellation policies and checked bags.

9. Recommendations

To implement the rewards program effectively, the following steps are recommended:

- **Personalized Communication**: Tailor marketing messages to highlight the most relevant perks for each customer segment.
- **Iterative Testing**: Continuously test and refine the rewards program based on customer feedback and engagement metrics.
- **Performance Monitoring**: Track the performance of the rewards program using KPIs such as sign-up rates, repeat business, and customer satisfaction.

10. Conclusion

The personalized rewards program, backed by solid data analysis, promises to enhance customer retention and loyalty at TravelTide. By understanding customer segments and their preferences, TravelTide can offer targeted rewards that resonate with customers, driving repeat business and long-term engagement.

SQL TRAVELTIDE

```
WITH filtered_users AS
( SELECT user_id, COUNT(*) from sessions s
 where s.session_start > '2023-01-04'
 GROUP BY user_id
 HAVING count(session_start)>7
 ),
session_base AS (select  s.session_id, s.user_id, s.trip_id, s.session_start,
s.session_end,
 EXTRACT(EPOCH FROM (session_end - session_start)) AS session_duration,
 s.page_clicks,  s.flight_discount, s.flight_discount_amount, s.hotel_discount,
s.hotel_discount_amount,
  s.flight_booked, s.hotel_booked, s.cancellation,  u.birthdate, u.gender, u.married,
 u.has_children, u.home_country, u.home_city, u.home_airport, u.home_airport_lat,
  u.home_airport_lon,u.sign_up_date,  f.origin_airport, f.destination, f.destination_airport,
 f.seats, f.return_flight_booked, f.departure_time, f.return_time, f.checked_bags,
  f.trip_airline, f.destination_airport_lat, f.destination_airport_lon,f.base_fare_usd,
h.hotel_name,
 CASE WHEN h.nights<=0 then 1 else h.nights END AS nights,
 h.rooms, h.check_in_time, h.check_out_time, h.hotel_per_room_usd AS
hotel_price_per_room_night_usd from filtered_users fu left join users u
on fu.user_id = u.user_id
left join sessions s
ON u.user_id = s.user_id
left join flights f
on s.trip_id = f.trip_id
left join hotels h
on s. trip_id = h.trip_id
where s.session_start > '2023-01-04'
AND s.user_id IN (SELECT user_id FROM filtered_users)
),
canceled_trips AS (
SELECT DISTINCT trip_id
FROM session_base
 WHERE cancellation = TRUE
),
```

```
not_canceled_trips AS(
SELECT *
FROM session_base
 WHERE trip_id IS NOT NULL AND trip_id NOT IN (Select trip_id FROM canceled_trips)
),
user_base_session AS (
 Select user_id,
  SUM (page_clicks) AS num_clicks,
  COUNT(distinct session_id) as num_sessions,
  ROUND(AVG(session_duration),2) as avg_session_duration
 From session_base
  GROUP BY user_id
),
user_base_trip AS (
 SELECT user_id,
  COUNT(DISTINCT trip_id) AS num_trips,
  SUM(CASE WHEN flight_booked = TRUE AND return_flight_booked = TRUE THEN 2
     WHEN flight_booked = TRUE THEN 1
     ELSE 0  END) as num_flights2,
  COALESCE(SUM(hotel_price_per_room_night_usd*nights*rooms)) AS
hotel_original_price,
  COALESCE((SUM((hotel_price_per_room_night_usd*nights*rooms) *
                  (1- (CASE WHEN hotel_discount_amount IS NULL THEN 0 ELSE
hotel_discount_amount END)))),0) AS discounted_hotel,
  COALESCE(SUM(base_fare_usd)) AS flight_original_price,
          COALESCE((SUM((base_fare_usd) *
                  (1- (CASE WHEN flight_discount_amount IS NULL THEN 0 ELSE
flight_discount_amount END)))),0) AS discounted_flight,
  COUNT(checked_bags) as checkedIn_bags,
  COUNT(seats) as num_seats,
  SUM(DATE_PART('day', departure_time - session_end)) AS days_until_departure,
  ROUND(SUM(EXTRACT(EPOCH FROM (return_time - departure_time)) / 86400)) AS
trip_length_days,
  COALESCE(SUM(base_fare_usd)) AS flight_without_discount,
  SUM(6371 * acos(
       cos(radians(home_airport_lat)) * cos(radians(destination_airport_lat)) *
       cos(radians(destination_airport_lon) - radians(home_airport_lon)) +
       sin(radians(home_airport_lat)) * sin(radians(destination_airport_lat))
    )) AS distance_km,
  SUM(nights) AS nights,
  SUM(rooms) AS rooms
FROM not_canceled_trips
          GROUP BY user_id
  ),
```

```
Persona_check AS (
SELECT b.*,
  EXTRACT(YEAR FROM AGE(u.birthdate)) AS age,
  t.nights,
          t.rooms,
  u.gender,
  u.married,
  u.has_children,
  u.home_country,
  u.home_city,
  u.home_airport,
  u.home_airport_lon,
  u.home_airport_lat,
  u.sign_up_date,
  t.num_trips,
  t.num_flights2,
  t.flight_original_price,
  t.discounted_flight,
  t.hotel_original_price,
  t.discounted_hotel,
  t.checkedIn_bags,
  t.num_seats,
          t.days_until_departure,
  t.trip_length_days,
  t.distance_km
FROM user_base_session b
LEFT JOIN users u ON b.user_id = u.user_id
LEFT JOIN user_base_trip t ON b.user_id = t.user_id ),
age_group_metrics AS(SELECT
    CASE
       WHEN age BETWEEN 18 AND 25 THEN '18-25'
       WHEN age BETWEEN 26 AND 40 THEN '26-40'
       WHEN age BETWEEN 41 AND 60 THEN '41-60'
       WHEN age >= 60 THEN '60+'
       ELSE 'Other'
    END AS age_group,
    COUNT(*) AS count
FROM
    persona_check pc
GROUP BY
    CASE
       WHEN age >= 18 AND age <= 25 THEN '18-25'
       WHEN age >= 26 AND age <= 40 THEN '26-40'
       WHEN age >= 41 AND age <= 60 THEN '41-60'
```

```
      WHEN age >= 60 THEN '60+'
      ELSE 'Other'
   END
ORDER BY age_group
),
persona_metrics AS (SELECT
   user_id,
   num_clicks,
   num_sessions AS num_sessions,
   avg_session_duration AS avg_session_duration,
   age,
   married,
   has_children,
   COALESCE(ROUND(AVG(nights)),0) AS nights,
   COALESCE(ROUND(AVG(rooms)),0) AS rooms,
   COALESCE(ROUND(AVG(num_seats))) AS num_seats,
   COALESCE(ROUND(AVG(checkedin_bags)),0) AS checkedin_bags,
   COALESCE(ROUND(AVG(trip_length_days)),0) AS trip_length_days,
   SUM(num_trips) AS num_trips,
   COALESCE(ROUND(AVG(num_flights2)),0) AS num_flights2,
   COALESCE(ROUND(AVG(flight_original_price)),0) AS flight_original_price,
   COALESCE(ROUND(AVG(discounted_flight)),0) AS discounted_flight,
   COALESCE(ROUND(AVG(hotel_original_price)),0) AS hotel_original_price,
   COALESCE(ROUND(AVG(discounted_hotel)),0) AS discounted_hotel,
   COALESCE(ROUND(AVG(days_until_departure)),0) AS days_until_departure,
   COALESCE(AVG(distance_km),0) AS distance_km,
      CASE
      WHEN AVG(trip_length_days) <= 5 THEN 'Short Trip'
      WHEN AVG(trip_length_days) BETWEEN 6 AND 10 THEN 'Medium Trip'
      ELSE 'Long Trip'
   END AS trip_length_category,
   CASE
      WHEN AVG(discounted_hotel) <= 500 THEN 'Low Spender'
      WHEN AVG(discounted_hotel) BETWEEN 501 AND 2000 THEN 'Medium Spender'
      ELSE 'High Spender'
   END AS hotel_spend_category,
   CASE
      WHEN AVG(num_trips) = 1 THEN 'Rare Traveler'
      WHEN AVG(num_trips) BETWEEN 2 AND 5 THEN 'Frequent Traveler'
      ELSE 'Very Frequent Traveler'
   END AS trip_frequency_category,
   CASE
      WHEN married = false AND has_children = false AND AVG(num_seats) =1 and
AVG(checkedin_bags)<=2 THEN 'Solo'
```

```
      WHEN married = true AND has_children = false AND AVG(num_seats) =2 and
AVG(checkedin_bags)>=2 THEN 'Couple'
      WHEN married = true AND has_children = true AND AVG(num_seats) >2 and
AVG(checkedin_bags)>1 THEN 'Family'
      WHEN married = false AND has_children = true AND AVG(num_seats) >2 and
AVG(checkedin_bags)>1 THEN 'Family2'
      WHEN AVG(trip_length_days) <= 5 AND AVG(num_seats) <2 and
AVG(checkedin_bags) <1 THEN 'Business'
      WHEN married IS NOT NULL AND AVG(num_seats) >2 and AVG(checkedin_bags) >2
THEN 'Groups'
        ELSE 'Normal Travelers'
        END AS persona_type
FROM persona_check
GROUP BY user_id, married, has_children, age, num_sessions, avg_session_duration,
num_clicks
ORDER BY hotel_spend_category desc
),
customer_perks AS(SELECT pm.user_id,
                CASE
                        WHEN persona_type = 'Solo' THEN 'No cancellation fees'
                        WHEN persona_type = 'Business' THEN 'free checked bag'
                        WHEN persona_type = 'Family' THEN 'Family bonding activities'
                        WHEN persona_type = 'Family2' THEN 'No cancellation fees and
family adventure tours'
                        WHEN persona_type = 'Groups' THEN 'One night free hotel stay with
meals'
                        WHEN persona_type = 'Couple' THEN 'couple massage'
                ELSE '10% discount on shopping'
       END AS perks
from persona_metrics pm)
select *
from persona_metrics pm
left join customer_perks cp
ON pm.user_id = cp.user_id
```