Executive Summary Report

TravelTide Customer Segmentation
& Rewards Program Analysis

Author: Md Arman Kaishar

Introduction:

TravelTide is a US-based E-booking startup in the online travel industry. It started operation in April 2021 on the strength of its data aggregation and search technology. It provides flights and accommodation across the world. TravelTide currently consists of around 1 million users.

Objective:

The objective is to design and execute a fantastic personalized rewards program, identify active users, and allocate the reward perks for each.

Overview of Dataset:

There are 4 data sets: flights, hotels, sessions & users

Time frame: 2021-04-01 to 2023-07-28

- Flights: 1901038 (>1.9 M) trip records (trip_id, departure, destination, location, etc.)
- Hotels: 1918617 (>1.9 M) trip records (trip_id, hotel name, location, check-in, check-out, etc.)
- Sessions: 5408063 (>5.4 M) session records (session_id, trip_id, user_id, booking info, etc.)
- Users: 1020926 (>1.0 M) user records (user_id, birthdate, gender, location, etc.)

As suggested by Elena (Head of Marketing), users with more than 7 sessions since Jan 4, 2023 are included. This led to 5998 users being eligible as Active Users.

Methodology:

For data extraction and exploration, SQL and for data visualization, tableau, and Google Sheets are used.
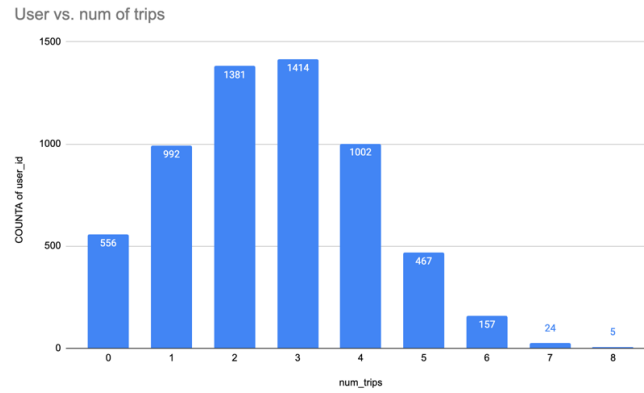
Customer Segmentation:
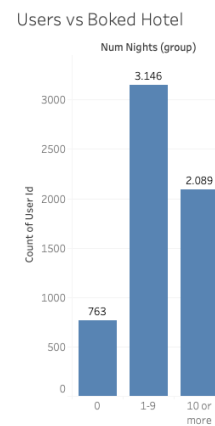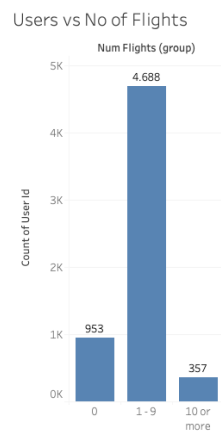
1. **Golden Guests**

Findings:

556 active users who have not traveled yet (9.40%)
5442 active users traveled at least once from whom 653 (10.89% of total) users traveled at least 5 times

User vs. num of trips

357 users booked flights at least 10 times (5.95%)
2089 users booked stayed hotel at least 10 times (34.83%)

Users vs No of Flights

Users vs Boked Hotel

Criteria: At least 5 trips, 10 flights and 10 nights of hotel stay
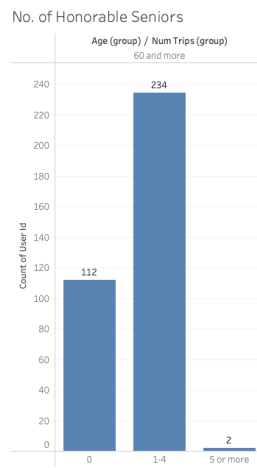
No. of Golden Guests
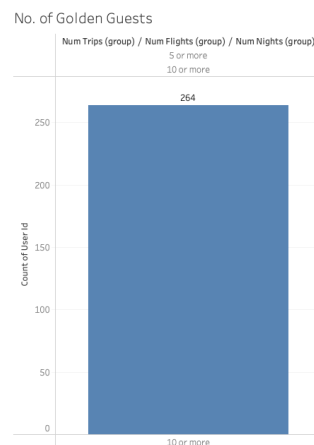
264 users (4.40% of total users) are allocated for this Group.

Rewards: VIP lounge at the airport and 1 night at the hotel with free meal voucher

2. Honorable Seniors

Findings:

No. of Honorable Seniors

Out of 5998 users, 348 (5.8%) users are of age more than or equal 60 years.

112 (1.87% of total users) – never booked
236 (3.93% of total users) – traveled at least once
1 senior user achieved the golden guest's reward

Criteria: At least 60 years of age and traveled at least once

235 (3.98% of total users)

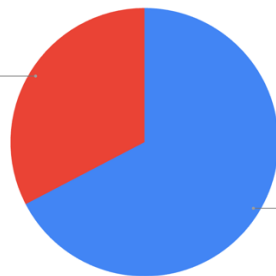**Rewards**: Complimentary meal at the airport / the hotel with travel guide
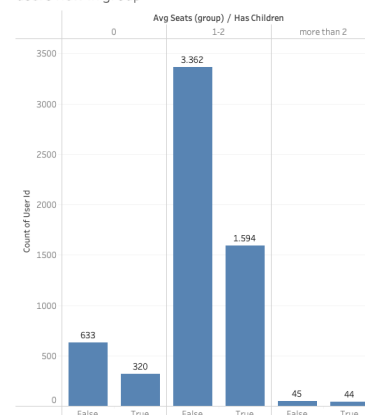
3. Family & Friends
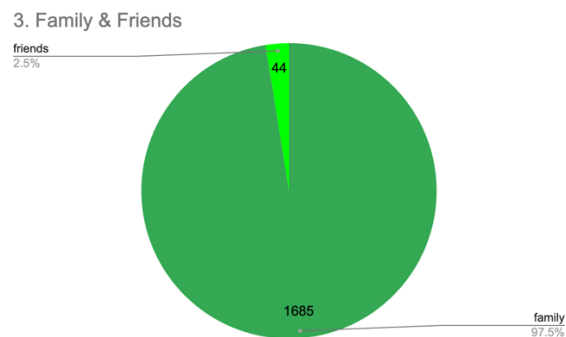
Findings:

users flew in group

- 1958 users are having children (32.68%)
- 182 users have not yet booked any travel (3.03% of total users)

- 1776 users traveled at least once (29.61% of total users)
- 45 users flew in groups without children (booked more than 2 seats on average)
- Among 1776 active family users, 55 user achieved "Golden Guests" and 36 users fulfiled "Honerable Seniors" critaria.
- Among 45 group users, 1 user is also in "Honerable Seniors" group

Criteria: Traveled at least once with children or booked flight seats more than 2

3. Family & Friends

friends
2.5%

44

1685

family
97.5%

Rewards: 1 free check-in bag with a complimentary meal

4. Valued Members

Findings:

no. of users vs. groups
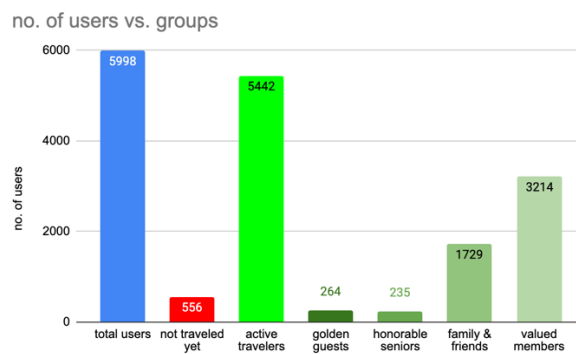
After allocating 3 prior active traveler groups, there are 3214 users who traveled at least once (53.58% of total users)

Criteria: Active users traveled at least once

Rewards: A Complimentary meal

5. Nobel Newcomers

Findings:

556 users who did not travel yet (9.27% of total users)

Criteria: Active App Users, who have not experienced any travel with TravelTide yet

Rewards: 30% off first travel

At a Glance

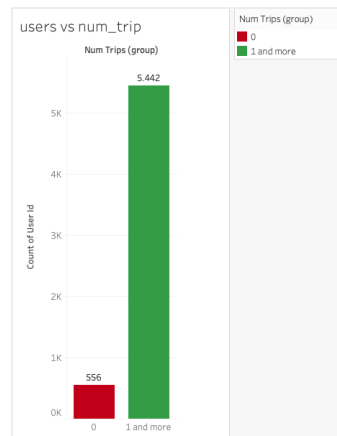| Groups	Perks	Number of Users
1.golden guests	VIP lounge at the airport and 1 night at the hotel with free meal voucher	264
2.honorable seniors	Complimentary meal at the airport / the hotel with travel guide	235
3.family & friends	1 free check-in bag with a complimentary meal	1729
4.valued members	Complimentary meal at the airport / the hotel	3214
5.nobel newcomers	30% off first travel	556
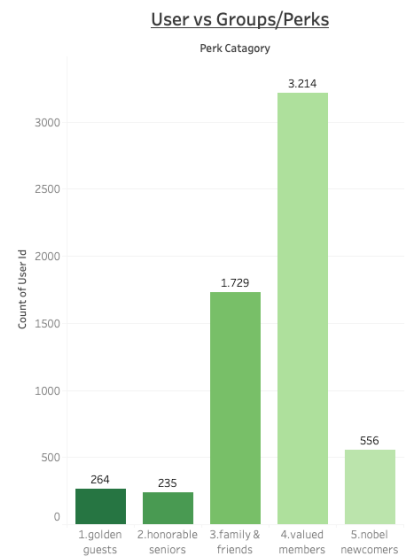
Recommendations

1. Mejority users are Female, 5292 out of 5998 (88.2% of total users). New perk can be introduced to attract male users
2. A/B Testing. Test different operational and marketing strategies tailored to each customer segment's demographics.
3. Machine Learning. All the methodoligies can be checked and the parks allocations can be proofed using ML.

SQL Code:

```sql
WITH -- limiting sessions on Elena's suggested timeframe (After Jan 4 2023)
  sessions_2023 AS (
    SELECT *
    FROM sessions s
    where s.session_start > '2023-01-04'
  ),

  -- all users with more than 7 sessions in 2023
  filtered_users AS (
    SELECT user_id, COUNT(*)
    FROM sessions_2023 s
    GROUP BY user_id
    HAVING COUNT(*) > 7
  ),

  -- Main session base table
  -- Each row is a browsing session on the TravelTide App
  -- Sessions have trips connected to them when they were booked or canceled
  session_base AS (
    select
      s.session_id, s.user_id, s.trip_id, s.session_start, s.session_end,
      EXTRACT(EPOCH FROM s.session_end - s.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
      sessions_2023 s
      left join users u on s.user_id = u.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.user_id IN (
        SELECT user_id
        FROM filtered_users)
  ),

  -- CTE returns the ids of all trips that have been canceled through a session
  -- This list is to filter all canceled sessions in the next CTE
  canceled_trips AS (
    SELECT DISTINCT trip_id
    FROM session_base
    WHERE cancellation = TRUE
```

```sql
    ),

    -- All sessions without trips, all canceled trips have been removed
    -- Each row represents a trip that a user did
    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
        )
    ),

    -- To aggregate user behaviour into metrics (a row per user)
    -- CTE contains metrics that have to do with the browsing behaviour
    -- ALL SESSION within our cohort get aggregated
    user_base_session AS (
      SELECT user_id, SUM(page_clicks) AS num_clicks,
        COUNT(DISTINCT session_id) AS num_sessions,
        AVG(session_duration) AS avg_session_duration
      FROM session_base
      GROUP BY user_id
    ),

    -- To aggregate user behaviour into metrics (row per user)
    -- CTE contains metrics that have the travel behavious
    -- Only rows with VALID trips within our cohort get aggregated
    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)
            OR (return_flight_booked = TRUE) THEN 1
            ELSE 0 END) AS num_flights,
          AVG(seats) AS avg_seats,
          SUM(nights) AS num_nights,
        SUM((hotel_price_per_room_night_usd * nights * rooms) * (
            1 - (CASE WHEN hotel_discount_amount IS NULL THEN 0
                ELSE hotel_discount_amount END))) AS hotel_expenses,
        SUM(base_fare_usd * (1 - flight_discount_amount)) AS flight_fare,
        AVG(EXTRACT(DAY FROM departure_time - session_end)) AS time_after_booking,
        AVG(haversine_distance(
            home_airport_lat,
            home_airport_lon,
            destination_airport_lat,
            destination_airport_lon)) AS avg_km_flown,
        AVG(checked_bags) as avg_bags
      FROM not_canceled_trips
      GROUP BY user_id
    ),
```

```sql
  -- User metrics table - the session metrics, trip metrics and general user information joined
  -- Using a left join, each row for each user from original cohort codition (7+ browsing
sessions in 2023)
  -- inner join avoided not to get rid of users that have not actually travelled
  user_metrics AS (
    SELECT b.user_id, b.num_sessions, EXTRACT(YEAR FROM AGE (u.birthdate)) AS age,
      u.gender, u.married, u.has_children, u.home_country, u.home_city, u.home_airport,
      COALESCE(t.num_trips, 0) AS num_trips,
      COALESCE(t.num_flights, 0) AS num_flights,
          ROUND(COALESCE(t.avg_seats, 0), 2) AS avg_seats,
          ROUND(COALESCE(t.flight_fare, 0), 2) AS flight_fare,
          COALESCE(t.num_nights, 0) AS num_nights,
      ROUND(COALESCE(t.hotel_expenses, 0), 2) AS hotel_expenses,
      ROUND(COALESCE(t.time_after_booking, 0), 2) AS time_after_booking
    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
  ),

  user_perk_metrics AS(
    SELECT *,
        CASE WHEN num_trips >= 5 AND num_flights >= 10 AND num_nights >= 10 THEN '1.golden
guests'
        WHEN num_trips > 0 THEN
                  (CASE WHEN age < 60 THEN
                            (CASE WHEN (has_children) OR avg_seats > 2 THEN '3.family & friends'
                            ELSE '4.valued members'
                            END)
                  ELSE '2.honorable seniors'
                  END)
        ELSE '5.nobel newcomers'
        END AS perk_catagory
        FROM      user_metrics
        )

SELECT
  user_id, perk_catagory,
  -- 1. golden guests (trips more than 3 and flights more than 5)
  CASE WHEN perk_catagory = '1.golden guests' THEN 'VIP lounge at the airport and 1 night at the
hotel with free meal voucher'
  -- 2. honored/senior customers (age more than or equal 60)
  WHEN perk_catagory = '2.honorable seniors' THEN 'Complimentary meal at the airport / the hotel
with travel guide'
  -- 3. family customers (with children)
  WHEN perk_catagory = '3.family & friends' THEN '1 free check-in bag with a complimentary meal'
  -- 4. valued member
  WHEN perk_catagory = '4.valued members' THEN 'Complimentary meal at the airport / the hotel'
  -- 5. nobel newcomers
  WHEN perk_catagory = '5.nobel newcomers' THEN '30% off first travel'
  END AS perk
FROM user_perk_metrics
;
```