Milestone 7 | GameJet Transactions

INTRODUCTION: In this Milestone, you'll help a mobile game company summarize and describe some general patterns in the app usage and spending for one of their games. Revenue from this game is supported by a traditional microtransaction model, where users can spend real money on in-app purchases for items that allow players to access more game features.

Monetization is a difficult problem for companies in the mobile games space, and they need to balance converting as many users from free users to paid users, without being so aggressive that they actually end up driving away more users and actually make less revenue. But before they can start developing strategies for improving the game from a business perspective, the team needs your help to summarize what has been done so far!

HOW IT WORKS: Follow the prompts in the questions below to investigate your data. Post your answers in the provided boxes: the **yellow boxes** for the queries you write, and **blue boxes** for text-based answers. When you're done, export your document as a pdf file and submit it on the Milestone page – see instructions for creating a PDF at the end of the Milestone. Please don't *ever* remove (paste your query below 👇) or (write your **answer** below 👇). These help your Evaluator!

– Data Set **Description**

The data for this Milestone (`game_jet.*`) describes activity for users of a mobile-based game with microtransactions. In the game, two currencies called 'passes' and 'gems' are used to unlock more game content and to purchase cosmetic items. These currencies can be earned slowly over regular play, but gems can also be purchased as in-app purchases. You will be working with three tables in this Milestone: `users`, `sessions`, and `iaps`.

The `users` table lists users who have downloaded the game app, one per row. The table contains six columns, of which include:

- **`udid`** - unique id for the user / device
- **`install_date`** - date of user installing the app
- **`lang`** - two-letter code for native language for device
- **`country`** - two-letter code for user's country

The `sessions` table records every session taken by a user where they opened the app. There are four columns in this table:

- **`udid`** - user / device id
- **`ts`** - timestamp for start of session
- **`date`** - pre-truncated date for session
- **`session_num`** - cumulative session number for the user

The `iaps` (In-app purchases) table records all purchases made by users, exchanging real money for in-game currency. There are six columns in this table:

- **`udid`** - user / device id
- **`ts`** - timestamp for purchase
- **`date`** - pre-truncated date for session
- **`prod_type`** - product type, can be gems, passes, or a value pack.
- **`prod_qty`** - quantity code that is proportional to each product (e.g. a quantity of 1 for gems might actually represent 10 gems)
- **`rev`** - revenue / price of the purchase, in cents

– **Task 1:** App usage behaviors

How long does a user spend with the app in terms of the number of sessions or the days that they stick with the app? This can have implications for how much time we have to 'hook' a player and give them a reason to spend money on the app.

A. Query the `users` data table to discover the total number of users represented in the data.

(paste your query below 👇)

```
Select Count (*) as total_users
From game_jet.users
```

How many users are there in the dataset?

(write your **answer** below 👇)

22576

B. Next, write a query to count the number of distinct users in the sessions data table.

(paste your query below 👇)

```
Select Count (DISTINCT udid) as total_users
From game_jet.users
```

Recalling your answer from Task 1A, how many users downloaded the app but didn't actually open it? HINT: A user who downloaded the app will always add a tally in the users table, but won't show up in sessions if they never use the app.

(write your **answer** below 👇)

32

C. Write a query that returns the number of sessions made by each user. Sort the output by the number of sessions made, from largest to smallest.

(paste your query below 👇)

```
SELECT
  s.udid,
  COUNT(*) AS session_count
FROM game_jet.sessions s
GROUP BY s.udid
ORDER BY session_count DESC;
```

How many sessions did the user with the most sessions have with the app?

(write your **answer** below 👇)

1939

D. We should also be interested in how many sessions were taken by a 'typical' app user. The average number of sessions with the app is 32.1. Modify your query from part C to only return users who have made more than that many sessions. HINT: Your query should include the HAVING keyword here!

(paste your query below 👇)

```
SELECT
  s.udid,
  COUNT(*) AS session_count
FROM game_jet.sessions s
GROUP BY s.udid
HAVING Count(*) >32
ORDER BY session_count DESC;
```

How many users are there? (Read this from the SQL app interface.) Based on that number, how well does the average represent a ‘typical’ player?

Try this prompt: The average number of sessions per user is 32.1, but my SQL query shows that only XXXX users exceed this number out of XX,XXX total users, meaning the majority of players engage far less. What does this tell us about engagement patterns in mobile games? Why might the average be a poor measure of the 'typical' user?

Based on ChatGPT’s response, what’s one reason the average may not represent the typical user experience? What other metric might better describe typical engagement?

(write your **answer** below 👇)

Even before asking chatgpt its important to understand the pareto effect (otherwise called the 80/20 rule) This distribution shows that there is a number of power users and the amount that these power users pay is skewing the distribution. In order to get a better measure of centrality we will have to account for this skewed distribution. Chatgpt confirmed that “the median or percentiles would better reflect ordinary user behavior.”

– **Task 2:** In-App Purchases

How large is our user base in terms of who is spending money? Is it a lot of users spending a little, or is it a few users who are spending a lot? Different types of users will likely be responsive to different types of tactics to get them to spend more money.

A. Using the `iaps` table, write a query to determine the number of users who have made at least one purchase. Remember that each time a user makes a purchase, a new row is added to the `iaps` table.

(paste your query below 👇)

```
Select Count (Distinct udid)
From game_jet.iaps
```

B. Both tables share values in the **`udid`** column. Write a query that left joins the `iaps` table to the `users` table and returns the total (sum) amount of money spent on in-app purchases, grouped by user id. Your query should return every user in the users table, regardless of whether or not the user made a purchase; the total spent for users with no record in the `iaps` table should be a *`null`* value.

(paste your query below 👇)

```
Select
  u.udid,
  SUM(i.rev) AS total_spent_cents
From game_jet.users u
Left join game_jet.iaps i
  on u.udid = i.udid
Group by u.udid
Order by total_spent_cents desc NULLS LAST
```

C. Modify your query in part B to create a new feature called **persona** that segments users into four purchasing personas:

- a “free player” who does not spend any money on in-app purchases,
- a “minnow” who spends less than $20 on in-app purchases,
- a “dolphin” who spends between $20 and $100 on in-app purchases,

- and a “whale” who spends at least $100 on in-app purchases.

Remember: revenue is recorded in terms of cents, so $20 is equal to 2000 cents and $100 is equal to 10 000 cents.

(paste your query below 👇)

```
Select
A.udid,
sum(b.rev) as total_spent_cents,
Case
When sum (b.rev) is null or sum (b.rev) =0 then 'free player'
When sum (b.rev) < 2000 then 'minnow'
When sum (b.rev) < 10000 then 'dolphin'
Else 'whale'
End as persona
From game_jet.users as a
Left join game_jet.iaps as b
On a.udid=b.udid
Group by a.udid
Order by total_spent_cents Desc
```

— **Task 3:** Analyzing Free-to-Play Personas

Now it's time to analyze the overall spending behavior of different purchasing personas. You will use your final query from Task 2 together with the WITH keyword for the remainder of this Milestone as you aggregate and analyze the data you’ve joined. For a refresher, rewatch “🍿The WITH Keyword” in SkillBuilder 6.

A. Write a query that returns the total number of each purchasing persona.

(paste your query below 👇)

```
With p_persona AS (Select
A.udid,
sum(b.rev) as total_spent_cents,
Case
When sum (b.rev) is null or sum (b.rev) =0 then 'free player'
When sum (b.rev) < 2000 then 'minnow'
When sum (b.rev) < 10000 then 'dolphin'
Else 'whale'
End as persona
From game_jet.users as a
Left join game_jet.iaps as b
On a.udid=b.udid
Group by a.udid
Order by total_spent_cents Desc)

Select persona,
Count(*) AS user_count
From P_persona
Group by Persona
Order by user_count DESC
```

Which persona accounts for the majority of users?

(write your **answer** below 👇)

Free Player

B. Write a query that returns the total amount of money spent, grouped by each purchasing persona. Since free players don't spend any money on in-app purchases, filter your query to show only the total revenue generated by minnows, dolphins, and whales.

(paste your query below 👇)

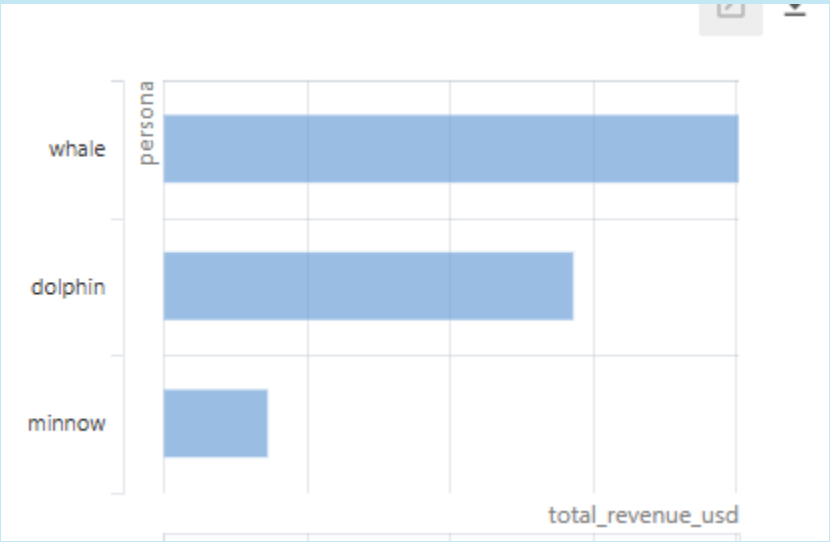
```
With persona_by_user as (
  Select
    u.udid,
    Sum(i.rev) as total_spent_cents,
    Case
      When Sum(i.rev) is null or Sum(i.rev) = 0 then 'free
player'
      When Sum(i.rev) < 2000 then 'minnow'
      When Sum(i.rev) < 10000 then 'dolphin'
      Else 'whale'
    End as persona
  From game_jet.users u
  Left Join game_jet.iaps i
    On u.udid = i.udid
  Group By u.udid
)
Select
  persona,
  Sum(total_spent_cents)/100.0 as total_revenue_usd
From persona_by_user
Where persona <> 'free player'
Group By persona
Order By total_revenue_usd Desc;
```

Does the player persona with the most users also generate the most revenue? What does this tell us about player spending behavior in free-to-play games?

(write your **answer** below 👇)

The more a player becomes a power user the more likely they are to be a player that buys more. Its about a 1:2 dollar spending habit per ascending tier. Whales bring in a total of 40,255.35 on total

revenue compared to Dolphins (28695.30)



– **Task 4:** Analyzing Player Engagement Patterns

The answers to the questions in the Milestone tasks above should paint a picture of the game app being mostly supported off of a relatively small, dedicated group of players. Most users only engage with the app for a short amount of time and end up not buying anything. Even a fair amount of users who do buy something don't make too many purchases.

A. Another angle of attack you might be interested in looking at is how long it takes for a user to make their first purchase, if they make any. You can do this in two steps. First, take the difference between the user's install date and the date of each purchase they made to get the amount of days between when they started playing, and when they made a purchase. Their first purchase will then be the purchase with the smallest difference in dates.

(paste your query below 👇)

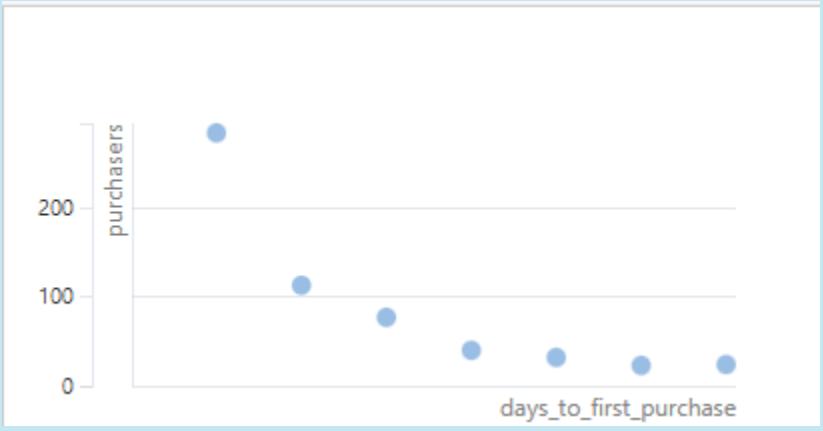
```
With first_purchase as (
  Select
    a.udid,
    Min(b.date - a.install_date) as days_to_first_purchase
  From game_jet.users a
  Join game_jet.iaps b
    On a.udid = b.udid
  Group By a.udid
)
Select
  Case
    When days_to_first_purchase = 0 then 'Day 0'
    When days_to_first_purchase Between 1 and 6 then 'Day
1-6'
    When days_to_first_purchase Between 7 and 13 then 'Days
7-13'
    When days_to_first_purchase Between 14 and 20 then 'Days
14-20'
    When days_to_first_purchase Between 21 and 27 then 'Days
21-27'

  End as bucket,
  Count(*) as purchasers
From first_purchase
Group By 1
Order By 1
```

How many users make a purchase on the first day they have the app? Are there any other periods where we could possibly entice game players to make purchases?

(write your **answer** below 👇)

A little less than 50% of purchases are made on the first day. If we look at the first week closely the distributions trend downwards

from the first day.



Enticement would be achieved on the second day with ranks second as far as frequency of sales.

B. Our previous discovery suggests that the onboarding process plays a significant role in driving early engagement and conversion. You'll explore what this means for improving the onboarding experience and discuss strategies that could encourage more first-time purchases and boost early user engagement.

Try this prompt: We discovered that many users who make a purchase do so on their first day using the app. What does this suggest about the importance of the onboarding experience?"

Based on ChatGPT's response, what's one suggestion for improving early engagement or encouraging first-time purchases?

(write your **answer** below 👇)

When we looked at the data, almost half of paying users spent money on the exact same day they downloaded the app. That

tells us everything about how critical Day 0 really is. If the onboarding feels exciting, smooth, and rewarding, people are ready to pull out their wallets right away. If it's clunky or boring, you probably lose them forever. The first impression isn't just about teaching the game — it's the make-or-break moment for engagement and revenue.

— Submission

Great work completing this Milestone! To submit your completed Milestone, you will need to download / export this document as a PDF and then upload it to the Milestone submission page. You can find the option to download as a PDF from the File menu in the upper-left corner of the Google Doc interface.