

GamesABC Product Revenue Analytics 2022

Power BI

by [Olena Petrova](#), Data Analyst





About the Project

The project goal is to create a Revenue Analysis dashboard for a Product team at a Games Development company.

There were **3 games**, which were presented at different time during **10 month**, were available for different timeframes and in **1-3 languages**.

A transaction-based revenue model was applied with **pay-per-use** pricing tactics.

The Product Team needs **to analyze Revenue data** to plan product-market fit improvements, as well as to set sales and marketing focus and KPIs. The dashboard shows data

- **at the end of the period** to gain insights about user groups and games performance,
- **in monthly dynamics** to review the sources of income and to define factors, impacting the Revenue changes.

The terms of the project:

February 15 - February 29 2024.

The technology stack:

- Initial dataset: two tables in **PostgreSQL** database,
- Dashboard: **Power BI**
- **Excel** - for testing purposes





Project Functionality

Requirements analysis and Design of the Reports

- **End of the period report:**
 1. **Summary metrics** totals (Revenue, User Count total and by status at the end of the period, ARPPU) should be compared with those for selected
 - Game (and language) and
 - User group.
 2. **Users should be grouped** by age, number of active months (LT), total spend (LTV).
 3. It should be possible to see data for selected groups and games relative to the entire scope of data, as well as absolute ones (i.e. **selected values filtering, as well as slicing should be available**).
- **Monthly analytics report:**
 1. Total Monthly Recurring Revenue (MRR) data should be provided for the 10-month period.
 2. MRR structure should be clear: revenue generated by New Returned and Retained Users.
 3. Churn rate and Retention rate change should be presented.
 - It should be possible to compare the Churn rate for a selected Game or user group with the total Churn rate
 4. Factors impacting Revenue (including components of its structure) and Churn Rate should include:
 - the change of User Count and
 - Games launch/ withdrawal dates
 5. It should be possible to see data for selected user groups and games relative to the entire scope of data, as well as absolute ones (i.e. **selected values filtering, as well as slicing should be available**).

```

--create function f_periodlastmonth()
--change
--end;

--with user_details as
--select
--  user_id,
--  user_name,
--  company,
--  mpa,
--  last_p_month,
--  last_g_month,
--  last_p_month_lagged between user and gpa
--extract month from last_p_month-extract month from first_p_month_l as if month, --number of calendar months it
--month part, -- number of months since user transaction
--timestamp, -- number of payments made by user during the period
--gpa
--as
--
--when last_p_month=periodlastmonth then 'Current'
--
--infer as
--e_current,
--
--when extract month from last_p_month-extract month from first_p_month_l month part then 'Released'
--
--else null
--as e_release,
--
--user_status
--from user_details

22 /* In this script we add to EXTRACTED table data from related user table */
23 /* For MONTH data calculation we need have first_p_month from SOURCE table,
24  * but for EXTRACTED table it is not needed */

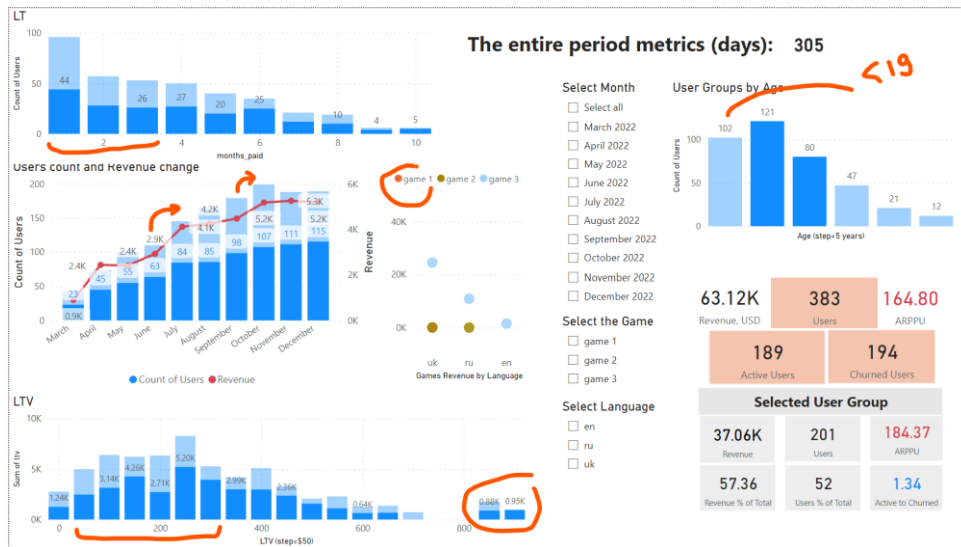
--with all_data as ()
--
--(select month_data as)
--
--(select
--  customer_id,
--  m_payment_name,
--  m_payment_date,
--  m_revenue_amount_usd,
--  mcp_month,
--  max(mcp_month_rev) as month_rev,
--  mcp_month as prev_month,
--  max(mcp_month_rev) as prev_month_rev,
--  mcp_month as next_month,
--  min(mcp_month_rev) as next_month_rev
--  from user_monthly_data as ud
--  left join user_monthly_data as prev on prev.pcp_monthing_month = interval '1 month' and prev.user_id=ud.user_id
--  left join user_monthly_data as next on next.pcp_monthing_month = interval '1 month' and next.user_id=ud.user_id
--  group by 1,2,3,4,5,6)

```

-
- The screenshot shows the AWS Glue console with a workflow diagram. The workflow consists of four jobs:
- newly_created**: A crawler job that triggers the **get_new_users** job.
 - get_new_users**: A Python job that triggers the **get_new_players** job.
 - get_new_players**: A Python job that triggers the **Warehouse Storage** job.
 - Warehouse Storage**: A Hive job that triggers the **newly_created** job.
- The jobs are connected in a cycle, with 'newly_created' being the starting point and 'Warehouse Storage' being the final job in the sequence.



End of The Period technical and logical details



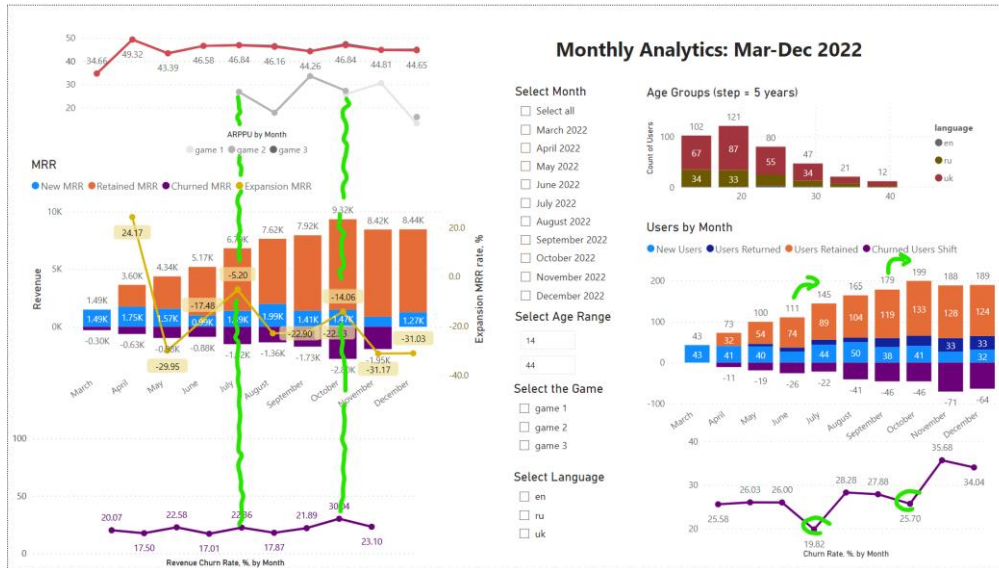
The PowerBI file with source dataset and SQL scripts may be found on my GitHub – [olenapetrova-da](https://github.com/olenapetrova-da)

The report represents a **baseline** for the first 10 months of product usage.

- Totals of Revenue, ARPPU, and User Count (with decomposition to Active and Churned Users at the end of the period) **ignore any filters**.
- “Selected User Group” shows **absolute and relative to total metrics** for the selected group of Users.
- **Users are grouped** in bins with 5-year step. A big group of teenagers was identified. The largest part of the revenue was generated by users younger than 29.
 - User distribution among LT with 1-month step. Paid months are used instead of the time difference between the first and the last payments (due to the revenue model).
 - LTV summed in boundaries of groups with the step of \$50. That allows to find groups of users with desirable LTV. Also, that shows anomalies.
- The scatter chart shows the Revenue amount in dimensions of **games and their languages**.
- **Change of user count and revenue by months**. That is the main subject of the project. It will be decomposed on the Monthly Analytics dashboard.

Monthly Analytics

technical and logical details



The PowerBI file with source dataset and SQL scripts may be found on my GitHub – [olenapetrova-da](https://github.com/olenapetrova-da)

The report represents **Revenue and Churn Rate changes** during the first 10 month of product launch.

- The **Revenue sources** were added to show the structure of MRR
- As **probable factors** of revenue change
 - the User Count by month was added. It consists of New, Retained, and Returned users measures,
 - ARPPU by Games was added. It shows the start and the end of each Game usage.
 - Users age groups were split by languages
- The **Revenue Churn Rate and Retention Rate** were added to plan the marketing and sales efforts.
- It is possible to see **ARPPU and Revenue Churn Rate** of selected User groups or Games **relative** to total numbers.
- All the charts may be **filtered**
 - by Month from any of the charts or from the slicer
 - by user Age and Language from binned age bar chart or from “between” slicer



Analysis results

- 80% of users are in age groups from 14 to 29 years old
- 27% of users are teenagers younger than 18. One of them has an anomaly large LTV.
- 60% of Users are engaged with the product from 2 to 6 months
- 56% of Users spent from \$50 to \$300 during the LT. Their LT has a distribution close to normal, the average LT is 3,5 months. They played game 3 in Ukrainian and Russian.
 - Revenue from those, who played the Ukrainian version had a stable significant growth till November (may be related to the Game 1 launch in October)
- An increase in Revenue correlates with the launch of new games. However, it looks like that does not have a long-term effect: the churn rate grows the next month.
 - ARPPU of new games is also lower than that of the initial one.
- English version of the product has a significantly higher Churn rate
- Revenue from existing users gets lower month by month, meanwhile, the Revenue Churn rate grows.



Pitfalls and Resolutions

Non-Technical issues I faced

- How to select metrics to show?
- What should be the design of reports?
- Which parameters to calculate in SQL before loading to Power BI?
- How to define factors impacting the Revenue change?

Business Analysis practices helped a lot:

- Study of the Revenue and Unit Economics metrics, and their decomposition (e.g. importance of Churn rate).
- Study and definition of applied Revenue model and pricing type (e.g. what data to use for LT calculation).
- Definition of possible business needs for data analysis in such a business and at such a stage (e.g. KYC, Product-Market fit).
- Definition of ER, scenarios of usage and Use Cases.

Technical issues I faced in Power BI

- How to group data?
- How to calculate number of users based on conditions?
- How to combine measures which calculate different User categories in one chart?
- How to ignore filters?
- How to move value on x-axis?
- How to ignore the first/ the last month?

Most of the difficulties were overcome by

- google search and
- parallel tests in SQL, Excel and Power BI,
- after the detailed decomposition of the metrics and tasks.