Final Report - Customer lifecycle management

Introduction

"Monster Hoarder," an AR-based mobile game launched by CEVER Technologies in 2016, has faced declining player numbers and revenue since 2021 due to market competition and seasonal engagement drops. The key challenges that the company is facing and trying to address are:

1. Understanding the four player types—Walkers, Socials, Grinders, and Miscellaneous—to tailor retention efforts based on their distinct behaviours and preferences.
2. Seasonality & Churn, to Identify engagement trends implementing strategies to reduce fall/winter churn, when player activity typically declines.
3. Monetization & Spending Behaviour, analysing in-game purchases to optimize revenue streams, in a freemium setting.
4. Churn Prediction, by developing a model to anticipate player churn by identifying high-risk users early.
5. Marketing Strategy and data-driven improvements in-game mechanics, in-game shop design, customer acquisition, and retention efforts.

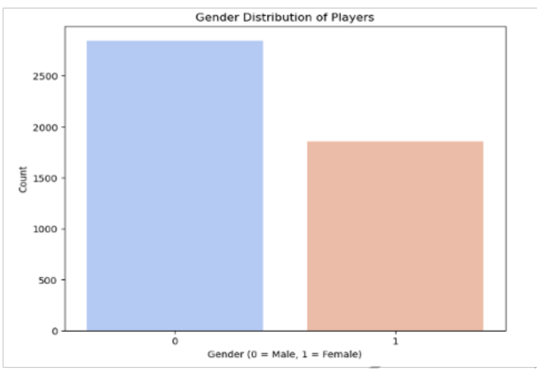
To tackle these challenges, we will analyse 5 datasets:

- Customer Data: Player demographics, registration date, income, and fall bonus eligibility.
- Financial Transactions in the Fall and Summer: Microtransaction details, including purchase amount and frequency.
- Play behaviour in summer and fall: Play sessions, experience gained, hunting/training interactions, social engagement, distance walked, and session duration.

This project will leverage these datasets to create a tailored strategy that optimizes engagement and enhances game profitability.

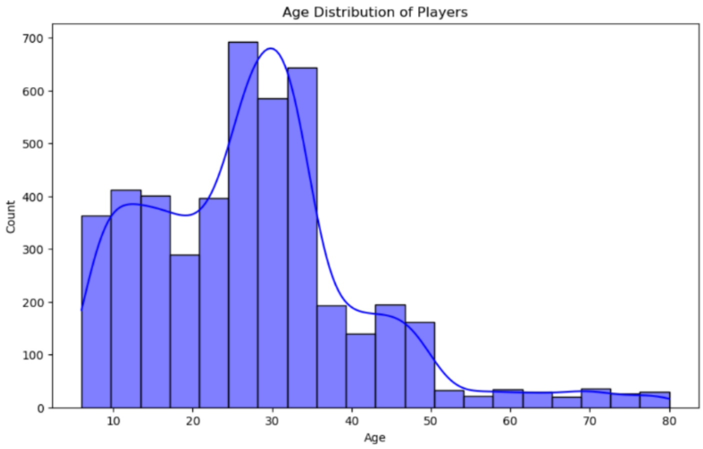
Task 1 - Demographics/behavioral description

To better understand the *Monster Hoarder* player base and give a clear idea of the customers for CEVER Technologies we started by analyzing key demographic variables, including gender, age, and income level.
During the summer the company recorded 4703 active customers, that played at least 1 session, 2992 of which never made a transaction. Therefore, we immediately notice a majority of free players, that might cost the company more than the revenue they generate.
Gender distribution is also important as it helps us understand the custoberbase we are dealing with.



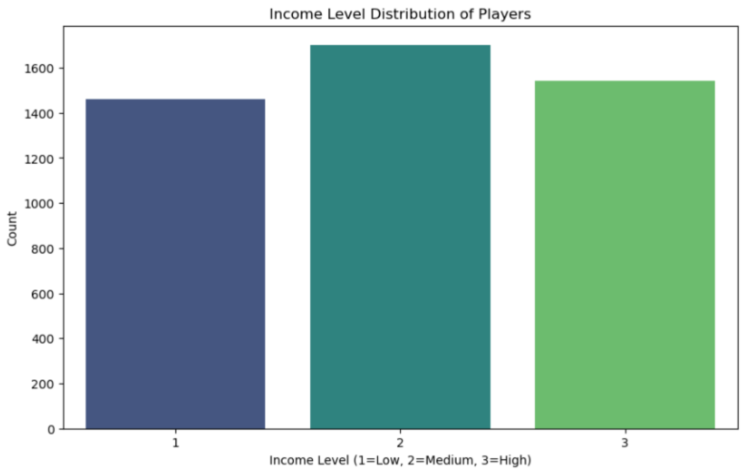
Our analysis indicates that male players consistently outnumber female players. However, there is still a notable proportion of female players, suggesting that while the game has a stronger male audience, it also retains engagement across both genders.
This is further confirmed by the lifecycle grid by gender (*Refer to Appendix Figure 1*).

Expanding our analysis, we now focus on the **age distribution of players.**



We discovered that there is a dominant concentration of players between 20 and 35 years old, with a peak around the late 20s. This suggests that *Monster Hoarder* primarily appeals to young adults, who are likely to have both the time and disposable income to engage in the game. However, there is also a significant presence of younger players under 18, indicating that teenagers form a secondary audience. The number of players declines significantly after the age of 40, showing that engagement among older demographics is relatively low.

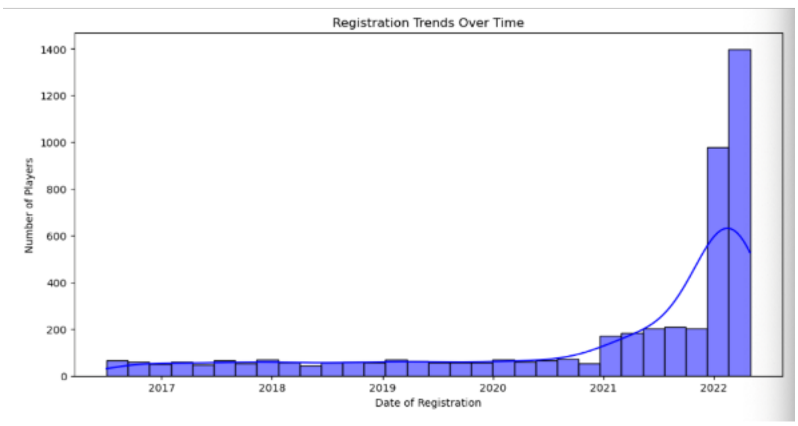
To understand the **financial availability** of our players we examined the income levels, and what we observe is a relatively balanced distribution across low, medium, and high-income players, with medium-income players making up the largest share of the user base.



This suggests that the game attracts a diverse economic audience. This balanced distribution of income levels is further reflected in the lifecycle grid analysis for income groups. While medium-income players remain the most numerous, confirming their dominance in the user base, this does not appear to translate into higher spending frequency or recency (*Refer to Appendix Figure 2*)

Further analysis of the **relationship between age and income levels** (*refer to Appendix Figure 3*) shows that all income groups share a similar median age, around the late 20s to early 30s, confirming that this is the core player demographic.

Another interesting thing to analyze is **how the customer base was built over time**.

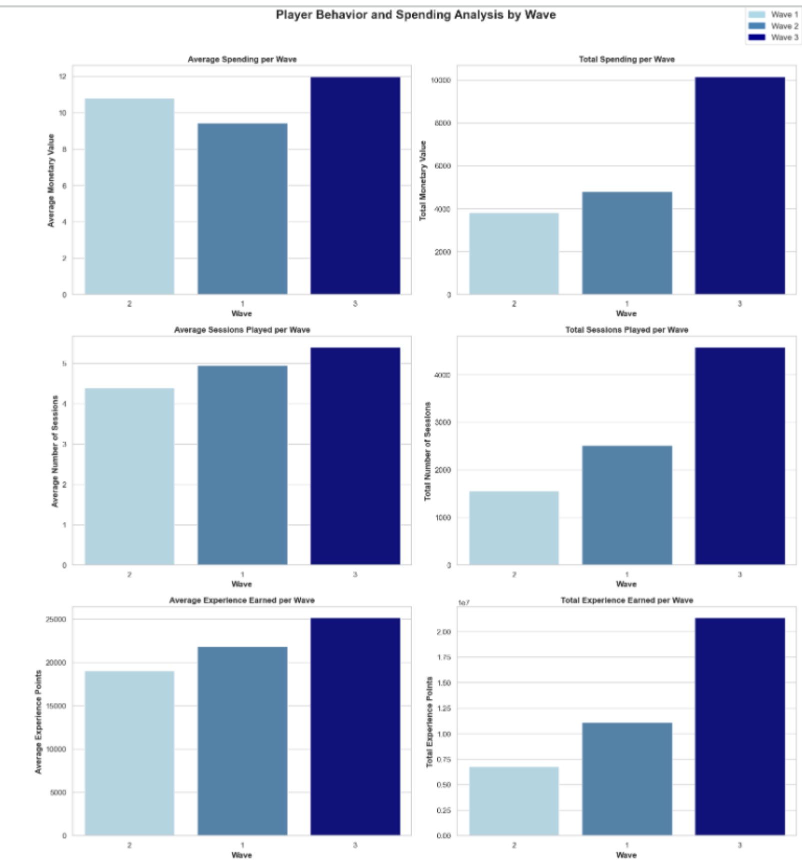


Caver technologies have launched the game in 2016. From its release *Monster Hoarder* steadily grew its user base. In 2021, the market for AR-based games became more competitive, leading to a decline in player numbers and profit for the first time in years. However, as shown in the registration trends, a major surge in new players occurred in late 2021 and peaked in 2022. This suggests that CEVER Technologies successfully adapted to the changing market conditions, potentially through targeted marketing campaigns, and significant game updates. While this growth is promising, it raises important questions about the retention of these new players, whether they will remain engaged long-term or contribute to increased churn.

How do they play and spend?

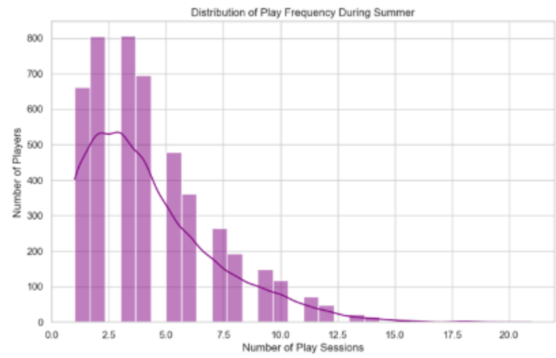
Now that we have established a clear demographic profile of the game player base, primarily male, aged 25-30, and spanning diverse income levels, we shift our focus to understanding how these players engage with the game.
Beyond whom they are, it is crucial to explore **how frequently they play, and how they spend.**

Our first focus is how the registration date impacts the players we categorize users into three waves based. **Wave 1** consists of early adopters who joined before 2021, likely forming the game's most loyal and engaged core. **Wave 2** includes those who registered throughout 2021, representing a phase of expansion driven by growing interest and promotional efforts. **Wave 3** comprises players who joined from 2022 onwards, marking the latest surge in user acquisition.



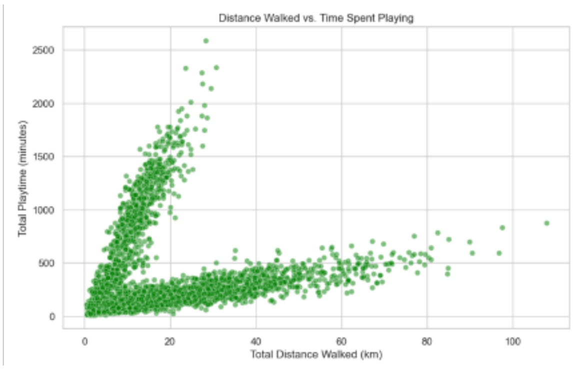
Across these waves, we observe a clear upward trend in both engagement and spending. Wave 3 players demonstrate the highest levels of interaction, playing more sessions, spending more, and accumulating more experience points. However, this could largely be attributed to the fact that Wave 3 also has the highest number of players, as seen in the registration trends. This means that their higher total engagement and spending may not necessarily indicate that newer players are more engaged or are being monetized more effectively, it could simply be a result of their larger population. Meanwhile, Wave 1 and Wave 2 players remain active but spend less on average, highlighting the importance of retention strategies tailored to long-term users.

A deeper look into **playtime and session frequency** reveals that most players engage in short and infrequent sessions.



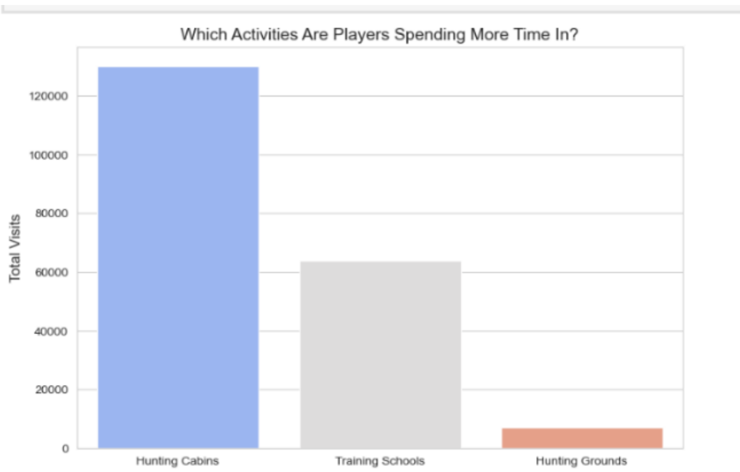
Most players participate in 2 to 5 sessions, while only a few exceed 10, indicating that casual engagement dominates. Similarly, total playtime (*refer to Appendix Figure 4*) follows a highly right-skewed distribution, with most players logging in for under 200 minutes, and only a small fraction exceeding 500 minutes. This suggests that while the game attract many users, long-term engagement remains a challenge.

Beyond playtime, **the distance covered during the play session** also varies.



Most users walk between 5 and 15 km, with only a few surpassing 40 km. This indicates that while movement is an integral part of the game, long distances are uncommon. There is a strong correlation between distance walked and playtime, as players who move more tend to play longer. However, some highly engaged users accumulate significant playtime with minimal movement, suggesting different playstyles, from active exploration to more stationary engagement.

To deepen our analysis of playtime, we examined **which in-game activities players engage in the most,** considering the core mechanics of *Monster Hoarder*. Players gain experience by catching wild monsters and defeating bosses in hunting grounds, while key locations include hunting cabins for resource collection, training schools for monster battles, and hunting grounds for encounters.

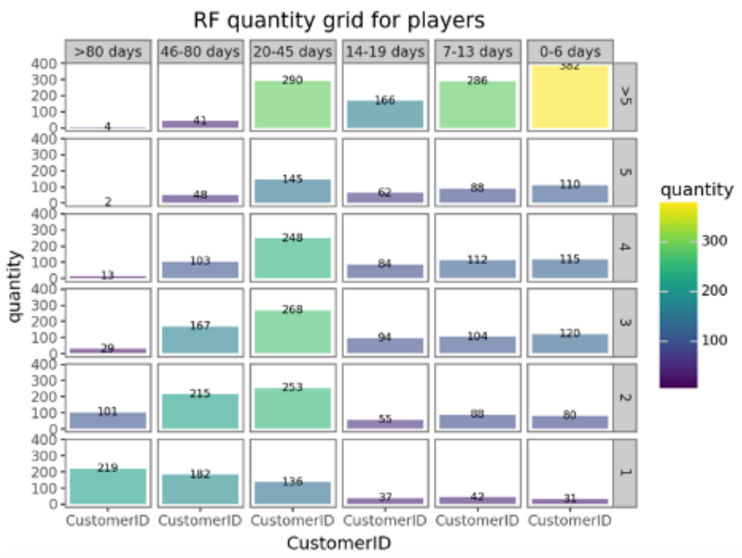


Our findings reveal that players engage the most in hunting cabins, followed by training schools and hunting grounds, suggesting a stronger focus on resource collection and monster training rather than direct combat.

In order to segment customers and understand common characteristics we have built an **RFM base table.**

However, since this is a gaming company, the focus is not only on how much players spend but also on their engagement, behavior, and playing patterns. For this reason, we also built a **RF base table for playing behavior.** The analysis obviously misses the monetary value, but it is useful to segment customers based on their habits allowing to target active player participation.

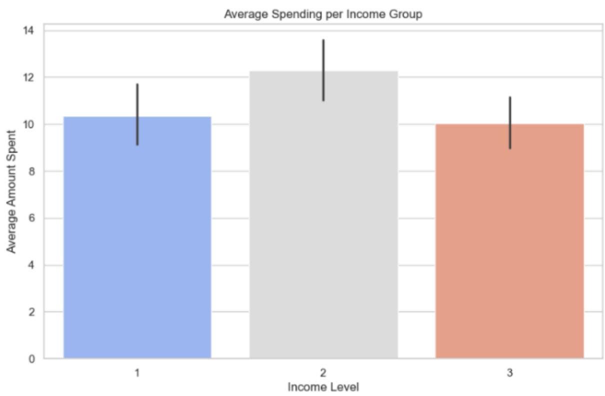
The first grid categorizes players into: **new ones** (low frequency-low recency), **one-time players** (low frequency-high frequency), **active players** (high frequency-low recency) and finally **former frequent players** (high frequency-high recency).



From the plot it is evident that one key issue is the high number of one-time players, which is not ideal, as it indicates that many users try the game once but do not return. Additionally, the number of new customers is also relatively low**,** suggesting challenges in attracting new players.

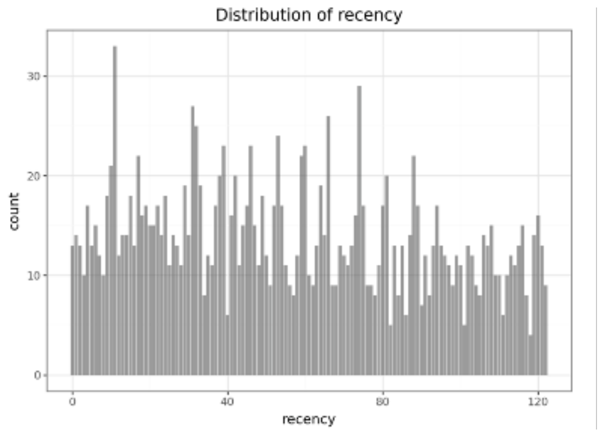
A significant portion of players fall into active customers, specifically with a very high recency, meaning that a relatively high group of customers have played within 0 to 6 days ago. Ensuring that these engaged users continue playing regularly is essential for sustaining long-term player retention.

Finally, the number of former frequent players is not as low as it should be. However, their recency is not too high, meaning they have not been inactive for too long. This presents an opportunity to re-engage them and convert them back into active players before they fully disengage.
Moving on, the next step was focused on analyzing the spending patterns of the summer customers and seeing how it relates with the playing behavior.

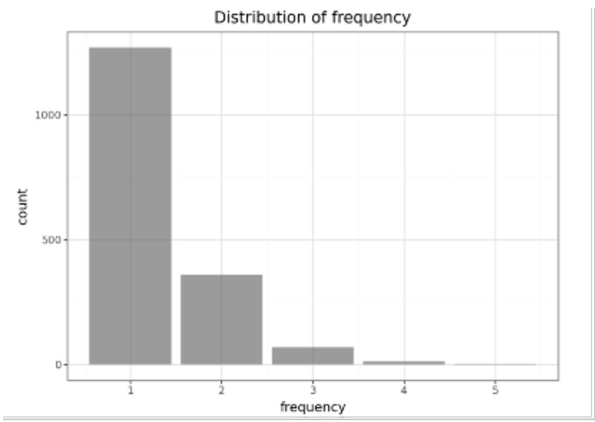


Starting from the spending behavior per income we can see that it is relatively balanced across different groups. Medium-income players tend to spend the most on average, but the difference is not significant, as low- and high-income players follow closely. This suggests that disposable income does not strongly dictate spending habits, indicating that other factors, such as engagement or in-game incentives, may play a larger role.

Before analyzing the RFM grids we wanted to understand the patterns of **recency** and **frequency.**

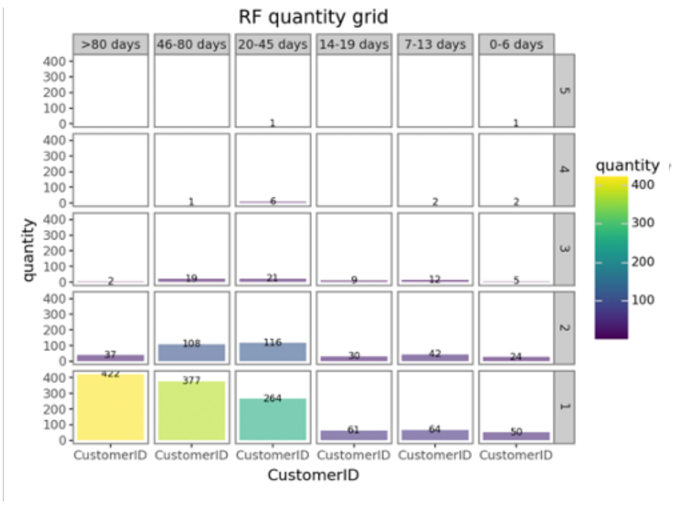


Recency measures the time since a player's last transaction, and as the plot shows it does not follow a clear pattern. Players engage in spending at highly varied intervals, with some clusters of highly active users making frequent transactions, while others have longer gaps between purchases. This suggests that spending behavior is inconsistent and does not follow a predictable cycle during the summer, highlighting the diversity in player engagement.

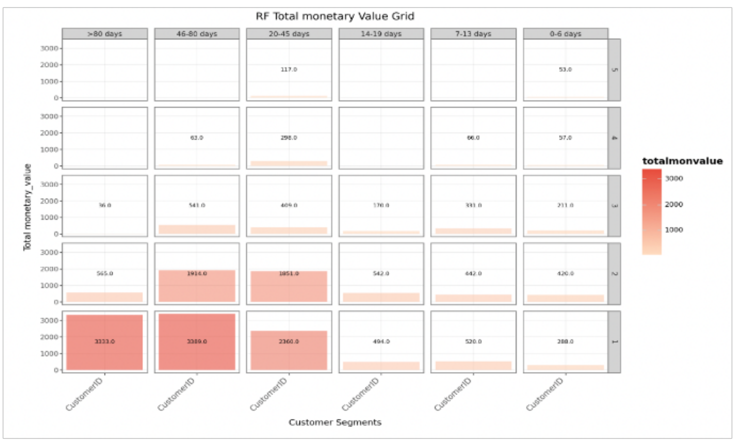


Frequency, instead, is the number of transactions made by each customer during the summertime period. Its distribution clearly highlights the typical trend of freemium models. Most of the players make one/two transactions and continue playing without spending. A detail that can be dangerous, because it can be sustainable only if the new customer segment is relatively big.
This is also clearly visible in the first RF grid, that counts the number of customers in the 4 quadrants: new customers, one-time buyers, active customers and former frequent buyers.

As we noticed before the most notable issue is the high number of one-time buyers, indicating that many customers make a single purchase and do not return.

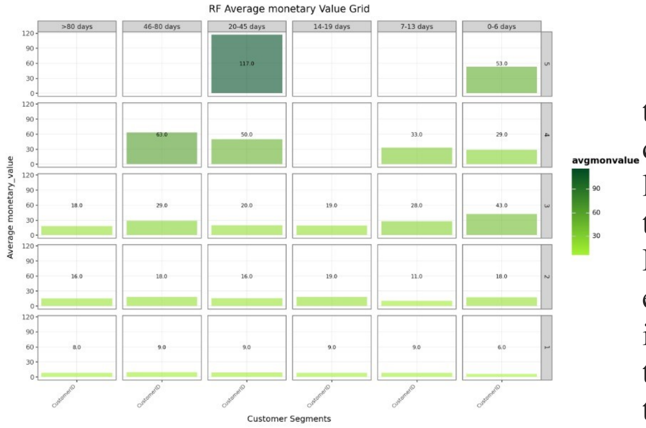


Additionally, the new customer segment is relatively small, suggesting that customer acquisition is not as strong as it could be. The active customer segment and the former frequent buyer segment are minimal, meaning that very few customers continue to engage in frequent, repeat purchases. There may be little incentive to make repeat purchases, as these do not directly impact gameplay. This could explain why some players do not continue spending.



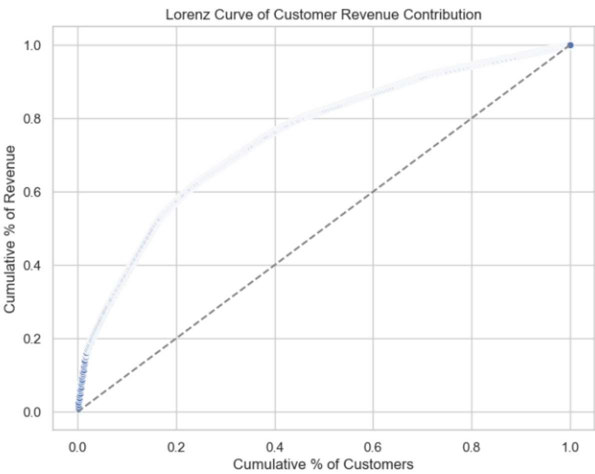
To further analyze the spending behavior of the customers we looked how different segments generate **monetary value**.

From this grid we can observe that a significant portion of the total revenue comes from one-time spenders. As we have seen, being very numerous, even if they make infrequent purchases, they generate the most value. Also, the new customers, again being the second most numerous segments, have moderate monetary value, significantly higher than frequent purchaser, as they are the least consistent group.



Moreover, from this last grid we can see that these latter, especially the former frequent customers, spending more frequently have a larger average monetary value. We can identify them as a small but highly valuable segment. However, we need to keep always in mind that, even if one-time buyers don't bring high value individually, they are a big segment, that all together generate the majority of Cever technologies revenue.

One interesting insight that we gathered is that active customers, despite being few, spend a lot compared to others. At the same time, they are not enough to make a relevant impact on the total monetary value.



This is confirmed by the **Lorenz curve**, which shows that even if a relatively small portion of customers account for the majority of the revenue (45%), our business model does not follow the **Pareto principle**. This implies that Moster hoarder relies heavily on onetime buyers and new customers to be profitable.

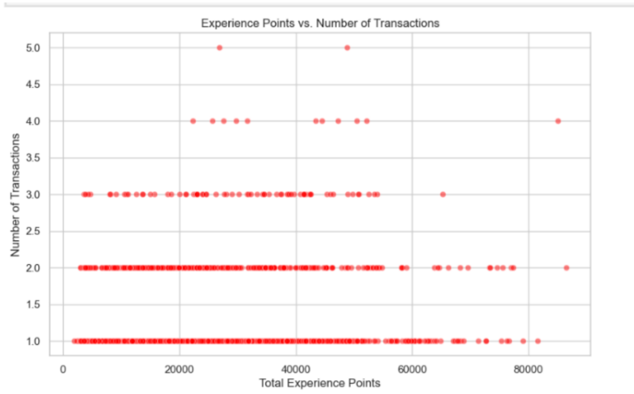
Comparing the results of the spending behaviour and playing behaviour, we can conclude that the majority of users play for free, making very rare transactions. To further investigate this relationship, we focused on the correlation between **gameplay engagement and transaction behaviour.** Specifically, we sought to determine whether the most active players are also the highest spenders or if spending patterns are influenced by other factors.

First, we delved deeper focusing on the month of first purchase, to see how it affects the spending behaviour.

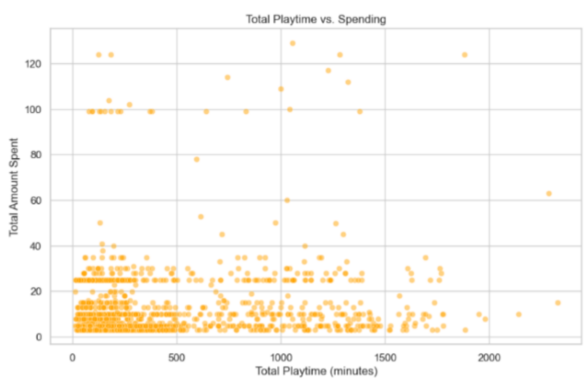
Customers who made their first purchase in May tend to have the highest average spending, which is expected given that they have had more time to engage with the game and make multiple transactions. In contrast, newer cohorts from July and August have had less time to accumulate purchases, resulting in lower spending levels. This suggests that new customers need time to develop engagement with the game and spend.
Moreover, we can also see that customers who made their first purchase in May tend to have a higher recency, while more recent first purchase dates show a lower recency. This suggests that customers, after having a peak of frequent spending, tend to slow down their purchases. The game mechanism probably leads to this customer lifecycle, since players can also spend coins earned playing. Therefore, higher experience means more earned coins, and a higher chance to advance without the need to pay.

After that, we wanted to see if there is any relationship between **experience, spending amount and frequency.**

When analyzing the plot most players spend less than $20 regardless of how many experiences points, they accumulate. While a few highly experienced players spend significantly more, there is no strong correlation between progress in the game and spending. This implies that advancing in the game does not necessarily encourage higher spending, as we have seen also by the grid above.



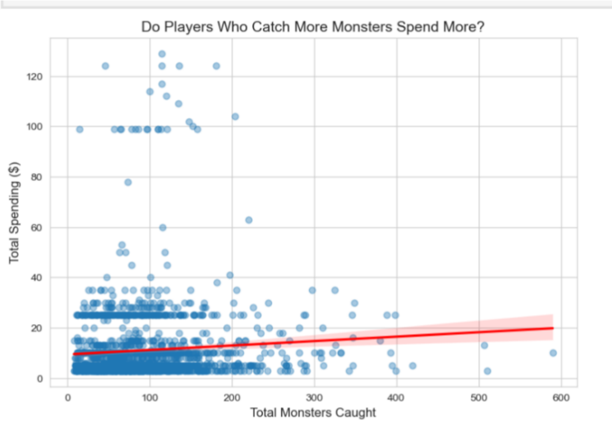
Similarly, examining the number of transactions in relation to experience, we see that most players make only one or two purchases, regardless of their experience level. A small group of engaged users conducts three or more transactions, but they remain a minority, confirming further our hypothesis.



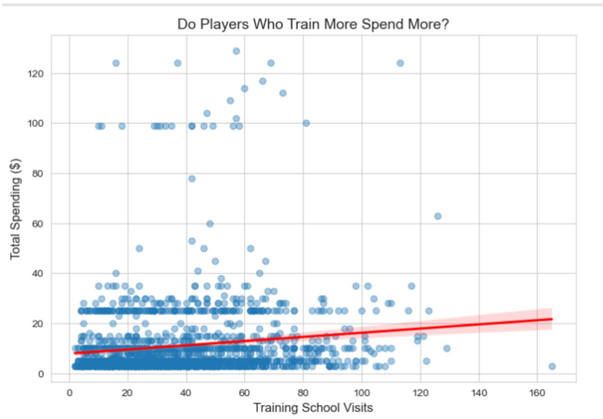
Again, the same pattern emerges when looking at total playtime and spending, most players spend low to moderate amounts, regardless of how much time they invest in the game. A few outliers with long playtimes do make higher purchases, but overall, extended playtime does not necessarily translate into increased spending.

To conclude, when combining experience levels, playtime and the number of transactions, the most evident pattern is that spending frequency remains low across all levels of engagement.

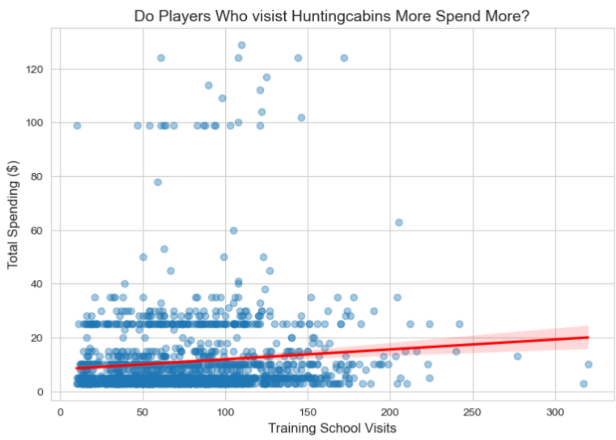
In addition, since in previous section we found out that during a session players spend more time in locations such us: Hunting cabins, Training schools, and Hunting grounds we decided to see **how they correlate with spending.**



If we look at the number of monsters caught and total spending we see a very weak positive correlation, meaning that while some players who catch more monsters do spend more, the effect is minimal. Most players, regardless of the number of monsters they catch, tend to spend under $20 as seen in the previous graph about spending and experience. A few outliers who have captured many monsters show higher spending, but overall, catching more monsters does not strongly predict higher spending.



The same accounts for training schools spending remains relatively low for most players. This suggests that training monsters does not significantly drive in-game spending.



Our last and third plot assesses whether players who frequently visit hunting cabins (where resources are collected) tend to spend more. Again, the trend line is slightly positive but weak, showing that increased visits to hunting cabins do not necessarily translate into higher spending. Most players still fall within the low-spending category, with only a handful of outliers contributing significantly higher amounts.

From this analysis we can assume that due to the freemium nature of the game is no need to spend in order to collect resources and catch monsters.

Therefore, our overall conclusion for this in-depth demographic analysis is that while Monster Hoarder successfully attracts a broad and engaged player base, the majority of users remain free-to-play, limiting the game's revenue potential. Understanding and addressing the barriers to spending will be crucial for CEVER Technologies to convert engagement into sustained monetization. For this reason, we will now delve deeper into the four player types, providing a more targeted understanding of monetization opportunities.

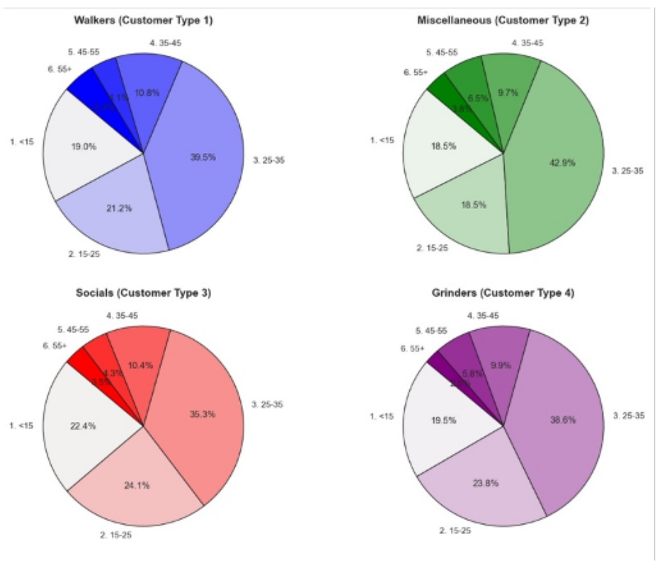
Task 2 – Analyse and visualize the differences among the four player types

As said before, one of the objectives of this analysis is to understand the different characteristics the 4 different player types **Walkers** (1), **Miscellaneous** (2), **Socials** (3) and **Grinders** (4). This would allow the company to maximize revenue, tailoring high profit strategies for the different groups.

To begin, we have plotted the distribution of the 4 segments, to see **if there are significant differences in their sizes** and to understand **the overall composition of our customer base.** This helps us identify which segments dominate and whether any require special attention in our analysis.

It emerges that the distributions are quite homogeneous, with Grinders and Walkers being slightly more numerous.

Next, we investigated the **age groups** that compose these 4 groups through customer crossing, compare how different age groups are distributed across your four customer segments to see if some age groups are overrepresented or underrepresented in any segment.



However, as we can see from the pie charts, all four player types exhibit a similar pattern of age distribution. The 25-35 and 15-25 age groups consistently make up the largest share across all segments, while older players (55+) are the least represented.

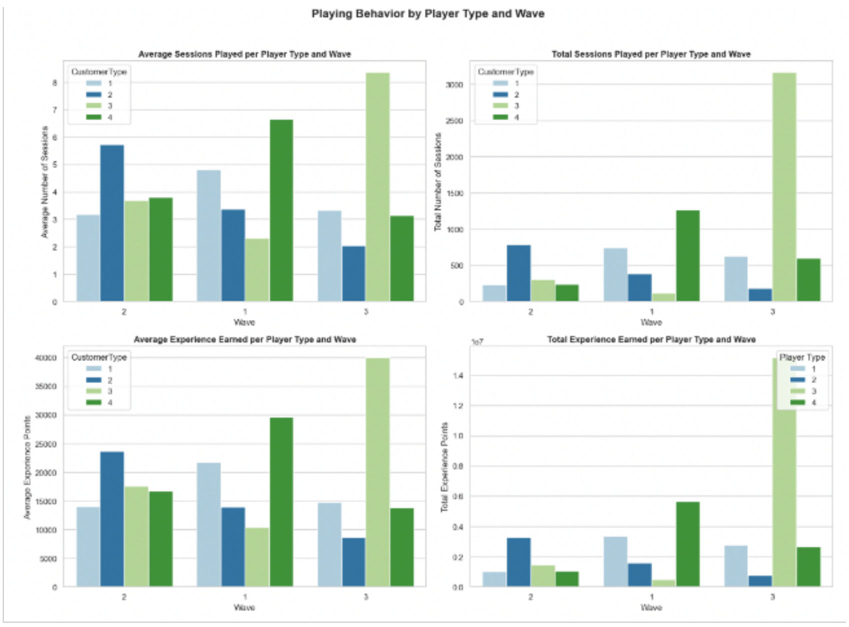
To conclude we have also seen the **purchasing power** of each group in order to see if it correlates with the actual spending behavior. Overall, we can say that socials and grinders have lower income level, while Walkers fall in the high-income category. Miscellaneous instead cover the 3 income levels, showing a higher variety of customers.

To proceed with the analysis, we moved on, trying to address **how the 4 segments differ in playing style and purchase behaviour.**

Building on our previous findings, we now examine how registration waves influence player behaviour, this time distinguishing between the four player types. As seen earlier, Wave 3 players demonstrated the highest engagement and spending levels, though this was partly driven by their larger population. However, when we analyse individual player types within each wave, we observe clear differences in engagement patterns that go beyond sheer numbers. Socials in Wave 3 stand out as the most active group, playing significantly more sessions on average and accumulating the highest total playtime. This suggests that these newer players are deeply engaged. Grinders maintain a moderate level of engagement, while Walkers and Miscellaneous players consistently show lower session counts, indicating a more casual approach to gameplay.

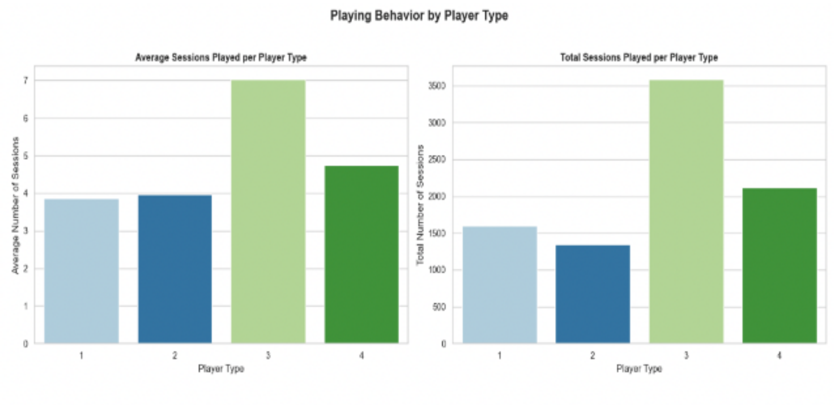
Looking at experience accumulation, Socials in Wave 3 once again dominate, earning the most experience points both on average and in total. This reinforces the idea that they are not just playing frequently but also maximizing their in-game progression. Meanwhile, Walkers and Miscellaneous players maintain consistently lower session counts and experience gains, indicating a need for targeted strategies to increase their engagement.

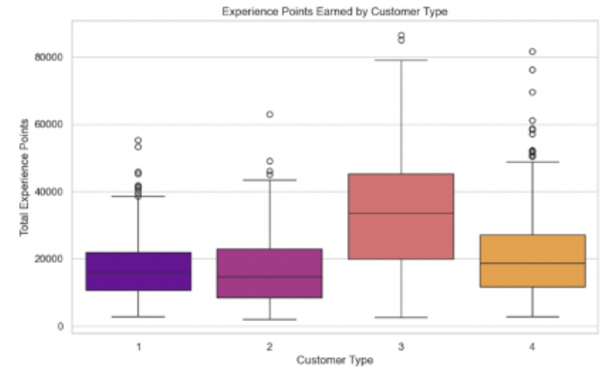
Overall, the data suggests that newer players, particularly Socials, are more engaged than previous waves, while Walkers and Miscellaneous players may need targeted strategies to boost their activity levels. Encouraging Grinders and Walkers to increase their engagement through tailored incentives or gameplay improvements could help create a more balanced and retained player base.



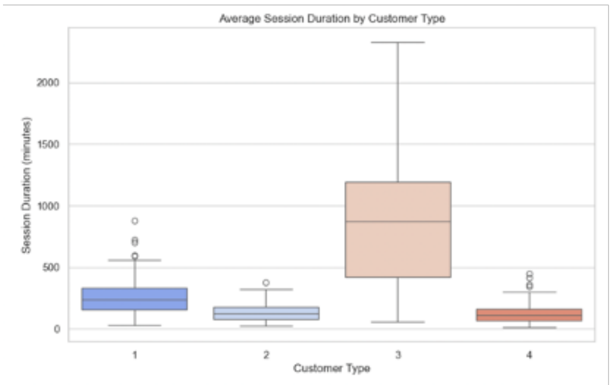
Building on our findings, analyzing total sessions played per player type, confirms that Socials are the most engaged, playing the most frequently. Grinders follow with sustained activity but fewer sessions. Walkers and Miscellaneous remain more casual. Given our observations, we expect Socials to lead across all engagement metrics, as they have already shown the highest activity levels. However, Grinders may outperform them in total monsters captured, as their playstyle is centered on maximizing collection.

The next analysis will verify whether these expectations hold true across **experience accumulation, session duration, social interactions, and monster captures.**





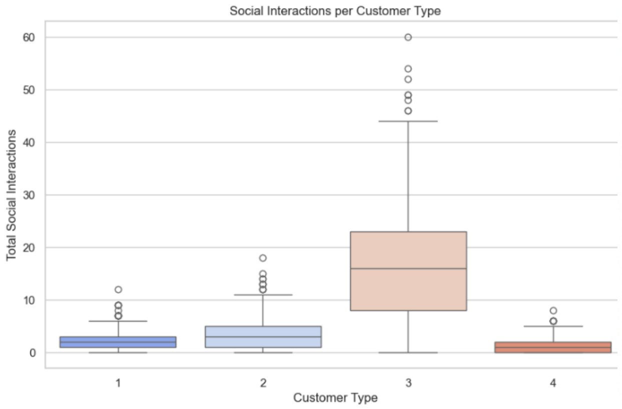
The **experience accumulation** graph highlights that Socials earn the most experience points overall, followed by Grinders.

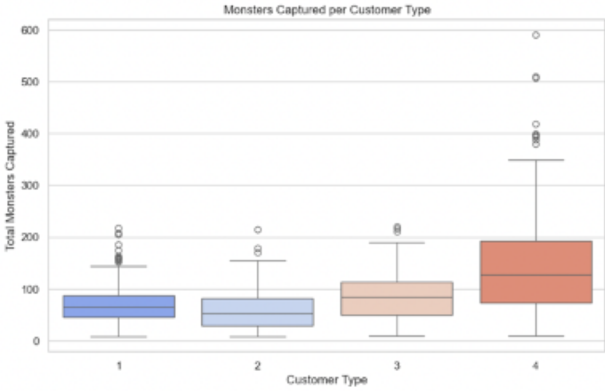


Also, the **average session duration** is higher for Socials, followed by Walkers, which may seem surprising given their lower engagement in previous graphs. However, being the walkers the one with longest distance, it could be perfectly explained by the correlation between session duration and distance covered.

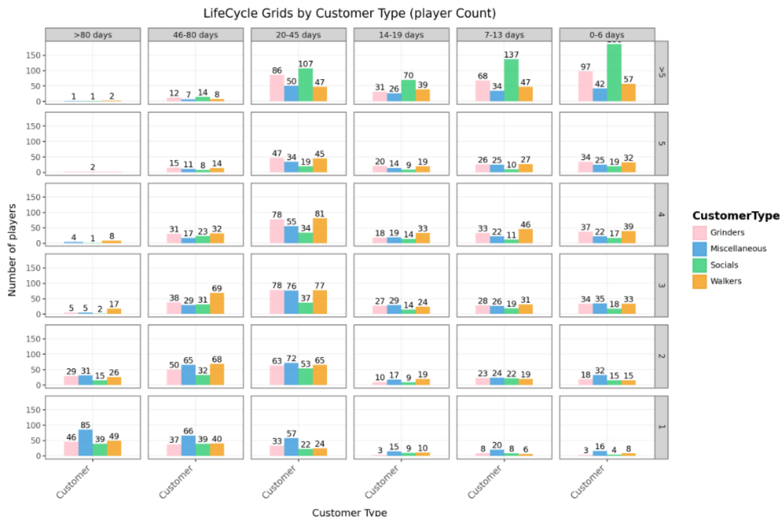
Finally, Socials boast also the highest number of social interactions, aligning perfectly with their playstyle, which revolves around engaging in multiplayer battles and interacting with other players. Interestingly, despite being highly engaged in previous analyses, Grinders have the lowest level of social interaction. However, this makes sense, as their gameplay is centered around individual achievement "*collecting as many monsters as possible*" rather than collaboration or competition.

Despite having no social interactions and low session duration, grinders lead in monster capturing as expected. Walkers and Miscellaneous players capture fewer monsters on average, meanwhile, Socials also rank notably high in monster capturing, indicating that while their engagement is primarily community-driven, they still actively participate in this core game mechanic.





To conclude, we have also built a **RF grid** to see **how frequently and recently the different customer groups play.**

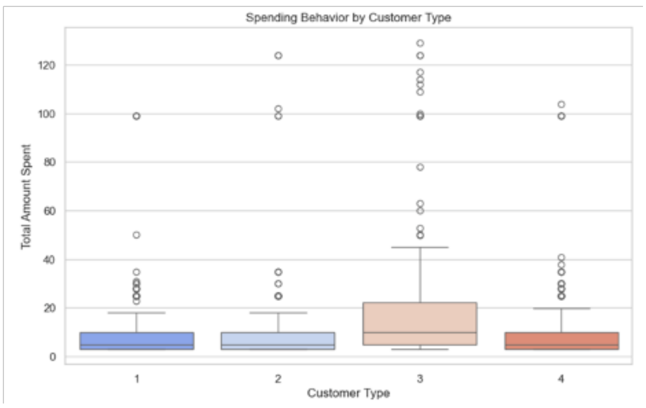


It shows that the 286 highest-frequency users we previously identified, predominantly belong to the socials category, confirming that during the summer, Socials are the most active players, with highest frequency and lowest recency, even if a consistent amount of them is moving towards the former active player quadrant.

After that we also see that grinders are also a consistent active group, while walkers and miscellaneous are mostly casual players.

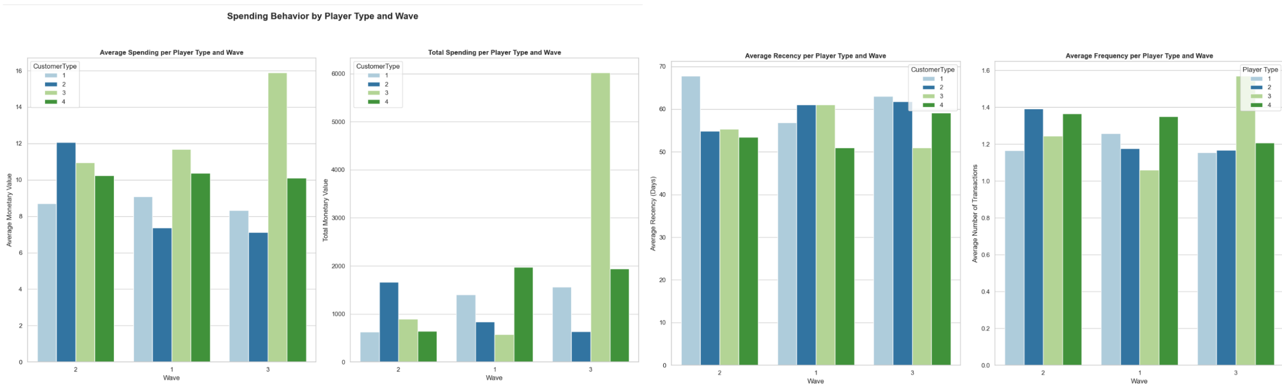
After getting a detailed picture of the paying behaviour, we analysed the spending patterns of the 4 groups.

Firstly, we see, that Socials, the most active customers, are also the ones with highest monetary value. Moreover, the plot shows few outliers with high spending that the business should care about.



Next, as we did before, we focused on 3 different waves, to see if there were any differences related to the 4 customer types. Indeed, the analysis highlights key differences in purchasing habits across player types. Socials (3) in Wave 3 emerge as the highest spenders. Walkers (1) and Miscellaneous (2) players show moderate spending across all waves, while Grinders (4) tend to spend the least, reinforcing their focus on gameplay rather than in-game purchases. Recency plot shows that players in Wave 3 have a slightly lower average recency, meaning they have played more recently compared to older waves. This suggests that newer players are more active and engaged, while players from previous waves may have higher dropout rates or less frequent activity.

However, the frequency of transactions remains relatively stable across all waves and player types, indicating that those who do spend tend to do so at a consistent rate, regardless of when they joined.



Finally, we moved onto **RFM analysis**, to see where the 4 customer types place themselves in the grid.

These grids don’t show anything new, proving once again that the few active customers are socials, and even if they contribute relatively to the total monetary value, they show the highest average monetary value. Moreover, we also notice that their recency is growing, signaling a potential risk.

Moreover, we see grinders follow the socials as high spenders, having also few instances in more recent quadrants.

Walkers and Miscellaneous players exhibit lower spending and activity levels, indicating that they might require additional incentives to encourage higher in-game purchases.

So, to conclude here are the profiles of the **four customer types**:

Walkers – Casual players who use the game during commutes or exercise. They have long session durations but low engagement. Despite being high-income, their spending is moderate. Targeted fitness-based rewards could increase spending.

Socials – The most engaged and highest spenders, playing competitively with friends. They lead in session frequency, experience, and social interactions, with some high-value outliers in spending. Exclusive multiplayer events could further boost revenue.

Grinders – Highly engaged but low spenders, focused on collecting monsters rather than social play. While they play frequently and capture the most monsters, they rarely make purchases. Offering rare collectibles or achievement-based rewards may increase spending.

Miscellaneous – A diverse mix of playstyles with moderate engagement and spending. They span all income levels, requiring personalized promotions based on in-game actions.

Task 3: Analyzing and Modeling Churn Behavior in Active Players During Fall

After getting the full picture of the customer base playing Moster Horder,the next step is identifying churning customers, investigating their playing behaviour and spending habits to understand why they left, and most importantly how to target them, given that it is worth retaining them.

However, in a setting where the service can be used without making any transactions, it is difficult to predict churn behaviors, as the transaction history alone does not necessarily indicate higher engagement.

Moreover, Monster Hoarder operates in a non-contractual setting, making the term churn relative, as a customer may simply take a break and return later, rather than permanently leaving the game. For the sake of simplicity, we will define churners through temporary inactivity.

To conduct a meaningful churn analysis, we first need a clear definition of what constitutes a churning customer. Since our primary focus is profitability, we have adopted the following definition from CEVER Technologies:

"Churn is defined as a player who executed at least one microtransaction during the summer period but did not make any transactions in the fall of 2022."

In simpler terms, a churning customer is someone who was previously monetized but has since stopped spending. Using this definition, we will predict churn behavior using data from summer 2022 and then we will analyze their in-game behavior, engagement levels, and possible indicators of disengagement. This will allow us to identify patterns, uncover key drivers of churn, and develop targeted strategies.

Since this analysis required data from fall game activity, we began by loading and inspecting the new datasets: fallfintrx (transaction data) and fallsesstrx (playing behavior data).

Our first step was to compare these datasets with their summer counterparts.

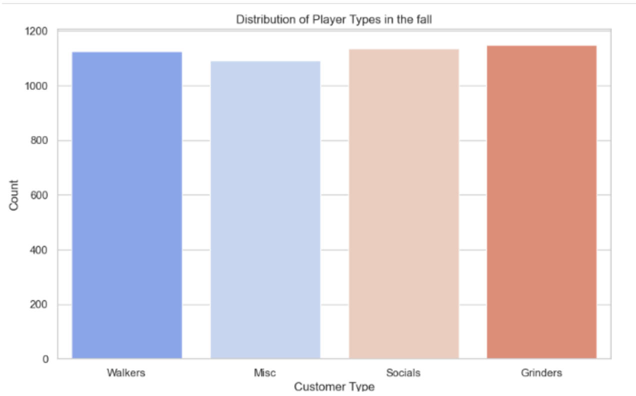
As expected, both the number of active players and spenders declined in the fall. Consequently, the average monetary value dropped by 17.84%, from $11 to $9, indicating a less favorable outcome than anticipated.

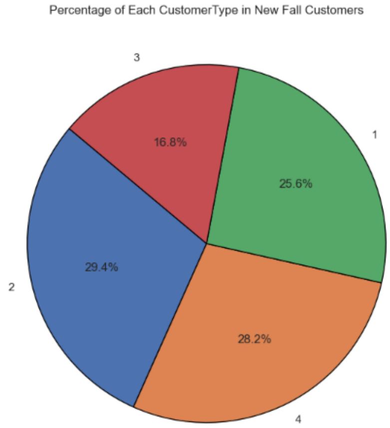
A possible explanation for this decline is the shift in player behavior. During the fall, 2,817 customers played at least once without making any transactions, while 1,684 players both played and made at least one transaction. This reinforces the trend that non-paying players continued to outnumber paying players, potentially impacting overall revenue.

However, despite the decline in total spenders, the percentage of paying users increased slightly from 36.38% in the summer to 37.41% in the fall. This suggests that while the overall player base shrank, those who remained active were more likely to spend.

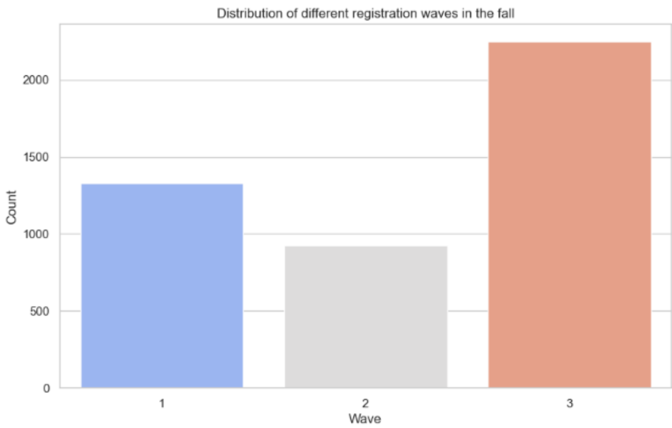
Additionally, all non-paying customers from the summer remained non-paying customers in the fall, indicating that all players continued engaging with the game without making purchases, while others may have churned entirely. This highlights the ongoing challenge of converting non-paying players into paying users.

In addition, the **distribution of customer types** has become more balanced, with all four groups being more evenly represented. Among them, Grinders became the most numerous.





This trend is also reflected in the distribution of newly acquired paying customers, where customer types are more homogeneously spread. In the fall, Monster Hoarders accounted for 18.76% of new paying customers. Notably, they did not originate from a single dominant group but rather from a mix of segments, with a slight majority coming from the Grinders and Miscellaneous categories.



Finally, we have also seen the d**istribution of the registration wave**, which shows no big changes with respect to the summer one.

After a quick analysis of the datasets we are working with, we built the base table needed for the churn predictions:

We labeled as 0 people who made at least one transaction both in the summer and fall, and categorized as 1 customers that were active spenders during the summer and did not continue spending during the fall, following the previously given definition.

The basetable contains 1711 customers and has 23 columns that might describe the churn behavior. These include:

'walkers','socials','fallbonus','grinders','Age','Gender','Churn','medium_income','frequency','monetary_value','low_income','recency','PlayID','Experience','Huntingcabins','Trainingschools','Huntinggrounds','Social','Monsters','Distance','third_wave','first_wave','Duration.

Before training and testing the predictive model, we split the dataset into train and test samples, with a test size of 30%.

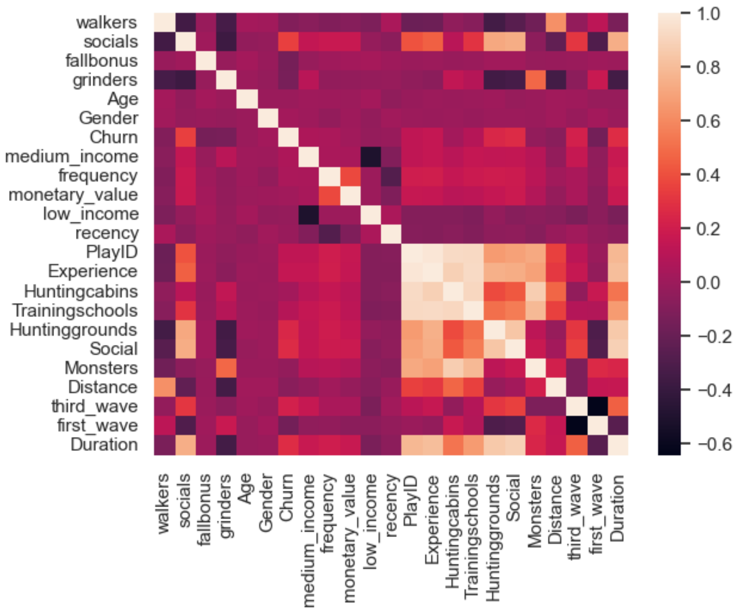
It was also check that both train and test datasets show an unbalanced proportion between churn and non churner:

```python
#is this a balanced dataset?
print('percentage churn in train dataset: ',np.round(np.mean(train['Churn'])*100,2))
print('percentage churn in test dataset: ',np.round(np.mean(test['Churn'])*100,2))
```

```
percentage churn in train dataset:  19.05
percentage churn in test dataset:  22.37
```

However, after trying to oversample the datasets using SMOTE[1] technique, we realized that it could lead to overfitting problems, especially when the minority class (churners) is artificially duplicated rather than representing real behavioral patterns. Therefore we decided to proceed without any solution, keeping in mind it can affect the predictive power of the models applied.

To avoid multicollinearity, we excluded variables that exhibited high correlation with others, ensuring that only the most essential information was retained for the prediction model.



The **correlation matrix** revealed that PlayID and Experience had a near-perfect correlation, as both serve as measures of player engagement. Additionally, Experience was strongly correlated with Huntingcabins, Trainingschools, Huntinggrounds, Social, and Monsters, which were also intercorrelated. Given this redundancy, we opted to retain only Experience as the primary engagement indicator.

Similarly, Distance and Duration showed a strong correlation, indicating that players who engage in longer sessions also tend to travel further. To minimize redundancy, we chose to exclude Distance. Finally, as expected, Frequency and Monetary Value were highly correlated. Since Monetary Value provides a more direct measure of player spending behavior, we decided to drop Frequency while preserving Monetary Value for analysis.
To summarize these are the variables selected to train our first prediction model:

Walkers, socials, grinders, Gender, Age, monetary_value, recency, Experience, Duration, fallbonus, first_wave, third_wave.

We first trained a simpler model, Logistic regression, to inspect the key predictors of churn status. However, the model summary showed that only 3 variables are significant (p-value<0.05)

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	-2.3946	0.413	-5.796	0.000	-3.204	-1.585
walkers	0.0488	0.288	0.169	0.865	-0.516	0.614
socials	1.8653	0.325	5.740	0.000	1.228	2.502
grinders	-0.2520	0.309	-0.815	0.415	-0.858	0.354
Gender	0.1832	0.165	1.111	0.266	-0.140	0.506
Age	0.0018	0.006	0.291	0.771	-0.010	0.014
monetary_value	-0.0075	0.006	-1.316	0.188	-0.019	0.004
recency	3.208e-05	0.002	0.013	0.989	-0.005	0.005
Experience	5.86e-06	1.35e-05	0.433	0.665	-2.07e-05	3.24e-05
Duration	-0.0006	0.001	-0.963	0.336	-0.002	0.001
fallbonus	-1.3199	0.249	-5.304	0.000	-1.808	-0.832
first_wave	0.1015	0.284	0.358	0.721	-0.455	0.658
third_wave	0.9404	0.286	3.288	0.001	0.380	1.501

[1] Synthetic Minority Oversampling Technique (SMOTE) is a statistical technique for increasing the number of cases in your dataset in a balanced way. The component works by generating new instances from existing minority cases that you supply as input. https://learn.microsoft.com/en-us/azure/machine-learning/component-reference/smote?view=azureml-api-2

To make sure we select the most relevant and informative factors we ran BidirectionalStepwiseSelection, a combination of forward selection and backward elimination, that identifies predictive variables that improve model performance.

It again selected third_wave, Socials and fallbonus.

The logistic regression trained with these 3 features shows these results:

Logit Regression Results

Dep. Variable:	Churn	**No. Observations:**	1197
Model:	Logit	**Df Residuals:**	1193
Method:	MLE	**Df Model:**	3
Date:	Sat, 22 Feb 2025	**Pseudo R-squ.:**	0.1540
Time:	14:42:56	**Log-Likelihood:**	-493.09
converged:	True	**LL-Null:**	-582.83
Covariance Type:	nonrobust	**LLR p-value:**	1.132e-38

	coef	std err	z	P>|z|	[0.025	0.975]
intercept	-2.2540	0.145	-15.586	0.000	-2.537	-1.971
socials	1.5980	0.167	9.570	0.000	1.271	1.925
fallbonus	-1.3198	0.248	-5.316	0.000	-1.806	-0.833
third_wave	0.7042	0.172	4.083	0.000	0.366	1.042

The model has a relatively low R-squared, indicating that the 3 predictors explains 15% of the variance in the Churn variable.

Moreover, the coefficients can provide us with information about the relationship between the independent variables and the dependent one.

For example, the positive coefficient 1.5 means that social players are more likely to churn than other types of players.

The same meaning for the coefficient of third-wave customers, who are likely to churn.

Finally, players who received a fallbonus, instead are less likely to churn than those who did not. The fall bonus was effective in retaining players. It reduced the likelihood of churn, suggesting that incentives play a key role in keeping players engaged.

To evaluate the model we have chosen 6 metrics, with a cutoff that minimize the costs, assuming a false negative cost of $94.21 and a false positive cost of $17.72, estimated in this way:

<u>Estimation False Positive Cost:</u>

Discount Value: Assuming an average in-game purchase (microtransaction) is approximately 87 dollars per player annually. [2]

Offering a 10% discount would equate to dollars 8.70 per player.

---

[2] https://en.wikipedia.org/wiki/Free-to-play?utm_source=chatgpt.com

Opportunity Cost: If the average revenue per user (ARPU) is dollars 90.16 annually, and offering unnecessary discounts reduces perceived game value, leading to a 10% drop in future spending, the opportunity cost would be dollars 9.02 per player. [3]

Retention Incentive Cost: In-Game Rewards: 1.00 personalized Discounts: 8.70, Communication Efforts: 0.30, Exclusive Content Development: 2.00, Community Management: 1.00

Total False Positive Cost: 8.70 (discount) + 9.02 (opportunity cost)+ 13 (Retention Incentive Cost) = 30 dollars per player.

<u>Estimation of False Negative Cost :</u>

Lost Revenue: With an ARPU of 90.16 dollars annually and an average player lifespan of 1 year, the CLV is 90.16 dollars. [4]

Acquisition Cost: The average cost per install (CPI) for mobile games is approximately 4.7 dollars on iOS and 3.4 dollars on Android. [5]

Assuming a 50/50 user split between platforms, the average CPI would be 4.05 dollars.

Total False Negative Cost: 90.16 (lost revenue) + 4.05 (acquisition cost) = 94.21 dollars per player.
The cutoff calculated, 0.18, allowed the model to minimize false negatives to retain high-value players, as the cost of losing a player far outweighs the cost of offering unnecessary discounts.

```
Accuracy Score :  75.88
ROC AUC score :  77.25
Precision score :  47.2
sensitivity score :  66.09
specificity score :  78.7
Log-Likelihood (MLE) : -231.8103
```

The metrics calculated shows that the model performs reasonably well in predicting churn, with an accuracy of 75.88% and an ROC AUC of 77.25%, indicating good discriminative power between churners and non-churners.

However, while it correctly identifies 66.09% of actual churners (recall), its precision is moderate (47.2%), meaning that many predicted churners do not actually churn. The specificity (78.7%) suggests it effectively avoids misclassifying non-churners, helping reduce unnecessary retention efforts. While the log-likelihood (-231.81) shows the model fits the data reasonably well, there is still room for improvement.

While logistic regression is simple and interpretable,we also trained a more advanced model. XGBoost is often a better choice for churn prediction due to its ability to handle non-linearity, interactions, missing data, multicollinearity, and imbalanced datasets, leading to higher predictive accuracy.

However, in this case, it showed nearly the same results of logistic regression, with the same selected feature and almost identical evaluation metric values, probably because of the similar cutoff values and similar, almost overlapping distribution of the predicted probabilities of the 2 models. This can be because the dataset does not contain strong non-linear patterns or complex

---

[3] https://scoop.market.us/gaming-monetization-statistics/?utm_source=chatgpt.com

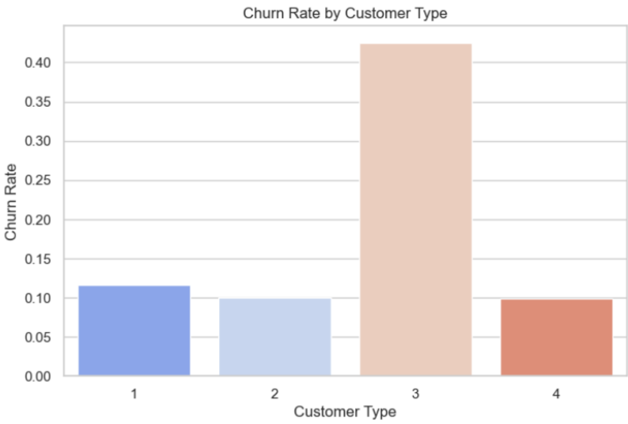
[4] https://scoop.market.us/gaming-monetization-statistics/?utm_source=chatgpt.com

[5] https://www.businessofapps.com/marketplace/user-acquisition/research/user-acquisition-costs/?utm_source=chatgpt.com

feature interactions, as a consequence, XGBoost does not gain an advantage over logistic regression.

Besides churn prediction, it is also important to investigate the player behavior and spending habits related to the churn status, in order to deeply understand the customer base and tailor a targeted strategy

First of all, the overall churn rate in the fall was 20.05%, a value that is still not bad for a freemium business since most Fto Play games lose 80% of their players within the first 7 days.



However, the customer category that churned the most might be a concern.

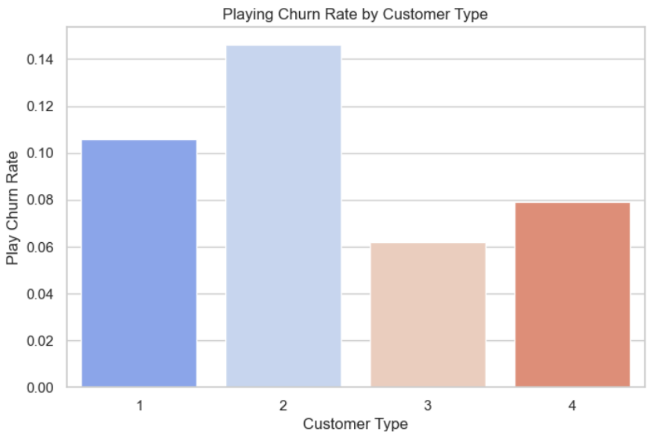
Socials have been shown to be big spenders, and active customers, but their recency slowly increased, and most of them stopped buying in the fall.

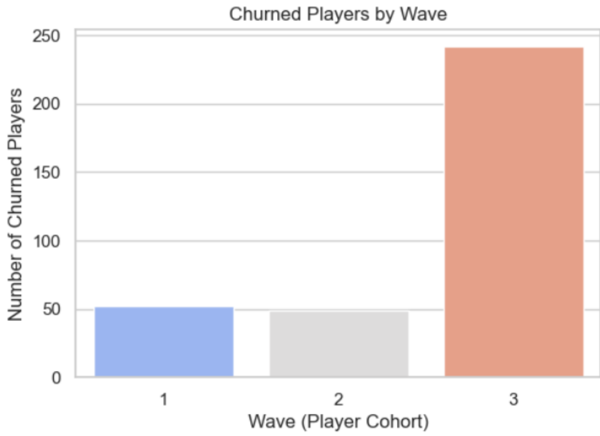
This phenomenon is due to a fall in the gaming mechanism that creates a customer lifecycle that inevitably leads to transaction churn soon.

It was demonstrated that new players start by exploring the game first without making any transactions. As their experience grows and time passes they start shopping coins. However, after gaining a good experience, the game allows them to earn enough coins, to continue playing and growing without making any transactions.

Since the game is a freemium business this problem needs to be fixed, also because 18.76% of fall's paying customers were new, suggesting that while the game attracted fresh spenders, it did not fully compensate for the loss of previous ones. This results in a small net decline in total paying customers, that, having high acquisition costs, must be fixed on the churn side.

Fortunately, even if most of the socials churned, we have noticed that they still actively play, making it easier to convert them back into active customers.





The predictive models also indicated that players from the third wave were more likely to churn, as the corresponding variable had a positive and significant coefficient. This prediction was further confirmed by actual data, which showed that the majority of churned players came from the third registration cohort.

This finding suggests that players who joined during the third wave were less engaged or had a lower likelihood of long-term retention.

Finally, to conclude our analysis, the customer lifetime value is an essential measure that must be calculated and considered to craft any strategy.
It allows the business to understand the money that a single customer brings to the company over their lifetime minus the costs for acquiring them. We estimated the latter to be around $10.

Being in a non-contractual setting, we could not calculate a proper retention rate, for this reason we firstly estimate the expected number of future purchases for each customer over a one-year period, followed by a Gamma-Gamma model to predict the monetary value associated with these future transactions.

To refine our insights, we subtracted a fixed customer acquisition cost from each CLV estimate to ensure a more realistic evaluation of customer profitability. Customers whose CLV resulted in a negative value were filtered out, as they would not contribute positively to long-term revenue. Finally, we segmented customers into four distinct categories based on their CLV distribution:

- Champions: The highest-value customers with strong engagement.
- Loyal Customers: Customers who contribute consistently but at a slightly lower level than Champions.
- Need Attention: Customers with moderate value who may require targeted engagement strategies.
- Hibernating: The least active customers who may have limited long-term potential.

As expected, we have noticed that the churn rate is lower for champions compared to the churn rate of Loyal Customers, suggesting higher investemnts to retain this group of profitable customers, that are likely the more experienced, and have enough earned coins to continue gaming for free.

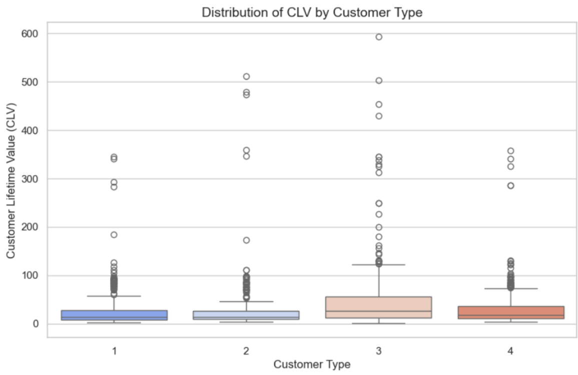
Furthermore, we also investigate the clv distribution across the 4 different customer types.

The findings gathered are:

Socials have the highest customer lifetime value, but they are also at higher risk of churning (42.5%)

Grinders are the best good long-term customers (lowest churn rate), with high monetary value and customer lifetime.

Finally, Walkers and Miscellaneous show similar results, with low average monetary value and customer lifetime value but with a few outliers that that should be nurtured.



Regarding outliers, we can see they are present in each group, confirming what we have already discovered about a few high value customers.

To conclude the game show key issues that need to be solved:

1. High Dependency on One-Time Buyers: The game does not encourage continued spending after the first purchase.
2. Freemium Model Limits repeated purchase incentives: players can progress significantly without paying
3. High customer lifetime value customers are at risk to churn
4. New customers take a while to begin making transactions
5. Player activity declines in Fall/Winter.

These are the proposed strategies for CEVER technologies:

1. Retention plan for Socials
2. Encourage Spending from Grinders
3. Seasonal Retention Strategy
4. Reduce time to first purchase
5. Incentivize repeated purchases
6. Encouraging Spending at Every Experience Level

Task 4: Developing a Data-Driven Marketing Strategy to Reduce Churn and Maximize Profitability

Retention plan for Socials

This first strategy aims at targeting the socials that have shown higher spending and purchase frequency but with high recency.

We know that they are still actively playing, this gives the advantage that they would read in game notifications and new events.

First, we suggest Introducing ranked tournaments where entry is free, but players can pay for premium brackets with higher rewards.

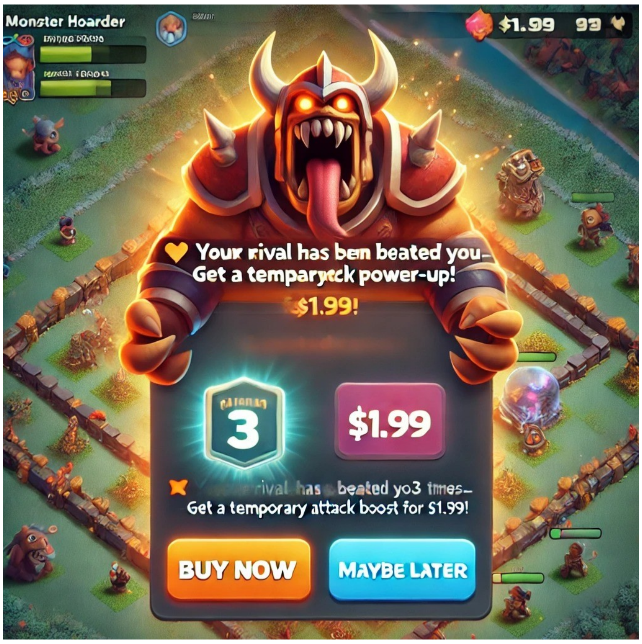
This takes advantage of their competitive side and will pay for status-based rewards.

Offer seasonal Battle Pass with exclusive social-themed rewards, it can Encourages groups to spend together, increasing overall purchase frequency.

Example: "*Unlock premium guild features & team XP boosts with the Social Pass—$5.99/month!*"

We can also leverage FOMO and loss aversion to offer a limited-time paid power-up to increase their odds of winning, after a player has been beaten by their rival a few times

Example: " *Your rival has beaten you 3 times—get a temporary attack boost for $1.99!*"



Finally, we can offer the option to make gifts for your friends.

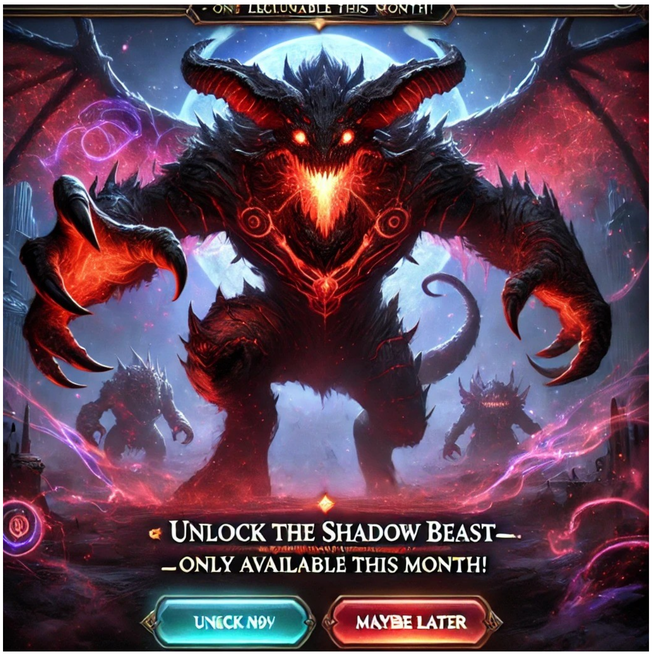
Socials usually play with their friends, and the possibility to buy each other different skins or booster, can be a great incentive to increase their spendings.

Encourage Spending from Grinders

We have seen that grinders are skilled players; however, they make transactions less often than socials.

Since they are eager to collect monsters, one possible solution could be introducing premium-only monster species that cannot be found through regular play.

Example: "*Unlock the Legendary Shadow Beast—Only Available This Month for $3.99!*"



Another solution that leverages their obsession with collecting monsters could be placing a cap on free monster storage and allow premium expansions for purchase.

Example: "*Your collection is growing! Expand your Monster Vault for $1.99*."

We also noticed that Grinders are skilled gamers that value efficiency, this is why we can introduce location-based premium events where certain monsters only appear for paying players. This small fee would optimize hunting.

Example: "*Activate the Ultra Lure—Guaranteed Rare Monster Spawn for $2.99!*"

Seasonal Retention Strategy

We have seen that people tend to play less during the fall winter due to the weather conditions. Walkers in particular play the game during their daily commute or exercise routine, which is difficult to do in cold temperatures.

For these reasons we suggest introducing a fall/winter pass with exclusive skins & power-ups for winter walking milestones. This encourages customers to reach long term goals like walking 5km in one week, that are feasible to reach at home or on a treadmill.

Moreover, since the fall bonus sent at the end of summer actually worked, it must be implemented more intelligently. Target high clv customers only, in order to have a lower cost than the customer clv. Moreover, send it to active spenders, especially the outliers we have visualized, present in each customer group.

Invest especially in a spring ‘come back’ bonus, to reactivate customers, paired with a guided "Catch-Up" mode, that would gradually reintroduce players to new game features.

Reduce time to first purchase

We have seen that customers take 1 to 2 months to reach the purchase level of the customer journey, implying that most of the new customers don’t contribute to the revenue.

To address this problem, we suggest creating early spending incentives that feel natural and rewarding, without making purchases feel mandatory.
Examples could be offering discounted features for the first 72 hours.

Incentivize repeated purchases

Since one-time buyers make up the majority of revenue, the goal is to convert them into repeat customers by making spending feel habitual, and necessary for continued engagement.
Introduce a First Purchase Bonus, to reward first-time spenders with an exclusive discount that would incentivize a second purchase.

Another effective strategy that many games are implementing is offering many Low-Commitment microtransactions and sending banners to advertise them. This would increase the number of transactions as players are more likely to spend small amounts first before committing to larger purchases.

Encouraging Spending at Every Experience Level

Highly experienced players, particularly Socials, stop making purchases because they accumulate enough in-game coins. The strategy introduces meaningful spending incentives that cannot be bypassed with earned currency, ensuring even experienced players find value in purchasing premium items.

For instance, premium battle modes or private tournament brackets where only paid entries are allowed, or simply premium skins, animations, and battle effects that cannot be purchased with earned coins—only through microtransactions.
In this way CEVER technologies would avoid the risk of losing highly engaged customers, as happened with Socials.

Conclusion

To conclude, we have analyzed in detail the players of Moster Hoarder and identified different problems that might impact the profitability of the game.

The first top 2 priorities are retaining Socials and keep engaging Grinders, encouraging to spend more. This is because their customer lifetime vavalue is high and would cover the investments.

We have also concluded not to present a specific strategy to address miscellaneous, as their heterogeneity and low clv and monetary value would not be a high-ROI effort.

Additionally, we have proposed potential changes required in game mechanics and in-game shop design, to maximize the game's revenue.

Group members

Chiara Chianese - 102335
Eleonora di Mario - 102336
Giulia Macis - 102347
Philippine Carpentier - 86070
YeonGyeong JO - 102341
Anagh Bhardwaj - 87429
Eklavya Singh - 87390

Appendix

Figure 1

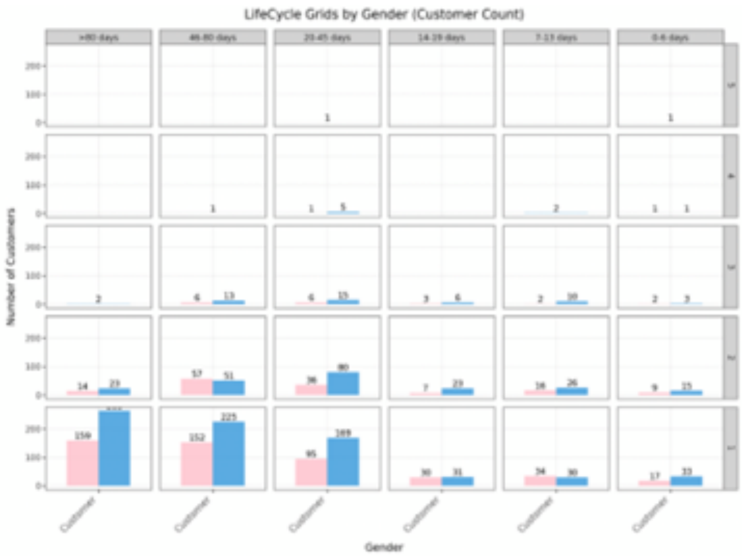


Figure 2

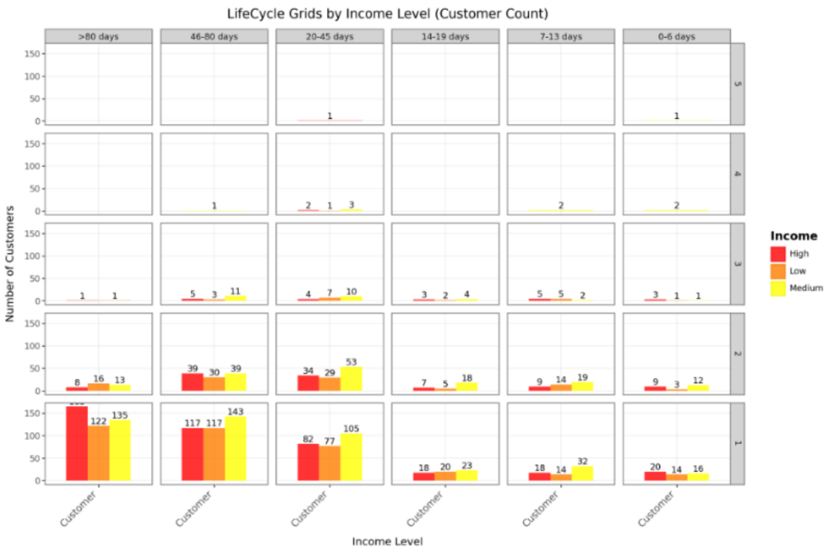


Figure 3

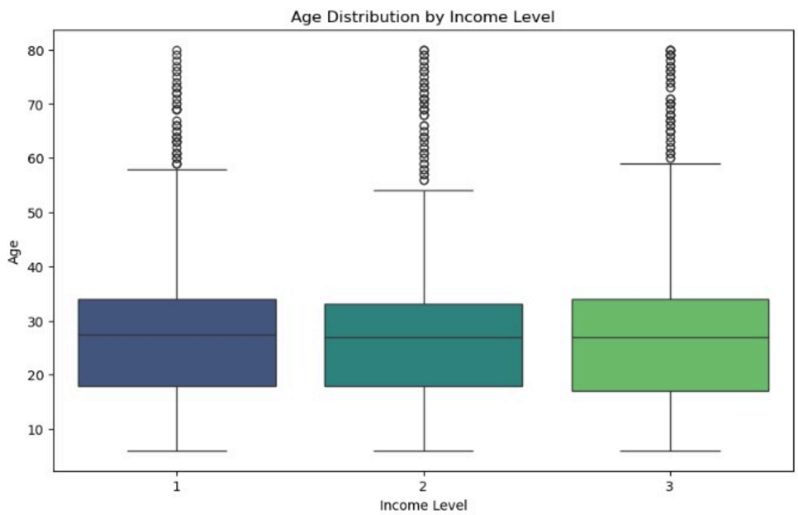
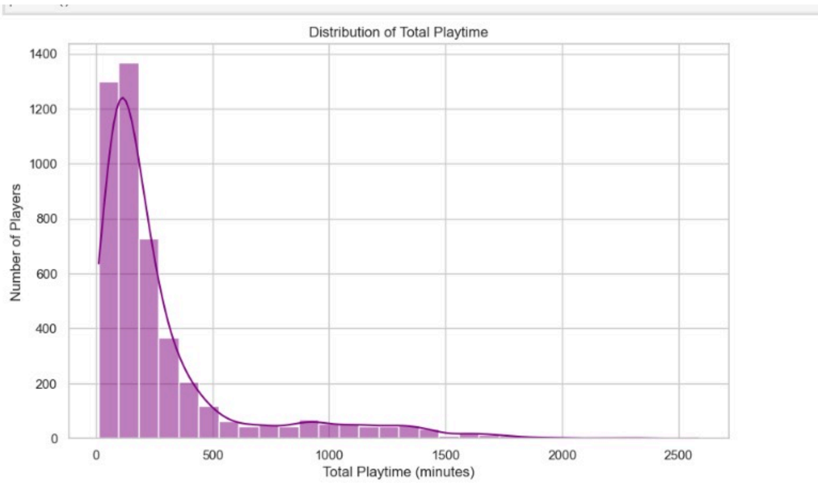


Figure 4


Distribution of Total Playtime
Number of Players
Total Playtime (minutes)
0
200
400
600
800
1000
1200
1400
500
1000
1500
2000
2500