Microtransaction is the Future of MMORPGs*

Tian Yi Zhang

09 May 2022

Abstract

Advertisement and monetization schemes are essential to obtaining a stable revenue stream and ensuring the continued surival of an online video game. We obtain data through a survey distributed in several online communities for the massively multiplayer online game known as Black Desert Online developed by Pearl Abyss and analyze it through several graphs, tables, and a model in this paper. Results show a positive correlation between a player's total playtime, rated importance of subscription items, and their total monetary spending in the in-game cash shop. Players with a female main character in the game spend on average 9.23% more real world currency on cosmetic items in-game. On average, upfront cost for the game accounts for only 5-6% of a player's total spending on the game. These findings suggest that MMORPGs should release with an upfront cost before transitioning to free-to-play later down the line. The introduction of more than one type of subscription items becomes marginally less profitable. Further analysis on the topic of whales is recommended. Keywords: VP, BK, OM, Whale, MMORPG, Subscription, Pearl Shop, Central Market, Black Desert Online, Pearl Abyss, Gender, Microtransaction

Contents

1 Introduction **2**

2 Data **3**

2.1 Survey Design 3

2.2 Dataset Cleaning and Description 3

2.3 Preliminary Analysis 5

3 Model **10**

4 Results **12**

4.1 Model Summary 12

4.2 Character Gender Inequality 13

4.3 Free Game Package vs Microtransaction 13

4.4 Subscription Items 14

*Code and data are available at: https://github.com/zhan7818/bdo_monetization

5 Discussion **15**

5.1 Free Game Versus Upfront Cost 15

5.2 Number of Subscription Items 16

5.3 Whales 16

5.4 Limitations 17

Appendix **18**

A Datasheet for the Dataset **18**

B Additional Details **24**

References **25**

1 Introduction

Up until a decade or two ago, singleplayer-driven content was the predominant focus in the video game industry. This was mostly a product of the time, as the infrastructures needed to support online-based video games were yet to be fully developed. More importantly, what this also meant was a lack of post-launch updates being delivered for a game, as developers also lacked means to deliver constant updates effectively into the hands of the consumers. As a result, most video games received little to no post-launch updates aside from one or two major expansions that contained sufficient content to be put on the shelves as add-ons to the original product. Despite this, consumers back then were generally content with the amount of content they received for their purchases.

The drastic shift towards online gaming started in the early 2000s, according to Chikhani (Chikhani 2015). During that time period, networking structures matured and internet services became more affordable. As such, more and more video game developers started focusing more on the online aspects of gaming. The number of multiplayer-focused video games in the market started to steadily increase. Now, in an era of live service online video games, developers are pressured to continuously produce and deliver new and exciting content on a consistent schedule in order to compete with other products. This need to create more content for a game means that development resources are constantly required, and online focused games can no longer afford to charge their consumers a simple upfront fee; live service video games must find an effective and profitable monetization scheme in order to fund their continued development.

The most prominent examples of live service online video games are the massively multiplayer online role-playing games, henceforth referred to as MMORPG, a genre of video games known specifically for its focus on multiplayer. Due to their focus on progression-based gameplay, developers of MMORPGs must constantly create new and refreshing content to add to their games. To subsidize the development costs that go along with it, developers have created a variety of monetization models to ensure a steady revenue stream. For example, World of Warcraft, one of the biggest MMORPGs, relies primarily on a monthly subscription from its players. Others, such as Lost Ark, adopt a mix of strategies with a monthly subscription that grants massive boosts in-game and an in-game cash shop that allows players to purchases items for a variety of purposes using real world currencies. If done properly, an analysis on the monetization schemes of MMORPGs could have meaningful implications for future MMORPG developers on how they should monetize their video games to maximize revenue.

This paper primarily focuses on the MMORPG known as `Black Desert Online`, commonly and henceforth referred to as BDO, as it is an MMORPG that utilizes a of mix of upfront cost , subscription and in-game cash shop. `Pearl Abyss`, the developers behind Black Desert Online, never published any official statistics

that touch upon the revenue of Black Desert Online. As such, the following analysis will rely on data collected by the author of this paper.

The remainder of this paper is organized as follows: Section 2 covers the survey design, sampling and preliminary analyses of the dataset used in this paper. Section 3 details the model used to find possible correlations between variables in the dataset, the results and implications of which will be covered in Section 4. Any limitations or possible extensions of this paper will be discussed in Section 5.

All tables and graphs presented in this paper are produced using the statistical programming language `R` (R Core Team 2020). The `knitr` (Xie 2021) and `kableExtra` (Zhu 2021) packages are used for the tables, and the `ggplot` (Wickham 2016) and `cowplot` (Wilke 2020) packages are used for the graphs.

2 Data

2.1 Survey Design

Due to the lack of APIs provided by Pearl Abyss that concerns the subject of this paper, convenience sampling is used to collect data from the general North American and European playerbase for the game. The survey is created in Google Forms and is distributed to Black Desert Online players on Steam discussion, Black Desert Online forum board, and the Black Desert Online subreddit on Reddit.[1] The survey consists of questions mostly regarding the respondents' amount of spending in the various monetization models employed by Black Desert Online, as well as some general background information regarding each respondent and their in-game characters. The survey questions could be largely divided into the following set of topics:

- Background characteristics such as age, gender and playtime
- Upfront spending for the game such as the purchase of game packages
- Spending for in-game subscription-esque items using either in-game or real world currencies
- Opinion on in-game subscription-esque items
- Spending for other general cash shop items in-game using either in-game or real world currencies

In the post containing the survey on each of the distributed sites mentioned above, efforts were made to be as transparent as possible with the participants: respondents were informed of the intended usage for the data collected and were assured that no information that can be used to identify them will be collected nor included in the survey and the final paper.

2.2 Dataset Cleaning and Description

A total of 212 responses were collected from the survey. The data was exported to Excel (Microsoft Corporation 2022-03-30) where the columns names were shortened to be readable. Two new columns were created from the responses:

- packages_usd: amount in USD spent on purchasing game packages
- packages_free: binary indicator of whether the respondent received the game for free

The dataset is then cleaned using the statistical programming language R (R Core Team 2020) and the `janitor` (Firke 2021) and `dplyr` (Wickham et al. 2022) package. The `pointblank` (Iannone and Vargas 2022) package is used to set up validation tests for the class and the content of the dataset. A new variable, `age_group`, is constructed based on the ages of the respondents. The final dataset consists of the following variables:

[1] A PDF copy of the survey can be found in the github repo for this paper in inputs/survey/bdo_survey.pdf.

- `age`: age of the respondent
- `age_group`: age group of the respondent
- `gender`: gender of the respondent; male, female or other
- `character_total`: respondent's total number of characters in Black Desert Online
- `character_female`: respondent's total number of female characters in Black Desert Online, excluding the character classes Tamer and Shai
- `playtime_total`: respondent's total playtime in Black Desert Online
- `playtime_main`: respondent's playtime on their primary character in Black Desert Online
- `gender_main`: gender of respondent's main character; male or female
- `packages`: the game packages purchased by the respondent
- `packages_usd`: the amount in USD spent purchasing game packages by the respondent
- `package_free`: binary indicator of whether the respondent received the game for free
- `importance_vp`: respondent's rating on the importance of Value Pack to their enjoyment of the game
- `importance_bkom`: respondent's rating on the importance of Blessing of Kamasylve and Secret Book of Old Moon to their enjoyment of the game
- cm_vp_count: number of Value Packs purchased on the central market by the respondent
- `cm_bkom_count`: number of Blessing of Kamasylve and Secret Book of Old Moon purchased on the central market by the respondent
- `ps_vp_usd`: amount in USD spent by the respondent purchasing Value Packs in the pearl shop
- `ps_bkom_usd`: amount in USD spent by the respondent purchasing Blessings of Kamasylve and Secret Books of Old Moon in the pearl shop
- `cm_outfits`: number of outfits purchased on the central market by the respondent
- `ps_cosmetic_usd`: amount in USD spent by the respondent purchasing cosmetic items in the pearl shop
- `ps_functional_usd`: amount in USD spent by the respondent purchasing functional items in the pearl shop

Typically, players in Black Desert Online will consider one character as their primary character and spend considerably more time on that character than their other characters. As such, playtime for the main character and total playtime were survyed separetely.

There are two methods to purchase premium items in this game, through the `central market` or through the `pearl shop`: The pearl shop is BDO's cash shop, accessible to every player, where players could use `pearls`, a currency converted from real world currencies, to purchase premium items for cosmetic and functional purposes. Furthermore, players who purchased premium items in the pearl shop could list these items onto the central market for sale. Other players could then use `silver`, the primary in-game currency, to purchase the said premium items. This system serves as a way for some players to obtain more silver using real world currencies, and for other players to obtain premium items without ever spending any real world currencies to do so.

There are 3 subscription-esque items in the game: `Value Pack`, `Blesing of Kamasylve`, and `Secret Book of Old Moon`, henceforth referred to as `VP`, `BK`, and `OM`. VP existed as an item since the launch of BDO, and BK and OM were introduced later on. Each of these 3 items offer unique bonuses and to players such as shortcuts and quality of life buffs to certain in-game systems to make them easier to use. Typically, VPs can be purchased for 30 days, 60 days, and 90 days, and BKs and OMs can be purchased for 15 days (exceptions occur in bundles and/or event sales including the said items). Because of this, these 3 items are usually bought once or twice per month so that a player can consistently maintain the buffs they provide. As such, they could be reasonably considered as a pseudo-subscription model employed by BDO.

Outside of the 3 subscription items, there are other premium items offered in the pearl shop, some of which are permanent while others are consumables. These items can be mostly categorized into two types: cosmetic (e.g., character outfits, home furnitures, dyes, etc) and functional (e.g., camping tent, energy potions, pets, etc). Most of the cosmetic items can be listed on the central market just like the subscription items, but functional items are generally only purchasable through the pearl shop.

Table 1: Summary of Average Playtime and Spending in Upfront and Cash Shop Categories in Black Desert Online

			Average Spending (USD)		
Age Group	Gender	Has Free Package	Pearl Shop	Game Packages	Avg Playtime (hrs)
13-22	Female	0	521.7	45.0	3262.7
13-22	Female	1	183.3	0.0	2883.3
13-22	Male	0	987.5	74.8	5374.3
13-22	Male	1	162.0	8.7	2411.1
23-32	Female	0	1798.3	83.3	6802.2
23-32	Female	1	420.0	70.0	15000.0
23-32	Male	0	2182.3	91.7	11258.5
23-32	Male	1	745.4	47.0	5623.9
33-42	Female	0	1437.5	127.5	9400.0
33-42	Female	1	400.0	80.0	2482.0
33-42	Male	0	1973.0	73.0	19018.1
33-42	Male	1	450.0	0.0	10000.0
43-52	Female	0	1750.0	160.0	1250.0
53-62	Female	1	100.0	0.0	1500.0
53-62	Male	0	278.5	30.0	8500.0
63+	Male	1	430.0	0.0	19990.0

2.3 Preliminary Analysis

Table 1 showcases a summary of the average spending of the respondents, categorized into age groups, gender, and whether the respondent obtained the free package for the game, in USD. Pearl Abyss, the developer and publisher of BDO, frequently employs events and/or utilize collaborations with other companies (e.g., twitch prime, anniversary event, etc) to hand out free packages for the game, which simply allows a player access to the game but grants no other benefits or items (Other package tiers grant players various items and benefits).

Already, we observe that there is a noticeably higher average for spending in the pearl shop compared to upfront package purchases. Note that this difference could be potentially misleading; The large majority of players do not spend much money in the pearl shop, and the exceedingly high average pearl shop spending is inflated by a select few players who have invested massive amounts of real world currency into the game. This issue is further elaborated upon in Section 5.

Due to having a variety of ways for consumers to spend real world currencies in BDO, a considerable variation of money spent on the game is expected for respondents of the survey. This is demonstrated in Figure 1, where we can see that the majority of respondents spend less than $1000 USD in the pearl shop, but there are also quite a few respondents that fall outside of that range by a substantial amount. Note that it is unwise to simply categorize these responses as simple outliers and exclude them from the dataset, as we must consider the phenomenon of "whales." This idea is also further explored in Section 5.

Figure 2 plots the rated importance of VP, and BK and OM. We can observe that there is a considerable number of respondents who consider VP as extremely critical to their enjoyment of the game. At the same time, we can see that a large number of respondents consider BK and OM largely unimportant to their enjoyment of the game. This shows that subscription-esque items are not universally regarded as equally valuable by the respondents.

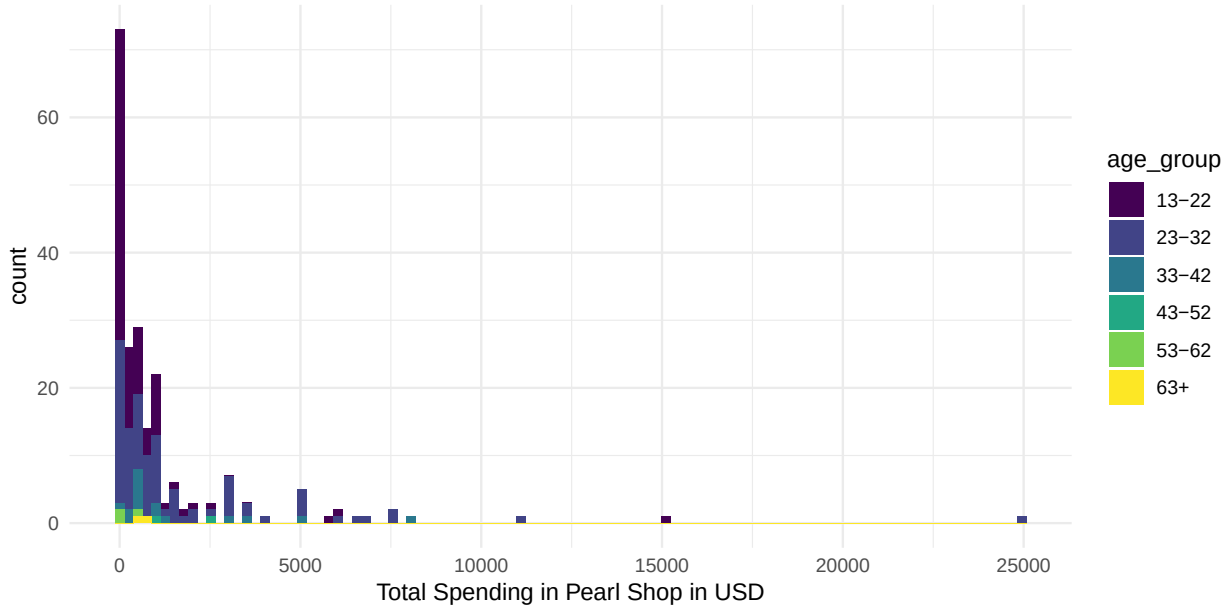

Figure 1: Spending in Pearl Shop in USD

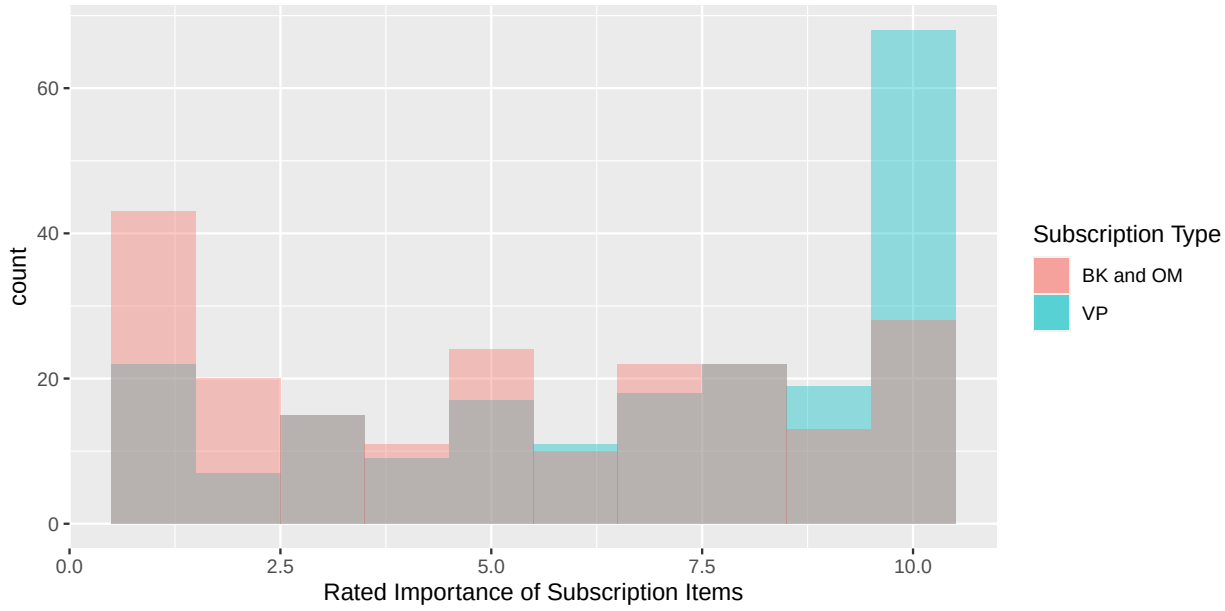

Figure 2: Rated Importance of Value Pack, and Blessing of Kamasylve and Secret Book of Old Moon

2.3.1 Supply and Demand

An important aspect of the central market in BDO is the issue of supply and demand: Despite being able to list most cosmetic items onto the central market, there is generally always more demand for these cosmetic items than supply. To combat the issue of price inflation, most items that can be listed on the central market in BDO–including value packs–have a price ceiling as dictated by the developers of the game. When that price ceiling is reached (i.e., the highest pre-order price placed for the item is the maximum price point), any further listings of the item will instead `randomly` choose a player with a max-priced pre-order placed and sell the item to the said player.

Because of the mechanic described above, an item pre-ordered at its price ceiling does not follow a first-come first-serve basis and the central market instead relies on randomness to decide which player to sell the item to. This means that for items with drastically higher demand than supply (such as a value pack), players are never guaranteed to obtain the item, regardless of how long they have placed the pre-order for the said item already. This often leads to cases where a player could place a pre-order for an item for an extensive period of time (months or years) yet never successfully purchases the item from the central market.

Figure 3 showcases this issue: For respondents who claimed that having a VP active is crucial to their enjoyment of the game, only a few reported having consistent success purchasing VPs from the central market. Because of this, respondents could be pressured into relying more on purchasing VPs through the pearl shop.

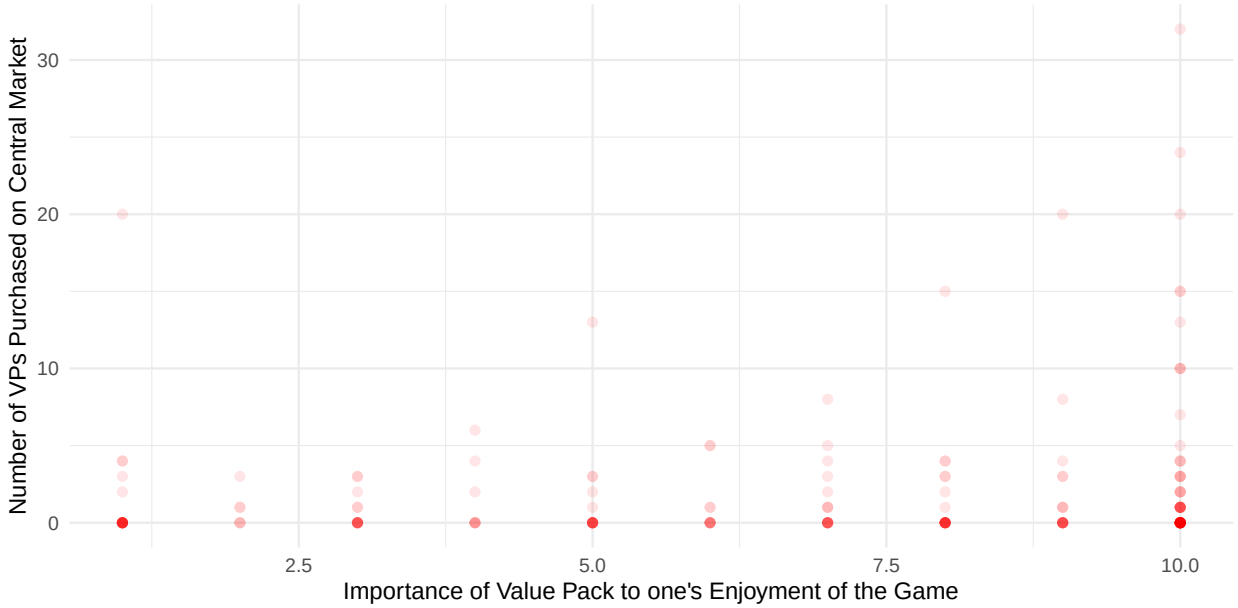

Figure 3: Rated Importance of Value Pack Versus Number of Value Packs Purchased on Central Market

This line of thinking is echoed by the observations in Table 4, where the vast majority of respondents report spending a considerable amount of real world currencies on purchasing VPs through the pearl shop. For reference, a 30-day value pack in the pearl shop costs 1500 pearls (the equivalent of $15.00 USD).

2.3.2 Gender of players and Characters

Gender-based appeal is a deep-rooted design pattern in most Korean MMOs. Most cosmetic options are designed to appeal to the general playerbase and often forfeits any semblance of practicality. This issue is very apparent in female armor designs in Korean MMOs, and Black Desert Online is generally not an

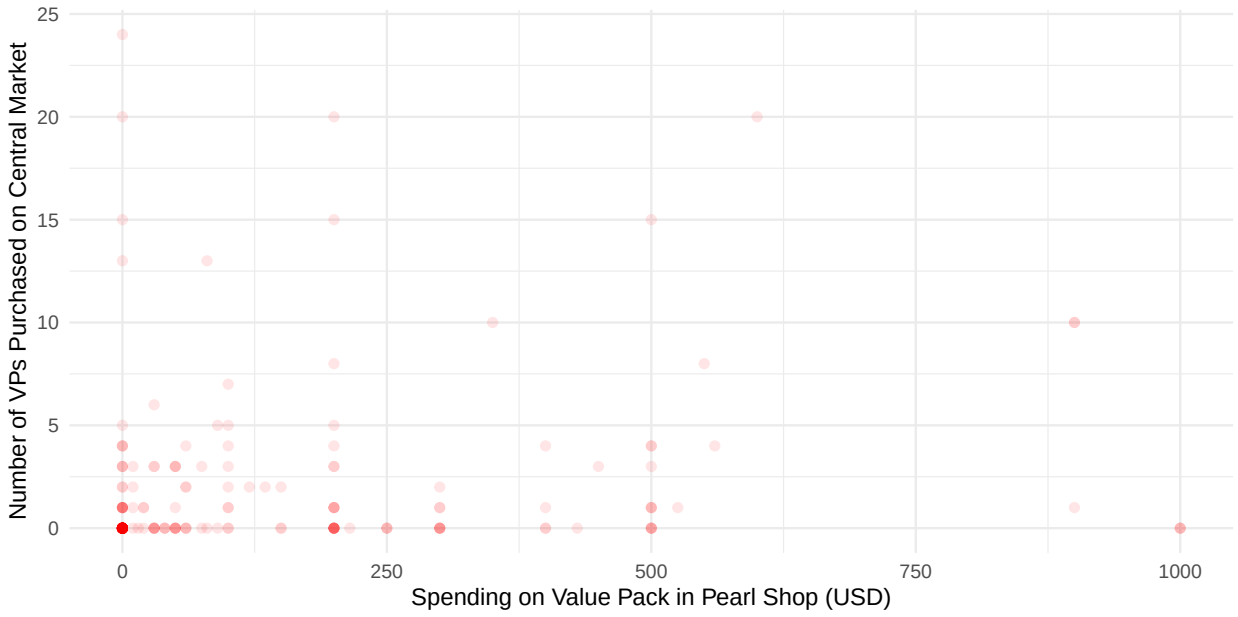

Figure 4: Spending on VPs in Pearl Shop Versus Number of VPs Purchased on Central Market

exception to this. There are 13 adult female characters in BDO and 9 adult male characters. To see if these appeal-based design decisions truly have any merit behind them, data was collected on the number of adult female characters that each respondent has. The intent is to compare it with their spending in cosmetic items, as briefly demonstrated in Figure 5.

It can be observed in Figure 6 that there are more respondents whose main character in BDO is female than male at all age groups. This statistic may not be surprising, given that male respondents similarly dominate all age groups, as shown in Figure 7. Nevertheless, this may grant some insight into the targeted audience of Black Desert Online.

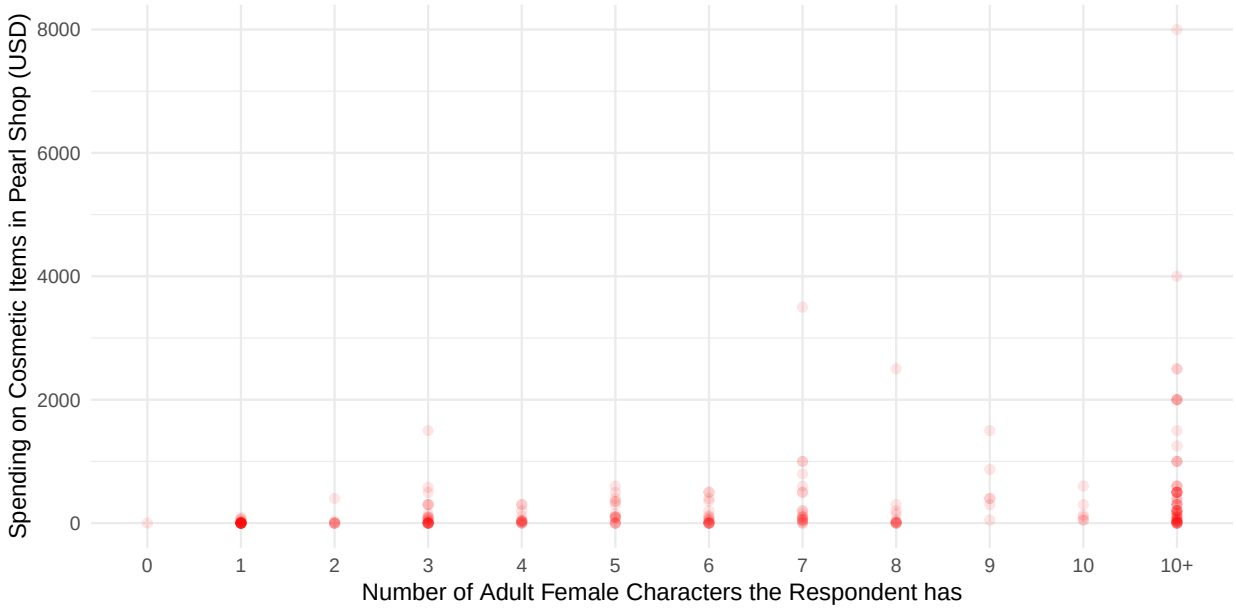

Figure 5: Number of Female Characters Versus Total Spending

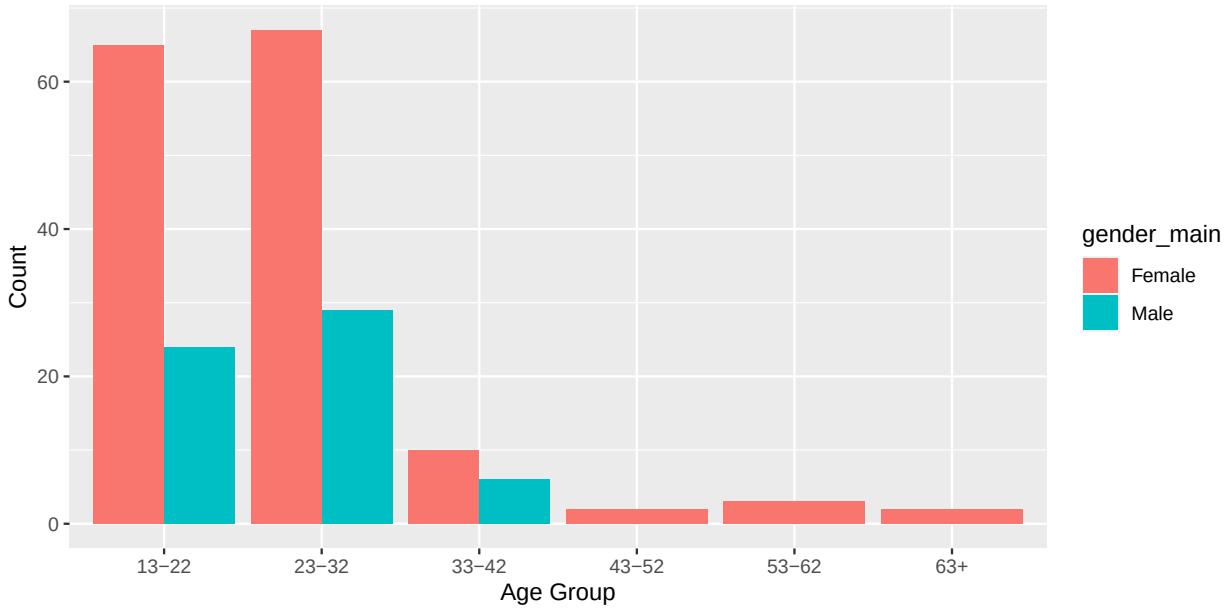

Figure 6: Gender of Main Character Across Age Groups

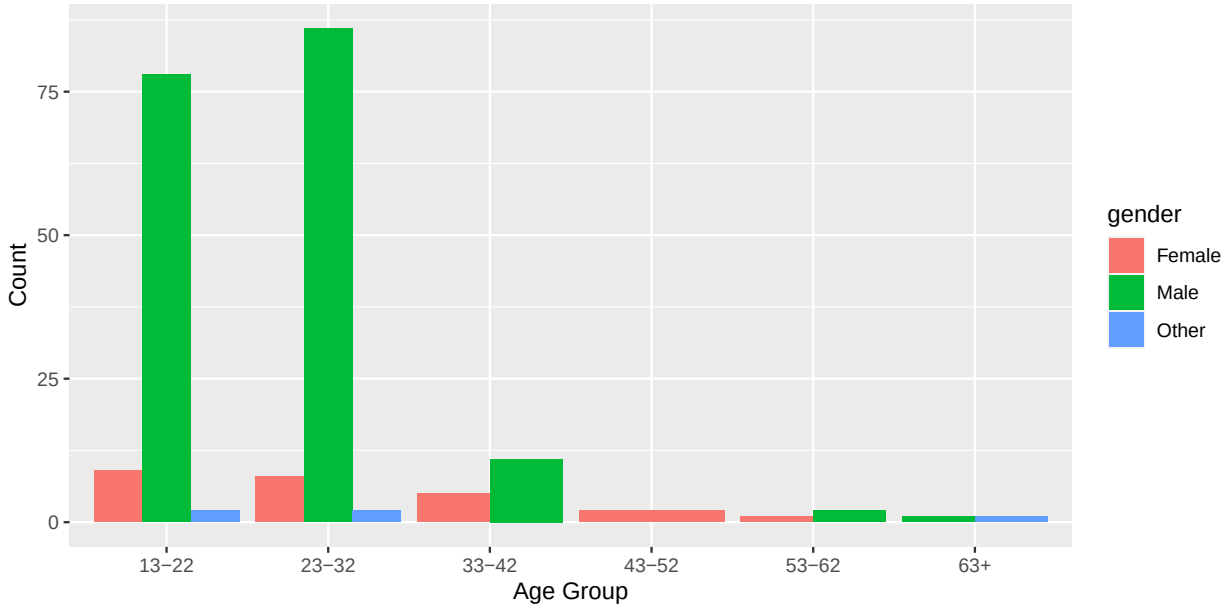

Figure 7: Gender of Respondent Across Age Groups

3 Model

Since the dependent variable is continuous, a linear model is considered instead of a logistic model. We will be considering both the player's rated importance of VP and BKOM, the number of female characters they have, and their total playtime as explanatory variables as part of a multiple linear regression model. For the analysis in the rest of this section, the dataset is split 80%/20% for training/testing purposes. The statistical programming language `R` (R Core Team 2020) is used to run this model and present the findings from it.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \tag{1}$$

The general assumption behind the consideration of this model is that as a player spends more time in BDO, the more likely they are to spend more real world currencies in exchange for in-game items. This could be due to their need to continue using subscription-esque items that grant them boosts, or due to the creation of more content in the game that requires real world currencies to purchase (such as cosmetic items). Given the context of gender-based appeal in Korean MMOs such as BDO, there could also be a connection between the number of female characters and a player's spending in the pearl shop. The model therefore seeks to explore whether there really is a statistically significant connection between the total amount of money spent in-game and the factors mentioned above. Equivalently,

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0 \tag{2}$$

Equation (1) represents the model used in this paper, where Y, the dependent variable, is a player's total spending in pearl shop, β_0 is the intercept, X_1 is the covariate of total playtime, X_2 is the covariate of the number of female characters the player has. X_3 is the covariate of the player's rated importance of VP. X_4 is the covariate of the player's rated importance of BK and OM. $\beta_1\beta_2, \beta_3, \beta_4$ are the coefficients of the variables.

The diagnostic plots of the model are shown in Figure 8. The fitted values vs. residuals graph (top-left) shows a roughly horizontal line without noticeable patterns, which indicated a linear relationship between the dependent and independent variable.

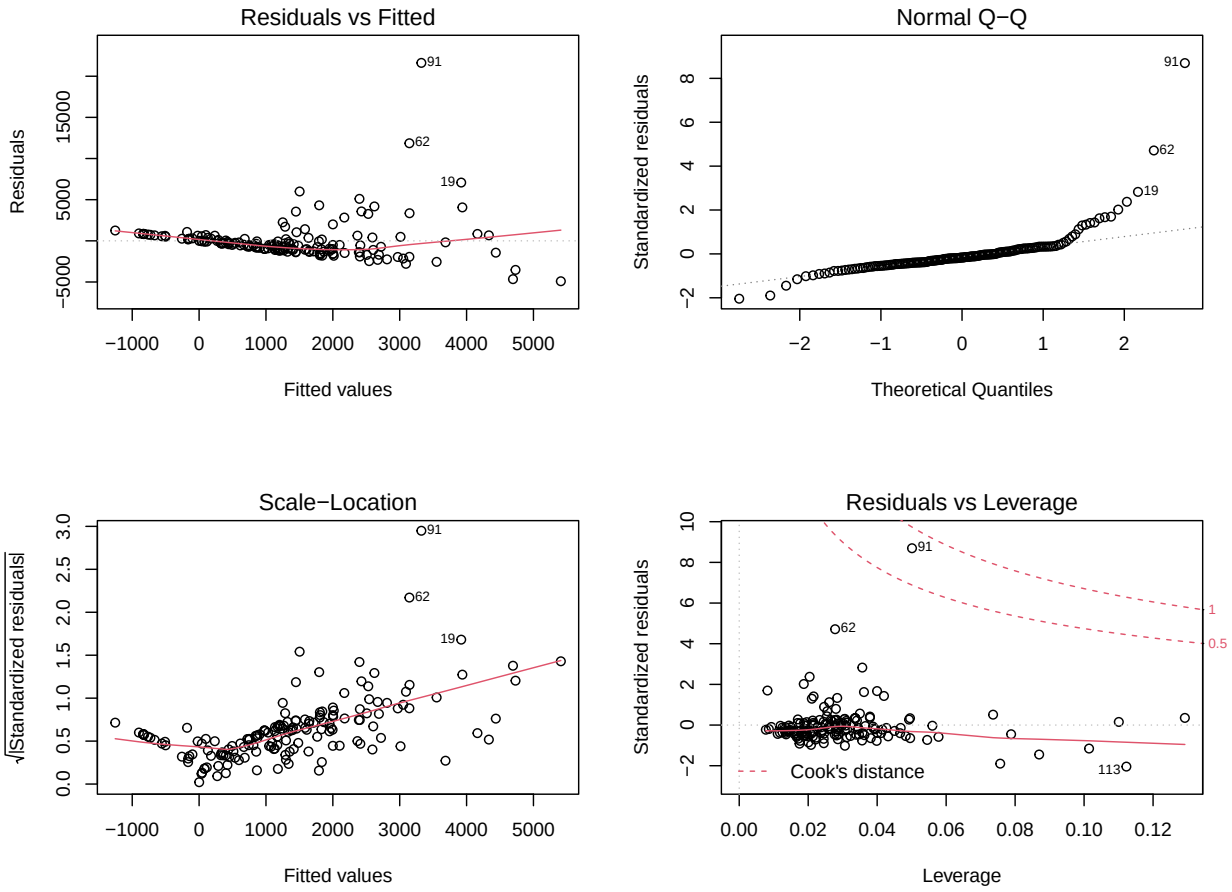

Figure 8: Model Diagnostic Plots

The Normal Q-Q plot (top-right) shows most points falling on the dashed line, though there is also a considerable number of points that fall off the dashed line on the right tail end, suggesting that the residuals may not be entirely normally distributed.

The Scale-Location plot (bottom-left) shows a red line that is not horizontal, meaning that the variance of the residual varies as the fitted value changes This suggests that there may be a violation of the homoscedasticity assumption which assumes that variance of residuals is the same for all X values. An alternative model with square-root transformation of the dependent variable was attempted, but heteroscedasticity remained.

Note that points 91, 62 and 19 are all marked as possible outliers in the three plots mentioned above, suggesting that their inclusion impacts the linear model greatly. As such, they were removed in the final model. The diagnostic conclusions remained similar to the ones above after the removal of the said points.

The residuals vs. leverage plot (bottom-right) identifies two outliers that greatly exceed the Cook's distance–points 91 and 62–and one influence point at point 113 The removal of point 66 did not greatly affect the overall trend shown by the red line in the same plot, thus it was kept in the final model.

The model is used to predict pearl shop spending using the test dataset and arrives at a mean squared error value of `967085.46`. This large MSE may be due to outliers similar to the ones removed during the model diagnostics section.

4 Results

4.1 Model Summary

A summary plot shown in Table 2 is generated using the `modelsummary` package (Arel-Bundock 2022) in `R` (R Core Team 2020). Model 1 is the model including the outliers and Model 2 is the final model.

Table 2: Model Summary of Pearl Shop Spending Based on Rated Importance of VP, BK and OM, Number of Female Characters, and Total Playtime

	Model 1	Model 2
(Intercept)	-1067.40	-785.22
	(538.45)	(307.75)
importance_vp	246.37	91.71
	(79.40)	(46.05)
importance_bkom	-79.55	69.28
	(75.03)	(43.63)
character_female	66.38	58.63
	(64.64)	(36.96)
playtime_total	0.08	0.06
	(0.02)	(0.01)
Num.Obs.	166	163
R2	0.203	0.279
R2 Adj.	0.184	0.260
AIC	3082.3	2843.4
BIC	3101.0	2862.0
Log.Lik.	-1535.146	-1415.714
F	10.271	15.253
RMSE	2550.96	1453.89

The coefficient estimates shown in Table 2 for the final model is equivalent to the following, shown in Equation (3):

$$\hat{Y}_i = -785.22 + 0.06X_{1i} + 58.62X_{2i} + 91.71X_{3i} + 69.28X_{4i} \tag{3}$$

Summary of the model shows that only the rated importance of Value Pack and the total playtime were statistically significant predictors at $\alpha = 0.01$ significance level. The rated importance of Blessing of Kamasylve and Secret Book of Old Moon, and the number of female characters were not statistically significant predictors.

4.2 Character Gender Inequality

As alluded to in Section 2.3.2, gender-based aesthetics are a major selling point of Korean MMOs such as BDO. Figure 9 plots the respondents' total spending in the pearl shop against the proportion of the said spending contributed to cosmetic items. The graph shows that on average, a player with a female main character spends a higher proportion of their pearls on cosmetic items. Specifically, on average, players with a female main character spends roughly an additional 9.23% percent of their total pearl shop spending when compared to players with a male main character. This implies that for games focused on adding more cosmetic content to their in-game cash shop, these cosmetics attract more customers if they were created for female playable characters.

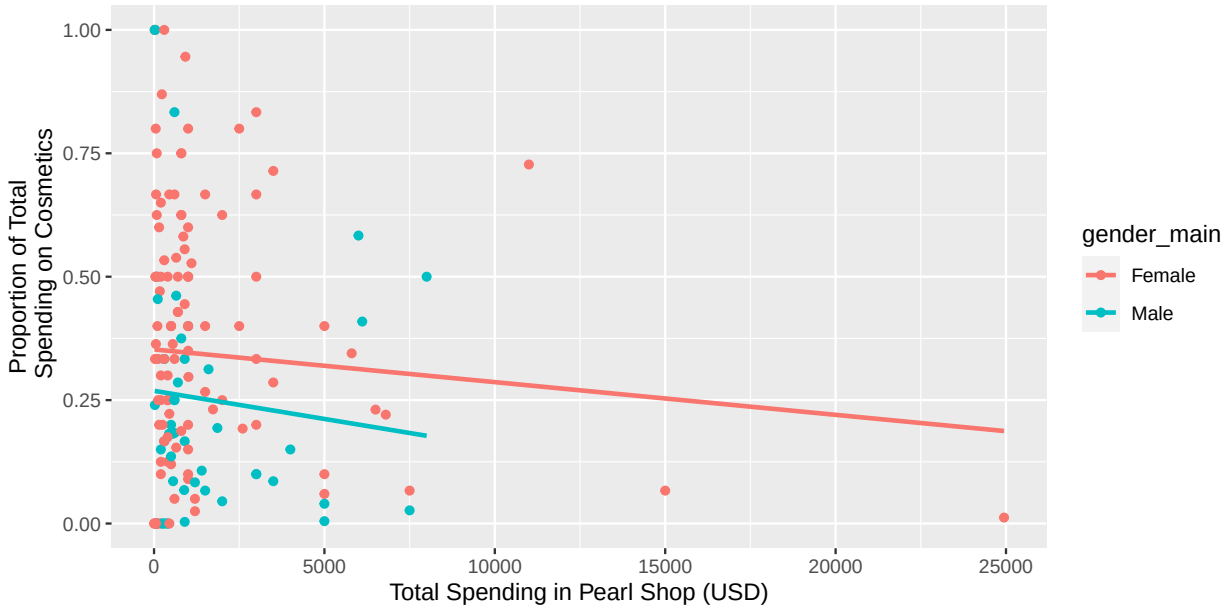

Figure 9: Cosmetic Spending Versus Functional Spending in Pearl Shop

4.3 Free Game Package vs Microtransaction

As shown in Table 1, some of the respondents had obtained the free package for the game, which were frequently given out through various methods by Pearl Abyss, as described in Section 2.3. Out of the 67 respondents who obtained the free package of the game, 23 of them remained completely free-to-play and did not spend any real world currency to purchase in-game items nor other game packages. The other 44 respondents–almost double those who remained completely free-to-play–purchased in-game items on the pearl shop and/or other game packages.

Table 3: Summary of Average Playtime and Spending in Upfront and Cash Shop Categories in Black Desert Online Based On Free Package Acquisition

Spending Status		Average Spending (USD)		Average Proportion of Total Spending		
Completely Free	Partially Free	Pearl Shop	Packages	Packages	VP	BKOM
no	no	1574.75	83.40	0.05	0.11	0.04
no	yes	631.48	39.77	0.06	0.15	0.05
yes	no	0.00	0.00	NaN	NaN	NaN

This is further elaborated upon in Table 3, which summarises the dataset based on whether the respondent's spending status is considered `completely free` if they obtained the free package for the game and did not spent money on other game packages nor in the pearl shop, or `partially free` if they obtained the free package for the game but spent money on other game packges and/or in the pearl shop. We can see from Table 3 that despite having a lower average spending in both the pearl shop and game packages, the "partially free" players still spent a substantial amount of money on further purchases for the game: On average, for "partially free" respondents, the amount of money spent on game package purchases accounts for only 6% of their total spending on BDO. We observe a similar statistic for players who did not obtain a free package for the game, whose spending on game packages accounts for only 5% of their total spending on BDO. In contrast, microtransactions in BDO account for roughly 95% of the total revenue, as shown in table 4, a staggeringly high percentage.

4.4 Subscription Items

As discussed in Section 2.3, Figure 2 suggests a difference in importance as regarded by the respondents. This is supported by the statistics shown in Figure 10, which show a stronger positive linear correlation between rated importance of VP and spending on VP when compared to the correlation between rated importance of BK and OM and spending on BK and OM. On average, respondents spend $148.0288 USD on purchasing VPs from the pearl shop, and spend $49.1202 USD on purchasing BKs and OMs from the pearl shop.

Table 3 shows that on average, VP accounts for 11% of a respondent's total spending on BDO (15% for partially free respondents). BK and OM account for 4% of a respondent's total spending on BDO (5% for partially free respondents).

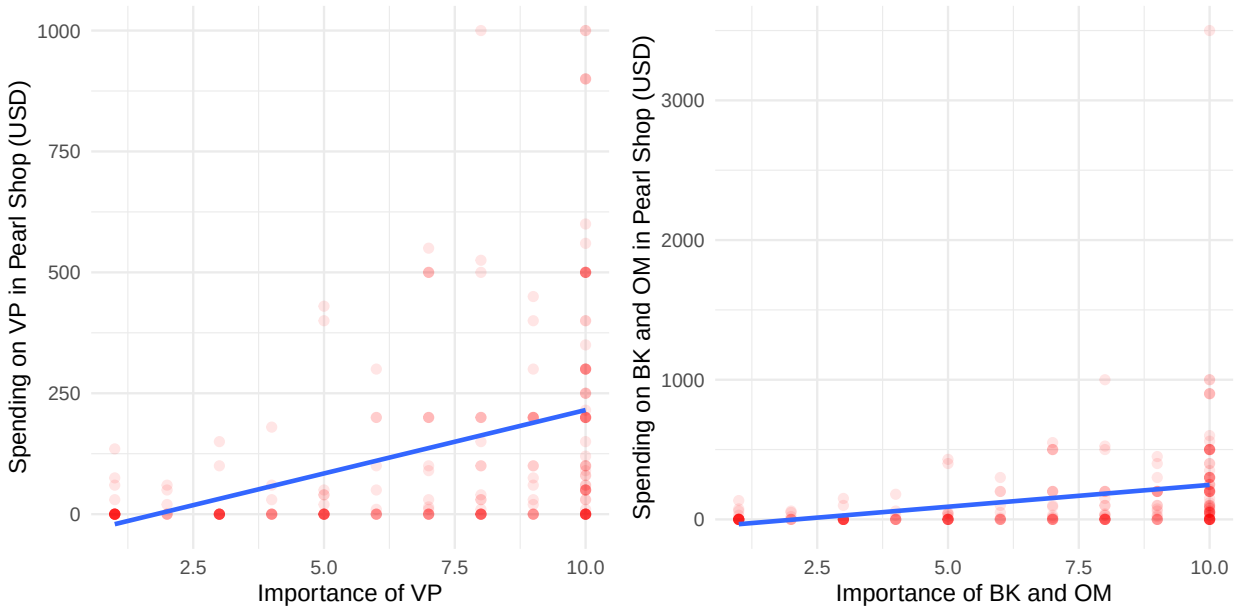

Figure 10: Rated Importance of Subscription Items Versus Spending on the Said Items

5 Discussion

5.1 Free Game Versus Upfront Cost

The statistics presented in Table 3 suggests that the amount of revenue obtained from game package purchases is relatively minuscule compared to other sources of income. This raises the question of whether publishers of MMORPGs should forgo charging an upfront fee for their games in order to draw in more players so that they can browse the much more profitable in-game cash shop instead.

From a pure-revenue standpoint, there is a debate to be made about whether or not a video game publisher should charge an upfront cost on consumers for access to their game. On one hand, giving concrete price tags to a game before its launch would grant the publisher the ability to let consumers pre-order the game, injecting a massive amount of funding into the company that could be used for future developments and investments on other projects well before they even need to finish and release the game. On the other hand, price tags too high may discourage consumer interest in the game, preventing them from ever engaging with the in-game cash shop which is much better suited for long-term revenue generation.

MMORPGs have a unique way to circumvent this dilemma by releasing the game with a price tag and removing the said price tag later down the line. Due to their progression-based nature and continuous development, consumer interest in MMORPGs is usually expected to remain consistently high past its initial launch period. As such, there is less worry about upfront price tags dissuading large amounts of potential players during the launch; High interest in the game post-launch should ensure continued influx of players who will have a chance to interact with the game's cash shop once the upfront cost is dropped.

As discussed in Section 4.3, roughly 66% of players who obtained the free package for the game ended up spending real world currencies in the pearl shop. This shows that the majority of players will spend money on a game even when it is free-to-play, providing further incentives for MMORPG publishers to abandon charging upfront fees.

Table 4: Summary of Average Percentage of Spending on Pearl Shop and Packages

Spending Percentage	
Pearl Shop	Game Packages
0.9487	0.0513

5.2 Number of Subscription Items

The three types of subscription items in BDO–Value Pack, Blessing of Kamasylve, and Secret Book of Old Moon–each offer a different set of bonuses relatively independent from each other. This means that some players may prefer certain bonuses over others and may not purchase all three subscription items at once. As discussed in Section 4.4, BK and OM are generally rated as less important than VP; BK and OM also account for less than half of the revenue generated by VP on average, as demonstrated in Table 3.

The lack of interest in BK and OM could be due to their status as subscription items introduced later than the VP. This means that the set of bonuses they grant must be separate from those granted by the VP, which was initially the only subscription item in the game. This need to avoid overlaps in bonuses may have caused BK and OM to be to designed to grant less generally useful bonuses as the vital ones are already covered by the VP.

Another potential reason to this disinterest is the difference in price between BK and OM, and VP: Without any discounts, 30 days of VP costs 1500 pearls (roughly $15.00 USD) while 30 days of BK/OM costs 2500 pearls (roughly $25.00 USD), an increase in price by roughly 66%. There is no plausible explanation as to why BK/OM are priced higher than VP except publisher bias; While most players regard VP as the more important subscription item (as shown in Figure 2), publishers of the game may simply think otherwise. If possible, further extensions on this study should inquire the players about their reasons for rating certain subscription items less important, be it the price or its bonuses.

While the statistics seem to suggest that the introduction of multiple subscription items for a single game is not an optimal strategy, this type of item does not require a lot of development resources from the developers to create and is a relatively effortless way to generate revenue. Therefore, the addition of new subscription items beyond the first one into a game is not unadvised but rather simply less profitable than the first one is.

5.3 Whales

According to Connick (Connick 2020), `whale` is originally a term in sales used to refer to clients that bring in substantially more business than an average client does. The video game industry has borrowed this termto refer to players who spend an exorbitant amount of real world currency in a game.

As alluded to in the beginning paragraphs of Section 2.3, we observer several outliers in the dataset with drastically more in-game spending than the rest of the respondents. These outliers are most likely the aforementioned whales.

Whales are often make up a major portion of a video game's revenue: For example, according to Activision Blizzard's (Activision Bllizard, Inc. 2022) 2021 fourth quarter revenue release, the company made a net revenue of $8.086 billion USD, $5.736 billion USD of which is from in-game transactions and subscriptions. A surprisingly high 70.94% of total revenue for the company is from sources other than upfront costs.

As Table 4 shows, if we do not exclude the whales as outliers from the dataset, then the average revenue generated from the pearl shop accounts for roughly 95% This clearly demonstrates the importance of whales to a video game company's profit, and why catering to these whales could often result in more profit than focusing on satisfying the average consumer.

Whales are rare but their impact have not gone unnoticed. In fact, catering to whales is not a new strategy for publishers by any means, as demonstrated by Jernström (Jernström 2016) , who is well versed in this type of monetisation and has created a thorough guide for other publishers to follow suit. It is apparent that publishers are already catching on to these whales and the best strategies to appeal to them. This should state the significance of whales as the highest potential source of income for a video game.

Unfortunately, due to the limited number of responses generated by the survey, an analysis on whales in particular was not able to be conducted in this paper. Possible extensions of this paper could focus on gathering a large enough dataset in the hopes that it contains a sufficient sample of whale respondents to fit a model to.

5.4 Limitations

As discussed in Section 5.3, the rarity of whales and the limited sample size of the dataset means that the analysis in this paper could not incorporate whales effectively into the model. Should a large enough sample size for whales be collected, a model specifically designed to analyze whale spending is a great extension for this paper.

As mentioned in Section 2.1, Pearl Abyss does not have any public APIs available for tracking user's spendings, playtime, and character counts, etc. This is understandable as it could be considered sensitive information. However, Pearl Abyss also does not include detailed and/or easily accessible information trackers for the players themselves. Players cannot access a detailed breakdown of the composition of their spending, nor can they calculate exactly how long they have played the game for save for a few milestone achievements at large intervals of playtime (500 days, 1000 days, etc). Respondents who played the game on the game distribution platform Steam has less issues with this as Steam offers them playtime trackers as well as access to their total spending on the game, though the composition of the spending is still not recorded. The problem caused by this is that respondents have to estimate their playtime and the composition of their spending to the best of their ability. As such, most monetary statistics in the dataset have some level of estimation involved in them except total spending in pearl shop and total spending on game packages. Should Pearl Abyss publish an API in the future for the game, future analyses will no longer need to rely on convenience sampling and could instead employ better sampling strategies that more accurately represent the true population.

Appendix

A Datasheet for the Dataset

Extract of the questions from Gebru et al. (2021)

Motivation

1. *For what purpose was the dataset created? Was there a specific task in mind? Was there a specific gap that needed to be filled? Please provide a description.*
    - This dataset was created to be used to analyze the various ways of monetisations that the video game Black Desert Online employs. Due to the lack of APIs and useful tracking tools provided by Pearl Abyss–the developers behind Black Desert Online–A survey was used instead to obtain this dataset through convenience sampling.
2. *Who created the dataset (for example, which team, research group) and on behalf of which entity (for example, company, institution, organization)?*
    - The survey used to collect the dataset was created by Tian Yi Zhang, a undergraduate student at the University of Toronto and the author of this paper.
3. *Who funded the creation of the dataset? If there is an associated grant, please provide the name of the grantor and the grant name and number.*
    - No organization funded the creation of the dataset.
4. *Any other comments?*
    - None.

Composition

1. *What do the instances that comprise the dataset represent (for example, documents, photos, people, countries)? Are there multiple types of instances (for example, movies, users, and ratings; people and interactions between them; nodes and edges)? Please provide a description.*
    - Each instance in the dataset represent a player of the video game Black Desert Online.
2. *How many instances are there in total (of each type, if appropriate)?*
    - There are 178 male respondents, 25 female respondents, and 5 respondents of other genders.
3. *Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set? If the dataset is a sample, then what is the larger set? Is the sample representative of the larger set (for example, geographic coverage)? If so, please describe how this representativeness was validated/verified. If it is not representative of the larger set, please describe why not (for example, to cover a more diverse range of instances, because instances were withheld or unavailable).*
    - The dataset is a sample of instances. It is obtained through convenience sampling from the larger set of Black Desert Online players on the North American and European servers. Since the dataset is obtained through convenience sampling, it is not representative of the total population in terms fof geographical coverage or any stratification. The choice of convenience sampling is due to the lack of APIs and useful tracking tools provided by Pearl Abyss, the developers behind Black Desert Online.
4. *What data does each instance consist of? "Raw" data (for example, unprocessed text or images) or features? In either case, please provide a description.*

- Each instance of data consists of a player's age, gender, total playtime in the game and on their main character, their rated importance of various subscription items in the game, and their total spending on various types of items in the game.

5. *Is there a label or target associated with each instance? If so, please provide a description.*
   - the gender is either male, female, or other.
6. *Is any information missing from individual instances? If so, please provide a description, explaining why this information is missing (for example, because it was unavailable). This does not include intentionally removed information, but might include, for example, redacted text.*
   - There is lack of female and other genders in certain age groups in the dataset due to insufficient sample size.
7. *Are relationships between individual instances made explicit (for example, users' movie ratings, social network links)? If so, please describe how these relationships are made explicit.*
   - There are no relationships between individual instances.
8. *Are there recommended data splits (for example, training, development/validation, testing)? If so, please provide a description of these splits, explaining the rationale behind them.*
   - None.
9. *Are there any errors, sources of noise, or redundancies in the dataset? If so, please provide a description.*
   - The sample size is unequal for female and male respondents. This unequal sample size does negatively affect the assumption of equal variances. Furthermore, lack of detailed statistic tracking tools and APIs provided by Pearl Abyss–the developer of Black Desert Online–means that playtime and most monetary questions in the survey rely on the respondents' best estimates.
10. *Is the dataset self-contained, or does it link to or otherwise rely on external resources (for example, websites, tweets, other datasets)? If it links to or relies on external resources, a) are there guarantees that they will exist, and remain constant, over time; b) are there official archival versions of the complete dataset (that is, including the external resources as they existed at the time the dataset was created); c) are there any restrictions (for example, licenses, fees) associated with any of the external resources that might apply to a dataset consumer? Please provide descriptions of all external resources and any restrictions associated with them, as well as links or other access points, as appropriate.*
    - The dataset is self-contained.
11. *Does the dataset contain data that might be considered confidential (for example, data that is protected by legal privilege or by doctor-patient confidentiality, data that includes the content of individuals' non-public communications)? If so, please provide a description.*
    - Breakdown of the monetary transcations of each player might be considered a confidential data by Pearl Abyss. However, the survey was designed to be as transparent with the respondents as possible, and no data that could be linked back to the respondents were included in the survey nor in the final paper.
12. *Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety? If so, please describe why.*
    - The dataset does not contain any offensive or insulting data.
13. *Does the dataset identify any sub-populations (for example, by age, gender)? If so, please describe how these subpopulations are identified and provide a description of their respective distributions within the dataset.*
    - The datasets are split by age groups and gender as self-identified by the respondents.

14. *Is it possible to identify individuals (that is, one or more natural persons), either directly or indirectly (that is, in combination with other data) from the dataset? If so, please describe how.*
    - It is entirely impossible to identify individuals from the dataset.
15. *Does the dataset contain data that might be considered sensitive in any way (for example, data that reveals race or ethnic origins, sexual orientations, religious beliefs, political opinions or union memberships, or locations; financial or health data; biometric or genetic data; forms of government identification, such as social security numbers; criminal history)? If so, please provide a description.*
    - The datasets do not contain data that might be considered sensitive in any way.
16. *Any other comments?*
    - None.

Collection process

1. *How was the data associated with each instance acquired? Was the data directly observable (for example, raw text, movie ratings), reported by subjects (for example, survey responses), or indirectly inferred/derived from other data (for example, part-of-speech tags, model-based guesses for age or language)? If the data was reported by subjects or indirectly inferred/derived from other data, was the data validated/verified? If so, please describe how.*
    - The data was collected through a survey distributed to the subreddit page of Black Desert Online, the official website's forum of Black Desert Online, and the steam discussion page of Black Desert Online. The data was reported by the respondents based on their best estimates.
2. *What mechanisms or procedures were used to collect the data (for example, hardware apparatuses or sensors, manual human curation, software programs, software APIs)? How were these mechanisms or procedures validated?*
    - The data was collected through a survey distributed to the subreddit page of Black Desert Online, the official website's forum of Black Desert Online, and the steam discussion page of Black Desert Online. The data was reported by the respondents based on their best estimates. Google Forms is used to produce the survey.
3. *If the dataset is a sample from a larger set, what was the sampling strategy (for example, deterministic, probabilistic with specific sampling probabilities)?*
    - Convenience sampling was used due to lack of detailed statistic tracking tools and APIs provided by Pearl Abyss.
4. *Who was involved in the data collection process (for example, students, crowdworkers, contractors) and how were they compensated (for example, how much were crowdworkers paid)?*
    - No one but the author is involved in the data collection process.
5. *Over what timeframe was the data collected? Does this timeframe match the creation timeframe of the data associated with the instances (for example, recent crawl of old news articles)? If not, please describe the timeframe in which the data associated with the instances was created.*
    - The data was collected over the span of four days from April 18, 2022 to April 21, 2022.
6. *Were any ethical review processes conducted (for example, by an institutional review board)? If so, please provide a description of these review processes, including the outcomes, as well as a link or other access point to any supporting documentation.*
    - No ethical review processes were conducted. Subreddit moderators were consulted for permission of posting the survey.
7. *Did you collect the data from the individuals in question directly, or obtain it via third parties or other sources (for example, websites)?*

- The data was collected through a survey distributed to the subreddit page of Black Desert Online, the official website's forum of Black Desert Online, and the steam discussion page of Black Desert Online.

8. *Were the individuals in question notified about the data collection? If so, please describe (or show with screenshots or other information) how notice was provided, and provide a link or other access point to, or otherwise reproduce, the exact language of the notification itself.*

    - Respondents were informed of the purpose of the data collection: "I wish to study the different monetization methods that BDO employs. Due to the lack of proper APIs provided by Pearl Abyss, I have resorted to surveys instead."
    - "To be clear, there will not be any information present in the final paper that can be used to link the data back to the original respondents."

9. *Did the individuals in question consent to the collection and use of their data? If so, please describe (or show with screenshots or other information) how consent was requested and provided, and provide a link or other access point to, or otherwise reproduce, the exact language to which the individuals consented.*

    - Since data collection is through a survey, consent was explicitly provided when respondents clicked the survey link.

10. *If consent was obtained, were the consenting individuals provided with a mechanism to revoke their consent in the future or for certain uses? If so, please provide a description, as well as a link or other access point to the mechanism (if appropriate).*

    - Respondents can revoke consent by simply exiting the survey.

11. *Has an analysis of the potential impact of the dataset and its use on data subjects (for example, a data protection impact analysis) been conducted? If so, please provide a description of this analysis, including the outcomes, as well as a link or other access point to any supporting documentation.*

    - N/A.

12. *Any other comments?*

    - None.

Preprocessing/cleaning/labeling

1. *Was any preprocessing/cleaning/labeling of the data done (for example, discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)? If so, please provide a description. If not, you may skip the remaining questions in this section.*

    - Certain variables names were shortened in Google Spreadsheet before the dataset is used for analysis in this paper.

2. *Was the "raw" data saved in addition to the preprocessed/cleaned/labeled data (for example, to support unanticipated future uses)? If so, please provide a link or other access point to the "raw" data.*

    - The raw data is accessable in the github link provided in this paper.

3. *Is the software that was used to preprocess/clean/label the data available? If so, please provide a link or other access point.*

    - statistical programming language R (R Core Team 2020) was used to preprocess the datasets.
    - Google Spreadsheet is publicly available for use by registering a Google account.

4. *Any other comments?*

    - None.

Uses

1. *Has the dataset been used for any tasks already? If so, please provide a description.*
    - At the time of writing this, only this paper has utilized the dataset.
2. *Is there a repository that links to any or all papers or systems that use the dataset? If so, please provide a link or other access point.*
    - The repository for this paper, which uses the datasets, are available at https://github.com/zhan7818/bdo_monetization.
3. *What (other) tasks could the dataset be used for?*
    - The datasets could be used to modeling or finding relationship between total in-game spending and age groups.
4. *Is there anything about the composition of the dataset or the way it was collected and preprocessed/cleaned/labeled that might impact future uses? For example, is there anything that a dataset consumer might need to know to avoid uses that could result in unfair treatment of individuals or groups (for example, stereotyping, quality of service issues) or other risks or harms (for example, legal risks, financial harms)? If so, please provide a description. Is there anything a dataset consumer could do to mitigate these risks or harms?*
    - None.
5. *Are there tasks for which the dataset should not be used? If so, please provide a description.*
    - The datasets should not be used in an effort to harass or discriminate against the general Black Desert Online community nor should it be used to misrepresent information about the said community.
6. *Any other comments?*
    - None.

Distribution

1. *Will the dataset be distributed to third parties outside of the entity (for example, company, institution, organization) on behalf of which the dataset was created? If so, please provide a description.*
    - Since there dataset is publicly available through the github link, yes.
2. *How will the dataset be distributed (for example, tarball on website, API, GitHub)? Does the dataset have a digital object identifier (DOI)?*
    - The dataset is distributed on Tian Yi Zhang's github page, along with this paper: https://github.com/zhan7818/bdo_monetization.
3. *When will the dataset be distributed?*

- The datasets used in this paper is distributed in 2022.

4. *Will the dataset be distributed under a copyright or other intellectual property (IP) license, and/or under applicable terms of use (ToU)? If so, please describe this license and/ or ToU, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms or ToU, as well as any fees associated with these restrictions.*
    - The dataset is distributed under a MIT license.

5. *Have any third parties imposed IP-based or other restrictions on the data associated with the instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms, as well as any fees associated with these restrictions.*
    - No restrictions are placed on the dataset.
6. *Do any export controls or other regulatory restrictions apply to the dataset or to individual instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any supporting documentation.*
    - None.
7. *Any other comments?*
    - None.

Maintenance

1. *Who will be supporting/hosting/maintaining the dataset?*
    - Tian Yi Zhang is hosting the dataset on the github link at https://github.com/zhan7818/bdo_monetization.
2. *How can the owner/curator/manager of the dataset be contacted (for example, email address)?*
    - The creator of the datasets, Tian Yi Zhang, can be contacted at tianyi.zhang6546@gmail.com.
3. *Is there an erratum? If so, please provide a link or other access point.*
    - As of writing this datasheet, no explicit erratum have been spotted.
4. *Will the dataset be updated (for example, to correct labeling errors, add new instances, delete instances)? If so, please describe how often, by whom, and how updates will be communicated to dataset consumers (for example, mailing list, GitHub)?*
    - Any updates will be posted on the github page used to host this paper which include the datasets.
5. *If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (for example, were the individuals in question told that their data would be retained for a fixed period of time and then deleted)? If so, please describe these limits and explain how they will be enforced.*
    - N/A.
6. *Will older versions of the dataset continue to be supported/hosted/maintained? If so, please describe how. If not, please describe how its obsolescence will be communicated to dataset consumers.*
    - Older versions of the datasets (if there will be any) will be kept as a record on github.
7. *If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so? If so, please provide a description. Will these contributions be validated/verified? If so, please describe how. If not, why not? Is there a process for communicating/distributing these contributions to dataset consumers? If so, please provide a description.*
    - One should contact Tian Yi Zhang, the author of the dataset.
8. *Any other comments?*
    - None.

B Additional Details

Code and data available at: https://github.com/zhan7818/bdo_monetization

- PDF copy of the survey used in this paper is available in inputs/survey/bdo_survey.pdf
- A simulation of the dataset was conducted in scripts/00-simulation.R which helped formulate plans for analysis involving the actual dataset.
- Dataset was cleaned in scripts/01-data_cleaning.R

References

Activision Bllizard, Inc. 2022. *Activision Blizzard Annouces Fourth-Quarter and 2021 Financial Results*. Activision Bllizard, Inc. https://investor.activision.com/static-files/92ffd159-f3ab-4f26-8dca-bab73cb457f6.

Arel-Bundock, Vincent. 2022. *Modelsummary: Summary Tables and Plots for Statistical Models and Data: Beautiful, Customizable, and Publication-Ready*. https://CRAN.R-project.org/package=modelsummary.

Chikhani, Riad. 2015. *The History of Gaming: An Evolving Community*. Tech Crunch. https://techcrunch.com/2015/10/31/the-history-of-gaming-an-evolving-community/.

Connick, Wendy. 2020. *What Is a Whale?* The Balance Careers. https://www.thebalancecareers.com/what-is-a-whale-2917364.

Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. "Datasheets for Datasets." *Communications of the ACM* 64 (12): 86–92.

Iannone, Richard, and Mauricio Vargas. 2022. *Pointblank: Data Validation and Organization of Metadata for Local and Remote Tables*. https://CRAN.R-project.org/package=pointblank.

Jernström, Torulf. 2016. *Let's Go Whaling: A Guide to Monetisation Through in-App Purchases*. Pocket Gamer. https://www.pocketgamer.biz/comment-and-opinion/63871/monetisation-lets-go-whaling/.

Microsoft Corporation. 2022-03-30. *Microsoft Excel*. https://office.microsoft.com/excel.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2022. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Wilke, Claus O. 2020. *Cowplot: Streamlined Plot Theme and Plot Annotations for 'Ggplot2'*. https://CRAN.R-project.org/package=cowplot.

Xie, Yihui. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.