

# Heroes of Pymoli

## Introduction

As a Lead Analyst for an independent gaming company, I have been assigned the task of analyzing the data for their most recent fantasy game *Heroes of Pymoli*. Like many others in its genre, the game is free-to-play, but players are encouraged to purchase optional items that enhance their playing experience. This report breaks down the game's purchasing data into meaningful insights.



Illustration: The fantasy artwork by Ian Miller

## Business Problem:

The project is based on the following business problem:

*What are the meaningful insights regarding - gender, age demographic, purchasing power and so on - hidden inside the purchase data of this fantasy game?*

## Target Audience:

This report and its analysis will be valuable to (a) business stakeholders of the company (b) sales team of the company and (c) virtual game designers planning their next product.

## Data source & cleaning:

1. The data for the analysis (`purchase_data.csv`) is available as a csv (comma separated value) file in the Resources subdirectory.
2. No discrepancy was found in the data types of each column. The columns "*Purchase ID*", "*Age*", "*Item ID*" were integers, "*Price*" was of data type float and remaining columns were of strings. Also, the dataset didn't have any missing data.

## Data Analysis:

The first 5 rows of the dataframe (*purchase\_data*) read from the csv file is shown below:

|   | Purchase ID | SN            | Age | Gender | Item ID | Item Name                                 | Price |
|---|-------------|---------------|-----|--------|---------|---|-------|
| 0 | 0           | Lisim78       | 20  | Male   | 108     | Extraction, Quickblade Of Trembling Hands | 3.53  |
| 1 | 1           | Lisovynya38   | 40  | Male   | 143     | Frenzied Scimitar                         | 1.56  |
| 2 | 2           | lthergue48    | 24  | Male   | 92      | Final Critic                              | 4.88  |
| 3 | 3           | Chamassasya86 | 24  | Male   | 100     | Blindscythe                               | 3.27  |
| 4 | 4           | Iskosia90     | 23  | Male   | 131     | Fury                                      | 1.44  |

The dataframe tabulates the *Purchase ID*, *Customer Name*, *Age*, *Gender*, *Item ID*, *Item Name* and *Price* of each item sold. It comprises of 780 rows and 7 columns.

| Total number of Players |     |
|-------------------------|-----|
| 0                       | 576 |

By counting the unique values in column “SN” we get the total number of players. There are **576 players** in the purchase data set.

Purchasing Analysis (Total):

As shown in the above table, 576 players made 780 purchases of 183 unique items bringing in a total revenue close to \$2379.77

|   | Number of Unique Items | Average Price | Number of Purchases | Total Revenue |
|---|------------------------|---------------|---------------------|---------------|
| 0 | 183                    | \$3.05        | 780                 | \$2379.77     |

Gender Demographics:

There are three gender categories in this dataframe, namely *Male*, *Female* and *Other/Non-Disclosed*.

Sales are strongly biased by gender. 84% of total players are Male as compared to only 14% being Female.

|                       | Total Count | Percentage of Players |
|-----------------------|-------------|-----------------------|
| Male                  | 484         | 84.03%                |
| Female                | 81          | 14.06%                |
| Other / Non-Disclosed | 11          | 1.91%                 |

Purchasing Analysis (Gender):

|                       | Purchase Count | Average Purchase Price | Total Purchase Value | Avg. Total Purchase per Person |
|-----------------------|----------------|------------------------|----------------------|--------------------------------|
| Gender                |                |                        |                      |                                |
| Female                | 113            | \$3.20                 | \$361.94             | \$4.47                         |
| Male                  | 652            | \$3.02                 | \$1,967.64           | \$4.07                         |
| Other / Non-Disclosed | 15             | \$3.35                 | \$50.19              | \$4.56                         |

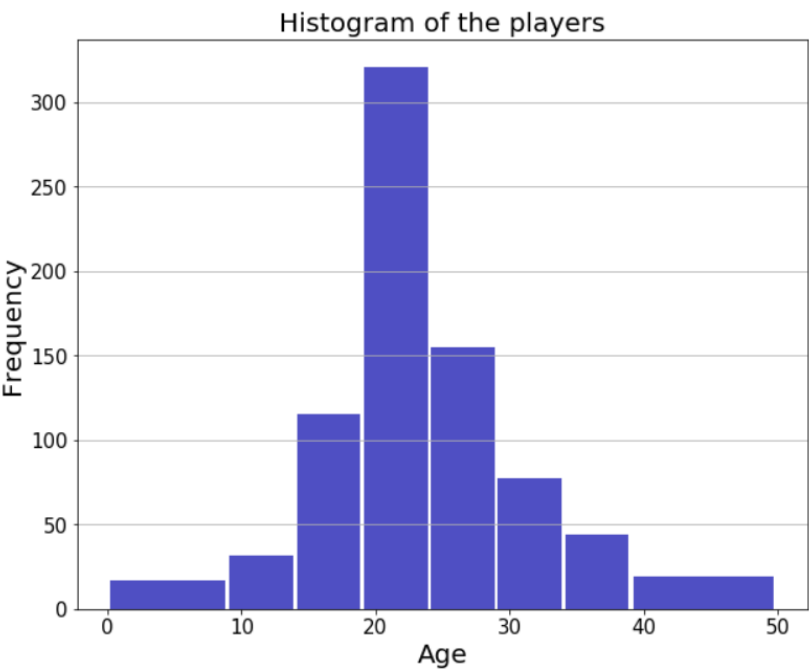
As we can see in the table above, though Male players are 6 times more in number than Female players, they spend only 5 times more than their female counterparts. Which implies that Female player paid more per person. Comparing average total purchase per person, Female players spend 10% more than their male counterparts.

Age Demographics:

Maximum and minimum ages of the player dataset are 45 and 7 respectively. The players are binned in 10 categories based on their ages. We can visualize the counts in each bin in the histogram shown here.

The *pd.cut* function was used to segregate the “Age” column into separate bins.

The new dataframe *df\_binned* is showed below:



|   | Purchase ID | SN            | Age | Gender | Item ID | Item Name                                 | Price | Age Ranges |
|---|-------------|---------------|-----|--------|---------|---|-------|------------|
| 0 | 0           | Lisim78       | 20  | Male   | 108     | Extraction, Quickblade Of Trembling Hands | 3.53  | 20-24      |
| 1 | 1           | Lisovynya38   | 40  | Male   | 143     | Frenzied Scimitar                         | 1.56  | 40+        |
| 2 | 2           | Ithergue48    | 24  | Male   | 92      | Final Critic                              | 4.88  | 20-24      |
| 3 | 3           | Chamassasya86 | 24  | Male   | 100     | Blindscythe                               | 3.27  | 20-24      |
| 4 | 4           | Iskosia90     | 23  | Male   | 131     | Fury                                      | 1.44  | 20-24      |

|            | Total Count | Percentage of Players |
|------------|-------------|-----------------------|
| Age Ranges |             |                       |
| <10        | 17          | 2.95%                 |
| 10-14      | 22          | 3.82%                 |
| 15-19      | 107         | 18.58%                |
| 20-24      | 258         | 44.79%                |
| 24-29      | 77          | 13.37%                |
| 30-34      | 52          | 9.03%                 |
| 35-39      | 31          | 5.38%                 |
| 40+        | 12          | 2.08%                 |

As we can see in the adjacent table, the age group (20-24) purchased the highest number (45%) of items.

In statistics, a normal distribution (often called the bell curve) is a probability distribution that occurs naturally in many situations. As we inspect the Histogram - Age demographics seem to fit a normal distribution which is symmetric around a central peak, namely, frequency of the age group (20-24).

## Purchasing Analysis (Age):

With the players grouped in bins based on their ages, we can now analyze their spending behavior:

| Age Ranges | Purchase Count | Average Purchase Price | Total Purchase Value | Avg. Total Purchase per Person |
|------------|----------------|------------------------|----------------------|--------------------------------|
| <10        | 23             | \$3.35                 | \$77.13              | \$4.54                         |
| 10-14      | 28             | \$2.96                 | \$82.78              | \$3.76                         |
| 15-19      | 136            | \$3.04                 | \$412.89             | \$3.86                         |
| 20-24      | 365            | \$3.05                 | \$1,114.06           | \$4.32                         |
| 24-29      | 101            | \$2.90                 | \$293.00             | \$3.81                         |
| 30-34      | 73             | \$2.93                 | \$214.00             | \$4.12                         |
| 35-39      | 41             | \$3.60                 | \$147.67             | \$4.76                         |
| 40+        | 13             | \$2.94                 | \$38.24              | \$3.19                         |

The **age group (20-24)** purchased the largest number of items and contributed a large fraction (~ 47%) of the total revenue. The **age group (35-39)** spend \$4.76 per person, which is more than any other age groups. This could be because more players in that group are employed and financially secure. Players less than 10 years old are impulsive buyers, securing the second highest “Avg. total purchase per person” position.

## Top Spenders:

Let's check the top spenders in the players dataset.

| SN          | Purchase Count | Average Purchase Price | Total Purchase Value |
|-------------|----------------|------------------------|----------------------|
| Lisosa93    | 5              | \$3.79                 | \$18.96              |
| Idastidru52 | 4              | \$3.86                 | \$15.45              |
| Chamjask73  | 3              | \$4.61                 | \$13.83              |
| Iral74      | 4              | \$3.40                 | \$13.62              |
| Iskadarya95 | 3              | \$4.37                 | \$13.10              |

By using the (.loc) function of Pandas, we can gain more information about the top three spenders.

Top-most spender:

|     | Purchase ID | SN        | Age | Gender | Item ID | Item Name                         | Price | Age Ranges |
|-----|-------------|-----------|-----|--------|---------|-----------------------------------|-------|------------|
| 74  | 74          | Lisosia93 | 25  | Male   | 89      | Blazefury, Protector of Delusions | 4.64  | 24-29      |
| 120 | 120         | Lisosia93 | 25  | Male   | 24      | Warped Fetish                     | 3.81  | 24-29      |
| 224 | 224         | Lisosia93 | 25  | Male   | 157     | Spada, Etcher of Hatred           | 4.80  | 24-29      |
| 603 | 603         | Lisosia93 | 25  | Male   | 141     | Persuasion                        | 3.19  | 24-29      |
| 609 | 609         | Lisosia93 | 25  | Male   | 40      | Second Chance                     | 2.52  | 24-29      |

Second-highest spender:

|     | Purchase ID | SN          | Age | Gender | Item ID | Item Name                          | Price | Age Ranges |
|-----|-------------|-------------|-----|--------|---------|------------------------------------|-------|------------|
| 290 | 290         | Idastidru52 | 24  | Male   | 147     | Hellreaver, Heirloom of Inception  | 4.93  | 20-24      |
| 490 | 490         | Idastidru52 | 24  | Male   | 148     | Warmonger, Gift of Suffering's End | 4.03  | 20-24      |
| 543 | 543         | Idastidru52 | 24  | Male   | 121     | Massacre                           | 1.60  | 20-24      |
| 676 | 676         | Idastidru52 | 24  | Male   | 111     | Misery's End                       | 4.89  | 20-24      |

Third-highest spender:

| Purchase ID |     | SN         | Age | Gender | Item ID |  | Item Name | Price | Age Ranges |
|-------------|-----|------------|-----|--------|---------|--|-----------|-------|------------|
| 222         | 222 | Chamjask73 | 22  | Female | 178     | Oathbreaker, Last Hope of the Breaking Storm |           | 4.23  | 20-24      |
| 520         | 520 | Chamjask73 | 22  | Female | 109     | Downfall, Scalpel Of The Emperor             |           | 4.76  | 20-24      |
| 564         | 564 | Chamjask73 | 22  | Female | 52      | Hatred                                       |           | 4.84  | 20-24      |

Top-most and second-highest spenders are Males of ages 25 and 24 years respectively.

Most Popular Items:

Here are the most popular items in the dataset.

| Item ID | Item Name                                    | Purchase Count | Item Price | Total Purchase Value |
|---------|--|----------------|------------|----------------------|
| 178     | Oathbreaker, Last Hope of the Breaking Storm | 12             | \$4.23     | \$50.76              |
| 145     | Fiery Glass Crusader                         | 9              | \$4.58     | \$41.22              |
| 108     | Extraction, Quickblade Of Trembling Hands    | 9              | \$3.53     | \$31.77              |
| 82      | Nirvana                                      | 9              | \$4.90     | \$44.10              |
| 19      | Pursuit, Cudgel of Necromancy                | 8              | \$1.02     | \$8.16               |



As seen in the table above, ***“Oathbreaker, Last Hope of the Breaking Storm”*** (Item ID# 178) is the most popular item in the data set. It was sold 12 times and brought in 23% more revenue than the second most popular item (Item ID# 145)

Item #s 145, 108 and 82 were sold same number of times. These three items brought different amount of revenues ranging from \$31.77 to \$44.10 because differences in their item prices.



## Most Profitable Items:

Here are the most popular items in the dataset.

| Item ID | Item Name                                    | Purchase Count | Item Price | Total Purchase Value |
|---------|--|----------------|------------|----------------------|
| 178     | Oathbreaker, Last Hope of the Breaking Storm | 12             | \$4.23     | \$50.76              |
| 82      | Nirvana                                      | 9              | \$4.90     | \$44.10              |
| 145     | Fiery Glass Crusader                         | 9              | \$4.58     | \$41.22              |
| 92      | Final Critic                                 | 8              | \$4.88     | \$39.04              |
| 103     | Singed Scalpel                               | 8              | \$4.35     | \$34.80              |

Item ID #178 ***“Oathbreaker, Last Hope of the Breaking Storm”*** is both the most popular and most profitable item. And two items in the previous list also appear in top three most profitable items list. Interestingly, Item ID# 108 (***“Extraction, Quickblade of Trembling hands”***) which was in top three “most popular” category but missing in the “most profitable” list. It sold more than ***“Final Critic”*** or ***“Singed Scalpel”*** but because of its low item price (\$3.53), it didn’t bring in as much revenue.



### Conclusions:

We conclude the report by noting the top three observable trends in the data set.

- 1. Majority of players are Male:** Among 576 total players, 84% are Male and only 14% of players are Female
- 2. 20 to 24-year-old players purchased most items:** The age group (20-24) bought 45% of all the items
- 3. 20 to 24-year-old players are the top spenders:** The age group (20-24) brought in 47% of the total revenue. Also, the age group (35-39) spend the highest per person.

### Postscript:

Few interesting findings from the report are:

1. Top-most spender was a 25-year-old Male followed by a 24-year-old Male in the second position. Even though the females make only 14% of the players, the third highest spender was a 22-year-old Female
2. Players in the age group less than 10 years old are only 3% of the total players, but they are the second highest spender per person (\$4.54)
3. Item ID# 108 has a purchase count of 9, making it the second most popular item. But due to its low item price, it didn't appear in the top five most profitable item list.