



TURTLE GAMES SALES CUSTOMER REVIEW TRENDS ANALYSIS REPORT

LSE_DA301



OCTOBER 1, 2023

ERDI OZYURT

Table of Contents

| | |
|--|----|
| 1. Introduction | 2 |
| 1.1. Introduction | 2 |
| 1.2. GitHub Repository..... | 2 |
| 2. Analytical approach..... | 2 |
| 2.1. Data Exploration and Wrangling | 2 |
| 2.2. Loyalty Points & Insights | 3 |
| 2.3. Different Target Market Segments & Grouping Using K-Means | 5 |
| 2.4. Analysing customer reviews to inform marketing campaigns | 7 |
| 2.4.1. WordCloud for Reviews:..... | 8 |
| 2.4.2. WordCloud for Summaries:..... | 8 |
| 2.4.3. Identify 15 most common words and polarity..... | 9 |
| 2.4.4. Review polarity and sentiment | 10 |
| Top 20 Negative Reviews:..... | 11 |
| Top 20 negative summaries: | 11 |
| Top 20 Positive Reviews: | 12 |
| Top 20 positive summaries: | 12 |
| 3. Product Exploratory Analysis based on Sales Data | 13 |
| 3.1. Statistical Analysis | 14 |
| Plot NA Sales: | 14 |
| Plot EU Sales:..... | 14 |
| Plot Global Sales:..... | 14 |
| 3.2 Sales Data Correlations | 15 |
| 4. Trends, Findings, and Further Recommendations | 15 |

1. Introduction

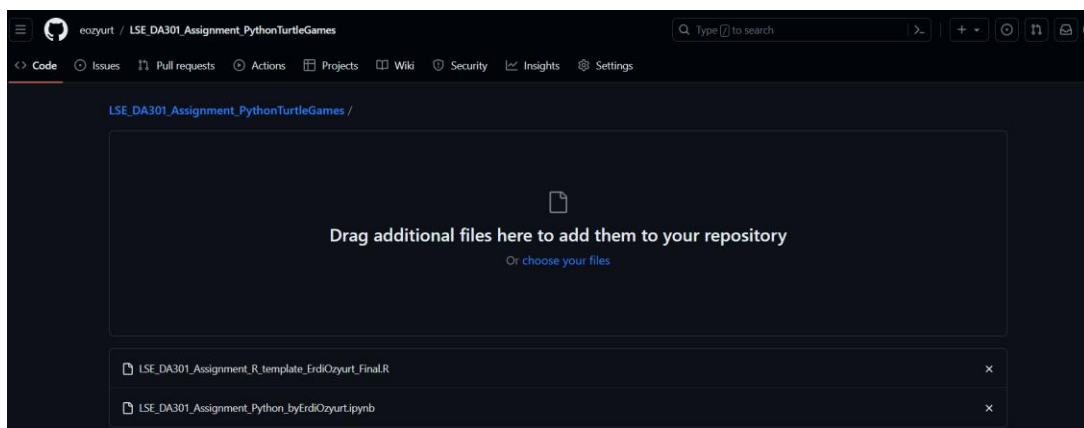
1.1. Introduction

Reflecting the Course 3, Advanced Analytics for Organisational Impact, performed the attached notebook file for a step-by-step analysis of sales and customer reviews. To complete this analysis, we will be analyzing two data sets, `turtle_reviews.csv`, and `turtle_sales.csv`, to identify insights using Python and R.

Turtle Games, a game manufacturer and retailer with a global customer base, is looking to improve overall sales performance by utilizing customer trends. There is an extraordinary growth is expected to continue for the foreseeable future in the gaming industry. This data analytics project aims to help Turtle Games understand:

- Loyalty points
- Customer groups
- Customer reviews
- Product sales
- Shape and correlation between sales regions

1.2. GitHub Repository



2. Analytical approach

2.1. Data Exploration and Wrangling

All the data exploration was completed using Python and R. The **customer review** was loaded into Python to clean, explore, and analyse. The data set was explored by the data types (Data was first

checked using the .info and. describe functions in Python and the view and summary functions in R.) and structure and searched for missing values to provide confidence in subsequent analysis. Unnecessary columns and duplicate observations were removed to aid performance and accuracy. Libraries used in Python were numpy, matplotlib, seaborn, statsmodels, sklearn, scipy, wordcloud, nltk, os, collections, and textblob. Packages used in R were tidyverse, moments, and psych. Further analytical approaches differed depending on the task and are discussed in the sections below.

1. turtle_reviews.csv Data Set

This data set has 2000 entries and 11 columns (inc.0 index), below is a summary of the data types and the number of null values in each column.

```
In [89]: 1 # View the Turtle reviews' metadata
        2 reviews.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2000 entries, 0 to 1999
Data columns (total 11 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                -
0   gender                                2000 non-null   object
1   age                                  2000 non-null   int64
2   remuneration (k£)                    2000 non-null   float64
3   spending_score (1-100)                2000 non-null   int64
4   loyalty_points                        2000 non-null   int64
5   education                            2000 non-null   object
6   language                             2000 non-null   object
7   platform                             2000 non-null   object
8   product                              2000 non-null   int64
9   review                              2000 non-null   object
10  summary                              2000 non-null   object
dtypes: float64(1), int64(4), object(6)
memory usage: 172.0+ KB
```

2. turtle_sales.csv Data Set

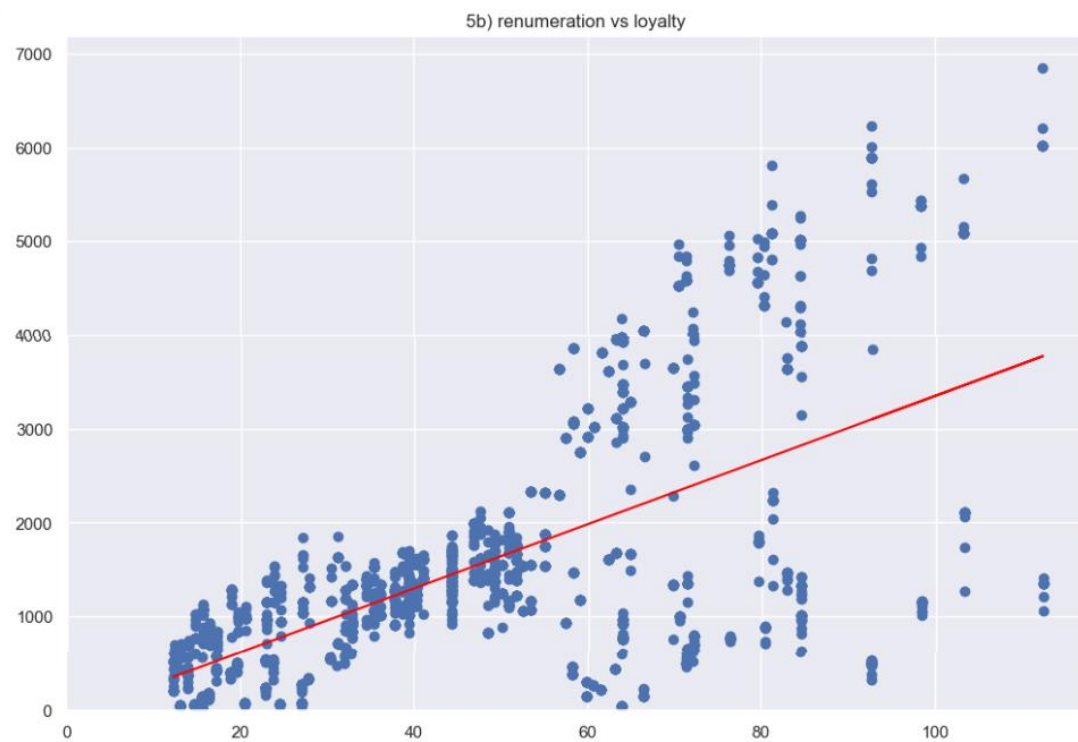
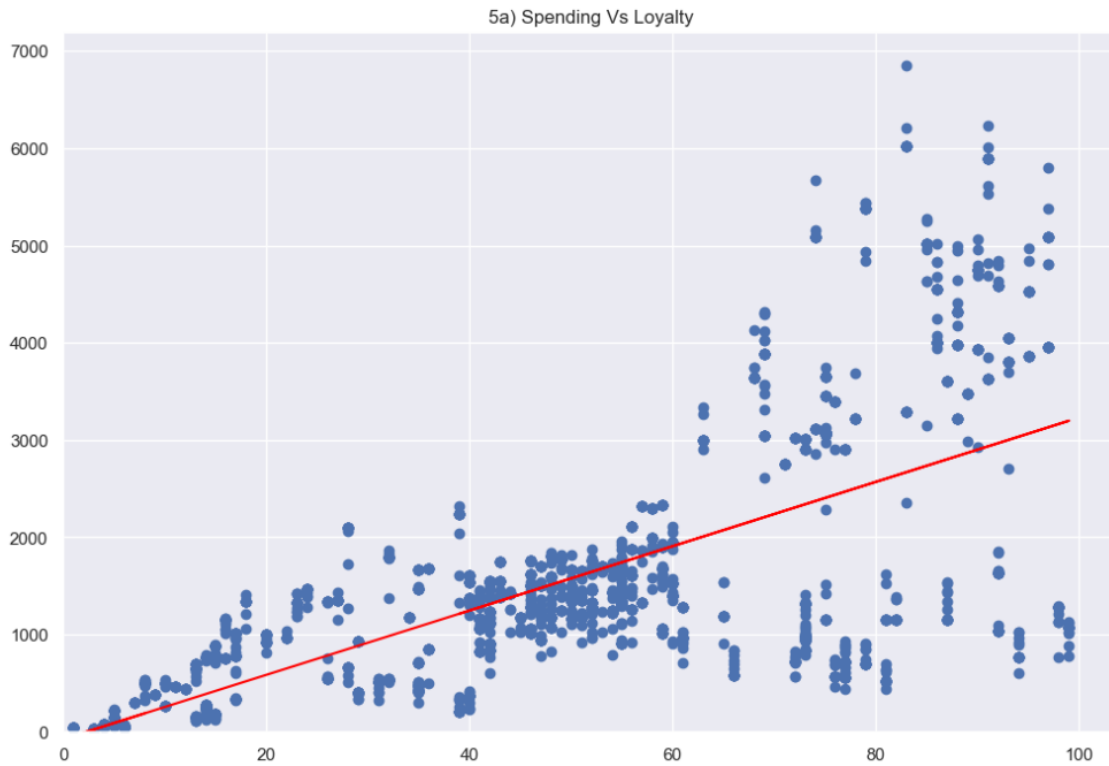
This data set has 352 entries and 5 columns, below is a summary of the data types and the number of null values in each column.

2.2. Loyalty Points & Insights

OLS regression methodology was used to model the linear relationship between loyalty points and other independent variables, chasing a predictor of loyalty points. Followed by reviewing the descriptive statistics for suitability of fit and generating a series of scatterplots.

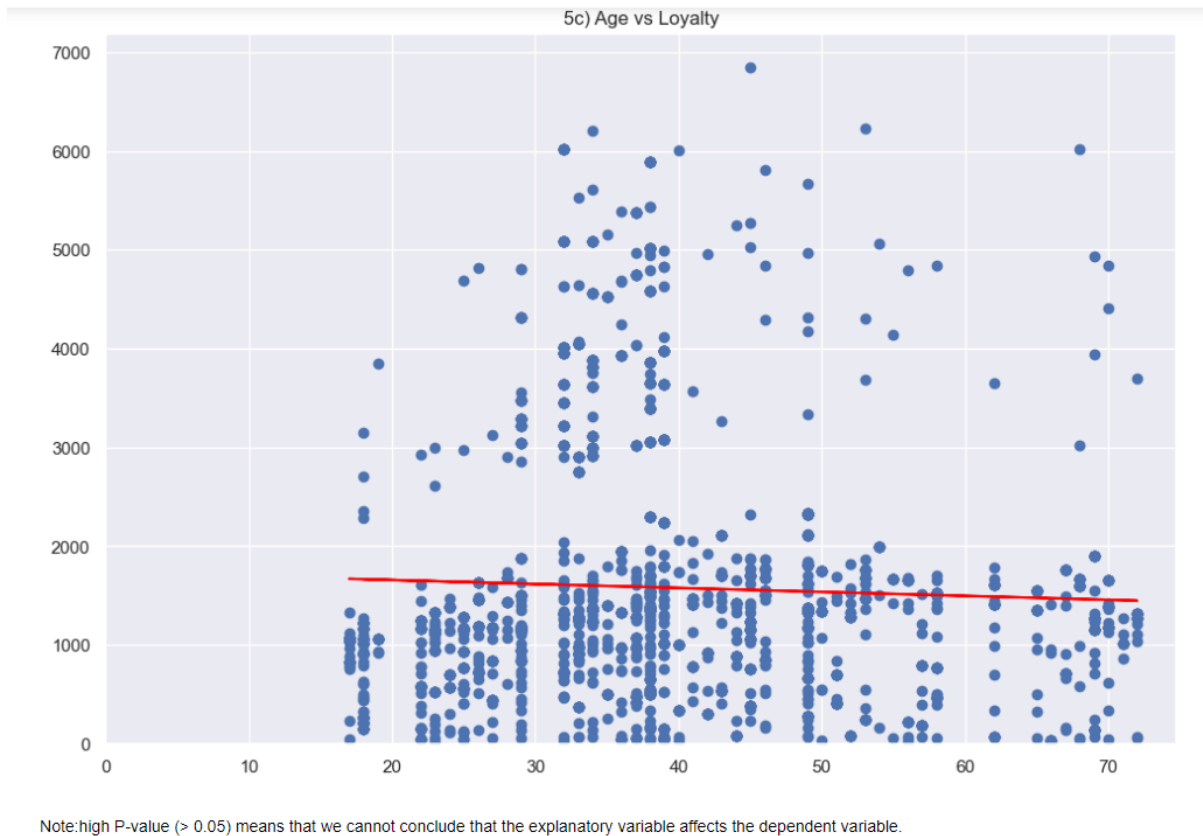
R^2 : 45% of the total variability of y (loyalty), is explained by the variability of X (spending).

P –value is smaller than 0.05 so the set of variables of the regression model is approved.



R^2 : 38% of the total variability of y (loyalty), is explained by the variability of X.

P – value is smaller than 0.05 so the set of variables of the regression model is approved.

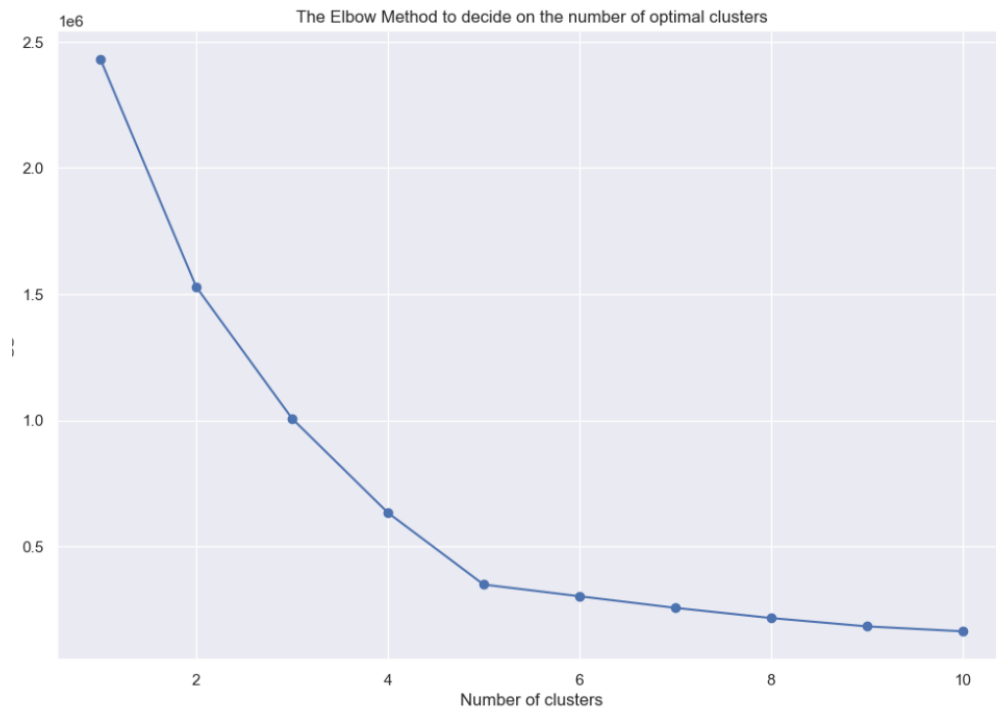


Both the spending vs. loyalty (5A) and the remuneration vs. loyalty (5B) models are significant but the age vs. loyalty model (5C) has a p-value bigger than 0.05 so it cannot be used.

2.3. Different Target Market Segments & Grouping Using K-Means

In this section, we will try to identify the different target markets and groups based on the remuneration and spending scores, within the customer base that can be used to target specific market segments.

Before k-means clustering can be used, the number of clusters must be determined, this is done using the elbow and silhouette methods, both of which suggested a value of k between 4 and 6.



After testing all three, $k = 5$ (five clusters) the most distinct clusters. Achieving this suggests that customers can largely be placed in 5 groups depending on their remuneration and spending score.



Using the k-means model where $k = 5$, cluster 0 appears to be the largest as it can be seen as the typical customer group with an average spending and remuneration level.

2.4. Analysing customer reviews to inform marketing campaigns

In this section, we will analyze the customer reviews. We will be using Natural language processing (NLP) methods, including sentiment analysis.

First, the dataset was prepared by transforming to lowercase, replacing punctuation, tokenization, and removing the stopwords. After these steps, we could reach the “Wordclouds” for reviews and summaries.

1. Load and explore the data
2. Prepare the data for NLP
3. Tokenize and create Wordclouds

2.4.1. WordCloud for Reviews:

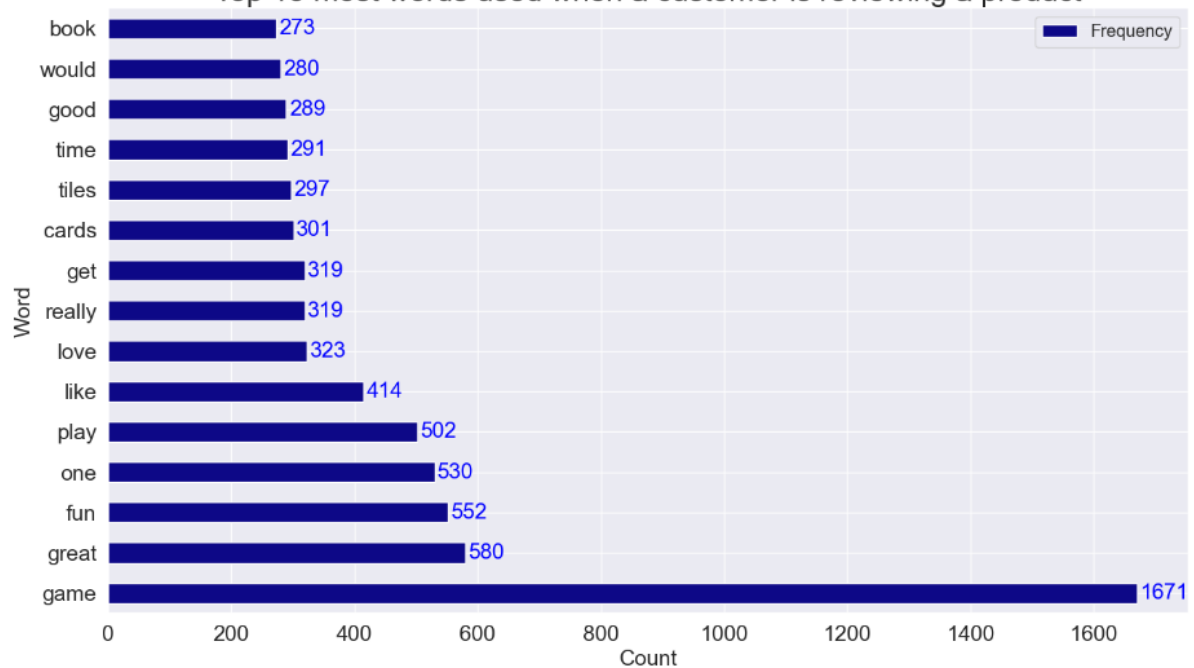


2.4.2. WordCloud for Summaries:

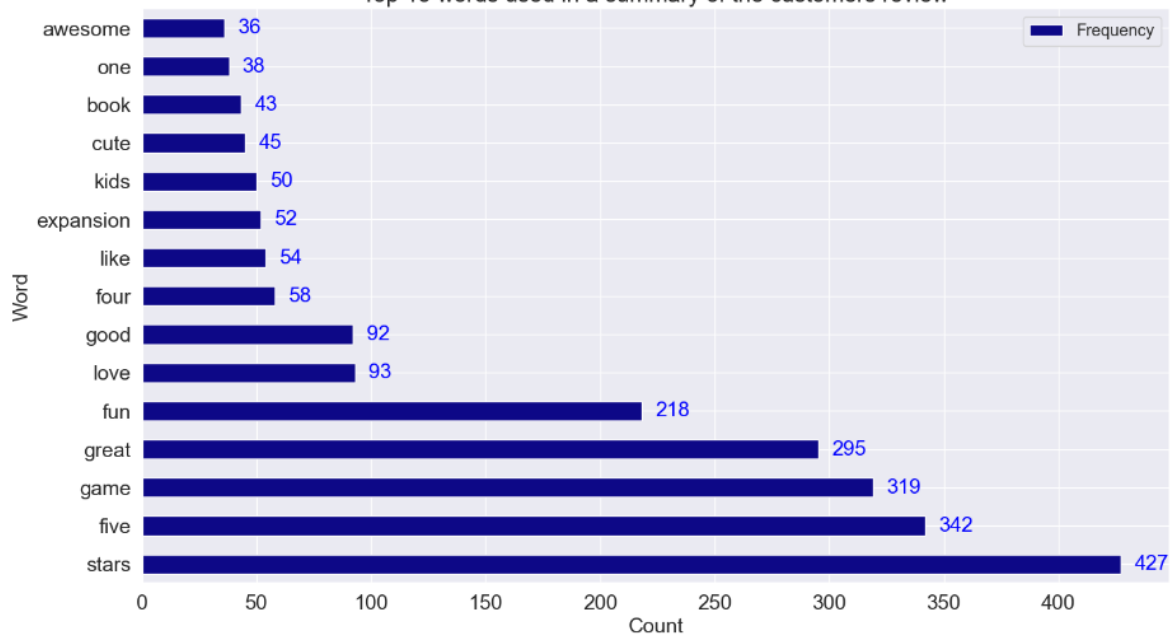


2.4.3. Identify 15 most common words and polarity

Top 15 most words used when a customer is reviewing a product

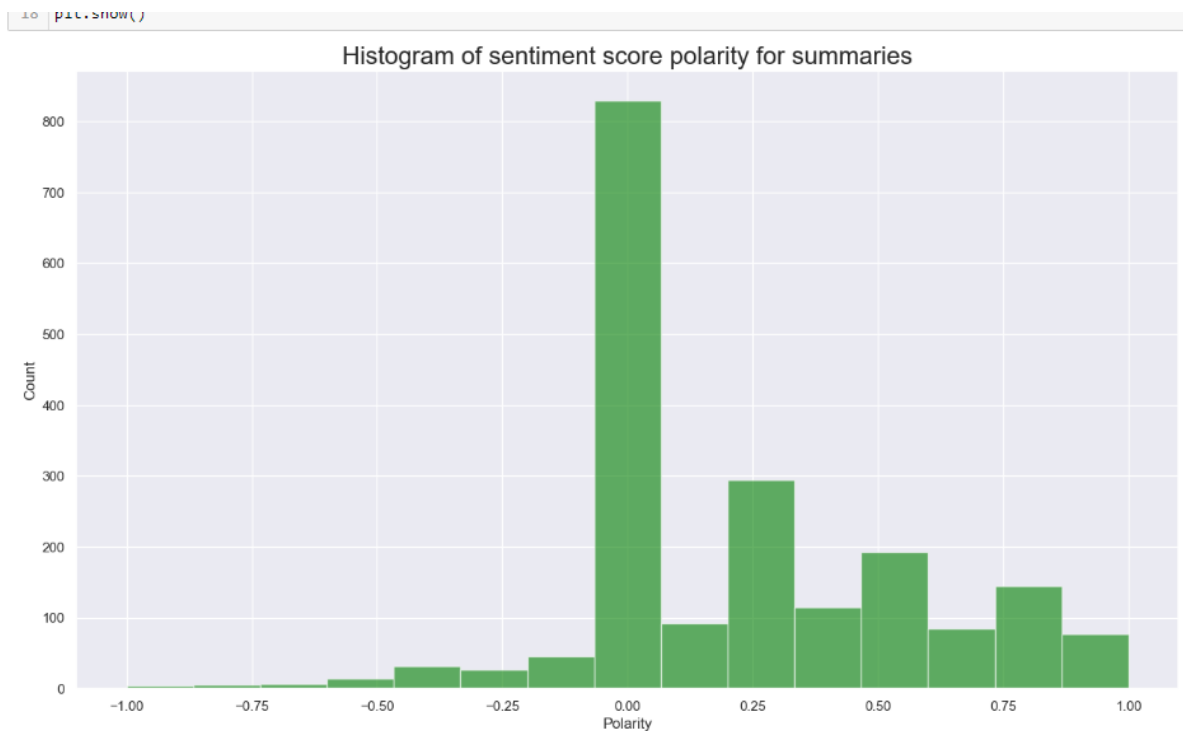
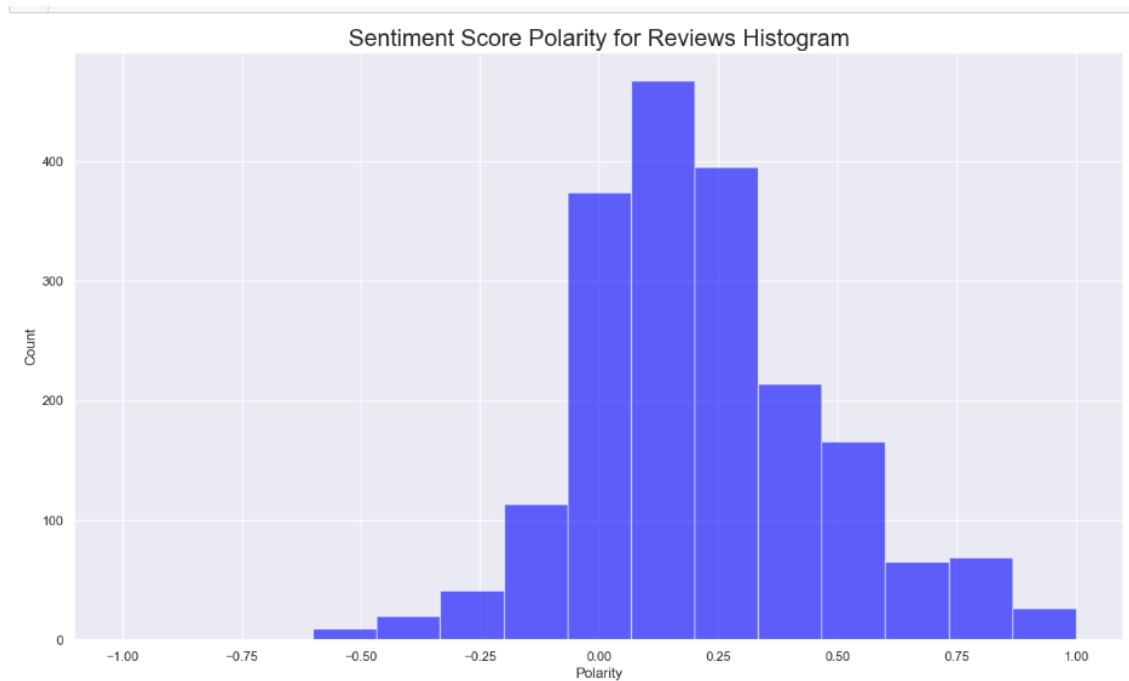


Top 15 words used in a summary of the customers review



First and second on the list are star and Five. I believe 'five stars' is a commonly used. Stars (427) - five (342) equals 85. 85-(58) Four words in comments equals 27 and this gives us a basic approach these comments mainly were 5 stars and 4 stars.

2.4.4. Review polarity and sentiment



We reached two different histogram models.

The histograms show a generally positive sentiment for both reviews and summaries with reviews especially having a higher frequency of results scoring positively for polarity. Summaries were also positive however a much larger number were neutral in sentiment.

Top 20 Negative Reviews:

Out[155]:

| | review | polarity_r |
|------|---|------------|
| 207 | booo unles you are patient know how to measure... | -1.000000 |
| 181 | incomplete kit very disappointing | -0.780000 |
| 1773 | im sorry i just find this product to be boring... | -0.583333 |
| 362 | one of my staff will be using this game soon s... | -0.550000 |
| 116 | i bought this as a christmas gift for my grand... | -0.500000 |
| 226 | this was a gift for my daughter i found it dif... | -0.500000 |
| 229 | i found the directions difficult | -0.500000 |
| 289 | instructions are complicated to follow | -0.500000 |
| 300 | difficult | -0.500000 |
| 1501 | expensive for what you get | -0.500000 |
| 173 | i sent this product to my granddaughter the po... | -0.491667 |
| 345 | my 8 yearold granddaughter and i were very fru... | -0.446250 |
| 531 | i purchased this on the recommendation of two ... | -0.440741 |
| 305 | very hard complicated to make these | -0.439583 |
| 421 | kids i work with like this game | -0.400000 |
| 430 | this game although it appears to be like uno a... | -0.400000 |
| 490 | my son loves playing this game it was recommen... | -0.400000 |
| 795 | this game is a blast | -0.400000 |
| 798 | i bought this for my son he loves this game | -0.400000 |
| 814 | was a gift for my son he loves the game | -0.400000 |

Top 20 negative summaries:

Out[156]:

| | summary | polarity_s |
|------|---|------------|
| 21 | the worst value ive ever seen | -1.000000 |
| 207 | boring unless you are a craft person which i am | -1.000000 |
| 819 | boring | -1.000000 |
| 1148 | before this i hated running any rpg campaign d... | -0.900000 |
| 1 | another worthless dungeon masters screen from ... | -0.800000 |
| 143 | disappointed | -0.750000 |
| 623 | disappointed | -0.750000 |
| 785 | disappointed | -0.750000 |
| 1591 | disappointed | -0.750000 |
| 361 | promotes anger instead of teaching calming met... | -0.700000 |
| 875 | too bad this is not what i was expecting | -0.700000 |
| 880 | bad qualityall made of paper | -0.700000 |
| 177 | at age 31 i found these very difficult to make | -0.650000 |
| 100 | small and boring | -0.625000 |
| 511 | mad dragon | -0.625000 |
| 797 | disappointing | -0.600000 |
| 1001 | disappointing | -0.600000 |
| 1099 | disappointing | -0.600000 |
| 1773 | disappointing | -0.600000 |
| 991 | then you will find this board game to be dumb ... | -0.591667 |

Top 20 Positive Reviews:

Out[157]:

| | review | polarity_r |
|------|---|------------|
| 7 | came in perfect condition | 1.0 |
| 164 | awesome book | 1.0 |
| 193 | awesome gift | 1.0 |
| 489 | excellent activity for teaching selfmanagement... | 1.0 |
| 517 | perfect just what i ordered | 1.0 |
| 583 | wonderful product | 1.0 |
| 601 | delightful product | 1.0 |
| 613 | wonderful for my grandson to learn the resurre... | 1.0 |
| 782 | perfect | 1.0 |
| 923 | awesome | 1.0 |
| 1119 | awesome set | 1.0 |
| 1150 | best set buy 2 if you have the means | 1.0 |
| 1159 | awesome addition to my rpg gm system | 1.0 |
| 1282 | its awesome | 1.0 |
| 1380 | one of the best board games i played in along ... | 1.0 |
| 1523 | my daughter loves her stickers awesome seller ... | 1.0 |
| 1580 | this was perfect to go with the 7 bean bags i... | 1.0 |
| 1684 | awesome toy | 1.0 |
| 1689 | it is the best thing to play with and also min... | 1.0 |
| 1695 | excellent toy to simulate thought | 1.0 |

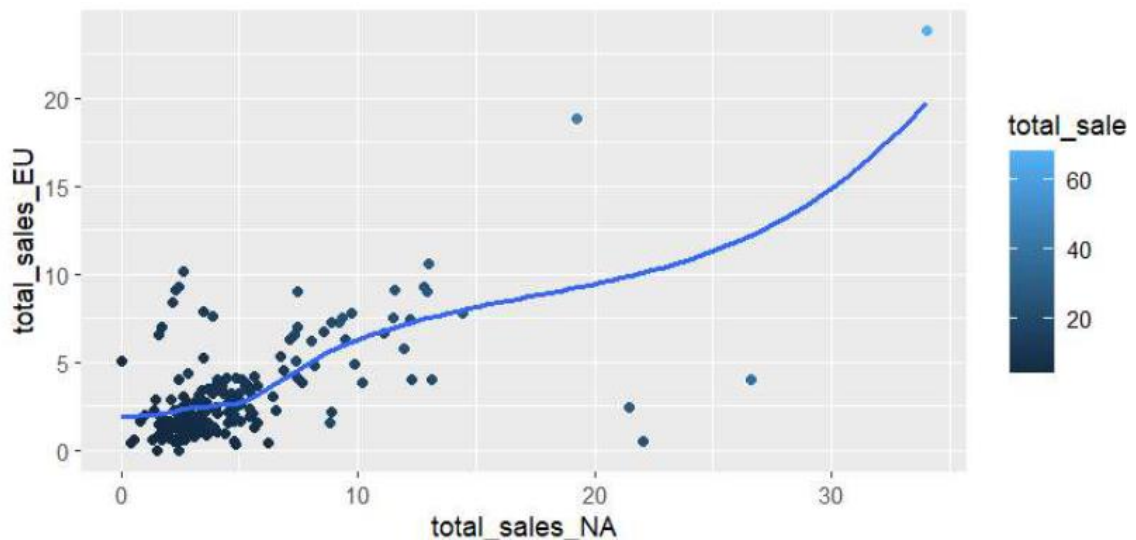
Top 20 positive summaries:

Out[160]:

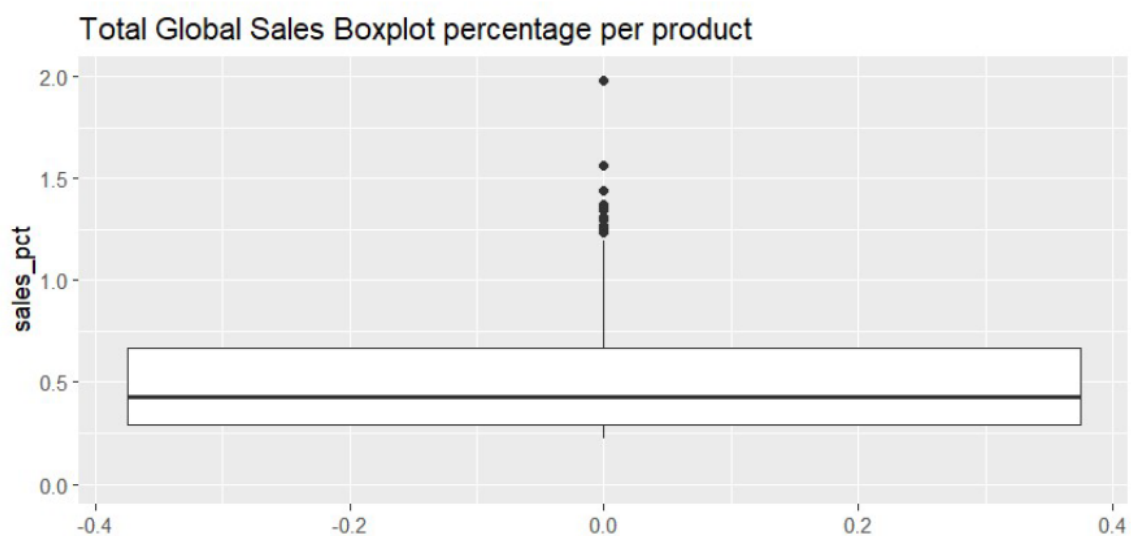
| | index | review | summary | tokens_review | tokens_summary | polarity_s |
|-----|-------|---|---|---|---|------------|
| 6 | 6 | i have bought many gm screens over the years b... | best gm screen ever | [i, have, bought, many, gm, screens, over, the... | [best, gm, screen, ever] | 1.0 |
| 28 | 28 | these are intricate designs for older children... | wonderful designs | [these, are, intricate, designs, for, older, c... | [wonderful, designs] | 1.0 |
| 32 | 32 | awesome my 8 year olds favorite xmas gift its ... | perfect | [awesome, my, 8, year, olds, favorite, xmas, g... | [perfect] | 1.0 |
| 80 | 80 | my daughter loves these little books theyre th... | theyre the perfect size to keep in the car or ... | [my, daughter, loves, these, little, books, th... | [theyre, the, perfect, size, to, keep, in, the... | 1.0 |
| 133 | 134 | this occupied my almost3 year old for nearly a... | perfect for preschooler | [this, occupied, my, almost3, year, old, for, ... | [perfect, for, preschooler] | 1.0 |
| 139 | 140 | i bought 8 of these for my 3 year old daughter... | awesome sticker activity for the price | [i, bought, 8, of, these, for, my, 3, year, ol... | [awesome, sticker, activity, for, the, price] | 1.0 |
| 160 | 161 | my 8 year old son loves this drawing book love... | awesome book | [my, 8, year, old, son, loves, this, drawing, ... | [awesome, book] | 1.0 |
| 162 | 163 | this was a christmas present for a nephew who ... | he was very happy with his gift | [this, was, a, christmas, present, for, a, nep... | [he, was, very, happy, with, his, gift] | 1.0 |
| 186 | 187 | great product took a little practice and time ... | awesome | [great, product, took, a, little, practice, an... | [awesome] | 1.0 |
| 209 | 210 | i was skeptical but my 9 year old has had so m... | awesome and welldesigned for 9 year olds | [i, was, skeptical, but, my, 9, year, old, has... | [awesome, and, welldesigned, for, 9, year, olds] | 1.0 |
| 412 | 418 | great resource for bhis care coordinators work... | perfect | [great, resource, for, bhis, care, coordinator... | [perfect] | 1.0 |
| 468 | 475 | this is a great anger management game for kidd... | excellent | [this, is, a, great, anger, management, game, ... | [excellent] | 1.0 |
| 536 | 543 | this is a great game to play with kids of diff... | excellent | [this, is, a, great, game, to, play, with, kid... | [excellent] | 1.0 |
| 541 | 548 | i am a therapist for children and this game is... | excellent therapy tool | [i, am, a, therapist, for, children, and, this... | [excellent, therapy, tool] | 1.0 |
| 572 | 580 | if you have the pigeon books in your elemental... | the pigeon is the perfect addition to a school... | [if, you, have, the, pigeon, books, in, your, ... | [the, pigeon, is, the, perfect, addition, to, ... | 1.0 |
| 591 | 599 | wow this product is incredible i was hesitant ... | best easter teaching tool | [wow, this, product, is, incredible, i, was, h... | [best, easter, teaching, tool] | 1.0 |
| 639 | 647 | beautiful we look forward to the easter story ... | wonderful | [beautiful, we, look, forward, to, the, easter... | [wonderful] | 1.0 |
| 643 | 651 | we have several different sets of these magnet... | all f the mudpuppy toys are wonderful | [we, have, several, different, sets, of, these... | [all, f, the, mudpuppy, toys, are, wonderful] | 1.0 |
| 649 | 657 | my son is 2 years and 4 months he is a little ... | awesome puzzle | [my, son, is, 2, years, and, 4, months, he, is... | [awesome, puzzle] | 1.0 |
| 654 | 662 | great puzzles for toddlers to put together by ... | not the best quality | [great, puzzles, for, toddlers, to, put, toget... | [not, the, best, quality] | 1.0 |

3. Product Exploratory Analysis based on Sales Data

In order to analyze this dataset; Scatterplots and histograms were used to investigate the sales of each product as well as other aspects. These show that most games tend to bring between zero and 5 million in sales. NA tends to be a stronger performer than the EU in terms of sales. there seems to be a positive correlation between NA sales, EU sales, and Global Sales.



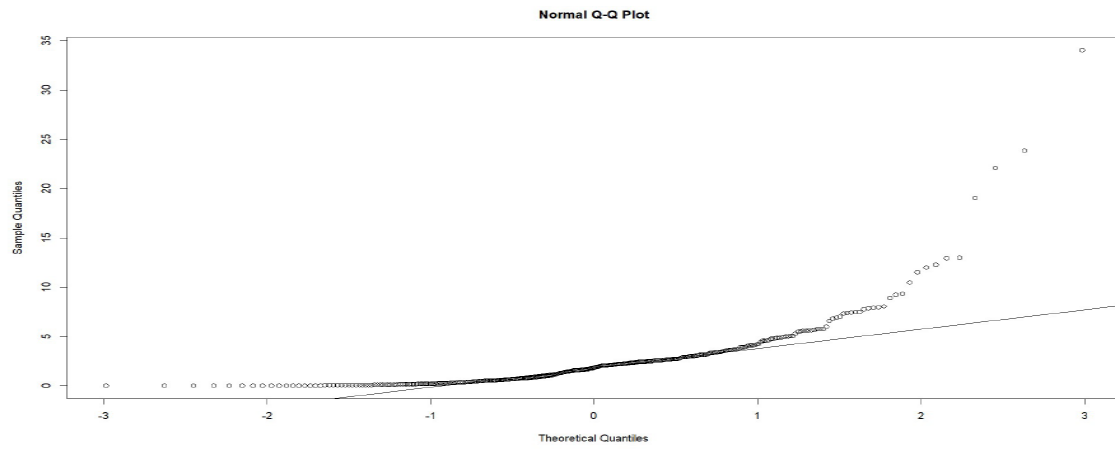
There are a few notable extreme outliers where product sales in one region have been much higher than the other.



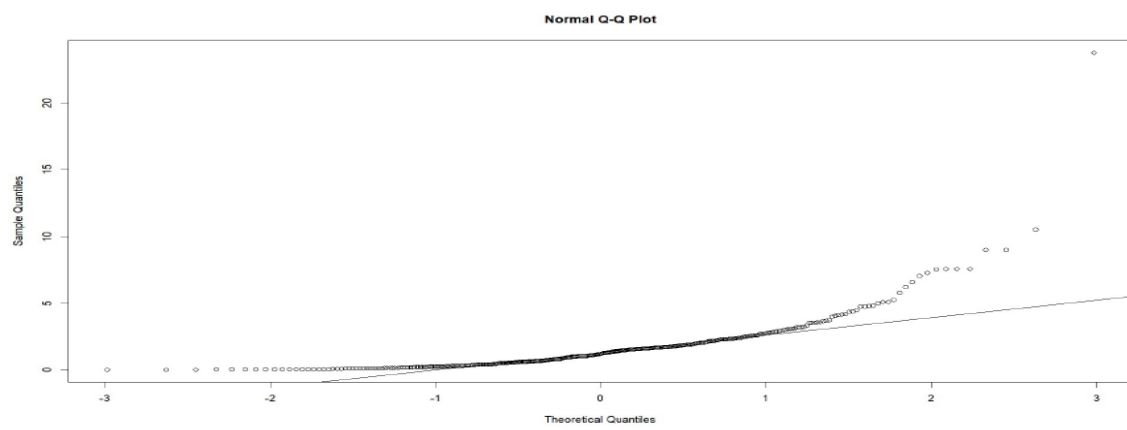
50% of the products are 0.3% to the 0.7% of the Global Sales.

3.1. Statistical Analysis

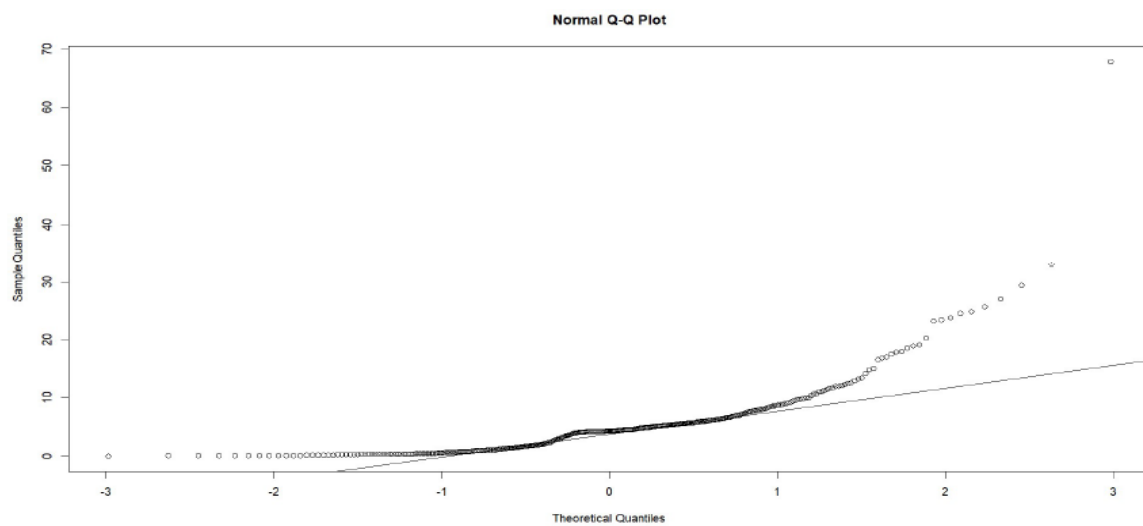
Plot NA Sales:



Plot EU Sales:



Plot Global Sales:



By looking at the three Q-Q plots, we can express that the data is not normally distributed.

3.2 . Sales Data Correlations

In Sales Data Correlations, we will explore the determine the relationship between Global Sales and EU and NA Sales.

The correlation plot indicated that most of the variables tend to have a positive correlation – for example, NA Sales & EU Sales (0.71), NA Sales & Global Sales (0.93) and EU Sales & Global Sales (0.88).

As a result of the multiple linear regression model, predicted values with the model and concluded that it was very accurate. Further analysis can be done to understand the factors that contribute to the high correlation between NA Sales & EU Sales, NA Sales & Global Sales, and EU Sales & Global Sales.

4. Trends, Findings and Further Recommendations

- Loyalty points accumulation shows a moderate correlation with customer spending (45%) and remuneration (38%), but not with age.
- Suggesting the segment of the customers into five groups and marketing can be targeted towards the high earners.
- Sentiment analysis reveals that negative comments often use words like "boring" and "disappointing", while positive comments tend to use words like "excellent", "perfect", "quality", and "awesome".
- Customer reviews showed overall sentiment towards the products are positive.
- Global sales are dependent on NA and Eu sales however sales on other areas are growing.

