Analysis for Turtle Games

Predict loyalty points, define customer segments and understand sentiment

Prepared by Lilliana Golob
LSE Data Analytics Career Accelerator

Why are we doing this analysis?

Increase sales

Insights and recommendations

Create targeted campaigns

Improve business operations

Before we begin, let's review why are we doing this analysis.

Turtle Games is a global manufacturer and retailer of games, books and toys.

The senior stakeholders want to increase sales, and are seeking insights and recommendations to create targeted and effective marketing campaigns, and improve business operations.

What do we want to know?

1. What contributes to loyalty points accumulation?
2. How can we predict loyalty points?
3. How can we segment and target customers?
4. How can customer reviews help the business?

Here are the four key questions that Turtle Games have.

In this presentation I'm going to answer each and provide recommendations as we go along.

Analytical approach

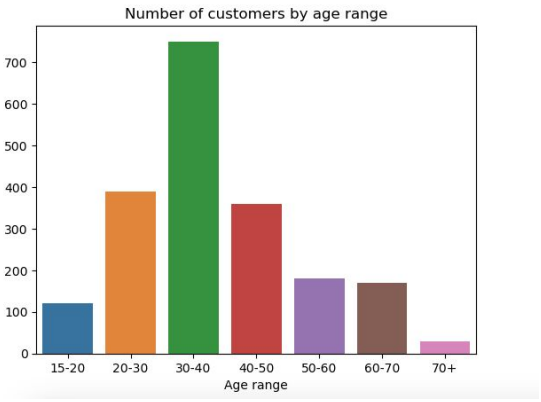
Using Python and R

- Data exploration
- Linear regression
- Decision trees
- K-means clustering
- Sentiment analysis

To conduct my analysis I used Python and R and performed data exploration, linear regression, decision tree, k-means clustering and sentiment analysis.

Let’s begin

Did you know 56% of UK adults play games?



Source: Ofcom Online Nation Report 2023

Let's start with a quick look at Turtle Games' customers.

38% of Turtle Games' customers are aged between 30 and 40 years old, and most have above basic education.

This chart confirms the research conducted by Ofcom that gaming is not just for young people. According to Ofcom's Online Nation Report 56% of UK adults play games on or offline, and 23% of 55-64-year-olds play games online.

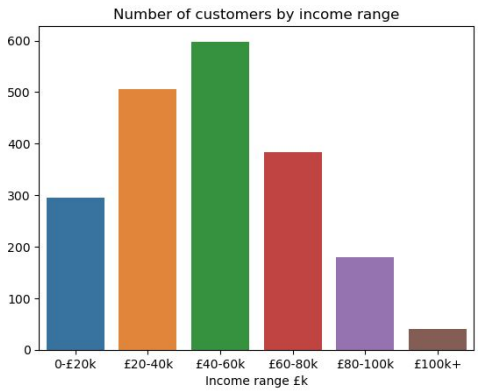
56% female

2,000 customers

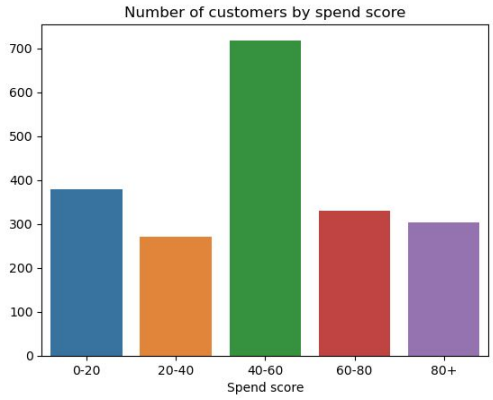
1,120 female

880 male

Looking at gender, the data set is pretty equally split with 56% of customers are female.



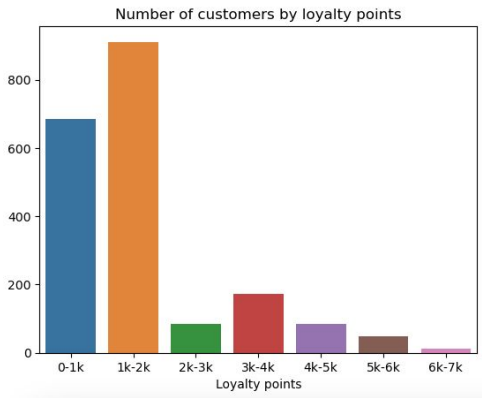
Income



Spend score

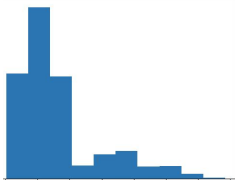
Looking at income and spend score, 55% of customers earn between £20k and £60k. And 36% of customers have a spend score between 40 and 60.

80% of customers have less than 2,000 loyalty points

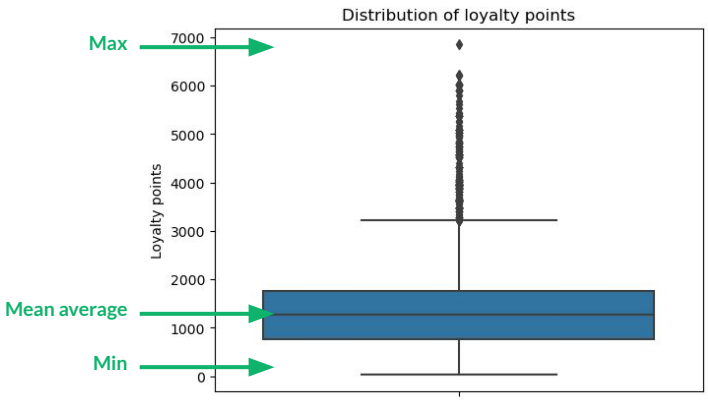


80% of Turtle Games' customers have between 0 to 2000 loyalty points, with 46% having 1000 to 2000.

Loyalty points are spread far and wide



Histogram of loyalty points distribution



The range of values for loyalty points is 25 to 6847, and the average value is 1578.

Data is skewed right, meaning most loyalty points are lower valued, as we can see from the histogram. We can also see that higher values appear as outliers on the boxplot.

Average by age, income and spend score

spend_score_bin	0-20	20-40	40-60	60-80	80+
loyalty_points	424.775132	1024.03321	1429.993036	2175.551515	3212.273927

income_bin	0-£20k	£20-40k	£40-60k	£60-80k	£80-100k	£100k+
loyalty_points	509.99661	982.409901	1623.792642	2291.788512	2877.357542	3641.7

age_bin	15-20	20-30	30-40	40-50	50-60	60-70	70+
loyalty_points	889.783333	1264.666667	2133.878667	1378.344444	1177.516667	1246.517647	1186.533333

As expected, customers with low spend scores and income have fewer average loyalty points.

Interestingly, customers aged 30-40 have an above-average number of loyalty points. And when we look at maximum loyalty points, those aged between 30 to 70 have the highest, indicating age could be important.

Q1:

What contributes to customers accumulating loyalty points?

Moving on to the first question: what contributes to customers accumulating loyalty points?

To answer this question, I used linear regression and decision tree regression. And I came to the conclusion *(drum roll please)*...

Income and spend score

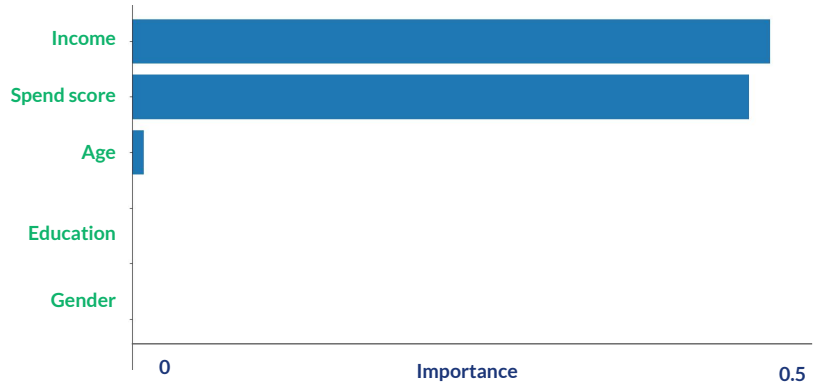
Together are strong predictors of loyalty points

Income and spend score together are strong predictors of loyalty points.

In terms of accuracy, the multiple linear regression model is 82.7% accurate. However, it did have a high mean absolute error, indicating that the model's predictions could be off by 424 from the actual values.

As you saw earlier the distribution of loyalty points was very wide with a high standard deviation so this was expected..

Important features



I also did decision tree regressor analysis and it came to the same conclusion that income and spend score are the most important features.

Interestingly, age is barely an important feature (contrary to our data exploration) and education and gender don't feature at all (no pun intended).

Recommendations

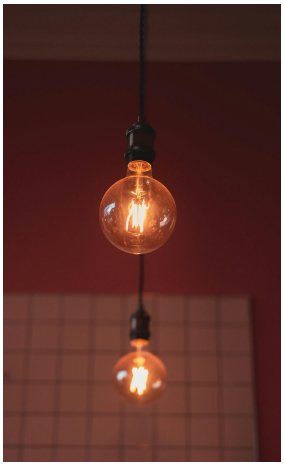
01 | Use multiple linear regression for predictions.

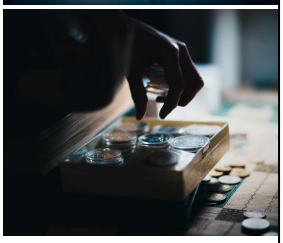
02 | Spend score and income as predictor variables.

03 | Additional data to help train the model and improve accuracy.

04 | Further analysis using classification decision trees.

05 | Look at categorical variables including product categories.





My recommendations are to use the multiple linear regression model with spend score and income as the variables to make your predictions.

I would like additional data to help train the model because the data set is only 2000 lines which is quite small. This could improve accuracy and reduce the error.

I would also like to do further analysis using classification decision trees so we can look at how categorical variables (such as product categories) affect loyalty points.

Q2:

How to predict loyalty points?

Question number two. To answer this I used the multiple linear regression model with spend score and income as the predictors.

And as expected *(drum roll please)*...

Loyalty point predictions

Spend score	£30k income	£50k income	£70k income
25	141	821	1501
50	964	1643	2323
75	1786	2466	3145
100	2608	3288	3968

The predictions from the multiple linear regression model show that spend score greatly impacts loyalty points accumulation for the same income.

For example: If you have a spend score of 75 and £30k income, the predicted loyalty points is 1786. However if you're spend score is only 25, you only get 141.

Recommendations

01 | **Focus on maximising customer's spend scores.**

02 | **Communicate the value of loyalty points.**

03 | **Explain how easy it can be to accumulate more.**

04 | **And what extra they can get.**

Based on these results, I would focus marketing efforts on maximising customer spend scores regardless of of income. I would communicate the value of loyalty points (e.g. the benefits) and how easy it can be to accumulate more.

Q3:
How to segment customers?

To answer this question around customer segmentation, I used k-means clustering (a popular unsupervised machine learning algorithm) using our two important features: spend score and income.

And from these results the answer is *(drum roll again please)*...

Five segments

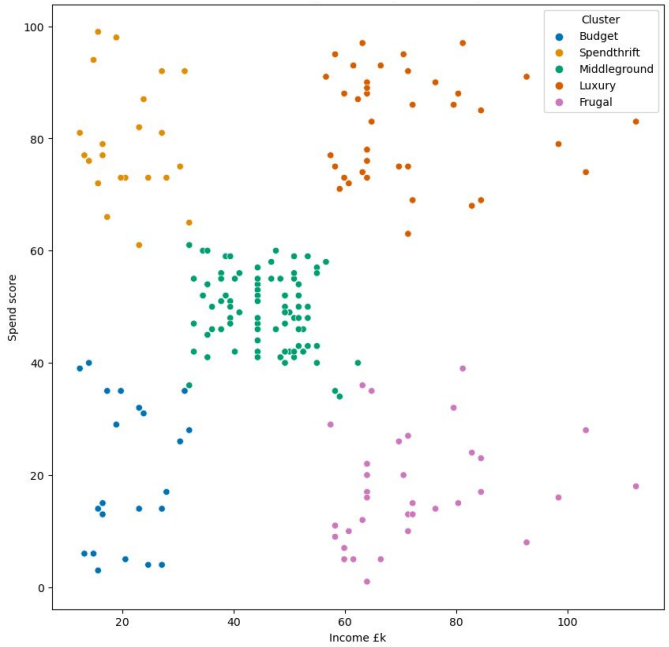
Budget
Low income, low spender

Spendthrift
Low income, high spender

Middleground
Moderate income, moderate spender

Luxury
High earner, high spender

Frugal
High earner, low spender



Turtle Games should segment their customers into five clusters, which I renamed based on their characteristics.

Segment	Income	Spend score	Marketing opportunity
Luxury	£60-100k+	60 - 100	Exclusive deals, premium products, premium loyalty deals, exclusive members club
Spendthrift	£10-40k	60 - 100	Exclusive deals, early access to new products, loyalty discounts, monthly subscriptions, payment plans
Middleground	£40-60k	40 - 60	Mid-range products, loyalty discount
Frugal	£60-100k+	10 - 40	Discounts, promote value for money, budget bundles
Budget	£10-40k	10 - 40	Discounts, end-of-line sales, payment plans, monthly subscriptions, budget bundles

This table shows the respective income, spend score and recommended marketing opportunities for each segment.

Recommendations

01 | Target Luxury and Spendthrift (higher spend scores).

02 | Communicate exclusivity and benefits.

03 | Test campaign at Frugal focussing on value for money.

04 | Further analysis on customers with higher loyalty points.

05 | Develop personas which will improve messages and results.

From the segment analysis my recommendations are to target Luxury and Spendthrift with higher spend scores, communicating exclusivity and benefits.

I would test campaign at Frugal focus on value for money, and apply the results from this campaign to the Middleground segment.

I would also like to do further analysis on customers with higher loyalty points to identify commonalities, and develop personas which help create personalised and appealing messages, improving results.

Q4:

How can customer reviews help the business?

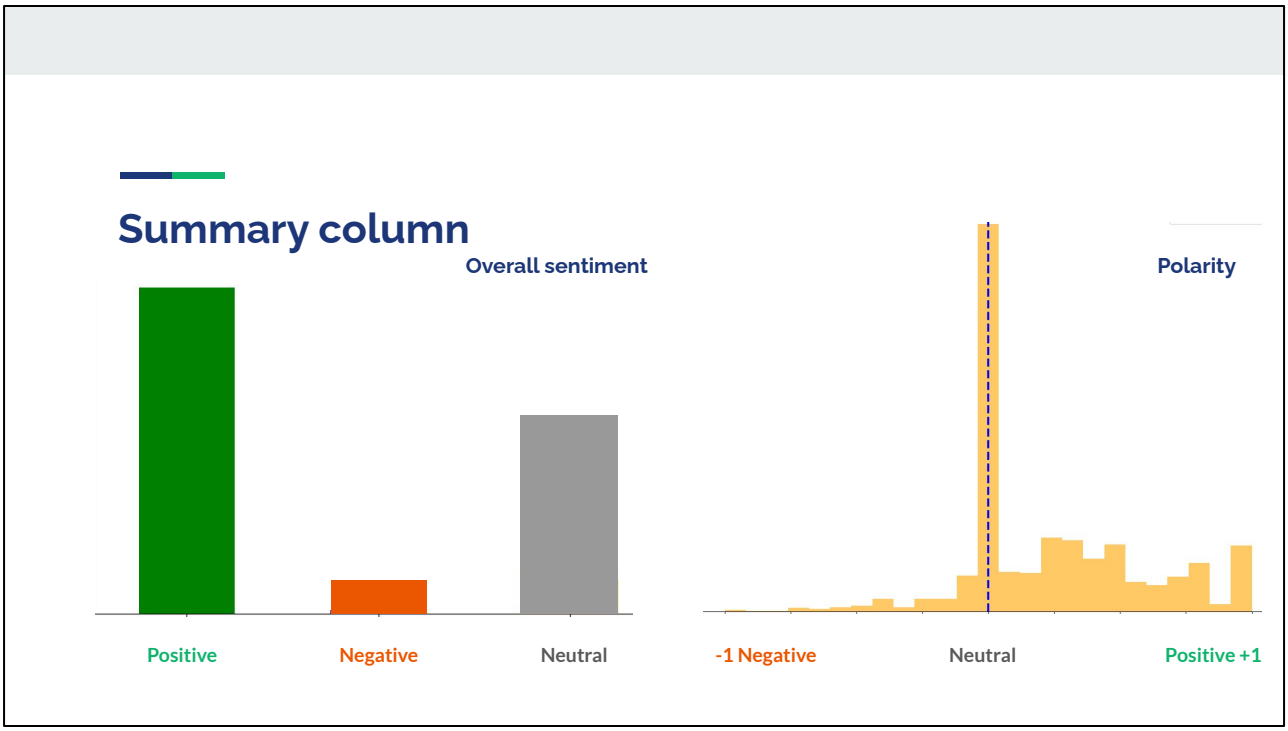
To answer how customers reviews can help the business I focused on VADER and TextBlob sentiment analysis, which use natural language processing (NLP).

These methods are useful to understand how customers are feeling, and identify issues and opportunities.

Before we go into the recommendations, let's first unveil how customers feel about Turtle Games *(final drum roll please)*...

Overall sentiment from customer reviews is positive

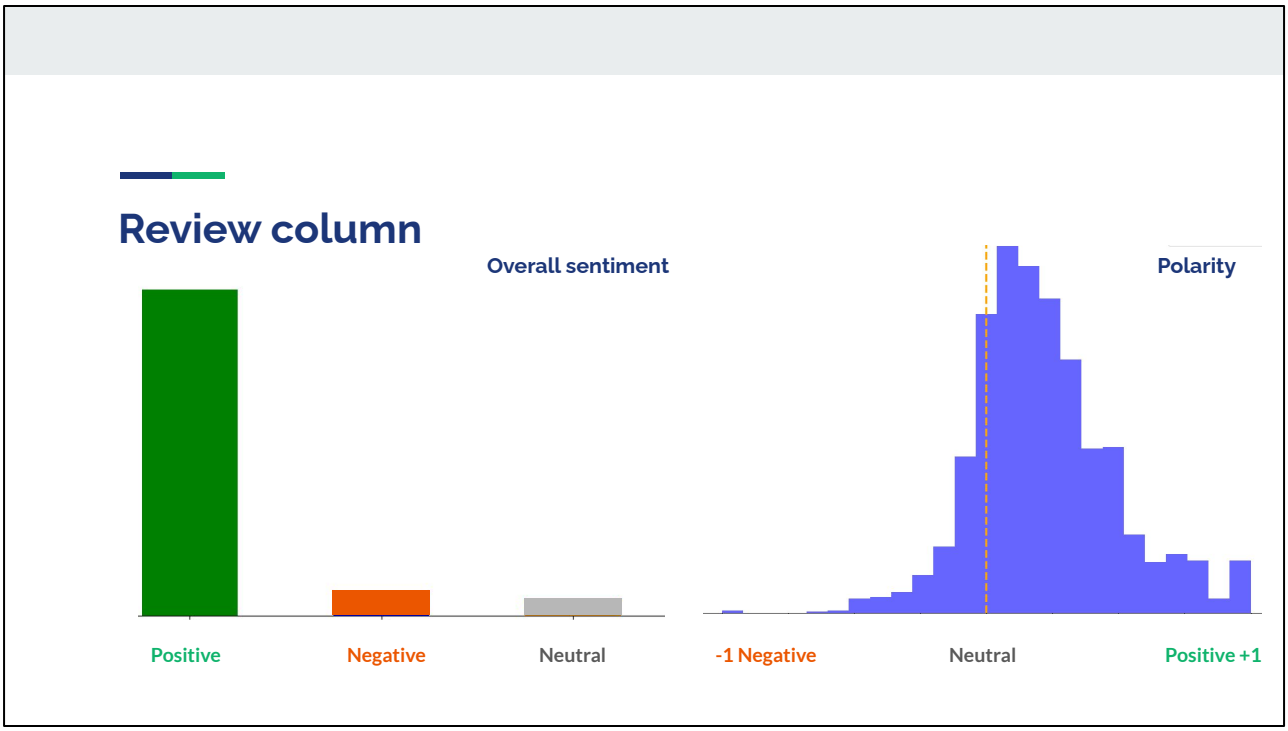
The overall sentiment for Turtle Games across both the Review and Summary columns is positive.



These are the results from the Summary column.

The chart on the left used VADER analysis which gives an overall sentiment score.

The chart on the right is from TextBlob which gives a polarity score from -1 (very negative) to +1 (extremely positive), with 0 as neutral.



Here are the results from the Review column.

As you can see from all four charts, there are more positive reviews than negative.

So while you can pat Turtle Games can pat themselves on the back for a job well done, we must remember that negative customer sentiment is one of the biggest risks to any business.

Recommendations

01 | Investigate negative reviews.

02 | Determine issues and common themes.

03 | Implement appropriate control measures.

04 | Turn threats into marketing opportunities.

05 | Communicate improvements to build trust.

06 | Promote and suggest alternatives that have high scores.

Going back to the question, how can you use sentiment to help your business?

For starters, you identify issues and turn them into marketing opportunities.

My recommendations include conduct further investigation into the negative sentiment to identify issues and common themes, and more importantly, develop and implement appropriate corrective measures. Whether that's discontinuing a product or contacting the customer, for example.

And let's try to flip those frowns upside down by turning negative sentiment (threat) into opportunities. For example communicate improvements or corrective measures to customers, and promote alternatives that have high sentiment. This will all go towards rebuilding trust and loyalty, and achieving your overall objective of more sales.

I'll leave you with this quote from Bill Gates, founder of Microsoft, which I feel best answers the final question.

"Your most unhappy customers are your greatest source of learning".

Get in touch

lilliana.golob@gmail.com | 07501 450 668