TURTLE GAMES

Marketing Case Study Using Predictive Analytics

By Sagar Varma

Date : 16-Dec-2024

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Turtle Games, a global game manufacturer and retailer, aims to improve sales performance by leveraging customer insights. This report analyses customer data to understand loyalty point accumulation, customer behaviour patterns, and market segments. Key findings include:

- Spending score, remuneration and age are the most significant predictors of loyalty points.
- K-means clustering identified 5 distinct customer segments based on spending score and remuneration.
- NLP analysis revealed common themes and sentiments in customer reviews.

These insights enable Turtle Games to tailor marketing strategies, optimize loyalty programs, and improve customer satisfaction.

INTRODUCTION

Turtle Games, a global manufacturer and retailer of games, books, board games, video games and toys, aims to improve its overall sales performance by understanding customer trends. This report focuses on analysing customer data to achieve the following objectives:

- Understand how customers accumulate loyalty points.
- Explore customer behaviour patterns using decision trees.
- Identify customer segments through clustering.
- Analyse customer reviews using Natural Language Processing (NLP).

The insights gained from this analysis will inform Turtle Games' marketing strategies, loyalty program design, and overall customer relationship management.

DATA AND METHODOLOGY

The analysis utilizes the 'turtle_reviews.csv' dataset, containing customer information such as age, gender, education, remuneration, spending score, and loyalty points. The dataset was cleaned and pre-processed to ensure data quality and consistency. For NLP analysis, customer reviews and summaries were extracted.

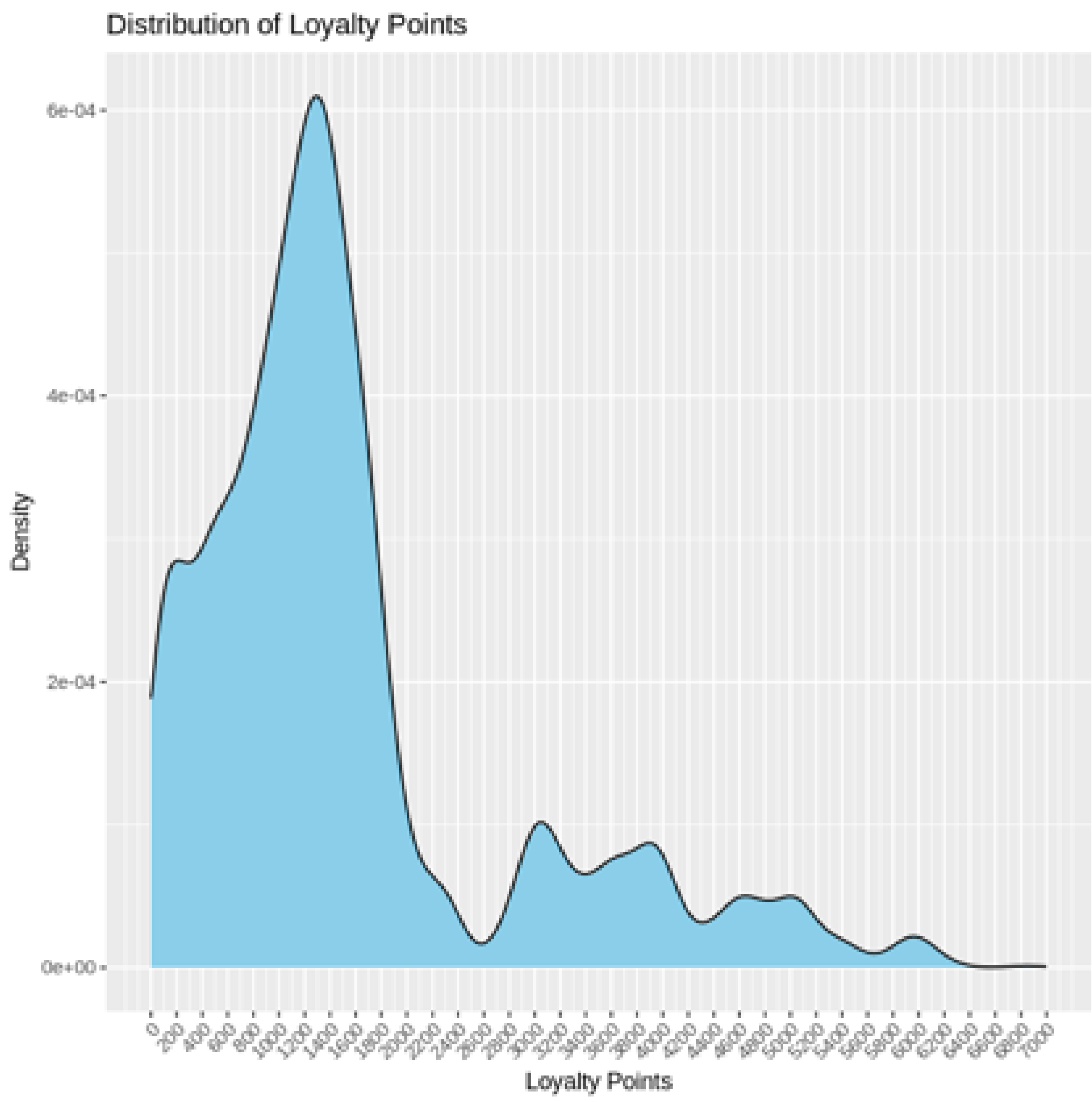
EXPLORATORY DATA ANALYSIS

DEMOGRAPHICS

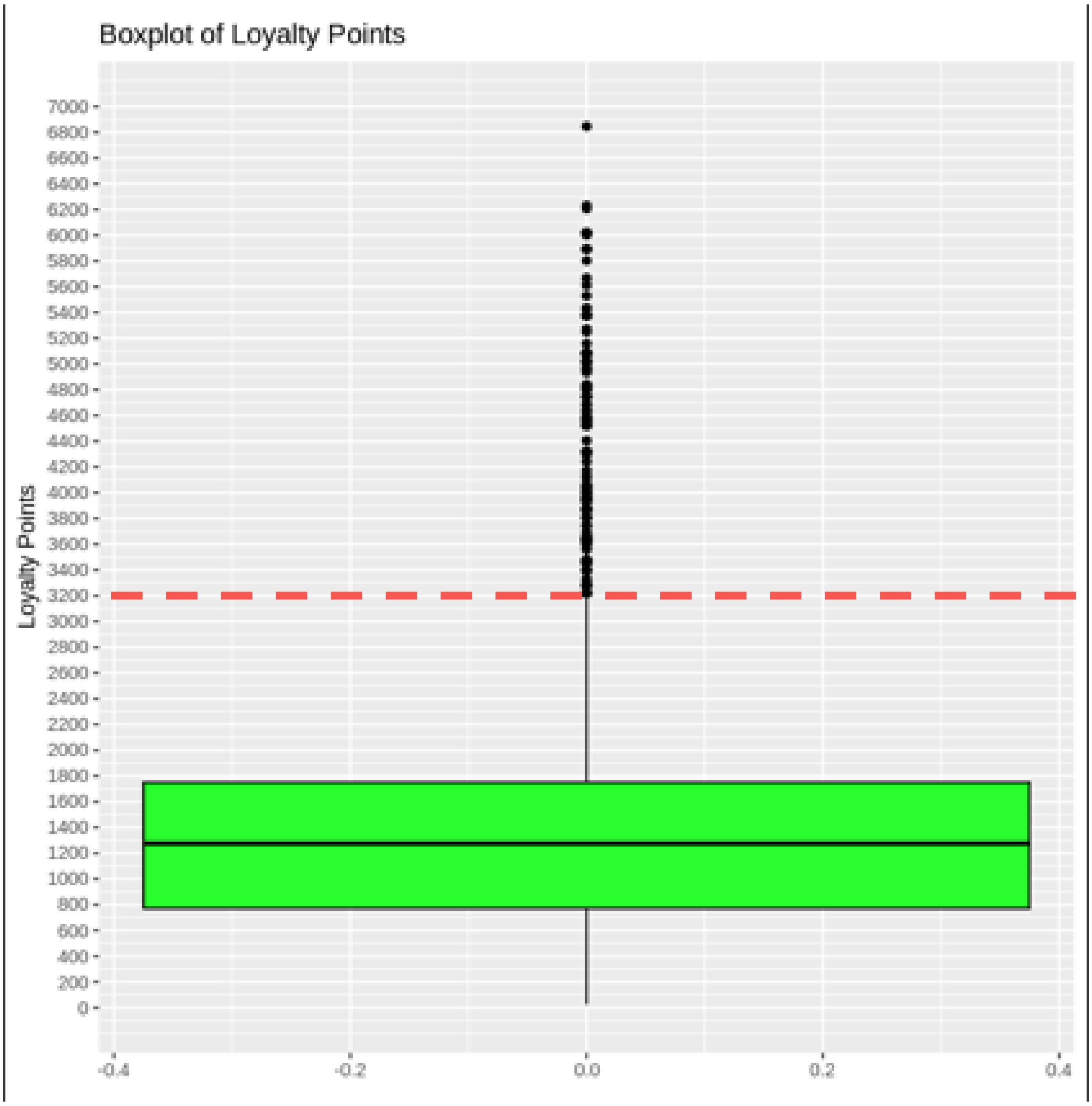
- The age of customers ranges from 17 to 72, with a mode of 21 and a mean of 39.5.
- The remuneration of customers ranges from $12.3k to $112.34k, with a mode of $63.14k and a mean of $48.08k.
- The spending score of customers ranges from 1 to 100, with a mode of 50 and a mean of 50.
- The loyalty points of customers range from 25 to 6847, with a mode of 1276 and a mean of 1578.
- There are more female customers than male customers.
- Most customers are graduates, followed by PhDs, then Postgraduates, then Diploma, and lastly, Basic.

UNIVARIATE ANALYSIS

- **Loyalty Points:** The distribution is right-skewed, with most customers having lower loyalty points and a smaller segment possessing significantly higher points (outliers above 3200 points).



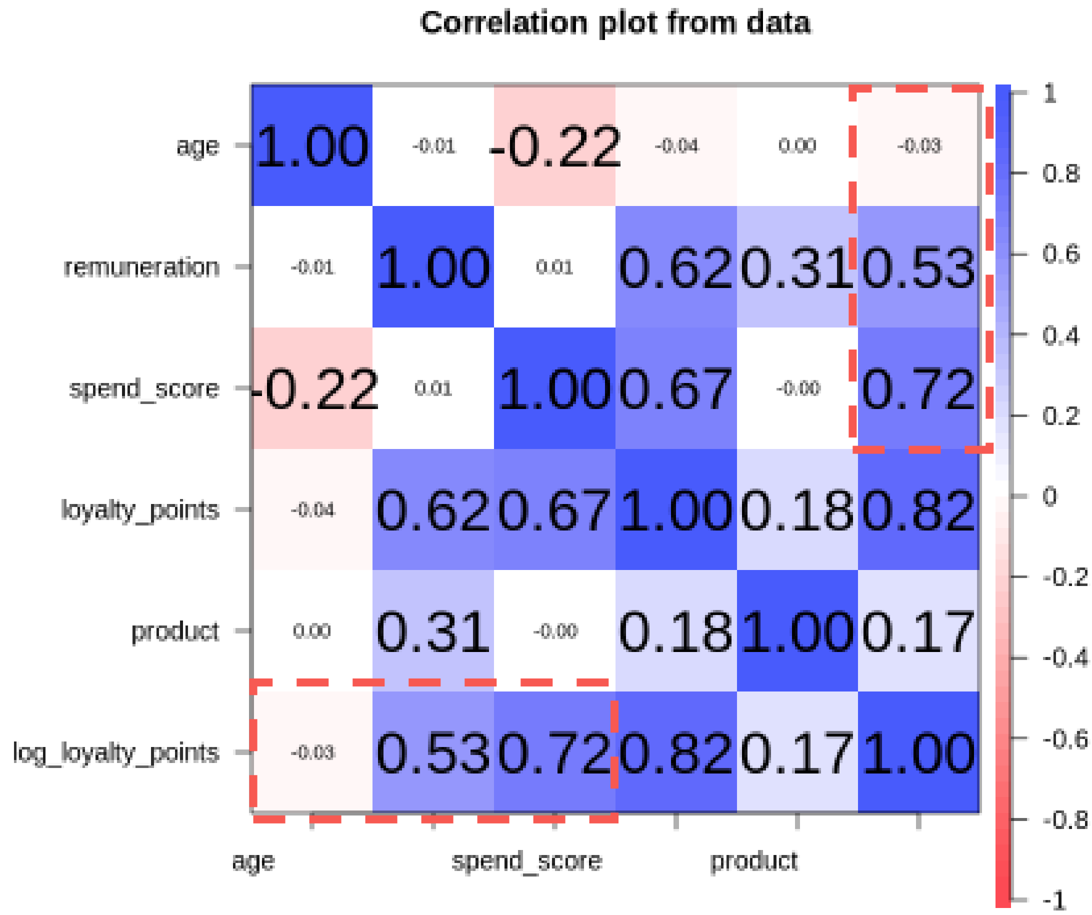
Distribution of Loyalty Points
Density
Loyalty Points
6e-04
4e-04
2e-04
0e+00



- **Spending Score, Remuneration, and Age:** These variables display relatively normal distributions.

BIVARIATE ANALYSIS

- **Loyalty Points vs. Spending Score:** Strong positive correlation.
- **Loyalty Points vs. Remuneration**: Moderate positive correlation.
- **Loyalty Points vs. Age:** Weak negative correlation.

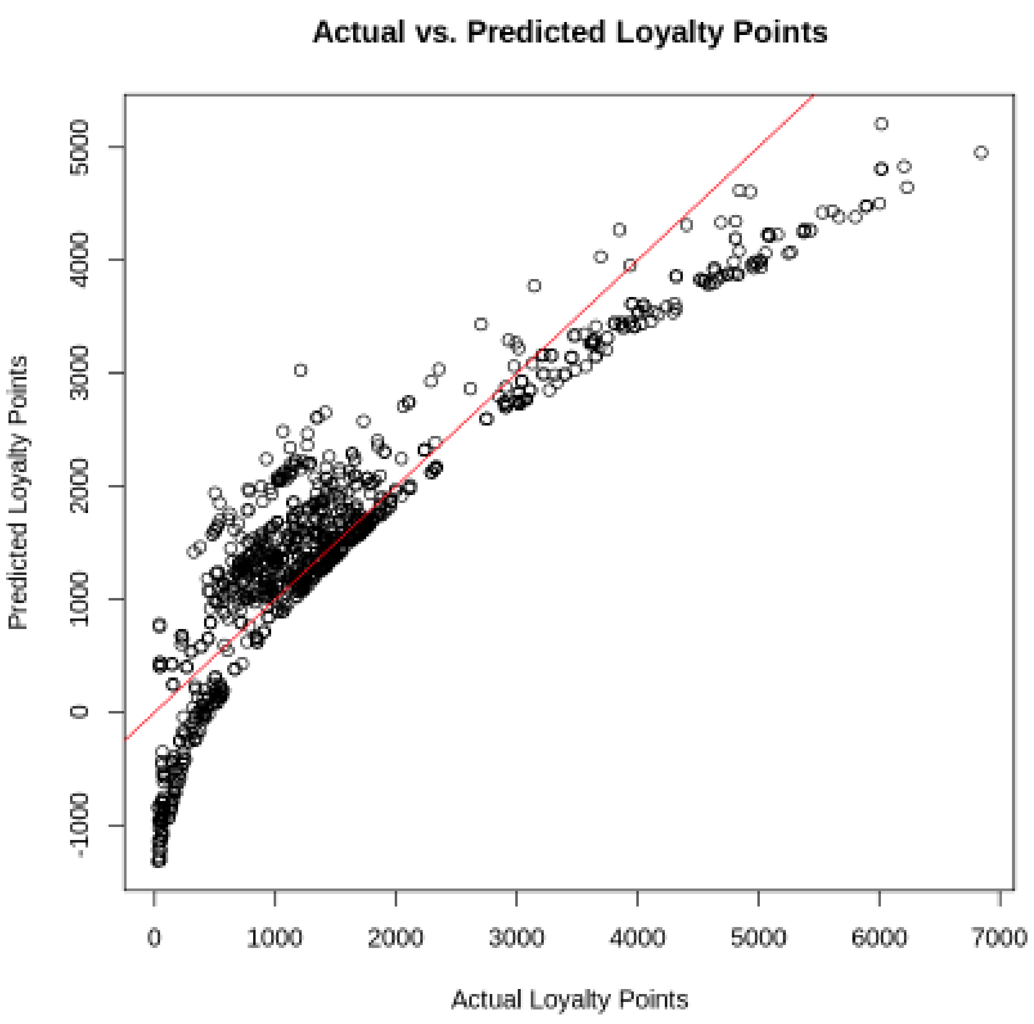

Correlation plot from data
age
remuneration
spend_score
loyalty_points
product
log_loyalty_points
1.00
-0.01
-0.22
-0.04
0.00
-0.03
0.01
0.62
0.31
0.53
0.67
-0.00
0.72
0.18
0.82
0.17
age
spend_score
product
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

STATISTICAL MODELING

MULTIPLE LINEAR REGRESSION (MLR)

- Several MLR models were constructed to predict loyalty points using different combinations of predictor variables (age, remuneration, spending score).
- The model (b) including age, remuneration, and spending score as predictors was selected as the most suitable.
- This model's adjusted R-squared value and p-values supported a moderately high prediction accuracy and relevance of these predictors.

```
Call:
lm(formula = loyalty_points ~ age + remuneration + spend_score,
    data = tr_nums)

Residuals:
     Min       1Q   Median       3Q      Max
-1819.11  -350.84     4.61   291.00  1894.62

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)  -2203.0598    52.3609  -42.08   <2e-16 ***
age             11.0607     0.8688   12.73   <2e-16 ***
remuneration    34.0084     0.4970   68.43   <2e-16 ***
spend_score     34.1832     0.4519   75.64   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 513.8 on 1996 degrees of freedom
Multiple R-squared:  0.8399,    Adjusted R-squared:  0.8397
F-statistic:  3491 on 3 and 1996 DF,  p-value: < 2.2e-16
```

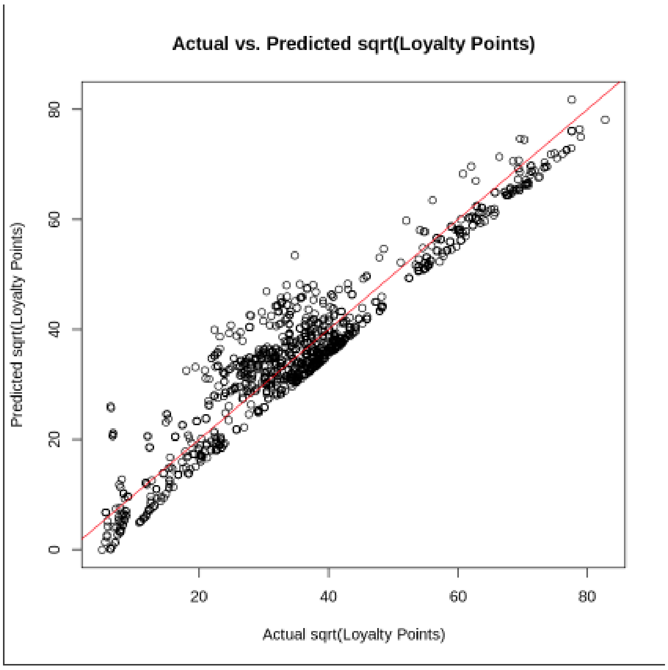


- Transformation using square-root on the target variable (loyalty points) improved the prediction accuracy further.

```
Call:
lm(formula = sqrt_loyalty_points ~ age + remuneration + spend_score,
    data = tr)

Residuals:
    Min      1Q  Median      3Q     Max
-19.696  -2.748   1.800   3.660   6.194

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)  -11.314014   0.480640  -23.54   <2e-16 ***
age            0.157902   0.007975   19.80   <2e-16 ***
remuneration   0.403941   0.004562   88.54   <2e-16 ***
spend_score    0.445083   0.004148  107.29   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.717 on 1996 degrees of freedom
Multiple R-squared:  0.907,     Adjusted R-squared:  0.9068
F-statistic:  6486 on 3 and 1996 DF,  p-value: < 2.2e-16
```

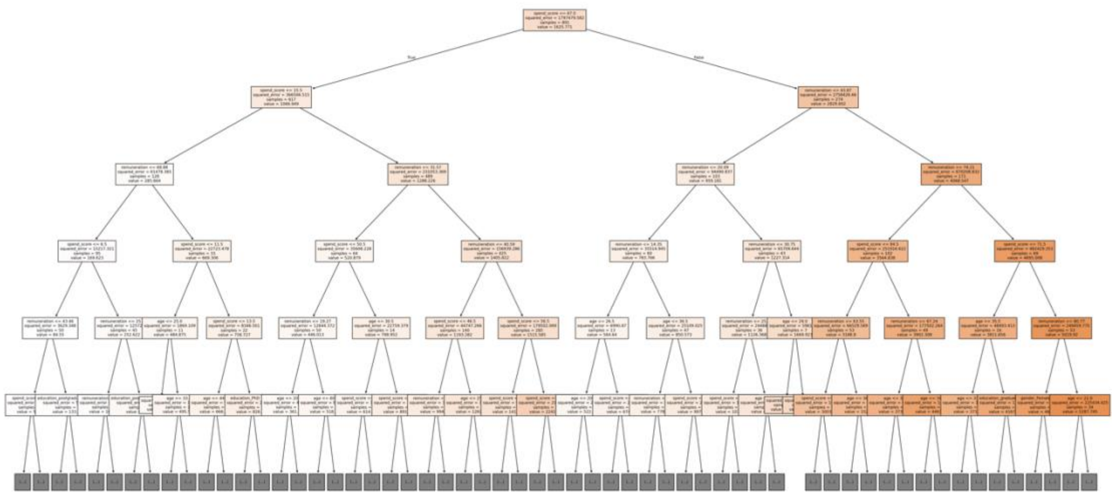
```
# Example prediction for new data
new_data <- data.frame(remuneration = 31.98, spend_score = 65, age = 26)
new_predicted_loyalty_sqrt <- predict(mlr_sqrt, newdata = new_data)
new_predicted_loyalty <- new_predicted_loyalty_sqrt^2
cat("\nPredicted Loyalty Points for new data:", new_predicted_loyalty)
```

```
Predicted Loyalty Points for new data: 1199.92
```



DECISION TREES

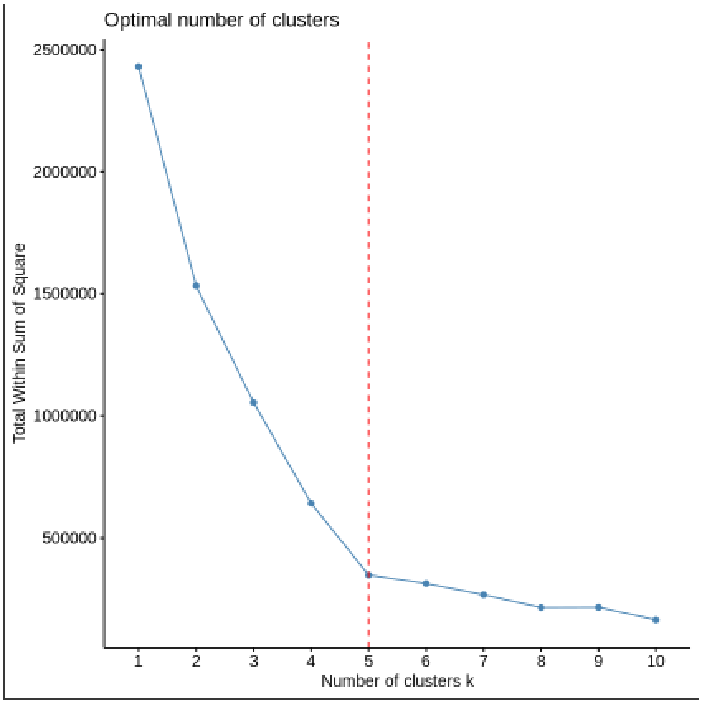
- A Decision Tree Regressor was employed to understand the structure and patterns in customer behaviour.
- The tree was visualized, and key decision rules were extracted to identify customer segments and understand the factors influencing loyalty points.
- Random Forest Regressor was used to compare accuracy and predictive capability.
- The decision tree revealed that spending score is the most important predictor of loyalty points, followed by remuneration and age.
- Customers with high spending scores and high remuneration tend to have the highest loyalty points.
- Random Forest Regressor model showed a slightly higher R-squared value on both the test and train data, indicating that it explains a slightly greater proportion of the variance in the target variable (loyalty points) compared to a single decision tree regressor.



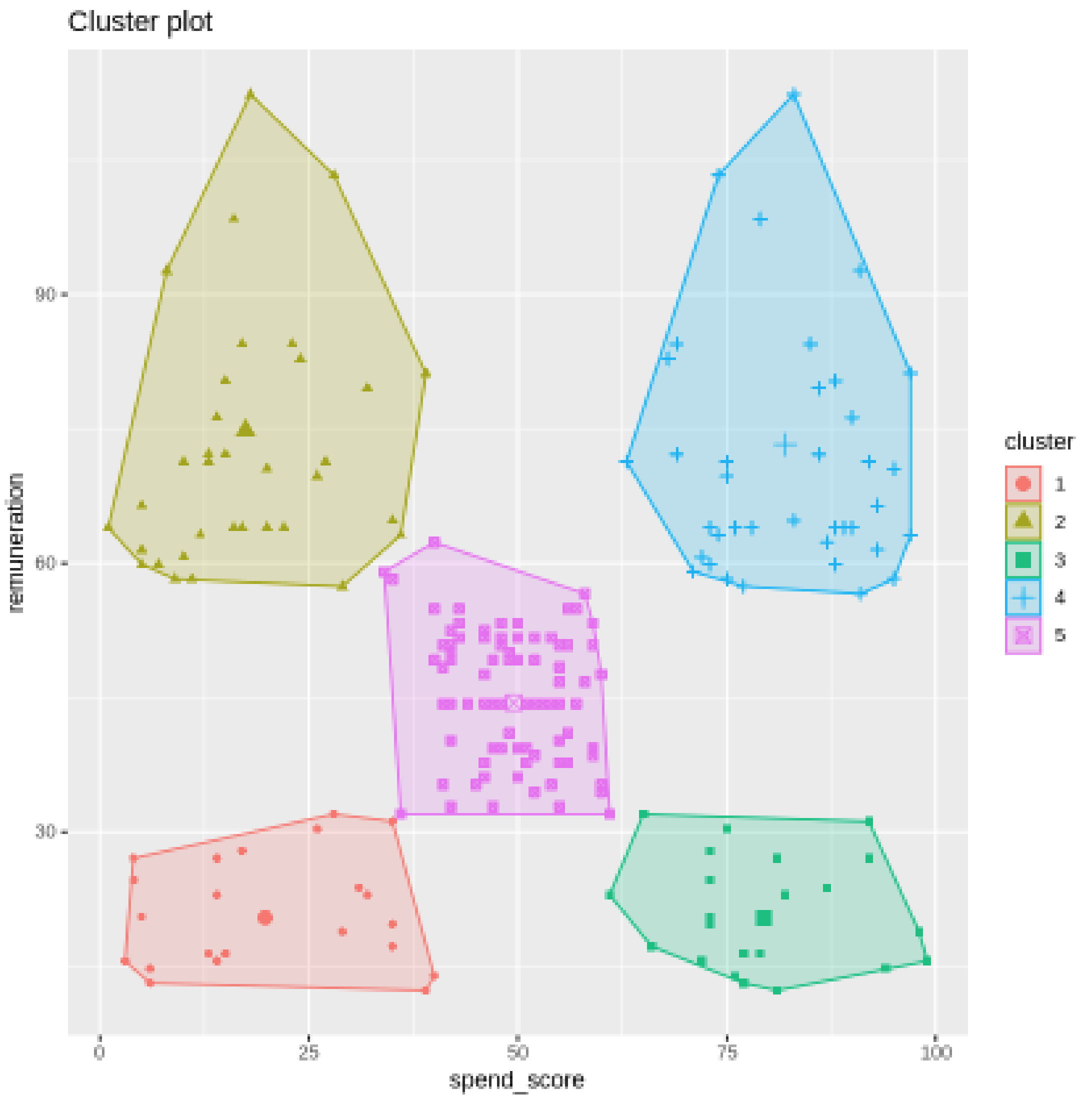
CUSTOMER SEGMENTATION

K-MEANS CLUSTERING

- K-means clustering was applied to segment customers based on spending score and remuneration.
- Five clusters were identified as optimal using the elbow method and silhouette analysis.



- Five distinct customer segments were identified based on spending score and remuneration, each with unique characteristics and loyalty point accumulation patterns.



cluster	size	mean_spend_score	mean_remuneration	mean_age	mean_loyalty_points
<int>	<int>	<dbl>	<dbl>	<dbl>	<dbl>
1	271	19.76384	20.42435	43.50554	275.0590
2	330	17.42424	74.83121	40.66667	911.7606
3	269	79.41636	20.35368	31.60223	971.9442
4	356	82.00843	73.24028	35.59270	3988.2388
5	774	49.52972	44.41879	42.12920	1420.3824

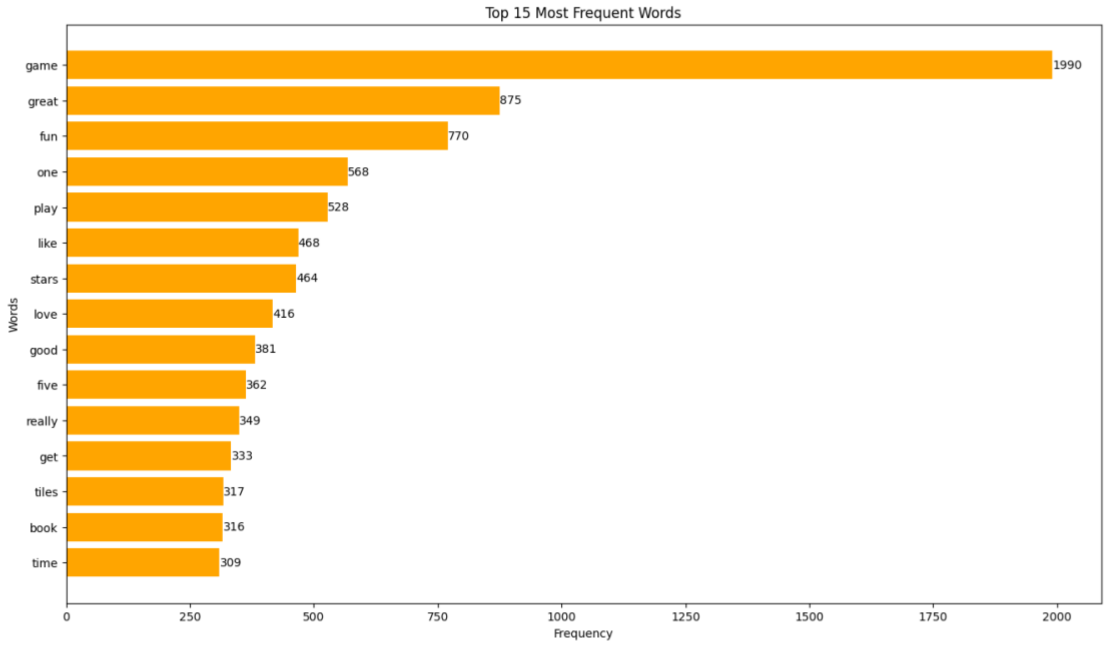
- Cluster Characteristics:

Cluster	Size	Spending Score	Remuneration	Age	Loyalty Points	Classification
1	14%	Low	Low	Middle-aged	Low	Price-Conscious
2	16%	Low	High	Middle-aged	Moderate	Affluent but Price Sensitive
3	13%	High	Low	Young	Moderate	Young Value Seekers
4	18%	High	High	Young	High	High Value Loyalists
5	39%	Moderate	Moderate	Middle-aged	Moderate	Average Consumer

NATURAL LANGUAGE PROCESSING (NLP)

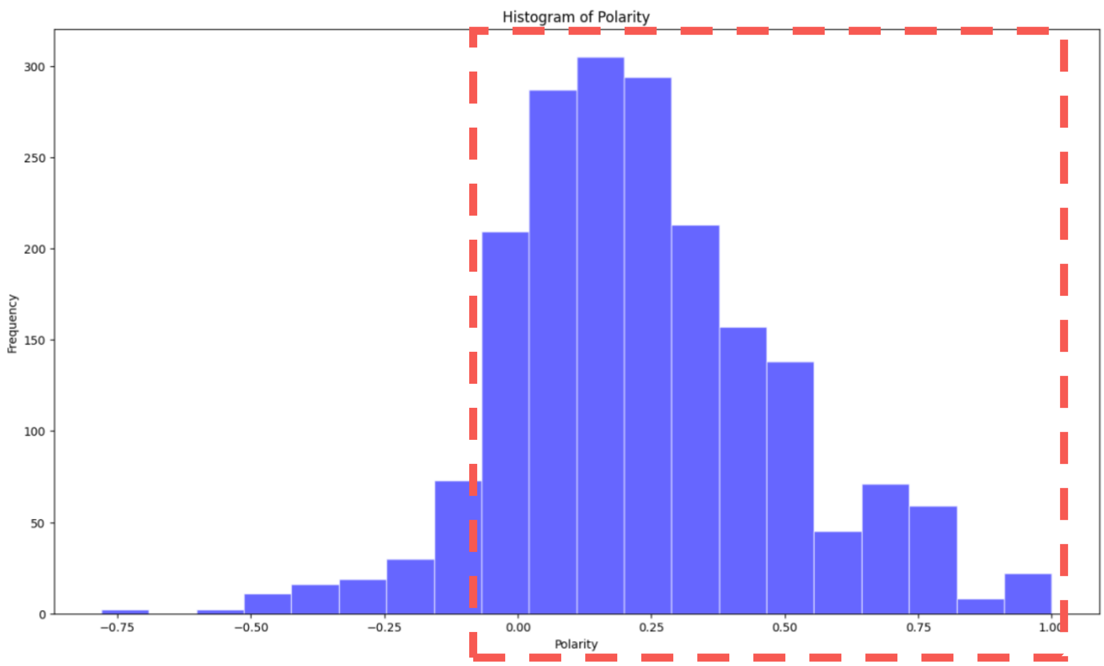
- NLP techniques were used to analyse customer reviews.
- Word frequency analysis was performed to identify the most common words used by customers.

- The most frequent words in customer reviews included "game," "fun," "great," "good," and "play."



- Sentiment analysis was used to identify positive and negative reviews, providing insights into customer satisfaction and areas for improvement.

- Sentiment analysis revealed a generally positive sentiment towards Turtle Games products.



OBSERVATIONS AND INSIGHTS

- The loyalty program effectively rewards customers with higher spending scores and, to a lesser extent, higher remuneration.
- Age is not a primary driver of loyalty point accumulation.
- Five distinct customer segments were identified, each with unique spending and loyalty characteristics.
- Targeted marketing strategies should be developed for each cluster to maximize engagement and spending.

RECOMMENDATIONS

- **Targeted Marketing:** Implement campaigns tailored to each cluster's preferences and needs.
- **Personalized Rewards:** Introduce personalized rewards based on individual customer behaviour and spending habits.
- **Loyalty Program Enhancements:** Consider tiered rewards and exclusive benefits for high-loyalty customers.
- **Customer Relationship Management:** Focus on building stronger relationships with all segments to enhance loyalty.
- **Further Analysis:** Investigate high-loyalty customers (outliers) to understand their behaviour and identify potential opportunities for program optimization.

SPECIFIC RECOMMENDATIONS BY CLUSTER

- **Cluster 1 (Price-Conscious Consumers):** Focus on value, highlight the value proposition, and introduce a gamified loyalty program.
- **Cluster 2 (Affluent but Price-Sensitive):** Offer premium offerings, exclusive rewards, and targeted communication.

- **Cluster 3 (Young Value Seekers):** Focus on trendy products, implement an engaging loyalty program, and utilize digital marketing.
- **Cluster 4 (High-Value Loyalists):** Offer personalized experiences, premium rewards, and focus on relationship building.
- **Cluster 5 (Average Consumers):** Develop targeted campaigns, offer personalized recommendations, and encourage loyalty program engagement.

CONCLUSION

This analysis provides valuable insights into customer behaviour and the effectiveness of Turtle Games' loyalty program. By implementing the recommendations outlined, Turtle Games can enhance customer engagement, loyalty, and ultimately drive business growth. Continuous monitoring and optimization are crucial for long-term success.