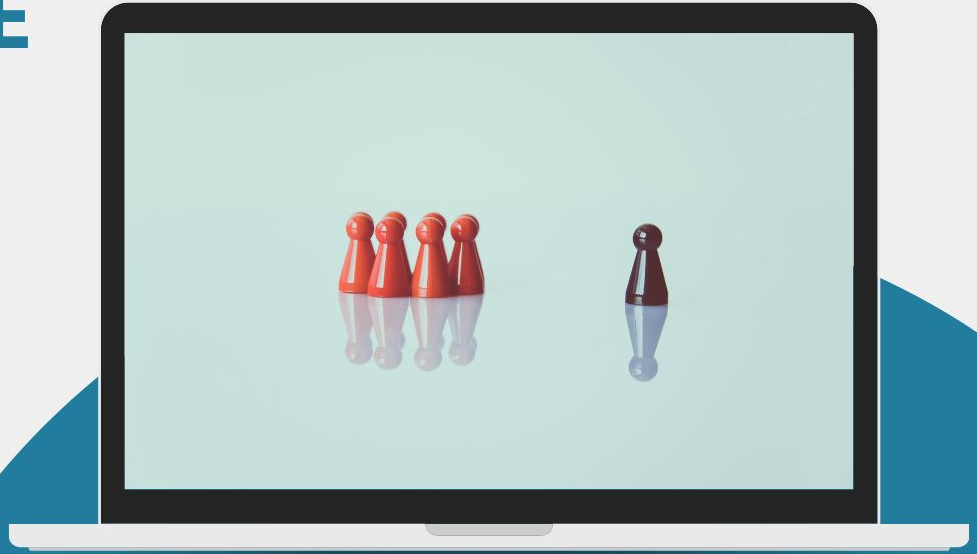
WHAT DO TURTLE GAMES CUSTOMERS NEED?

Pescod_Henry_DA301
Assignment Presentation

September 2024



"WE'RE NOT COMPETITOR OBSESSED, WE'RE CUSTOMER OBSESSED.

WE START WITH WHAT THE CUSTOMER NEEDS

AND WE WORK BACKWARDS"

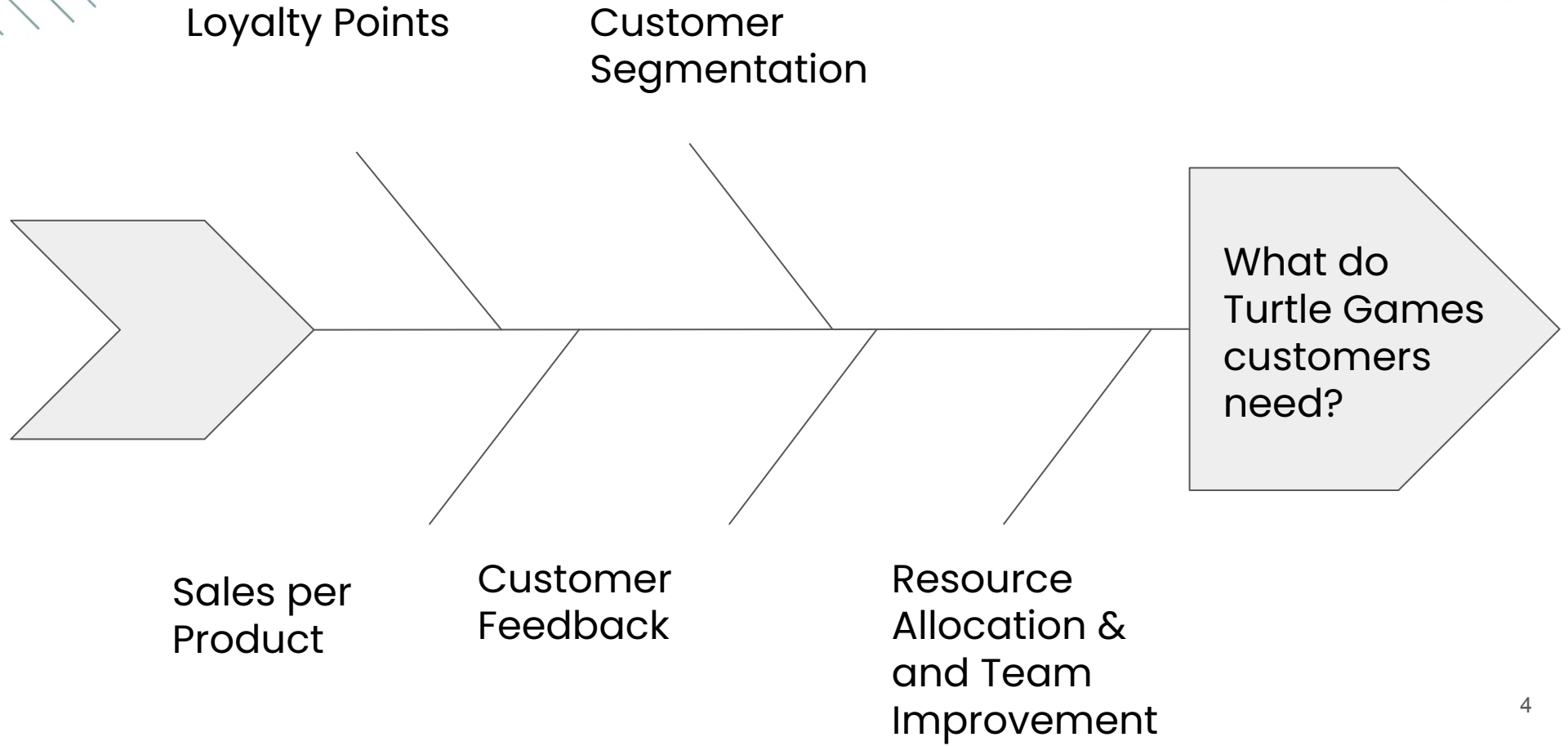
Jeff Bezos, CEO Amazon

Here's Lina, a new Turtle Games customer

What do we know about Lina
Age = 36
Annual income = 74,000
Spending score = 82

What do we want to know?
1) How will she behave?
2) How many loyalty points is she likely to accrue?
3) what does she need?

Breaking down the problem

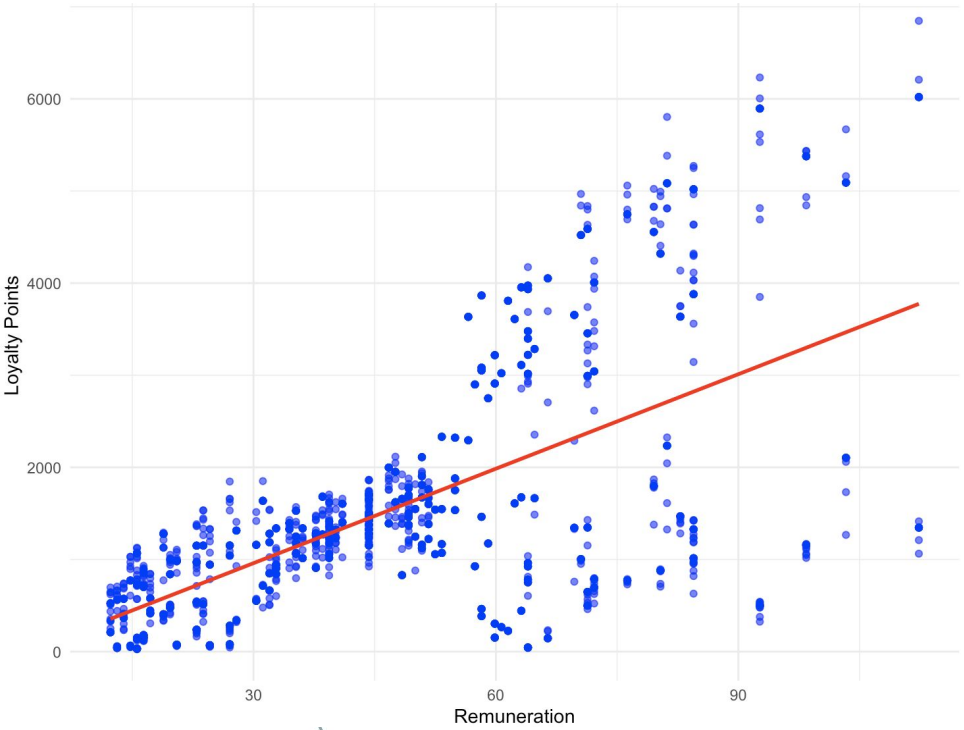


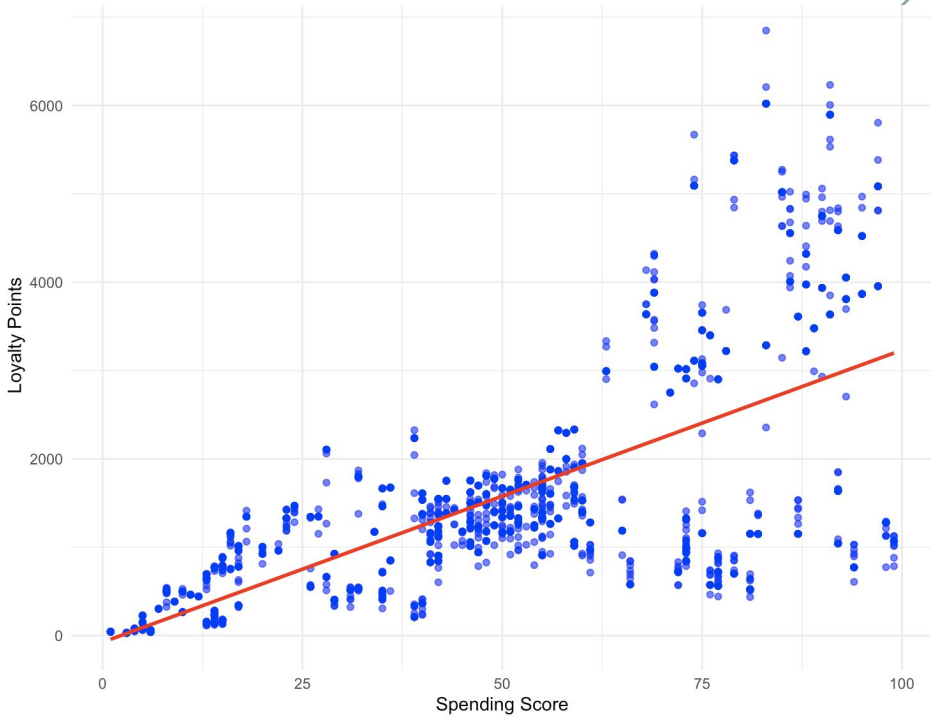
Analysis approach

Question	Problem Component	Analysis approach	Why is this a problem?
1. What drives Loyalty points of Turtle Games customers?	Loyalty Points	Linear Regression Correlation Plot Decision tree Regression Multiple linear Regression	Customers may not be accumulating points in a way that drives engagement and repeat sales.
2. How does this help to understand Turtle Games customers?	Customer Segmentation	K means clustering EDA by grouping of clusters	Lack of clear understanding of how different customer segments behave.
	Customer Feedback	NLP and sentiment analysis by clusters	Customer reviews may not be adequately informing team practices.
	Sales per Product	Initial EDA using R	Inefficient strategies may be leading to lower sales for certain products.
3. How does this help Turtle Games?	Resource Allocation & Team Improvement	Insights from analysis	Uncertainty about where to allocate resources for maximum impact on loyalty and sales.

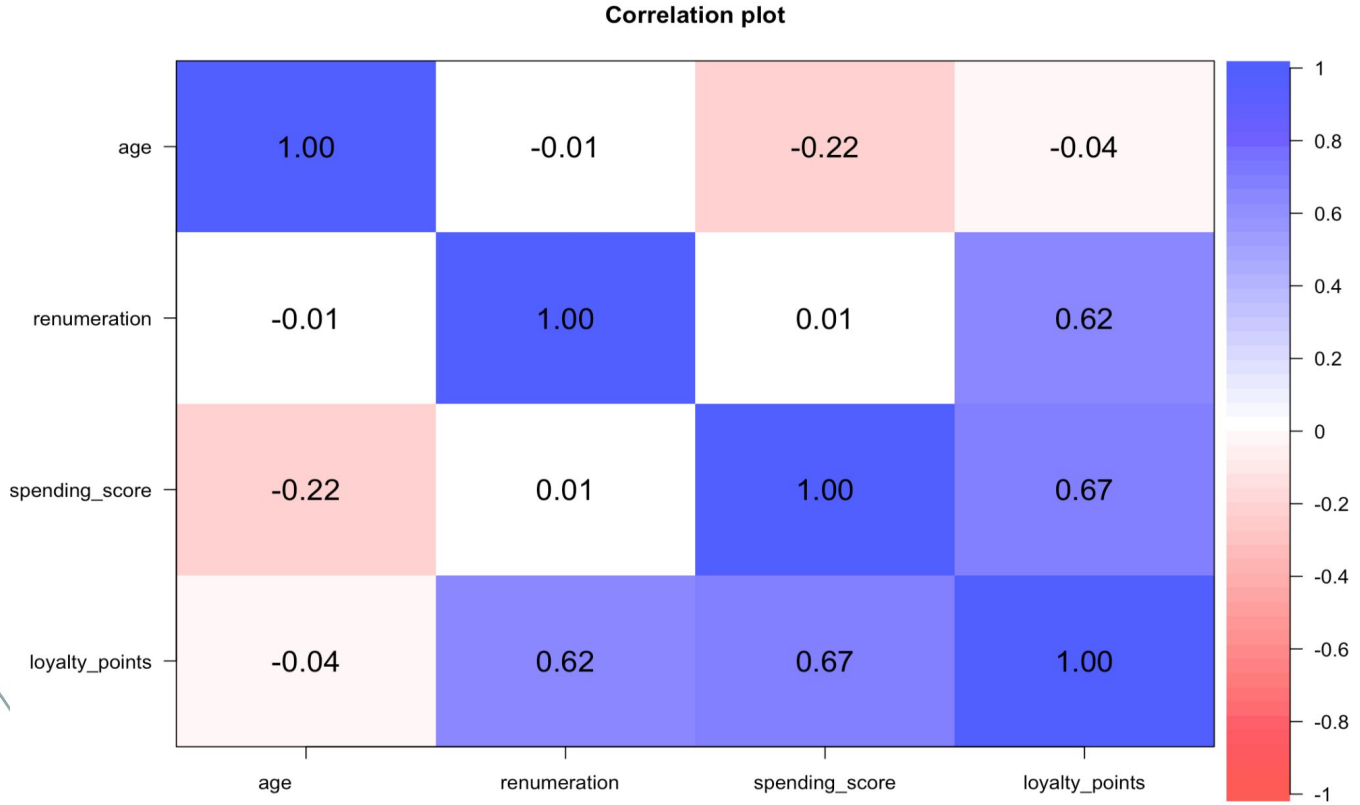
1) WHAT DRIVES LOYALTY POINTS OF TURTLE GAMES CUSTOMERS?

1) Linear regression analysis of loyalty points, Remuneration and spending score





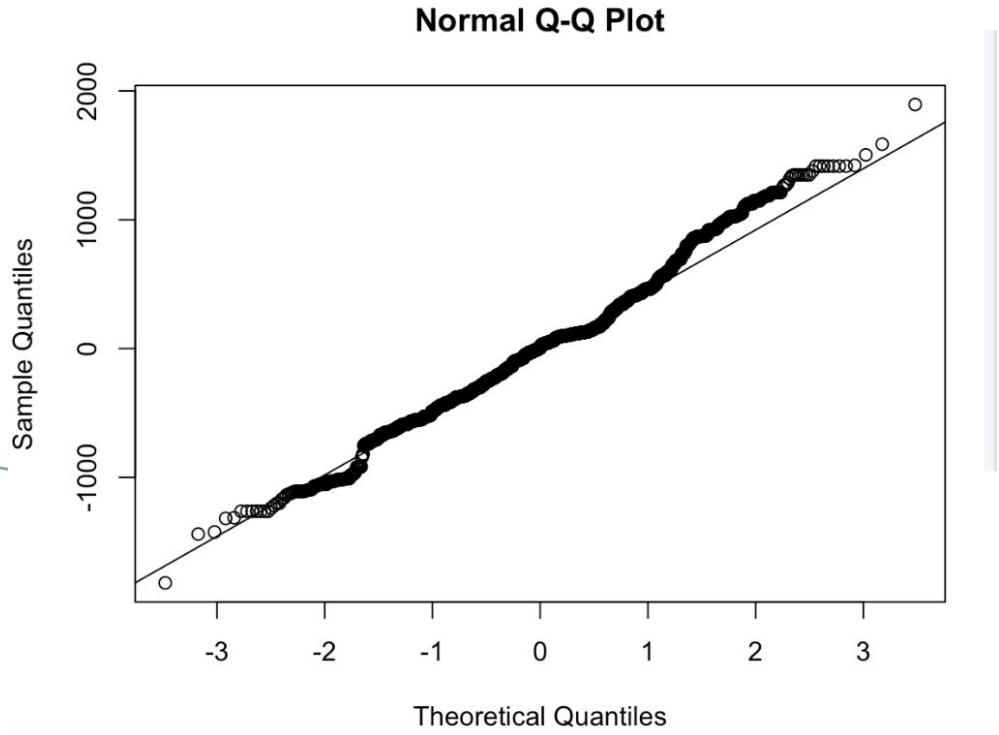
2) Correlation plot of loyalty points, age Remuneration and spending score



3) Decision Tree Regression Analysis of loyalty points, based on Remuneration and spending

Model	Mean Squared Error	Root Mean Squared Error	Mean Absolute Error	R-squared
Unpruned Decision Tree	5,442.09	73.77	23.96	0.9969
Pruned Decision Tree (Depth 3)	146,819.57	383.17	279.90	0.9170
Pruned Decision Tree (Depth 5)	75,027.12	273.91	182.27	0.9576
Pruned Decision Tree (Depth 7)	33,921.68	184.18	117.08	0.9808
Pruned Decision Tree (Depth 9)	20,983.43	144.86	86.64	0.9881

4) Multiple linear regression based on loyalty points, Remuneration and spending score

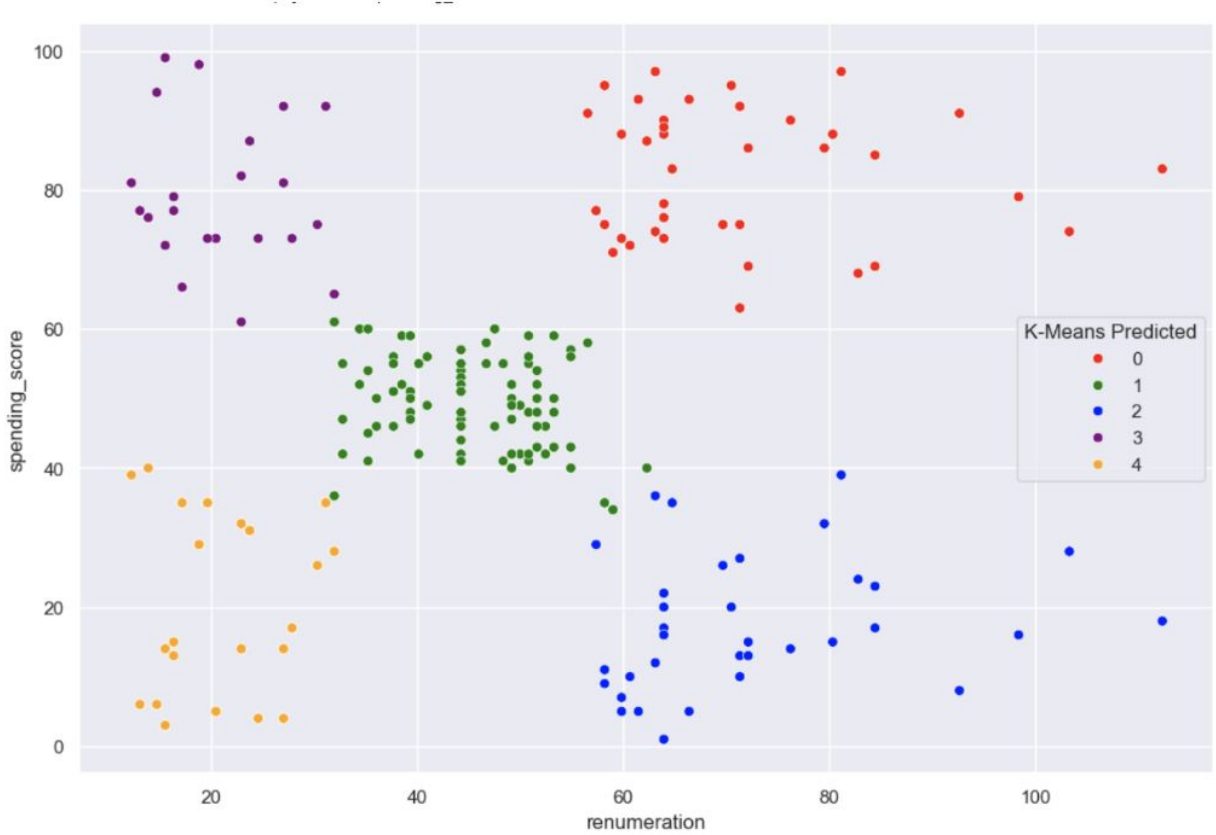



```
Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    -2203.0598    52.3609  -42.08   <2e-16 ***
renumeration      34.0084     0.4970   68.43   <2e-16 ***
spending_score    34.1832     0.4519   75.64   <2e-16 ***
age               11.0607     0.8688   12.73   <2e-16 ***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

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
```

2) HOW DOES THIS HELP TO UNDERSTAND TURTLE GAMES CUSTOMERS?

1) K Means clustering analysis



› Cluster 0 - **“high spenders”**

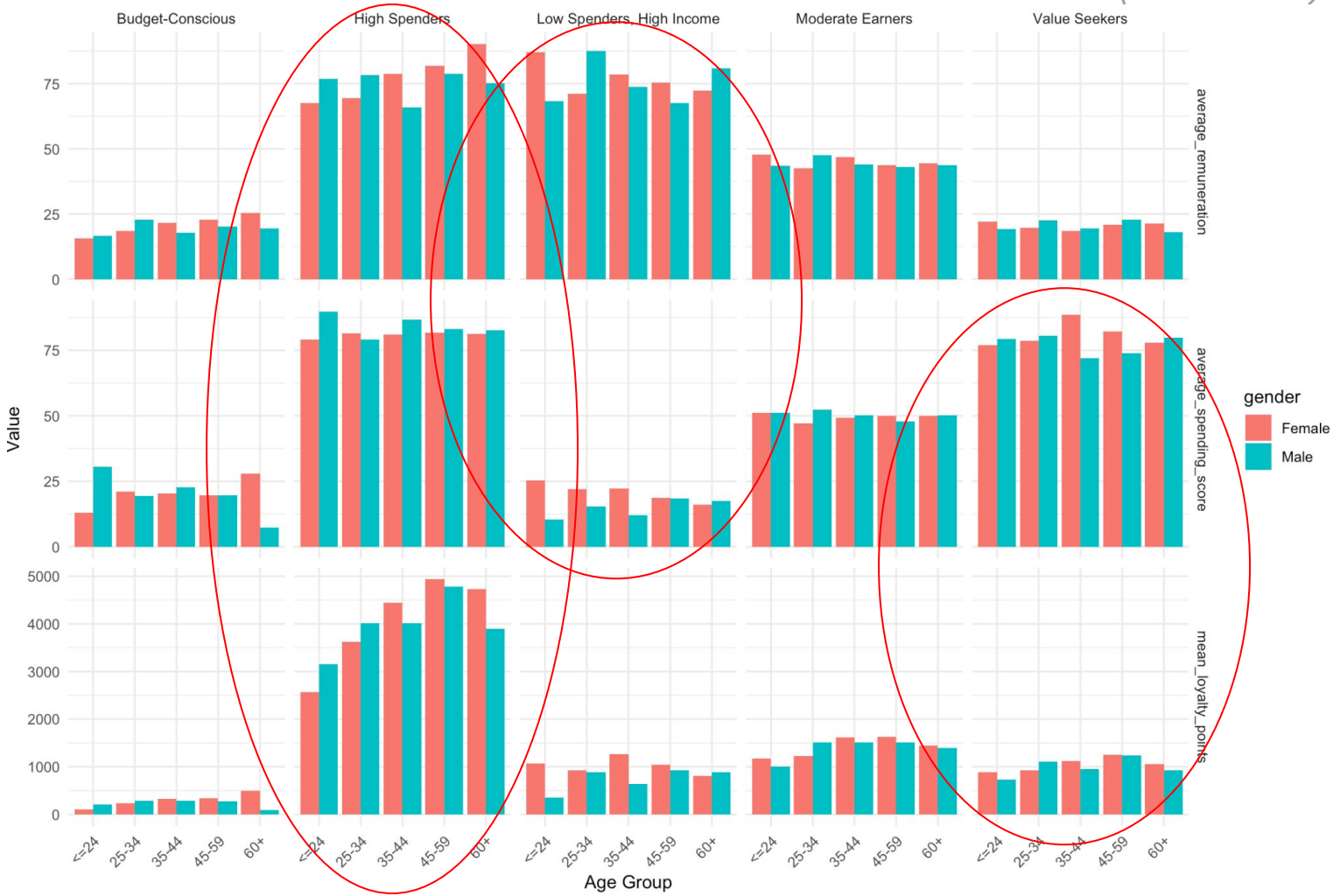
› Cluster 1 - **“moderate earners”**

› Cluster 2 - **“low spenders, high income”**

› Cluster 3 - **“value seekers”**

› Cluster 4 - **“budget conscious”**

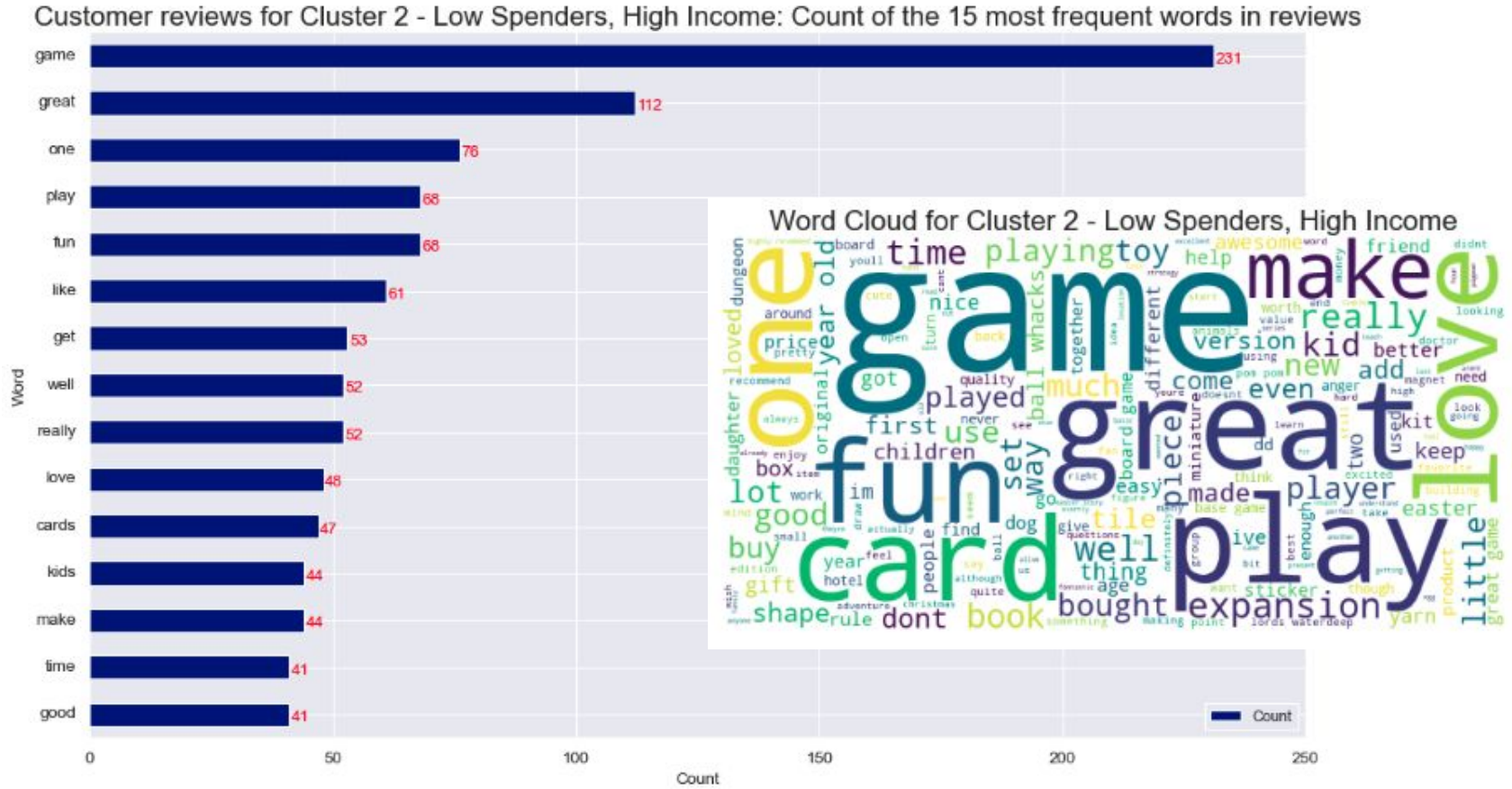
2) Mean Loyalty points, Spending Score and Remuneration by Age Group, Gender and Segment



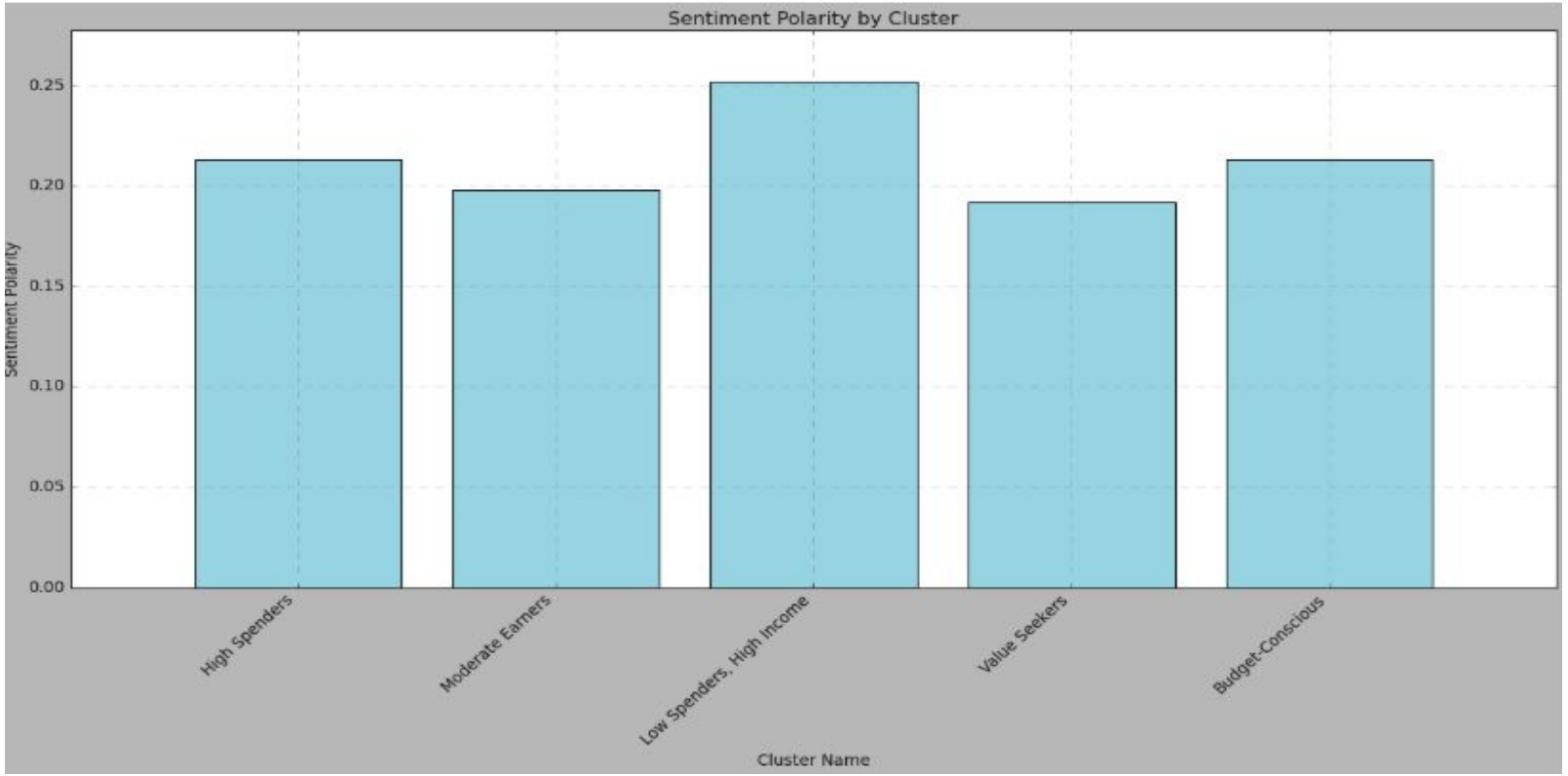
Insights about customer behaviour based on exploratory data analysis

Customer cluster	Insight
"High Spenders	Highly valuable customers
"Moderate earners"	largest portion of the sample data set. Consider certain products that appeal specifically to their profile
"Low Spenders, High Earners"	Suggests potential for growth - encourage spending among this segment.
"Value Seekers"	seem to seek value
"Budget Conscious"	might be more price sensitive?

Customer reviews analysis - Low spend, High income



Customer review sentiment polarity



Insights about customer behaviour based on reviews

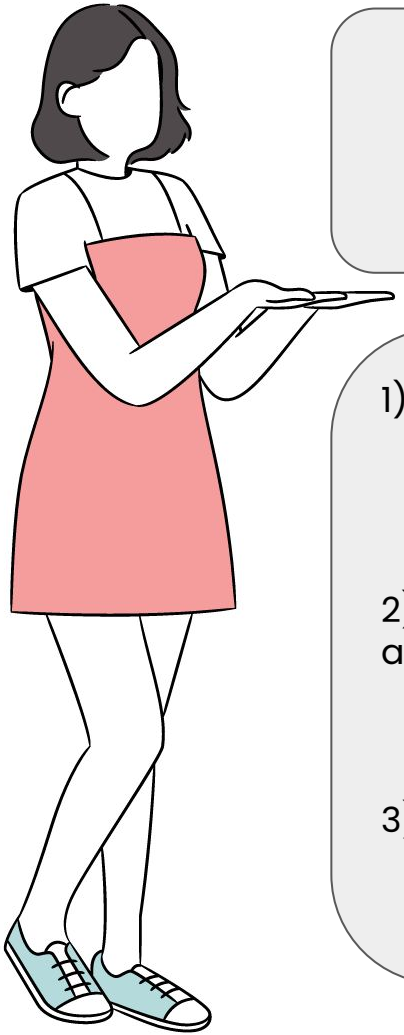
Customer cluster	Insight based on reviews
"High Spenders	Possibly family-oriented customers, given the words "kids" and "love."
"Moderate earners"	Possibly board game enthusiasts who enjoy traditional elements based on frequent words like "tiles," "cards," and "boards"
"Low Spenders, High Earners"	The words "make," "kids," and "well" could suggest a focus on games that are well-designed and engaging for children.
"Value Seekers"	Mention of "book" and "tiles" might indicate that this cluster prefers more strategic or intellectually engaging games.
"Budget Conscious"	Presence of words like "kids," "time," and "set" might suggest family-oriented customers who value easy to set up games that are fun for children

3) HOW DOES THIS HELP TURTLE GAMES?

Insights for Team Improvement and Resource Allocation

1) Target marketing strategy to match customer profiling
    - E.g. Do the “high spenders” require more targeted marketing and offers to reward their loyalty?
2) Target Product Strategy to match customer behaviour
    - E.g The “low spenders, high income” have potential for growth if can be encouraged to spend more with the right product
3) Use multiple linear regression to predict future loyalty points
    - Can funnel new customers into marketing categories

Returning to Lina, what can we now say about her?



What do we know about her
Age = 36
Annual income = 74,000
Spending score = 82

1) How will she behave?
- high spender - most valuable segment of customer to turtle games
- Possibly family-oriented

2) How many loyalty points is she likely to accrue?
- Based on our model approx 3500 loyalty points

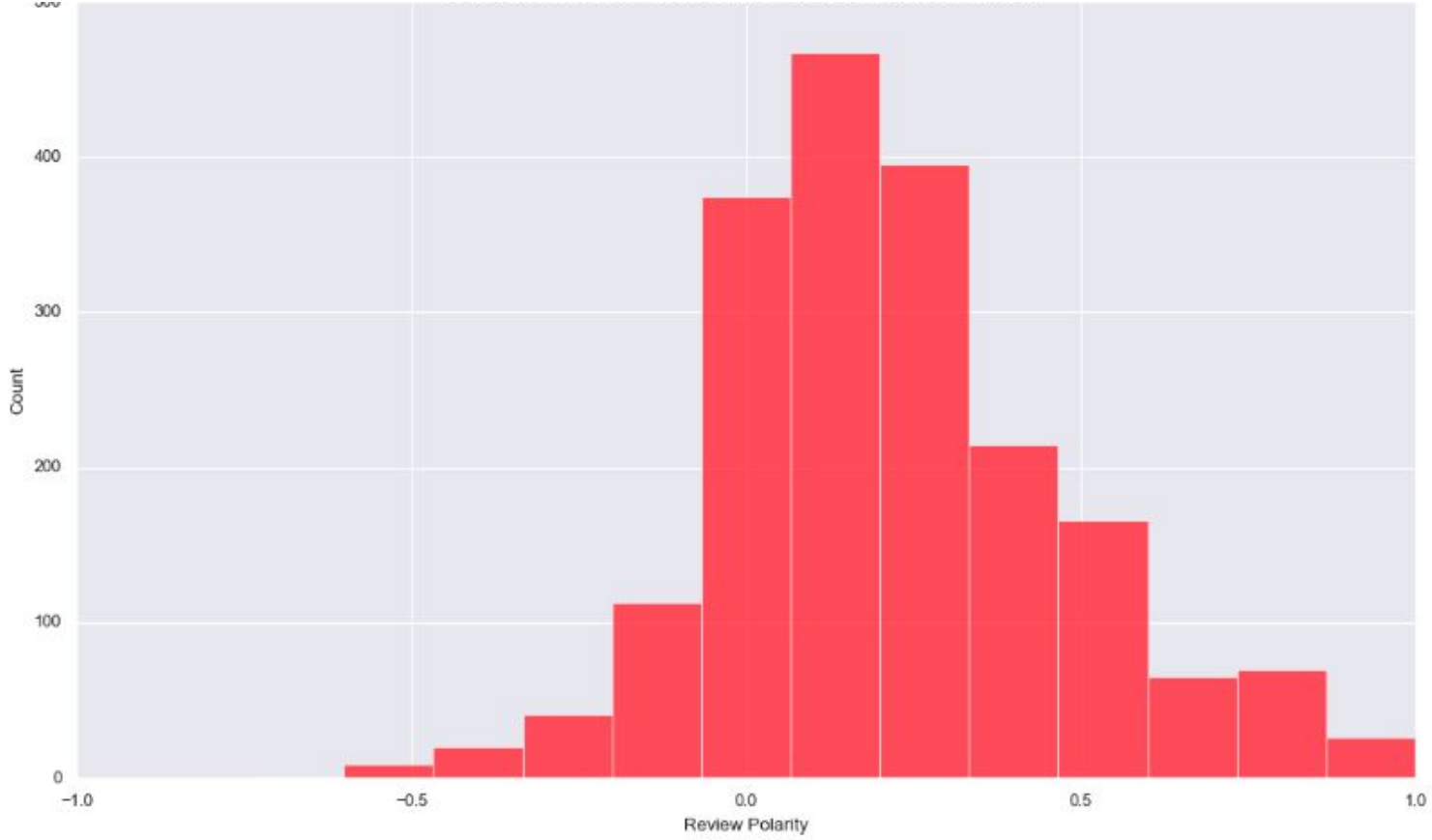
3) what does she need?
- exclusive rewards or incentives to ensure we retain Lina
- Family orientated product range

Recommendations for further exploration

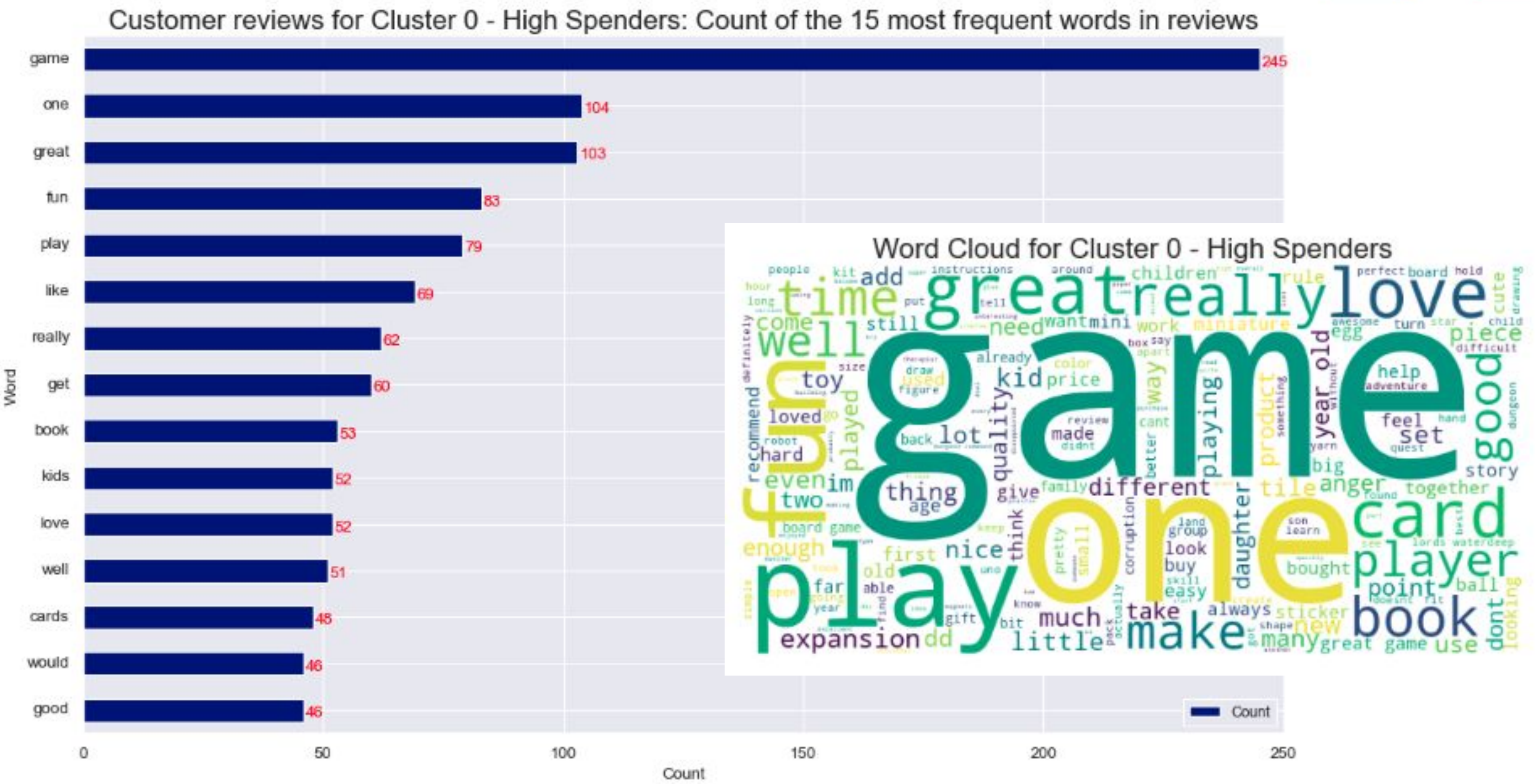
i) Integrate more detailed product data
ii) Further model training on larger datasets
iii) Model validation using new customer data

THANK YOU

Customer review sentiment polarity



Customer reviews analysis - High Spenders



CONCLUSIONS

1) LOYALTY PROGRAMME EFFECTIVENESS

3) CUSTOMER FEEDBACK AND TEAM IMPROVEMENT

2) UNDERSTANDING CUSTOMER BEHAVIOUR

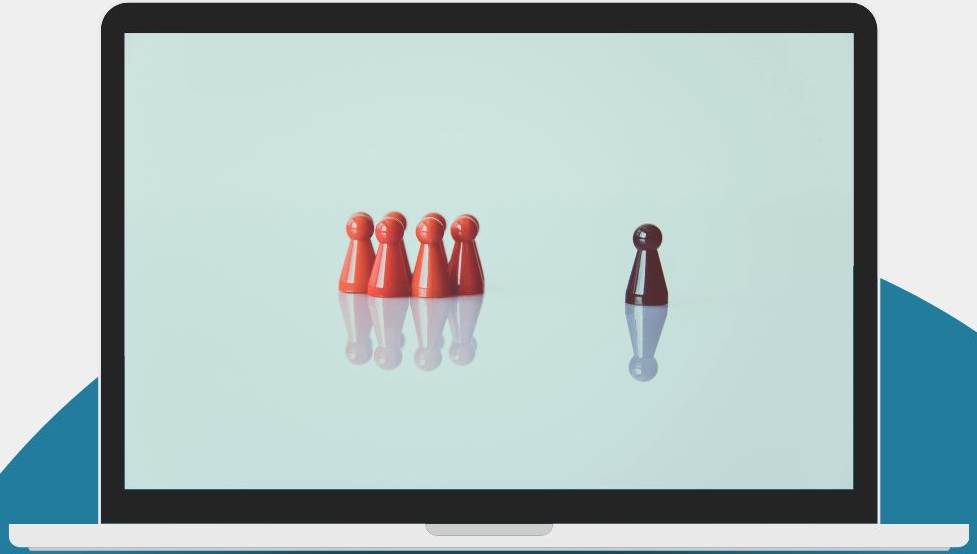
Distribution of total loyalty points 1/2

TURTLE GAMES

CUSTOMER ANALYSIS

Pescod_Henry_DA301
Assignment Presentation

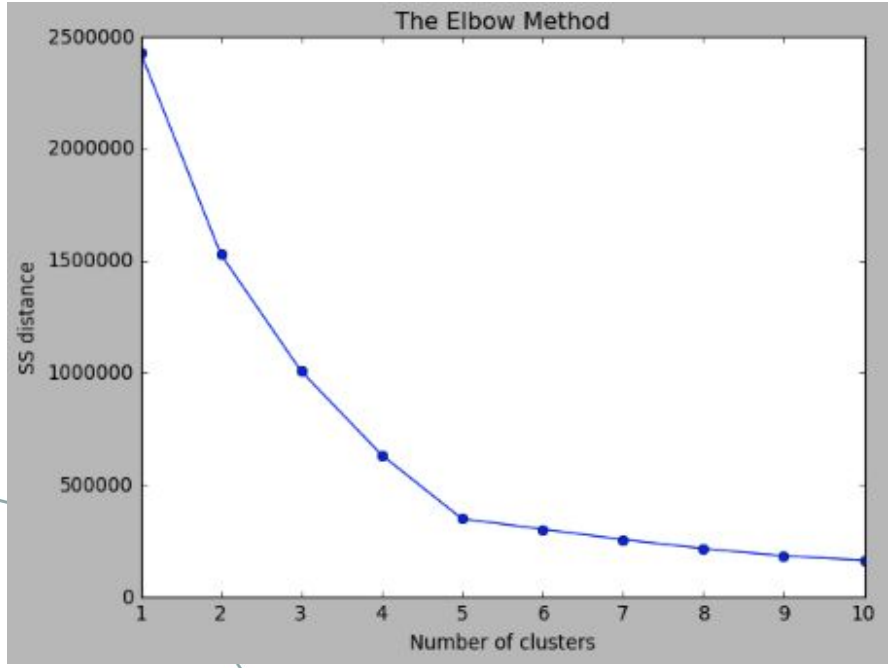
September 2024

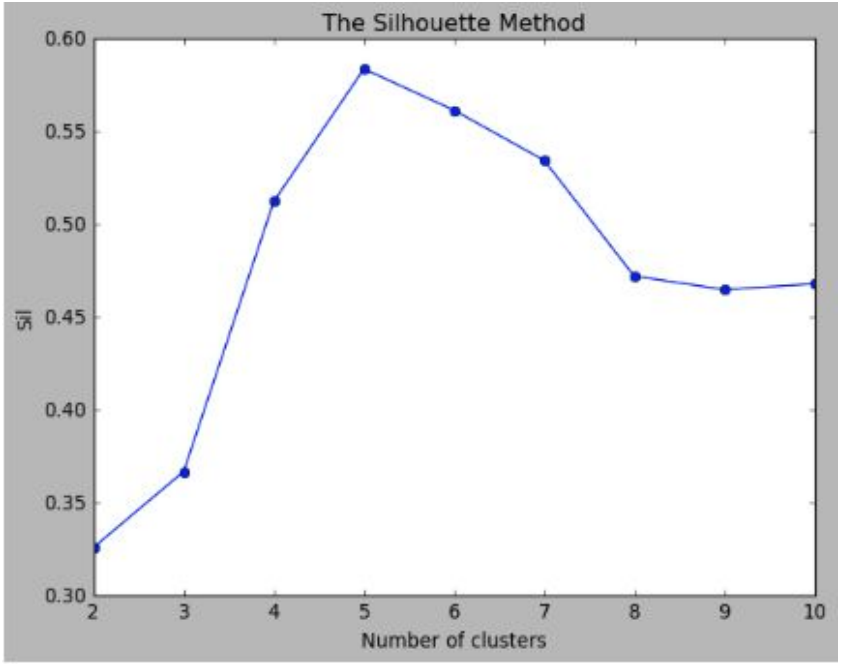


Insights about loyalty programme effectiveness

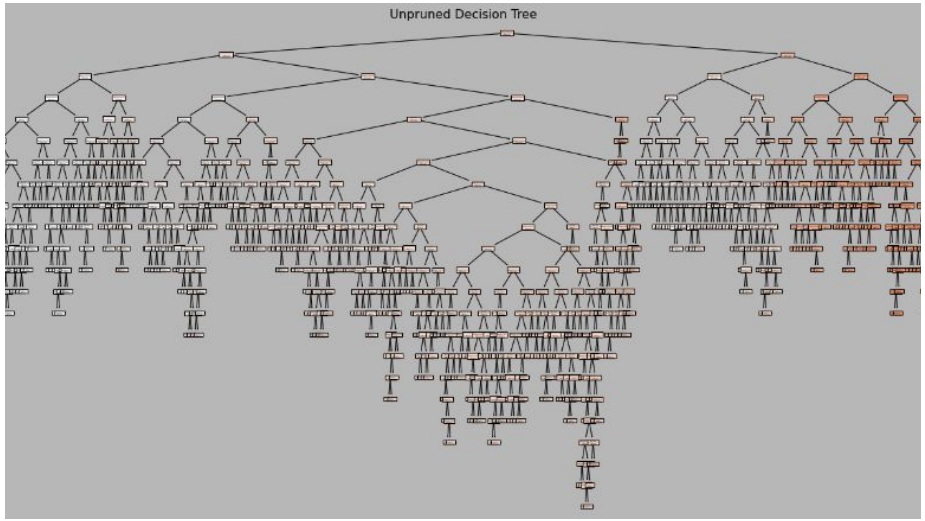
- Spending score and remuneration big drivers of loyalty point accumulation
-

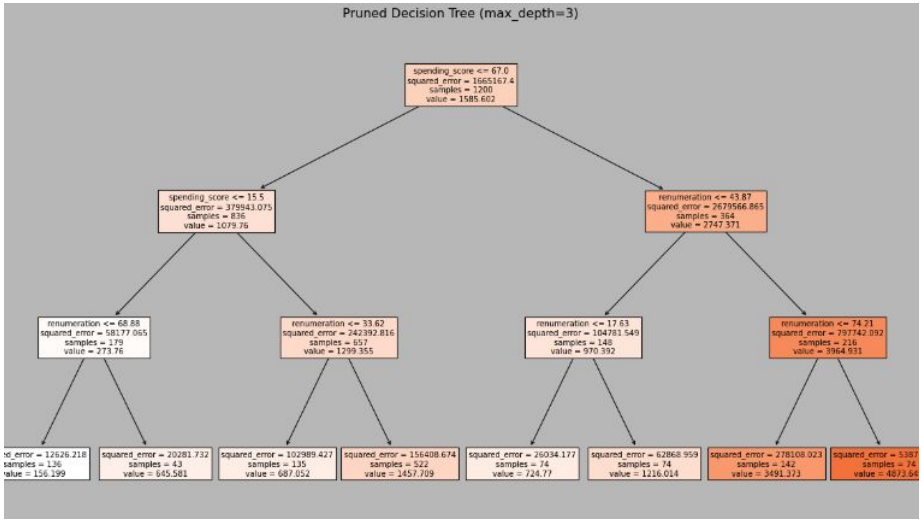
K Means clustering Analysis (1/3)



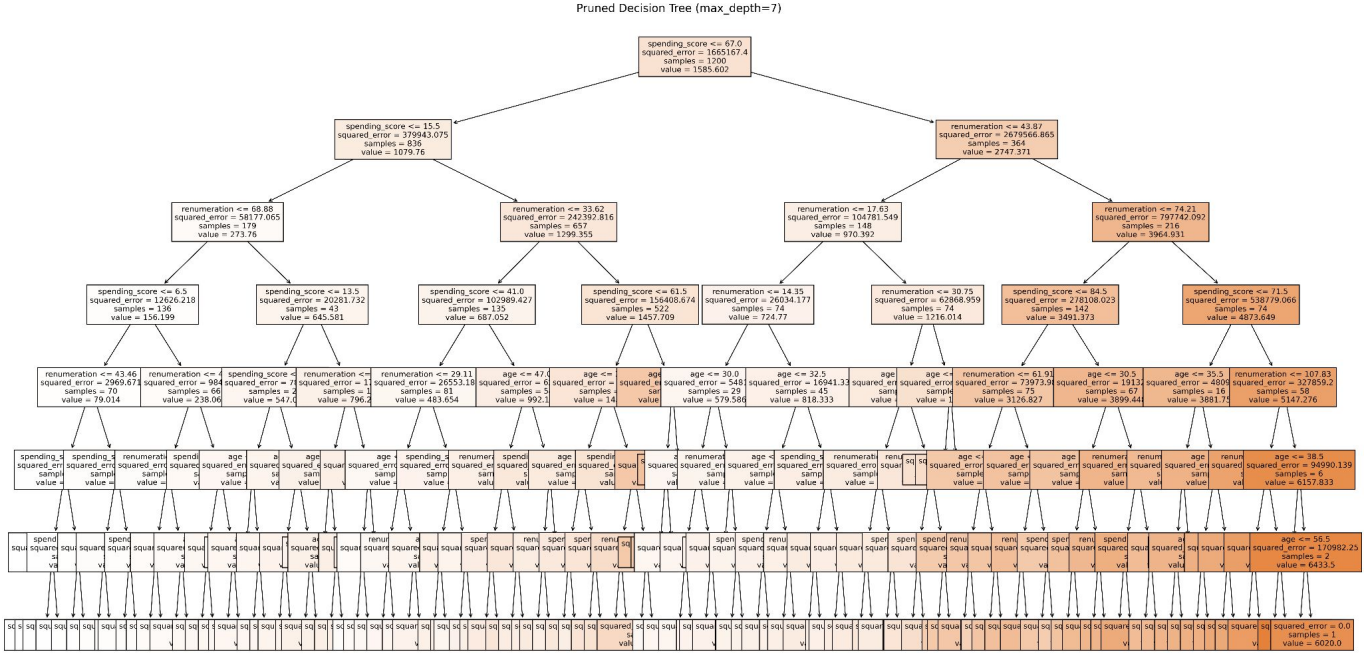


Decision Tree Regression Analysis (2/2)

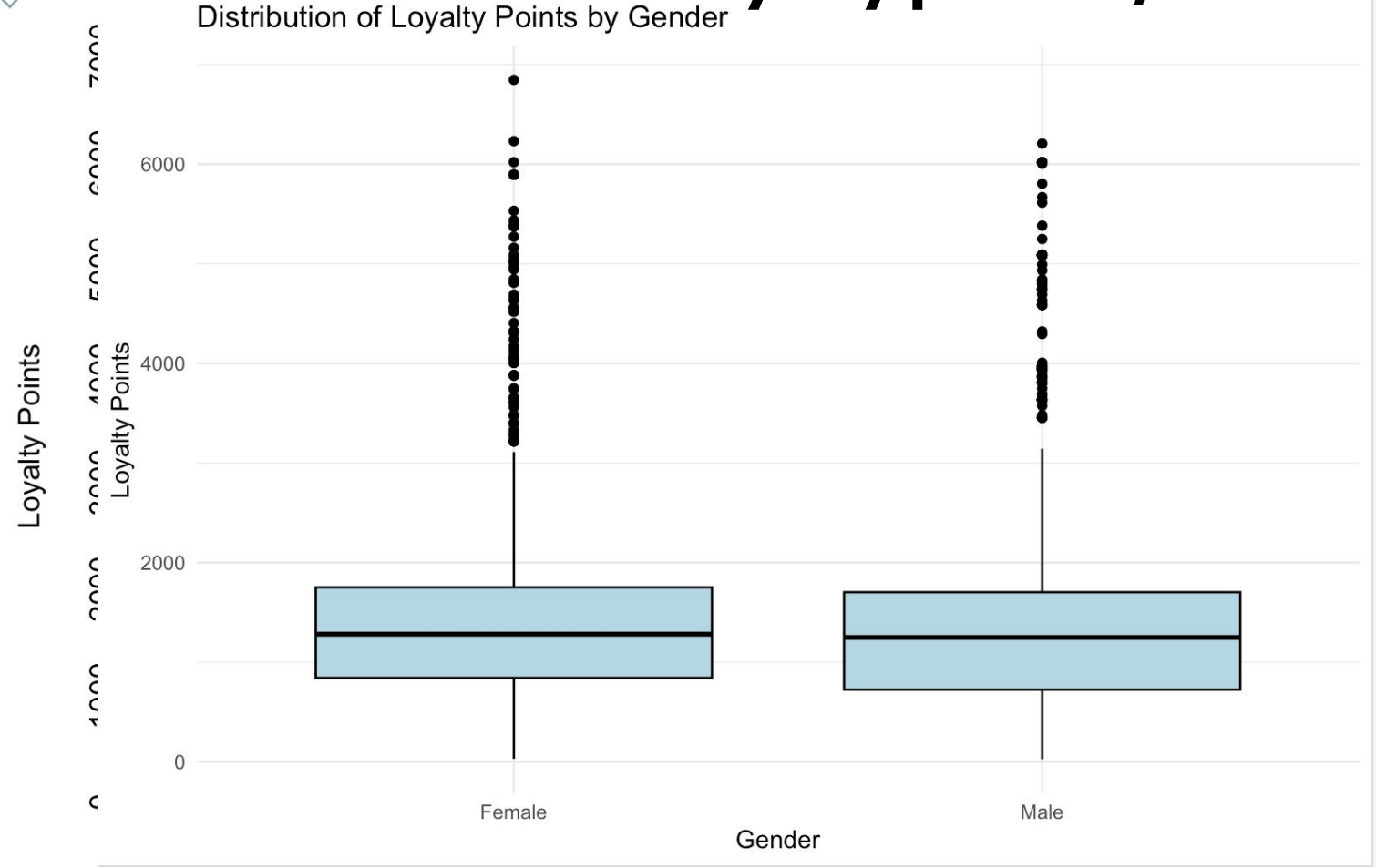




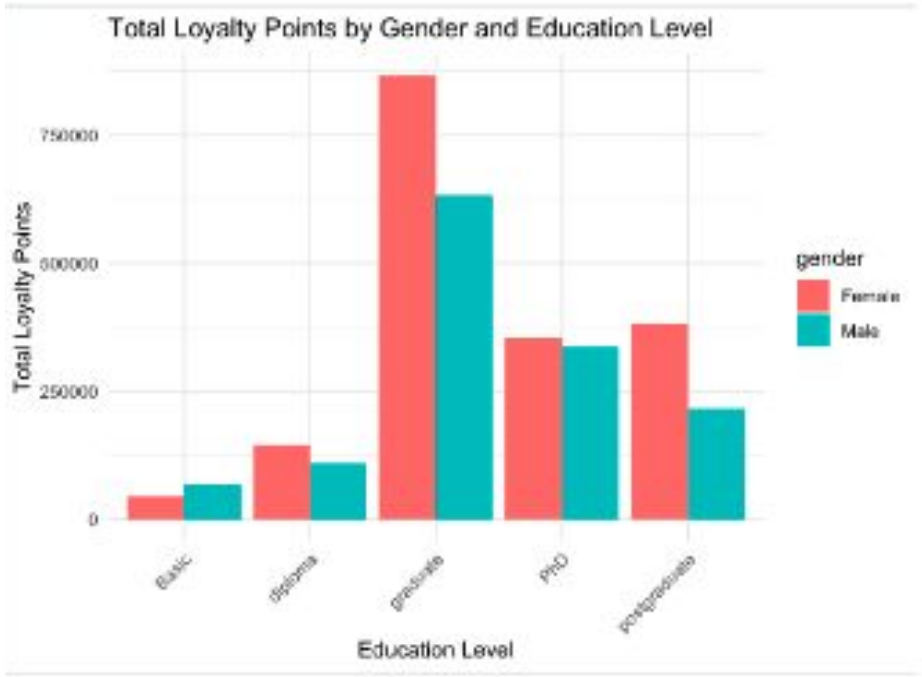
Decision Tree Regression Analysis (2/2)

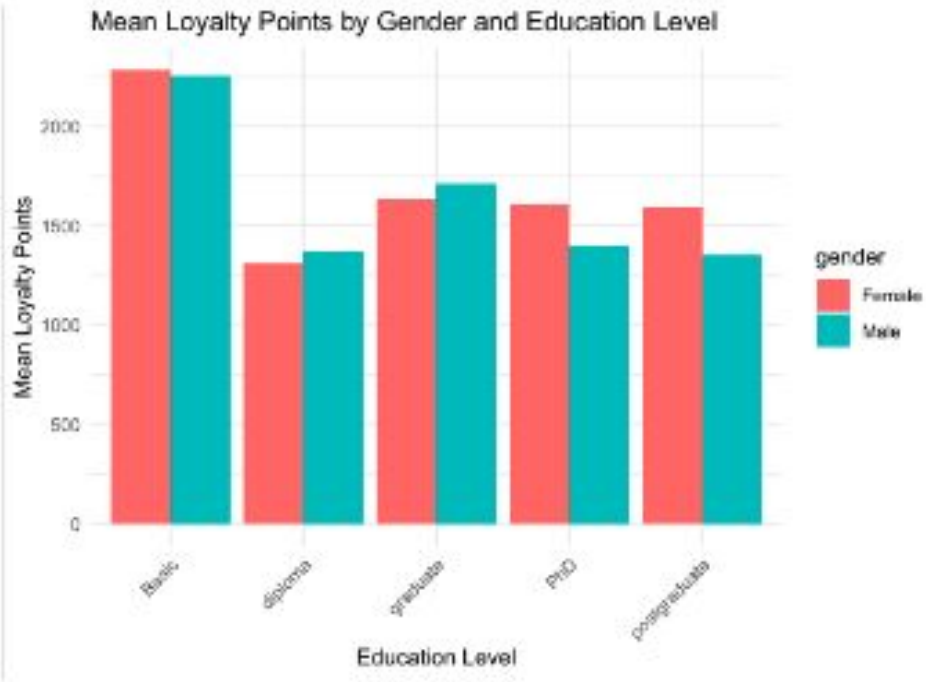


Distribution of total loyalty points 1/2



Loyalty points split by Education and Gender

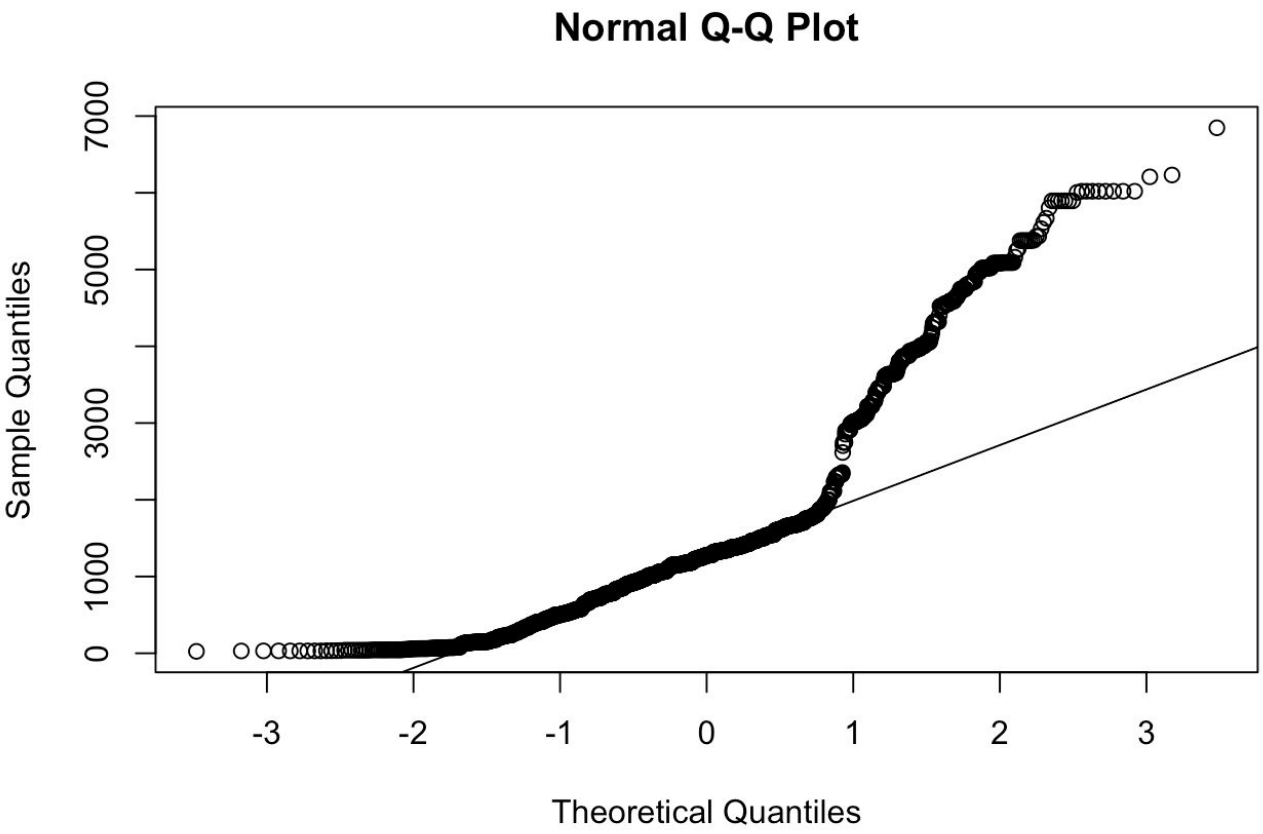




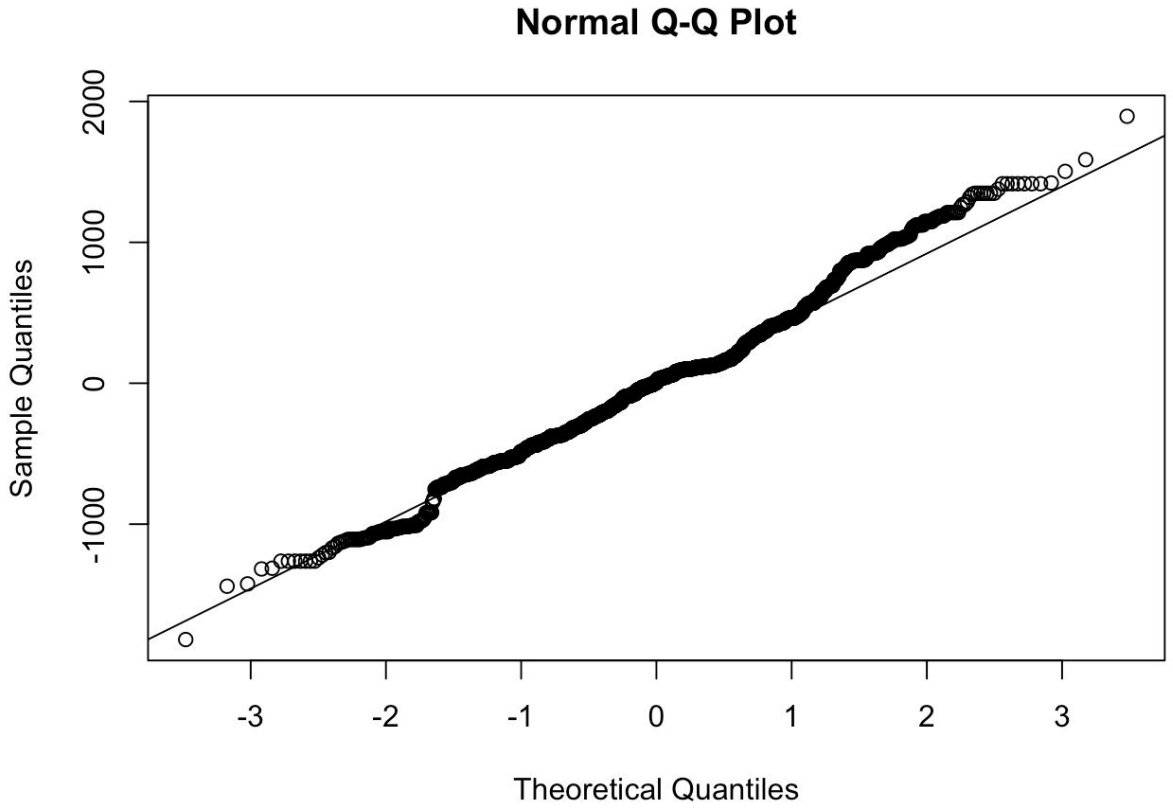
Insights about customer behaviour based on clusters

> - Cluster 0 - "High Spenders" combination of high spending and remuneration indicates that these customers are valuable.
> - Cluster 1 - "Moderate earners" make up the largest portion of the sample data set (774 customers). Should this pattern repeat on larger data set, Turtle Games may want to consider positioning certain products that appeal specifically to their profile when conducting product sales
> - Cluster 2 - "Low Spenders, High Earners" - low spending score but high average remuneration suggests potential for growth if Turtle games marketing strategies can encourage spending among this segment.
> - Cluster 3 - "Value Seekers" high spenders but lower remuneration, could suggest these customers seem to seek value, possibly making purchases during promotions or discounts.
> - Cluster 4 - "Budget Conscious" - Low spending score and remuneration indicating they might be more

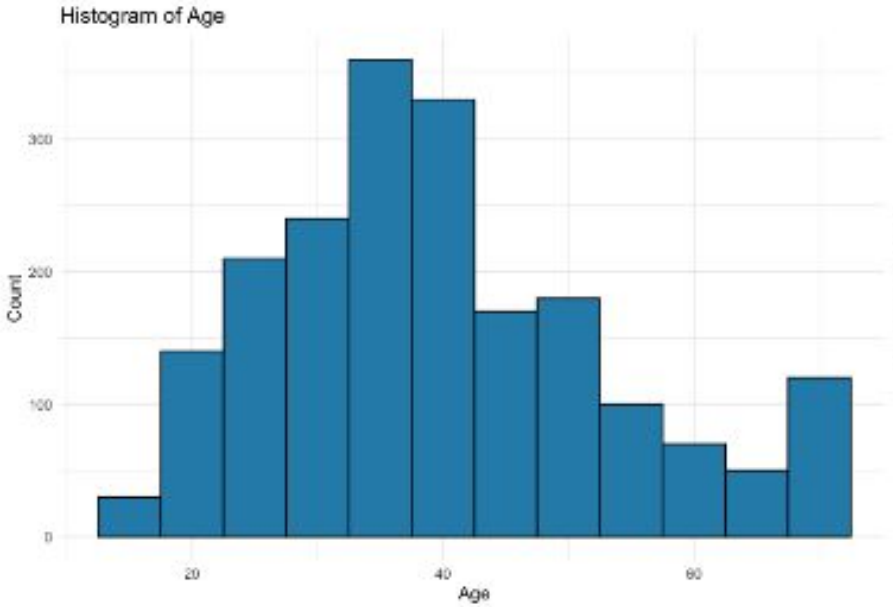
Creating a predictive model using a multiple linear regression model - pre refinement

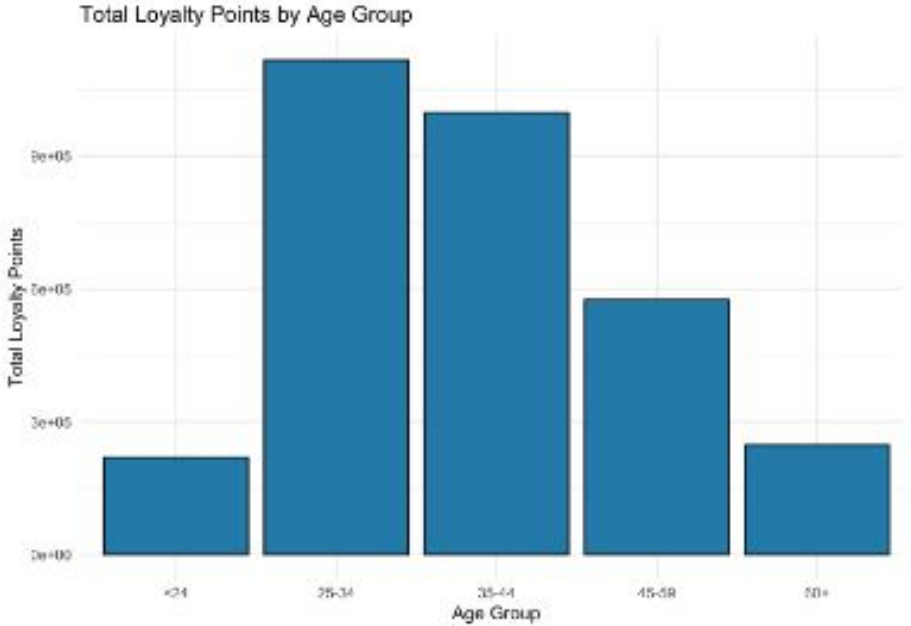


Creating a predictive model using a multiple linear regression model - post refinement

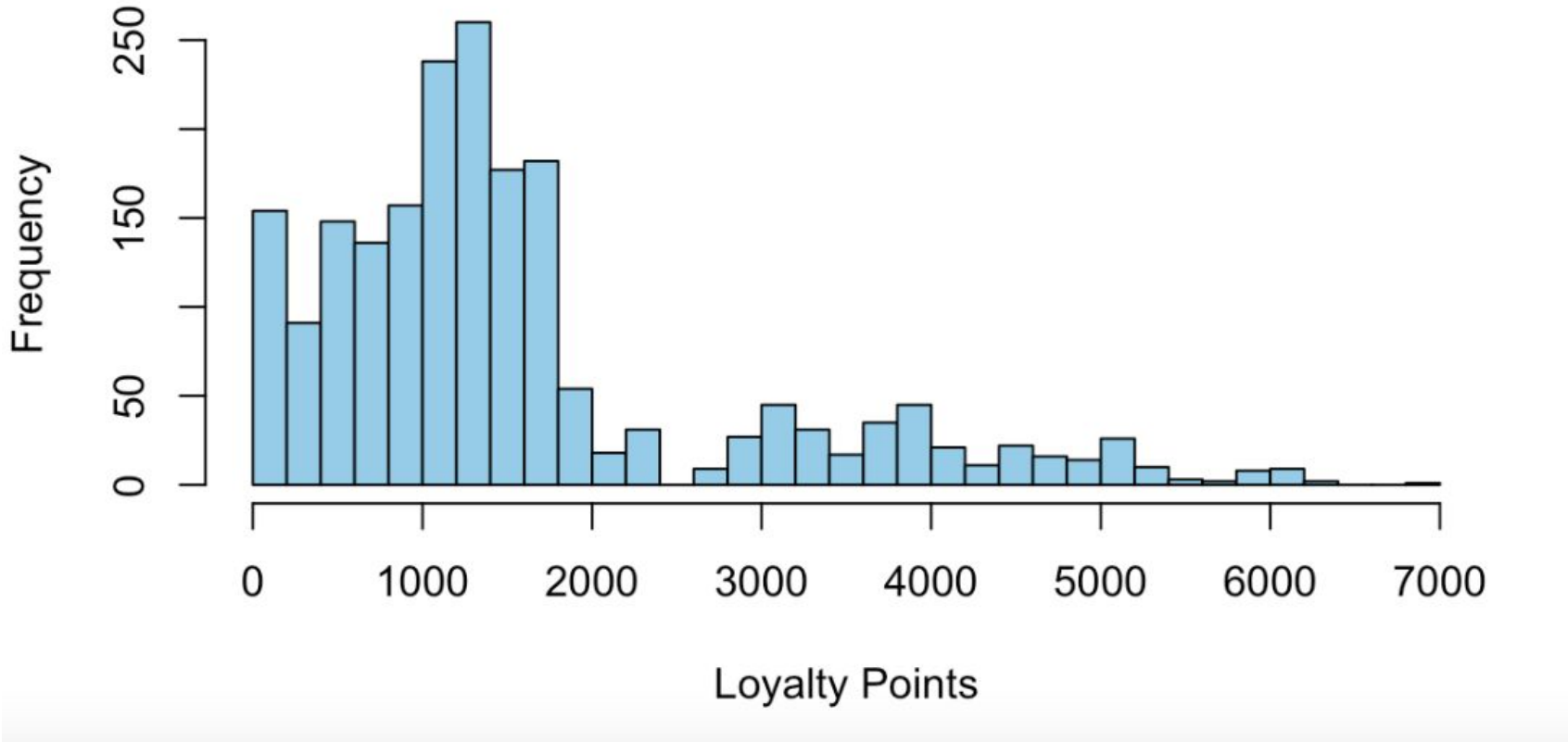


Distribution of Age against loyalty points





Distribution of total loyalty points 1/2



Review of supplied data

Column	Sample value	Interpretation of columns
gender	male or female	The gender of the customer.
age	number	Age of the customer in years.
remuneration (k£)	number	Total income per customer per year in pounds, where k=1000.
spending_score (1-100)	number	A score is assigned to the customer by Turtle Games based on the customer's spending nature and behaviour. The value ranges between 1 and 100.
loyalty_points	number	A score based on the point value of the purchase, converting the monetary value to point value, and the point value of an action (purchase).
education	graduate	Level of qualification that the customer obtained. For example: Diploma: completed school Graduate: undergraduate degree Postgraduate: postgraduate degree PhD degree.

TG Resources information?

language	EN	All the reviews were in English.
platform	Web	All the reviews were obtained from the website
product	number	Unique code allocated to product based on the item description.
review	When it comes to a DM's screen, ...	Online reviews submitted by customers who purchased and used the products.
summary	The fact that 50% of this ...	Summary of the customer's review

Matrix of customers segments based on 5 clusters

Cluster	cluster name	avg age	Avg spending score	Avg Remuneration	Number of customers	total loyalty points	Avg loyalty points
0	High Spenders	36	82.01	73.24	356	1419813	3988
1	Moderate Earners	42	49.53	44.42	774	1099376	1420
2	Low Spenders, High Income	41	17.42	74.83	330	300881	912
3	Value Seekers	32	79.42	20.35	269	261453	972
4	Budget-Conscious	44	19.76	20.42	271	74541	275

Breaking down the problem