

1. Business Background

Turtle Games, a manufacturer and a retailer of games and toys, aims to improve the overall sales performance. The scope of analysing the profitability of the company can be divided into three parts: reviewing product features and customer characteristics with current pricing strategy to maximise sales, determining the underperforming products to raise sales potential, and evaluating customer satisfaction to meet customer demand. The products of Turtle Games are segmented into 3 major types: 1. Lego, 2. board games and toys, and 3. video games.

Details of the analysis are in various aspects separately, business questions include

- a. Which age group submits the most reviews (in case of Lego)?
- b. What is the most expensive Lego set purchased by customers who are at least 25 years old?
- c. What price should be set for our target Lego product?
 - i. with 8,000 Lego pieces
 - ii. with 8,000 Lego pieces and for age 30
- d. What is the general sentiment of customers across all products? What are the top 20 positive and top 20 negative reviews.
- e. Determine the optimal global sales (in millions) for the next financial year for the video games.

2. Methodology

With a broad range of products sold, the information gathered would be various in nature based on the strong features of particular product type. Therefore, a wide range of tools and techniques are involved.

For Lego products, they are highly standardised and categorised. As a result, the collected data is specific, numeric and structured. Data of Lego products are relatively easier to handle and there are no restrictions on the choice of tools.

For games and toys with unique features of every single product, the direct feedback in words from our customers is valuable in suggesting the strengths and weaknesses of our products. And a word cloud is easy to understand the related words or feelings in them. Many well-

developed libraries and corpus in Python, like WordCloud, nltk, TextBlob, makes Python the best choice in processing the comment analysis.

Lastly, for video games, with the great variety of platforms and genres, focus would be on the sales instead of the numerous amount of products. To avoid conflicts, the analysis on video games is based in R.

Assumption: The results are under the assumption of interpreting the age categories of products instead of the ages of reviewers.

3. Data Cleaning

The data of Lego products are well-prepared for analysis. In the games sales data set with 15,000 observations, there are two rows with information mislocated in the columns, which are removed directly in a quick and easy manner and would have insignificant effect on the analysis.

Much effort is done to the game reviews data set in the data cleaning process. It is easy to get repeated comments from different customers if the comments are short. In order to retain the greatest value of the reviews data set, duplicated comments are checked with extra caution. Using the EXACT function in Excel, the reviewer ID, name and comments are all compared to filter the true duplications.

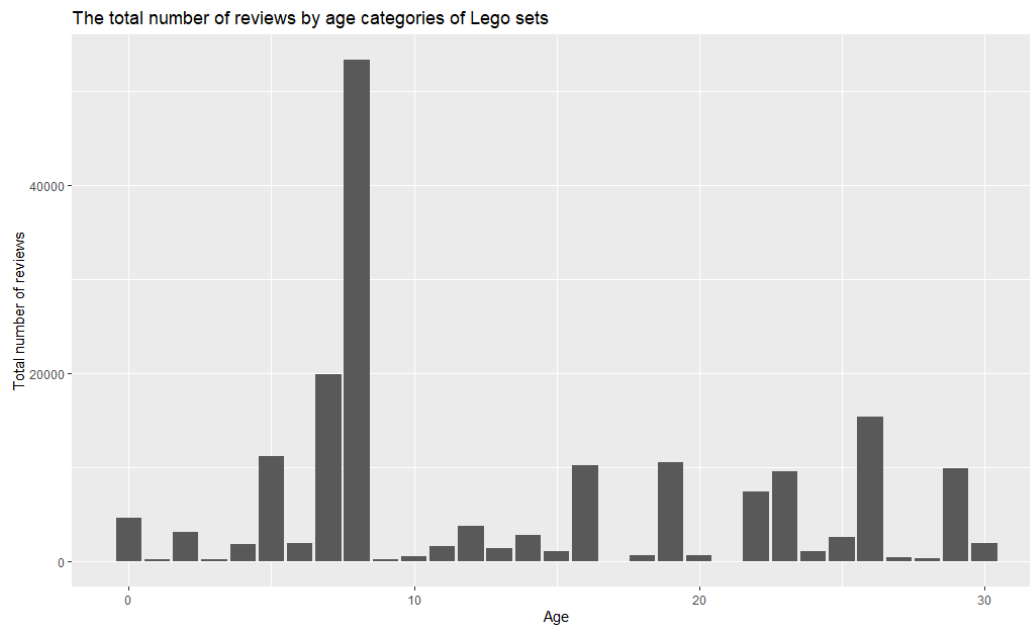
	D	E	F	I
	reviewerID	reviewerName	reviewText	UID
1146	ABWSQI9992Q29	Gary P.	nice	2
1147	ABWSQI9992Q29	Gary P.	nice	2
1156	A115TZEJ3U3AER	bebo995	On-time and nice item.	2
1157	A115TZEJ3U3AER	bebo995	On-time and nice item.	2
1268	AN82BJBT7QNT7	molfar	This set is slightly worse than earlier ones (in the old s	2
1269	AN82BJBT7QNT7	molfar	This set is slightly worse than earlier ones (in the old s	2
1995	ARNRPHG5F38GC	Jimmy	Gift	2
2007	ARNRPHG5F38GC	Jimmy	Gift	2
2585	A2534XA4DU1H4Y	Amazon Customer	All Good	2
2586	A2534XA4DU1H4Y	Amazon Customer	All Good	2
5883	A3MMWTSO936YX	Central Stores	This was purchased by our SGA for the students to pla	2
10655	A3A8UKN21UYLS1	Kurt	Big hit with the gamers	2
10656	ACNMHIMYEQX4Z	Joseph Vaughn	Beautifully written and bound. Full of the quirky humc	2
10805	A2R92WY3LPKNTF	Texas Kinder Mom	We really love the little story and elf, it is so adorable a	2
10806	A2R92WY3LPKNTF	Texas Kinder Mom	We really love the little story and elf, it is so adorable a	2
10926	A3A8UKN21UYLS1	Kurt	Big hit with the gamers	2
10932	ACNMHIMYEQX4Z	Joseph Vaughn	Beautifully written and bound. Full of the quirky humc	2
12117	A2D4OOJ0Y0T4J3	L. Scot Gowers	A+	2
12172	A2D4OOJ0Y0T4J3	L. Scot Gowers	A+	2
12966	A3MMWTSO936YX	Central Stores	This was purchased by our SGA for the students to pla	2
14062	AV2F45TXDRWJB	Stacy Hammel	Groovy	2
14122	A2EVAEJU340BWV	Patricia	Gift for 5 yr old grandson	2
14403	A3771WYFXR8ZQG	Peggy Perkins	Cute	2
14404	A3771WYFXR8ZQG	Peggy Perkins	Cute	2
14419	A2EVAEJU340BWV	Patricia	Gift for 5 yr old grandson	2
14490	A8RSCZ3GO9BIW	M's Mom	My son loves the book "How Rocket Learned to Read"	2
14492	A8RSCZ3GO9BIW	M's Mom	My son loves the book "How Rocket Learned to Read"	2
14519	AV2F45TXDRWJB	Stacy Hammel	Groovy	2

Screenshot of checking for duplicated comments Excel, formula used `=SUM(--(EXACT(D2:D15001&E2:E15001&F2:F15001, D1146&E1146&F1146)))`

4. Findings

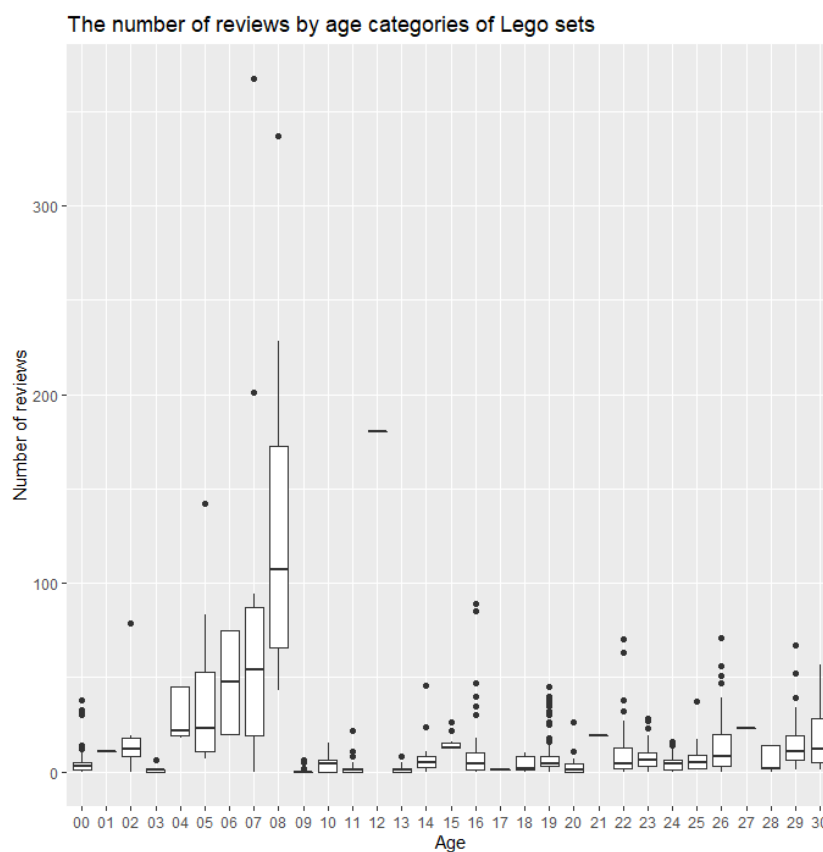
- a. Which age group submits the most reviews (in case of Lego)?

We have received particularly high number of reviews from the Lego product for age 8 in total (plot 1) and on average (plot 2). For a single product, however, a product for age 7 got the highest number of reviews and a product for age 8 comes next. (plot 2).



Plot 1:

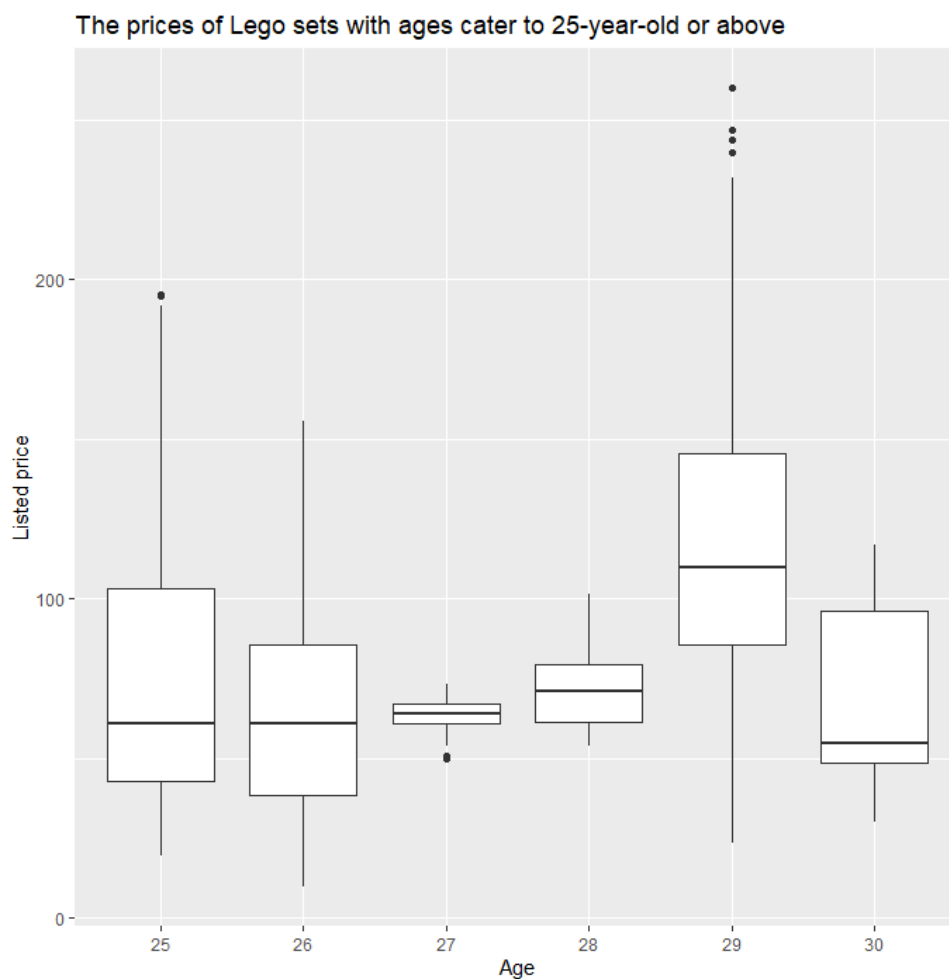
Plot 2:



- b. What is the most expensive Lego set purchased by customers who are at least 25 years old?

The most expensive Lego sets for age 25 or above is US\$ 259.87. However, this product only received 7 reviews, which is not very popular. Several products for age 29 are with higher price when compared to those for age 25 or above (plot 3). Among these expensive Lego sets, there is a product priced US\$ 246.87 with 52 reviews, which means more customers bought this set. Further analysis of differences between the products should be done to enhance sales of expensive sets.

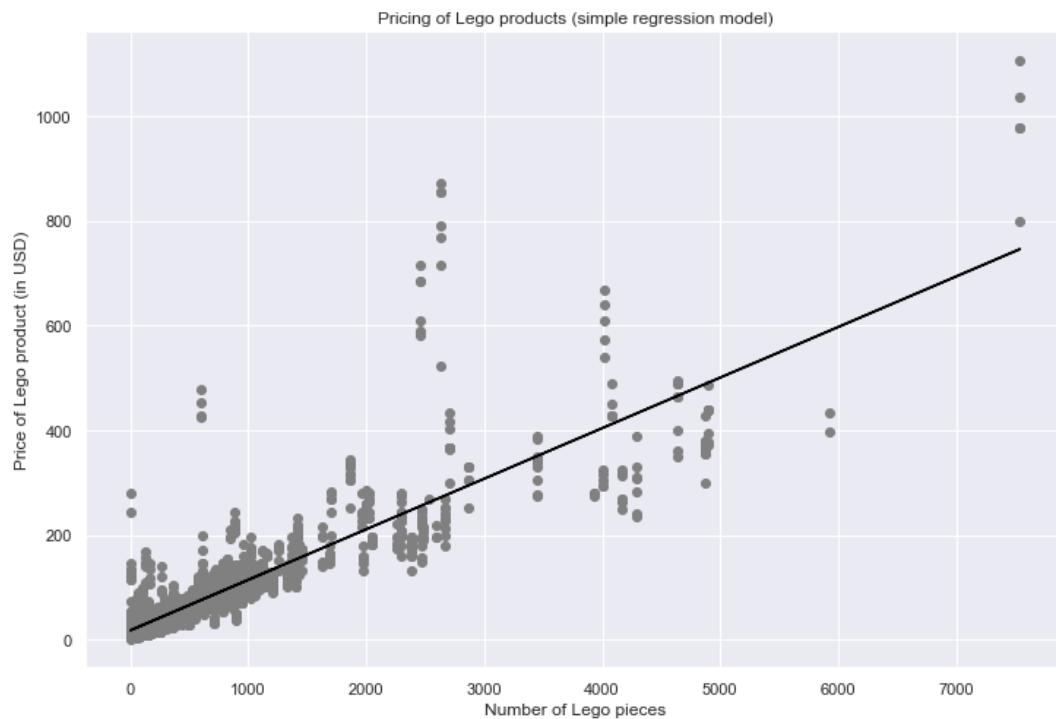
Plot 3:



- c. What price should be set for our target Lego product?

This part is done by linear regression model. The results are US\$ 789.83 for Lego with 8000 pieces and US\$ 808.99 for Lego with 8000 pieces and categorised for age 30. The optimal prices obtained from the models are lower than the listed price of our similar products (plot 4). The R-squared values of the models are below 80%, validating their explanatory power of selected factors could not fully represent the price. Any possible changes in price should be aligned with monitoring of sales and comments from customers.

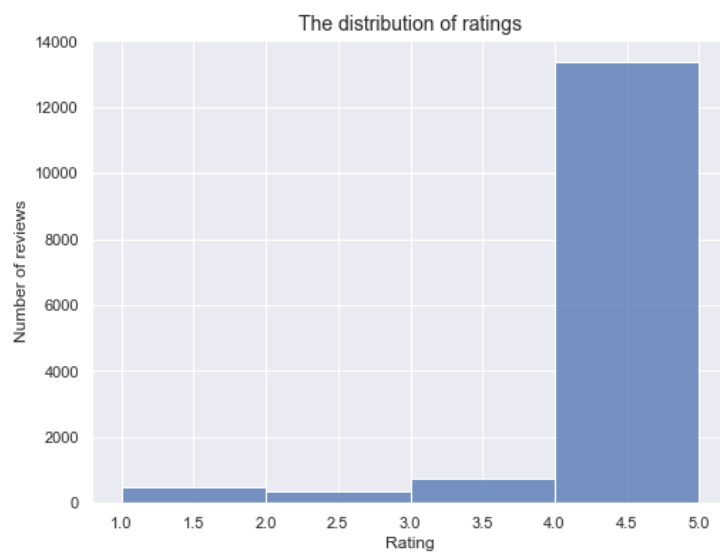
Plot 4:



- d. What is the general sentiment of customers across all products? What are the top 20 positive and top 20 negative reviews.

Using a word cloud can be easily spot most of the top words are with positive feedback (plot6) . The overall ratings are mostly with 5, i.e. very satisfied (plot 5). And with the polarity mainly distributed in the positive side (plot 7).

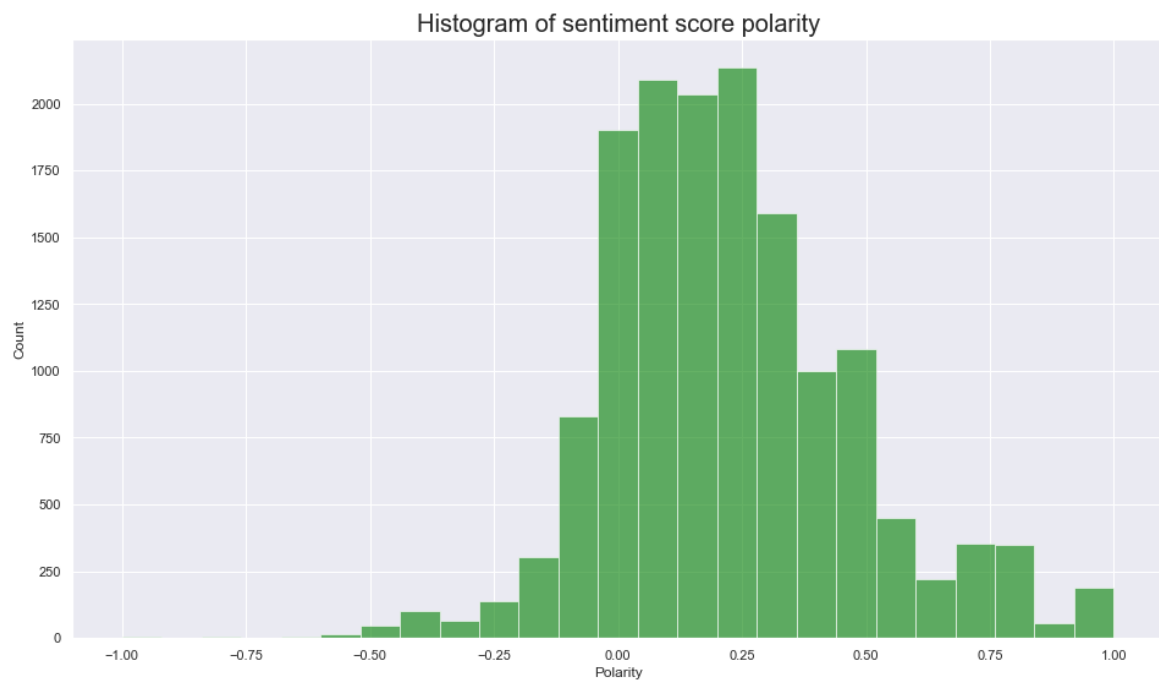
Plot 5:



Plot 6: The word cloud



Plot7:



List 1: The top 20 positive negative reviews.

	overall	reviewTime	reviewText	polarity
7	5	2016/12/17	came in perfect condition	1.000000
165	5	2015/1/12	awesome book	1.000000
194	5	2017/3/4	awesome gift	1.000000
496	5	2015/6/7	excellent activity for teaching self-management skills	1.000000
524	5	2014/8/2	perfect just what i ordered	1.000000
591	5	2018/4/2	wonderful product	1.000000
609	5	2018/3/15	delightful product	1.000000
621	5	2017/4/27	wonderful for my grandson to learn the resurrection story	1.000000
790	5	2016/1/20	perfect	1.000000
933	5	2017/11/22	awesome	1.000000
1037	5	2014/12/21	awesome	1.000000
1135	5	2017/1/12	awesome set	1.000000
1166	5	2015/3/6	best set buy 2 if you have the means	1.000000
1175	5	2015/2/18	awesome addition to my rpg gm system	1.000000
1298	5	2018/1/1	its awesome	1.000000
1398	5	2015/4/7	one of the best board games i played in along time	1.000000
1547	5	2015/4/24	my daughter loves her stickers awesome seller thank you	1.000000
1606	5	2014/10/15	this was perfect to go with the 7 bean bags i just wish they were not separate orders	1.000000
1712	5	2015/1/4	awesome toy	1.000000
1717	3	2014/12/25	it is the best thing to play with and also mind -blowing in some ways	1.000000

List 2: The top 20 negative reviews.

	overall	reviewTime	reviewText	polarity
208	1	2017/1/17	booo unles you are patient know how to measure i didnt have the patience neither did my daughter boring unless you are a craft person which i am not	-1.000000
2039	1	2017/1/15	kids did not like it thought it was boring	-1.000000
3352	1	2016/6/2	some of the suggestions are disgusting	-1.000000
8306	1	2016/12/15	awful we did not receive what was advertised we paid \$30 for the boxes set with book we got the elf in a bag without the book	-1.000000
7975	3	2017/12/18	was the elf on the shelf but it didnt have the dvd i was very disappointed	-0.975000
9497	1	2014/10/8	i havent even taken it out of the box yet but its already falling apart i contacted customer service and never even got a response i am very disappointed in this product	-0.975000
8745	1	2015/12/3	i hate the holidays bcuz of the elf he was disgusting i hate him with my life he doesnot leave the shelf alone	-0.866667
13161	4	2013/2/16	i do not under stand how you keep score or read the scoring i i do not like that at all i can never play score with anyone at all i hate that i cant play points	-0.800000
9130	2	2015/1/7	cliche and stupid i should not drink and amazon	-0.800000
9247	1	2014/12/27	just stupid	-0.800000
182	1	2017/7/17	incomplete kit very disappointing	-0.780000
14262	4	2013/1/19	i like this product for my daughter she is into the bad kitty book collection so it was an added bonus	-0.700000
4267	2	2017/5/29	ordered for my sons birthday opened it up today to play and the board is damaged before we even take it out of the box the game is already falling apart very disappointed	-0.687500
4299	1	2017/4/30	id like to upload a photo of the condition of the game boxit looks like its been used as a soccer ball 2 corners of the box are smashed in and on is even ripped how am i supposed to give this as a gift without it looking like i bought this on clearance very disappointed	-0.687500
11953	1	2014/12/11	horrible and incomplete flash cardsdo not buy not helpful i was too late to return them	-0.650000
2142	2	2016/7/25	this was a bit disappointing my students find it boring and the letters are hard to understand	-0.630556
11451	1	2013/1/15	boring did i mention boring well its boring pass on this one there are a lot better games out there	-0.625000
13962	1	2015/4/22	had no idea the extent you have to go through to put this together hundreds and i mean hundreds of pieces - that dont snap together it will take my teen age son and i months to put this stupid thing together horrible plan horrible	-0.622500
8229	1	2017/1/4	i received a small paperback bookfor \$3000 the picture shows an elf hardcover book and box that it all comes in very disappointed for the student we bought this for	-0.612500
4977	5	2016/4/8	want to hate your friends and family get this game	-0.600000

- e. Determine the optimal global sales (in millions) for the next financial year for the video games.

We can see a positive correlation between the sales in North America, Europe and globally, a multiple linear regression model is created for prediction of the global sales. The optimal global sales is 8340.804 in millions of units for the next financial year.

5. Suggested future actions

In the analysis, there is an extra information found on the Lego products. There is a group of products with lower ratings and high review difficulties. Also, a lot of the negative comments mentioned the missing parts of our products (List 2). Improving these under rated products could help gain customer trust and satisfaction, and so could enhance the sales revenue.