Table of Contents

Background information

Lego is the largest toys manufacturer with almost 100-year history, which is now presented in more than 130 countries. Through the year Lego managed to remain one of the key players in a very competitive toys market, which was valued at $92.2 billion in 2019 [6], and still rapidly growing.

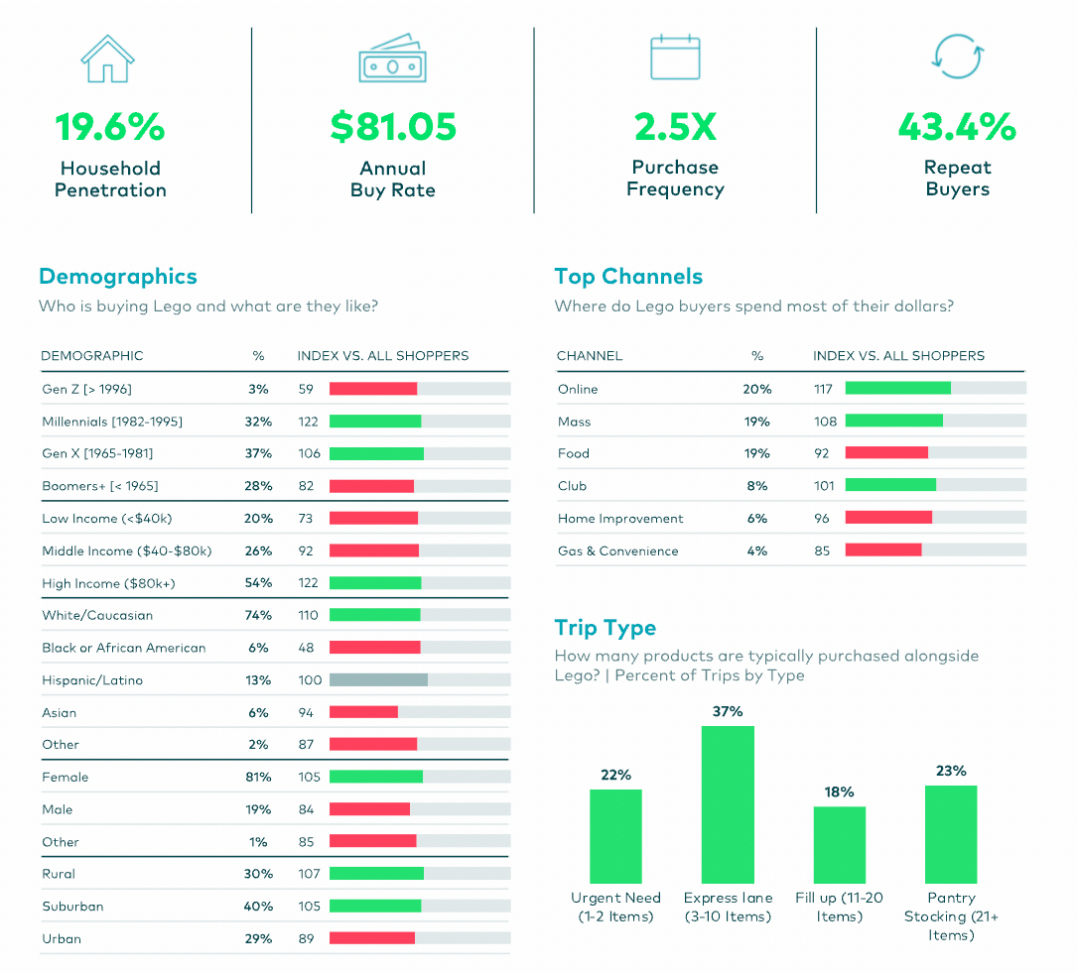

Figure 1 Lego Brand Snapshot

Figure 1 demonstrates the snapshot of the Lego brand with some outstanding numbers. It has almost industry best percentage of repeated purchases and customer satisfaction rate. Product of the company is highly popular all over the globe, among different cultures and age groups.

Business Problem

Although Lego has been a leader of its industry for many years, the rapid changes that occur in our society make the company always keep its finger on the pulse. In the era of digitalization children start to interact with electronic devices from the very young age, which dramatically decrease the demand for regular physical toys. Although Lego is not capable of stopping the progress, it still can grow and develop with it. Recently, Lego presented their new toys with augmented reality functionality, which enabled children to play with both their devices and physical toys simultaneously. To conduct such successful experiments, it is important to know where the trends of the audience are headed and follow them before other competitors. Those trends can be revealed through the market research.

Having a very powerful and well-known brand helps Lego to hold the leading position in the market. The current trend for sustainable and eco-friendly corporations is important for Lego too. It has a great potential to become one of the largest green companies which takes a great care of the future generations. However, it is important to know how exactly the customers will perceive the changes that the company is adopting.

Finally, Lego is still growing and expanding, enlarging their offline retail store network. Thus, it is utterly important to understand and locate any regions with potential points of growth or identify poor experience customers may have due to the bad location of the shops.

Research Aims

There are three general research aims that will be covered in this proposal. First, is to identify which trajectory of development Lego should follow to be in demand among today's children and upcoming generations. To achieve it, the company should find out more about the current trends in the entertainment industry at the children's side. We suggest several hypotheses to test to understand today's kids' preferences in their entertainment. The first hypothesis is that the younger generation of parents gives their children more time with gadgets than the older one. Another hypothesis is that children starting from the age of 5 prefer digital toys over the physical one.

The second aim of the research is to identify what group and proportion of the customers should be affected by marketing strategies promoting recycled plastic. In another words, on which segment of the customers the campaign should focus mostly. Here we suggest two hypotheses. The first one is that customers with children care more about recycling. The second one is that only younger generation (up to 25 years) considers recycling important.

The last aim is to identify whether the potential Lego customers have a need for more retail stores. Here the hypothesis is that there is a significant proportion of customers who do offline shopping and would make more purchases if the store had better location to them.

Research Design

Given the several hypotheses that should be tested based on the people's perception of the Lego products and toys in general, this research should be designed as Conclusive.

Survey method

Due to the nature of this research and given aims, we avoid using quantitative observation techniques and focus on survey. Different survey methods are separated from each other by the contact method used. We will describe each of them in details to justify the final choice.

Telephone interviews are frequently used because it is a fast and easy way to survey a large sample. However, in the given case this method will not be so valuable since we are focusing our research on the customers of Lego and their nearest competitors. Thus, the response rate in such case would be very small and it would take more time and finance to gather enough data to fill the required sample sizes.

Mail interviews are like the telephone method both having the same benefits and disadvantages. Similarly, they are inefficient and make it harder to get to our target audience, which would add more noise to the collected data and may a negative impact on the quality of the results.

Electronic surveys, on the other hand, give more flexibility as the questions can be dynamically changed and some additional information can be embedded into the response, for example the country of respondent can be extracted from the IP address when he submits the survey. Given that the target audience of the Lego company are young parents with above average income [3], electronic surveys are easier to target on this specific group of people. This will be discussed in more detail in the data collection paragraph.

Another survey method that we suggest is personal interviews. Since the client proposed that the customers of the Lego competitors should be included in the research, we can simply put surveyors at the exits of large toy retail stores. Although it is the most expensive contact method, it is the most accurate and easiest way to get responses from the target audience and competitors' clients.

Sampling method

Before considering the sampling technique, it is worth mentioning in the details the distinctive features of the Lego company and specifically the research itself. Although Lego has its factories in only 5 countries, the company's products are sold in more than 130 countries. Thus, the sampling frame includes customers of Lego and other toy retailers all over the globe.

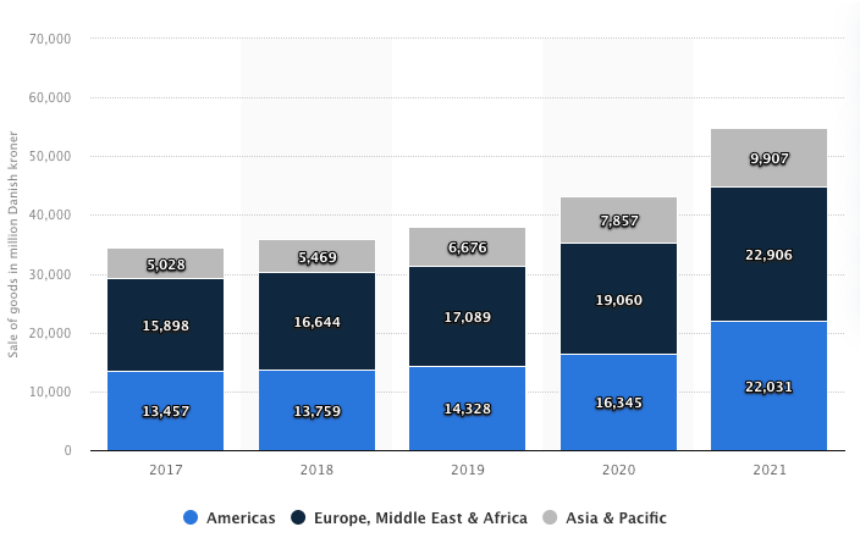

Figure 2. Lego revenue by region

On the Figure 1 presented Lego revenue for the last four years split by the region of sales. As it can be seen, there are three main regions with Americas and EMEA having almost equal shares, and Asia & Pacific accountable for 20% of the revenue. Since the research aims include topics related to the specific locations of the business, for instance, the decision to expand the retail network, all of these regions should be included into the sample. Thus, just a random sampling of retail stores or countries is not applicable here.

The first sampling method that suits this specific case is a disproportionate stratified sampling. In this method the sample size from each stratum depends on the relative size of the stratum and the standard deviation of the underlying distribution. Regions which were described earlier will be used to construct strata. However, since the market of EMEA is more saturated than the market of Americas, it is reasonable to take a large sample size from that region to account for more diversified population inside. Then a simple random sampling can be implemented to choose a country of the research in case of online questionnaire, and a specific location of a retail store in offline case.

Since we can assume that inside of one region the countries are homogeneous, another sampling method can be proposed: multistage sampling with mixed techniques. At the first stage there would be proportionate stratified sampling with strata being the region, then each stratum is subdivided by countries into the clusters, which are then used with probability proportionate to its size.

Sample size

First, to determine the required sample size we should start with the level of confidence in our research. This value can be described as the probability that the estimation of some statistical parameter that we get from the survey is also true for the whole population. In general, there are three main values to the level of confidence, which are 99%, 95% and 90%. 99% probability is an unnecessary accuracy for the research, which would require much larger samples, increase the time and the cost of the research, thus we will have a 95% confidence level.

For simplicity we will focus on calculating the sample size for categorical data. The formula is the following:

$$N = p*(100-p) * z^2 / err^2$$

Where N is the sample size, p is the percentage occurrence of a variable, z is the critical value to the chosen confidence level and err is the maximum percentage error which is often chosen to be 5% for social research. We retrieve the margin of error from the following table, given the maximum variance of population (we are interested in the upper bound) and the population size.

TABLE I: SAMPLE SIZE BASED ON DESIRED ACCURACY

SOURCE: (GILL ET AL., 2010)

	Variance of the population P=50%					
	Confidence level=95% Margin of error			*Confidence level=99% Margin of error*		
Population Size	5	3	1	5	3	1
50	44	48	50	46	49	50
75	63	70	74	67	72	75
100	79	91	99	87	95	99
150	108	132	148	122	139	149
200	132	168	196	154	180	198
250	151	203	244	181	220	246
300	168	234	291	206	258	295
400	196	291	384	249	328	391
500	217	340	475	285	393	485
600	234	384	565	314	452	579
700	248	423	652	340	507	672
800	260	457	738	362	557	763
1000	278	516	906	398	647	943
1500	306	624	1297	459	825	1375
2000	322	696	1655	497	957	1784
3000	341	787	2286	541	1138	2539
5000	357	879	3288	583	1342	3838
10000	370	964	4899	620	1550	6228
25000	378	1023	6939	643	1709	9944
50000	381	1045	8057	652	1770	12413
100000	383	1056	8762	656	1802	14172
250000	384	1063	9249	659	1821	15489
500000	384	1065	9423	660	1828	15984
1000000	384	1066	9513	660	1831	16244

Figure 3. Sample size based on desired accuracy

Given that we are calculating the required sample size in each stratum, the population size inside each of them is over 1 million. Finally, we calculate the sample size: N = 0.25 * $1.96^2/0.05^2$ = 384. On the Figure 2 [4] we can see that our calculated sample size is equal to the correct size in the table. Given that in the previous point we suggest sampling three strata, we propose having 400 respondents in each stratum, which leads to 1,200 people in total. However, given that the client is suggesting using 3,000 Lego customers, we can propose to use more strata and subdivide each region into three smaller one. This would not change the sampling technique but given 400 respondents in each stratum this will lead to 3,600 respondents in total.

Fieldwork & Data Collection

We begin describing data collection and the way the survey will be conducted by focusing on the electronic surveys. Since Lego has a large online retail channel, we can use it to distribute the questionnaire. There are several options that we suggest using such as sending emails to customers or a banner on the website.

LEGO has emails of each customer that had an order from their website or simply created an account there. Thus, the survey can be sent to the customers personally, considering the country of the customer (this is important because we have stratified and clustering sampling), and it would require almost no effort. Additionally, the responses can be then joined from the account information from the LEGO website, revealing details about client's preferences and spending.

We can also put a banner on the LEGO website, which would lead to the page with the survey. This way, we can control from which countries or cities users will be able to participate in the research.

As for the face-to-face interviews, we suggest having a surveyor at the big malls or near crowded toy stores. As described earlier, the location of such stores depends on the sampling stratum or the cluster. It is reasonable to choose larger shops to get more responses per hour. They will be interviewing customers or parents with their children, thus providing an opinion of those customers that buy LEGO and those that prefer other competitors. Once the surveyor received enough answers, he sends the completed questionnaire to the main office to get the results.

Timing and Cost

In general, the proposed survey consists of two parts: online and offline. Since the online survey includes mostly Lego customers and will make up to one-third of the sample size, which equals to 1,200 people. Previously, we described the distribution system which will allow easily to send questionnaire and receive the responses. Thus, adding a banner to the website or sending emails to customers might take 3 working days of the developers, which can be evaluated as 1,000$.

Since we would have to interview 2,400 more people offline, and each store location would require at least 30 responses, which can be collected in 6 hours. Let's assume we pay 15$/hour to the surveyor, and we should have 2,400/30 = 80 surveyors. In total, we pay for 6 hours shift for 80 workers and get a budget of 15*6*80 = 7,200$.

We should also add the cost of searching for such surveyors in different countries, giving them chancery and collecting the results. Thus, the approximate cost of the whole research is nearly 12,000$. As for the timing of the research, we estimate the online survey to be completed in two weeks, and the face-to-face interviews to be completed in one month.

Questionnaire

1. **Do you have children?**

   [ ] Yes [ ] No

   a. **If yes, how many children do you have?**

      [ ] 1 [ ] 2 [ ] 3 [ ] >3

2. **What is the age of your children?**

   [ ] < 4 [ ] 5-6 [ ] 7-10 [ ] >10

3. **How many hours per day they spend playing with the gadgets?**

   [ ] < 1 [ ] 1-2 [ ] 3-4 [ ] >4

4. **Do they prefer physical toys over digital games?**

   [ ] Yes [ ] Mostly yes [ ] Mostly no [ ] No

5. **Do you support recycling?**

   [ ] Yes [ ] Partially [ ] No

6. **Are you comfortable with buying toys made from recycled materials?**

   [ ] Yes [ ] No [ ] Cannot decide

7. **As a consumer, do you prefer brands that are sustainable and eco-friendly?**

   [ ] Yes [ ] Rather yes [ ] Rather no [ ] No [ ] Does not matter

8. **How many times a year do you make purchases at toy stores?**

   [ ] < 3 [ ] 4-6 [ ] 7-10 [ ] > 10

9. **Do you prefer buying toys offline or online?**

   [ ] Online [ ] Offline [ ] Both

10. **Would you do more shopping if toy shops were closer to you?**

    [ ] Yes [ ] Probably yes [ ] Does not matter [] Probably no [] No

11. **On the scale from 1 to 5 (where 1 is the worst, 5 is the best), how do you rate the location of the stores?**

    [ ] 1 [ ] 2 [ ] 3 [ ] 4 [ ] 5 [ ] Don’t know

12. **Indicate the city where you are taking the survey**

    _____

13. **What is your gender?**

    [ ] Male [ ] Female [ ] Other [ ] Prefer not to mention

14. **How old are you?**

    [ ] <18 [ ] 18-24 [ ] 25-30 [ ] 31-40 [ ] >4

Data Analysis

The first hypothesis to consider is to identify whether there is a relationship between the age of the parents and the time their children spend with the gadgets. We will use cross-tabulations and Chi-square test to verify the statistically significant relationship between these two variables. Another hypothesis whether the children from the age of 5 prefer digital toys over the physical ones can be tested using the logistic regression. If the coefficient in the regression will be statistically significant, we could make a conclusion about this trend.

Then, we could use cluster analysis to understand deeply the audience that supports the recyclable toys and would love to see the company eco-friendlier. To test the hypothesis that the customers with the age from 18 to 25 consider recycling more important than the ones that are older we could do an ANOVA test to verify whether there is a statistically significant difference between these two groups.

As for understanding of why some customers prefer offline shopping and dissatisfied with the location or other features of the shops, we could run a factor analysis. This would reveal what factors have the most effect on customers perception.

Further Research

In the further research the quantitative observation techniques can be tried. Lego is a huge company with hundreds of retail stores, which enables it to collect enormous amount of data about their products and customers. Joined with location data of purchases, this may add additional hints on in which regions the demand for toys is higher that could be a potentially a worthy place to expand.

Additionally, Lego may conduct a phycological research on how modern devices influence children behavior. This may help to understand which vector of the development Lego should take to be relevant for future generations.

Bibliography

1. https://www.encyclopedia.com/economics/economics-magazines/lego-group
2. https://www.statista.com/statistics/1103051/lego-net-sales-worldwide-by-region/
3. https://www.numerator.com/snapshots/brands/lego
4. https://hal.archives-ouvertes.fr/hal-02546796/document
5. Figure 1
6. https://www.alliedmarketresearch.com/toys-market-A08309

Word count: 2800