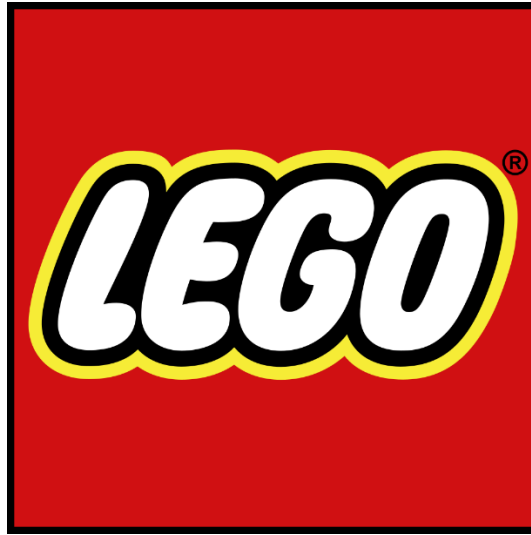# The Lego Group (Lego)

Market Research Proposal

Statistical Methods

Done by: Prisca Lim Shi Zhen

Executive Summary

The Lego Group is one of the successful largest toy makers for its interlocking plastic brick production for over 60 years. Despite the advancements in technology, it is still one of the international bestsellers with sustained interest among children and further expanded its range to cater to teenagers and adults. However, due to rapid shift in economic and environment such as rise in global warming and the cost of plastic materials as well as internet users due to technological advancement, it causes Lego to re-think their business operations, expansion, and innovations. Furthermore, due to dynamic shift in the toy industry with many new and existing competitors adapting to such shift, Lego is financially uncertain in maintaining growth and in increasing its market shares.

The purpose of this study carried out by the market research agency is to identify, examine and determine any patterns or association from data collection and analysis that may help to solve the current issues of Lego stated above. This allows the company to facilitate accurate decision-making, lower business cost or losses and increases its competition in a dynamic environment compared to competitors that lacks market research. In this research proposal, the primary focus on each research aims, question and objectives has been highlighted respectively to address Lego's concern in investigating type of products (physical, digital, or blended) that would appeal to children in the next 5 to 10 years, types of market strategies that Lego could deploy to promote use of recycled plastic bottles in brick production and deciding whether to expand their retail store network and at which location.

The methods proposed to gather data is through online surveys and questionnaires for quantitative data using stratified sampling and 100% sampling from online articles, and online focus-group for qualitative data using convenience sampling technique. The research proposed will involve both primary data for online surveying and discussion, and secondary data for online research articles or journals as source of data collection. This allows Lego to collect accurate, legitimate, and real-time information, leading to smooth adaption in a dynamic market. Additionally, timescale and budgets for a four-month window requested by Lego has been proposed in the research proposal to provide estimated cost and time spent on execution of market research. This will allow Lego to determine if the research proposal is within their financial ability and time before execution.

Recommendations has also been proposed to include data intelligence tools such as machine learning algorithm to extract in-depth insights for profiling, as well as increasing customer's relationship and providing promotions to attract potential or deter existing customers to switch to other competitors. This will improve Lego's competition and sustainability in long run.

Table of Contents

1. Introduction

1.1. Background

The Lego Group was founded in 1932 by Ole Kirk Christiansen who initially produce wooden toys as their primary business was manufacturing household goods. Currently, it is a private owned company based in Denmark, which began to produce its first interlocking plastic toy bricks for children in 1949[1] and has started expanding its customer segment to teenagers and adults. In 2020, the company's revenue amounted to approximately $7.2 billion with more than 200 employees, 700 store and several Legoland amusement parks globally[2]. With dynamic shift in market environment and competitions towards digital economy, the company seeks to gather insights and adapt to such market to prolong its sustainability.

1.2. Business Issues

Presently, there are 13.8 billion active connections users in Internet of Things (IoT) and forecasted a rise to 30.9 billion active users by 2025[3] with overall expected spending of 1.1 trillion U.S. dollars in 2023[4]. Furthermore, climate change has worsened as global temperature increased to 1.01 degree since 1883[5] which partly caused by plastic pollution accounting for 20% of oil consumption by 2050[6], resulting in rapid shift in dynamic market. Despite its record sales and profits in 2020, Lego is uncertain if investment in technology and business expansion will be commercially successful to maintain its growth and further increase its market share in the toy sector in the long run to sustain its competition in a dynamic market.

---

[1] The Lego Group's background information taken from: https://en.wikipedia.org/wiki/The_Lego_Group

[2] Statistic information on Lego taken from: https://www.statista.com/chart/24021/revenue-and-net-profit-of-the-lego-group/ and https://www.lego.com/en-us/aboutus/news/2020/september/interim-results/

[3] Statistics information on IoT taken from: https://www.statista.com/statistics/1101442/iot-number-of-connected-devices-worldwide/

[4] Worldwide spending on IoT taken from: https://www.statista.com/statistics/668996/worldwide-expenditures-for-the-internet-of-things/

[5] Global climate change information taken from: https://climate.nasa.gov/

[6] Information on Plastic pollution taken from: https://www.weforum.org/agenda/2022/01/plastic-pollution-climate-change-solution/

1.3. Research Aims

To analyze the business objectives, three research aims and their respective questions and objectives has been identified and addressed in this research proposal aimed to solve Lego's issues above. The following three research aims proposed are:

I.   R.A.1: To accurately understand what products (physical, digital, or blended) would appeal to children in the next 5 to 10 years.

II. R.A.2: To research what marketing strategies, if any, they could deploy to promote the use of recycled plastic bottles in brick production.

III. R.A.3: Using market research to help Lego to decide whether to expand their retail store network and, if so, where.

1.4. Research Questions

To accurately understand individual research aims, research questions have been identified for each research aims to reduce inaccurate data collected. The following proposed research questions for individual research aims are:

To address R.A.1:

I.   R.Q.1: How does the annual household expenditure on Lego and appeal rating on physical products differ between demographic groups of consumers and brand names?

II. R.Q.2: What is the extent of the impact on the level of internet usage that affects customer's purchasing preference towards digital products?

III. R.Q.3: What type of innovation does customer expects in Lego?

To address R.A.2:

I.   R.Q.4: What is the customer's perception towards environmental issue regarding the use of plastic bottles?

II. R.Q.5: What are the marketing techniques that can be used to attract large audience of customers on environmentally friendly campaign?

To address R.A.3:

I.   R.Q.6: What are the most important customer value proposition?

II. R.Q.7: How does customer satisfaction rating differ towards types of shopping experience?

1.5. Research Objectives

To further enhance the effectiveness of insights gathered from respective research aims, research objectives have been proposed for each individual research aims. The following research objectives for individual research aims are:

To address R.Q.1:

   I. <u>R.O.1</u>: Examine the relationship of customer's annual household expenditure on Lego across different demographical variables.
   II. <u>R.O.2</u>: Investigate the extent to which appeal rating on physical products differs across different brands and age-group of children.

To address R.Q.2, R.Q.3 and R.Q.4 respectively:

   I. <u>R.O.2</u>: Examine the relationship between the level of internet usage and customer's purchasing preference towards digital products.
   II. <u>R.O.3</u>: Investigate the key innovations that contribute to customers' expectation in Lego.
   III. <u>R.O.4</u>: Examine factors that can influence current lifestyle and behaviour of customers towards recycling of plastic bottles.

To address R.Q.5:

   I. <u>R.O.5.1</u>: Examine types of advertisement that effectively attracts customer's awareness.
   II. <u>R.O.5.2</u>: Examine the effect of new advertising campaign on customer's concern on recycling plastic bottles.

To address R.Q.6 and R.Q.7 respectively:

   I. <u>R.O.6</u>: Identify customer's value proposition that are important in contributing to decision on expansion of retail store.
   II. <u>R.O.7</u>: Examine the extent to which customer satisfaction rating differs across different type of shopping experience.

2. Methodology

2.1. Research Design and Approach

In this research proposal, 3 research design being descriptive, exploratory, casual and 2 approaches being qualitative and quantitative research are allocated to each research questions to facilitate data collection to solve Lego's issues above. Firstly, descriptive design has been proposed to segment and identify target audience and their purchasing pattern in the toy industry. Followed by exploratory design to investigate the key innovations that contribute to customer's expectation in Lego and lastly, casual design to examine the before/aftereffects of new advertising campaign on customer's concern on recycling plastic bottles which allow us to understand their awareness level towards recycling plastic[7].

Online Survey Questionnaires and Online Focus-Group is proposed to gather quantitative and qualitative data. To minimize non-response error, dual-moderator will be included in Focus-Group discission to facilitate smooth flow session and ensures issues addressed, and monetary incentives will be given out in exchange for their participation.

2.2. Timeline

As Lego provided four-month window period to execute market research, multiple cross-sectional design is proposed as it is cheap and easy to gather initial data from different groups of participants simultaneously to identify any differences and correlations in their purchasing preference and behaviour than longitude design where it surveys same characteristic over time which is time-consuming in a dynamic market.

2.3. Sources of Data

Both primary and external secondary data will be used as the source of data collection for Focus-Group and Online Surveys to provide more accurate analysis. In section 3.1 and 3.2 below provides details on individual variables and their data types.

_____

[7] Correlation vs Causation and Explanatory taken from: https://www.scribbr.com/methodology/correlation-vs-causation/#causal-research

2.4. Data Analysis

The following proposed statistical techniques generated from SPSS is used to analysis the data collection for each research objective is:

     I.   R.O.1.1 – Multiple Linear Regression
     II.   R.O.1.2 – 2-way ANOVA
    III.   R.O.2 – Contingency Table
    IV.   R.O.3 – Not Applicable (Focus-Group)
     V.   R.O.4 – Discriminant Analysis
    VI.   R.O.5 – Contingency Table
   VII.   R.O.6 – Paired-t test
  VIII.   R.O.7 – Not Applicable (Focus-Group)

3. Data & Sampling Design

3.1. Primary Data

S/N	Variables	Data
1	Demographics: I. Annual Household Income II. Region of residence III. Age IV. Age-Group V. Gender VI. Annual Household Expenditure on LEGO VII. Household Size	I. Continuous II. Categorical Nominal III. Continuous IV. Categorical Nominal V. Categorical Nominal VI. Continuous VII. Continuous
2	Ratings-related score: I. Appeal rating on physical product II. Attitude score towards current price of plastic bottles III. Importance rating attached to current environmental-friendly campaigns. IV. Customer's level of concern before and after new advertising campaign. V. Customer satisfaction rating	Continuous – Likert Scale (1 to 5)
3	Preference, Level and Frequency-related score: I. Purchasing preference towards digital products II. Level of internet Usage III. Perception level towards recycling plastic bottles IV. Frequency of monthly encounters of advertisement.	Categorical Ordinal

Table 1: Primary Data

3.2. Secondary Data (External)

S/N	Variables	Data
4	Brand names	Categorical Nominal
5	Types of advertisement encountered	
6	Types of shopping experience	

Table 2: Secondary Data

3.3. Sampling Frame

Due to the pricing strategy of Lego's product, stratified random sampling has been proposed for non-focus group studies to divide target population into samples based on specific region that has the highest per gross domestic product (GDP) capita[8]. This will provide systematic and accurate analysis in determining the sample size to be used for research proposal that is relevant to Lego's issue. The following proposed stratified regions are:

G7 countries	Non-G7 countries
I. France	I. India
II. Germany	II. Korea
III. Canada	
IV. Italy	
V. Japan	
VI. United States	
VII. United Kingdom	

Table 3: G7 and Non-G7 Countries

Followed by convenience sampling which will be used on focus-group studies to investigate customer's expectation on innovations in Lego and identify their value proposition that is important to them. This facilitate data collection to be readily available and it is inexpensive to create due to low cost and lesser restrictions to follow in a dynamic environment[9].

_____

[8] Level of GDP per capita and productivity taken from: https://stats.oecd.org/index.aspx?DataSetCode=PDB_LV

[9] Definition, Advantages and Example of Convenience Sampling taken from:
https://www.questionpro.com/blog/convenience-sampling/

3.4. Sampling Techniques

The proposed sampling techniques will be conducted through online for both Surveys Questionnaires and Focus-Groups as it is more convenient and efficient to reach out to global participants. Online Surveys Questionnaires will be done through email as it provides a degree of personalization at low cost that would likely reduce non-response rate[10]. For Online Focus Group, it will be done through video conferencing. This allows access to broad participants easily at low cost which improves quality of discussion, participate rate and communication due to comfort at the expense of their time.

3.5. Sample Size

Below is the calculation on minimum sample size that is required for sampling where United States has been used as a baseline reference for sample size calculation. According to study, an estimated of 41.94% completion rate with least 15 questions[11]. Thus, assumption is made on 95% confidence level with margin error ($e = 0.05$) and an estimated of 60% survey completion rate ($\pi = 0.6$) due to monetary incentives provided that increase its effectiveness by 18%[12]. The calculation is represented by:

$$n_k \geq \frac{\left(Z_{\frac{\alpha}{2}}\right)^2 \pi(1-\pi)}{e^2}$$

$$n_c = \frac{nN}{N + n - 1}$$

Where n_k and n_c represents the minimum requirement on number of participants and its finite population corrector respectively. Due to non-replacement of 3,000 proposed sample population by Lego, finite population correction (n_c) is recommended as adjustment would reduce the required sample size for small population[13]. 1,000 (33.33%) proposed participants will be selected as Non-Lego customers and remaining 2,000 (66.66%) are Lego customers.

_____

[10] Distribution of Online Survey taken from: https://www.satrixsolutions.com/blog/top-five-ways-distribute-online-survey/

[11] Survey Completion Rate over 80% taken from: https://survicate.com/surveys/survey-completion-rate/

[12] Survey response rate and completion taken from:
https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-017-0353-1

[13] Finite Population Correction Factor taken from:
https://stats.libretexts.org/Bookshelves/Applied_Statistics/Book%3A_Business_Statistics_(OpenStax)/07%3A_The_Central_Limit_Theorem/7.04%3A_Finite_Population_Correction_Factor

3.5.1. Non-Lego customers

The calculation for Non-Lego customers is represented by:

$$n_1 \geq \frac{(1.96)^2 0.6(1-0.6)}{(0.05)^2} \approx 368.79 \approx 368 \ (rounded \ down \ to \ nearest \ whole \ number)$$

$$n_c = \frac{368(1,000)}{1,000 + 368 - 1} \approx 269.20 \approx 269 \ (rounded \ down \ to \ nearest \ whole \ number)$$

Hence, final adjustment made on minimum sample size required for Non-Lego customers to be 269.

3.5.2. Lego customers

To identify potential Lego customers, sample size has been statistically adjusted to include the incidence rate of $\frac{1}{0.5*0.6} = \frac{1}{0.3}$, based on estimated of 50% adult population in United States that has possible income[14]. The calculation for Lego customers is represented by:

$$n_2 \geq Incidence \ rate \ * \ \frac{\left(Z_{\frac{\alpha}{2}}\right)^2 \pi(1-\pi)}{e^2}$$

$$n_2 \geq \frac{1}{0.3} * \frac{(1.96)^2 0.6(1-0.6)}{(0.05)^2} \approx 1229.31 \approx 1229 \ (rounded \ down \ to \ nearest \ whole \ number)$$

$$n_c = \frac{1229(2,000)}{2,000 + 1229 - 1} \approx 761.46 \approx 761 \ (rounded \ down \ to \ nearest \ whole \ number)$$

Hence, the final adjustment made on minimum sample size required for Lego customers is 761. Additionally, 100 of Non-Lego and 200 Lego customers will be allocated to Focus-Group studies.

4. Data Analysis and Findings

4.1. Research Objective 1.1

Multiple linear regression is proposed to examine the relationship of customer's Annual Household Expenditure on Lego across different demographical variables. The dependent variable: Annual Household Expenditure on Lego, and the independent demographic variables: Annual Household Income, Region of Residence[15].

_____

[14] Statistics of US Age Distribution from: https://www.statista.com/statistics/270000/age-distribution-in-the-united-states/

[15] Refer to section 3.1 on their respective data types for each variable specified in this research proposal.

The following dummy values for independent categorical variables are:

Region of Residence		Gender	
• United States	0	• Female	0
• United Kingdom	1	• Male	1
• Japan	2		
• Italy	3		
• Canada	4		
• Germany	5		
• France	6		
• India	7		
• Korea	8		

Table 4: Independent categorical dummy variables

The formulated hypothesis is:

- H_0 : There is no relationship between annual household expenditure on LEGO and following demographic variables.
- H_1: There is a relationship between annual household expenditure on LEGO and following demographic variables.

The following multiple linear regression equation is:

$$Annual\ Household\ Expenditure = \beta_0 + \beta_1 AnnualHHIncome + \beta_2 Region + \beta_3 Age + \beta_4 Gender + \varepsilon_i$$

$$Annual\ \widehat{Household\ Expenditure} = \widehat{\beta_0} + \widehat{\beta_1} AnnualHHIncome + \widehat{\beta_2} Region + \widehat{\beta_3} Age + \widehat{\beta_4} Gender$$

- $\widehat{\beta_0}$: Intercept value of Annual Household Expenditure when variables are zero.
- $\widehat{\beta_j}$: Changes in Annual Household Income when X_j changes by a unit while other independent variables values are held constant. Subscript j, used to identify respective independent variables which takes a value of 1 to n number of independent variables.

Table 5 below shows the proposed t-test, P-value (Sig.) and 95% confidence interval for B(Beta) for measuring individual's statistical significance to the model. This will provide insights in identifying variables that significantly influences Annual Household Expenditure on Lego and often researcher uses standardized coefficients for different units to avoid significant changes to the model.

In table 6, Adjusted R^2 that adjusts for non-significant variables[16] will be used to measure the fit of model to determine significant changes of individual variables that affects Annual Household Expenditure. This will allow Lego to understand customer's lifestyle and purchasing patterns and facilitate market segmentation.

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std.Error	Beta			Lower Bound	Upper Bound
1	(Constant)							
	AnnualHHIncome							
	Region							
	Age							
	Gender							

a. Dependent variable: Annual Household Expenditure.

Table 5: Coefficient and t-test table

Model Summary

Model	R	R Square	Adjusted R square	Std. Error of the Estimate
1	^{a}			

a. Predicators: Constant, AnnualHouseholdIncome, Region, Age, Gender

Table 6: R^2 and Adjusted R^2 table

_____

[16] Definition of Adjusted R^2 taken from: https://www.statisticshowto.com/probability-and-statistics/statistics-definitions/adjusted-r2/

4.2. Research Objective 1.2

A 2-way ANOVA is proposed to investigate any differences exist among the independent variables: Brand Names and Age-Group, on dependent variables: Appeal rating on physical product. The following formulated hypothesis is:

- H_0 : Appeal rating on physical product is the same across brand names and age-group.
- H_1: Appeal rating on physical product is different brand names and age-group.

The levels of each independent categorical variables are:

Brand names	Age-Group
• LEGO • Bandai Namco • Hasbro • Mattel	• 1-7 years old • 8-15 years old • 16-23 years old • 24 years and above

Table 7: Categorical independent variables.

In table 8 below reflects the Tests of Between-Subjects Effects table generated by SPSS. The proposed F-test is used to identify any significant differences exist in the population mean of Brand names and Age-Group, which is supported by the P-value (Sig.). Furthermore, Interaction term has been added to investigate the effect of independent variables have on different categories of dependent variable and partial eta squared is used to quantify the strength of an effect by independent variable on dependent variable. This facilitates comparison on product type among competitors thereby allowing Lego to understand their competitive position.

Tests of Between-Subjects Effects

Dependent Variable: Appeal rating on physical product

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	[a]					
Intercept						
Brand names						
Age-Group						
Brand names*Age-						
Error						
Total						
Corrected Total						

a. R Squared = ___ (Adjusted R Squared = ___)

Table 8: Test of Between-Subject Effects Table

4.3. Research Objectives 2

Cross-Tabulation (Contingency) generated by SPSS software is proposed to determine if Level of Internet Usage (independent variable) differs by Purchasing Preference towards digital goods (dependent variable). The formulated hypothesis is:

- H_0 : There is no association between Level of Internet Usage and Purchasing Preference towards digital goods.
- H_1: There is an association between Level of Internet Usage and Purchasing Preference towards digital goods.

Level of Internet Usage * Purchasing Preference towards digital goods Crosstabulation

Purchasing Preference

Count		Definitely Not	Less Likely	Neutral	Somewhat Likely	Most Likely	Total
Level of Internet Usage	Never						
	Rarely						
	Sometimes						
	Often						
	All the time						
	Total						

Table 9: Crosstabulation - Level of Internet Usage and Purchasing Preference towards digital goods.

Table 10 below reflects the Chi-Square Tests of association between dependent and independent variable. The null hypothesis is rejected only if the P-value (2-sided) is <0.05, indicating that there is an association between level of internet usage and purchasing preference towards digital goods. The statical test result is reliable if Superscript a. has cells with expected count less than 5 and meet the minimum requirement of expected count.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	a		
Likelihood Ratio			
Linear-by-Linear Association			
N of Valid Cases			

a. __ cells have expected count less than 5. The minimum expected count is __.

Table 10: Chi-Square Tests of association

Table 12 below shows the Symmetric Measure where Kendall's Tau-b will be used to measure strength and direction of association due to its 5x5 categories; Gamma will be used to measure improvement on prediction accuracy of independent and dependent variables. The guidelines for the values and its strength of association are stated in table 11, where negative coefficient indicates negative association

and vice versa. This allows Lego to understand the behaviour of digital and non-digital users and decide whether to expand into digital goods in future.

Strength of Association	Value of Measure of Association
None	0.00
Weak	±0.01 to 0.09
Moderate	±0.10 to 0.29
Strong	±0.30 to 0.99
Perfect	±1.00

Table 11: Strength of Association

Symmetric Measure

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Ordinal by Ordinal	Kendall's tau-b				
	Gamma				
N of Valid Cases					

a. Not assuming the null hypothesis

b. Using the asymptotic standard error assuming the null hypothesis

Table 12: Symmetric Measure of Ordinal Variables

4.4. Research Objectives 3

A qualitative online focus-group will be conducted to investigate the key innovations that contribute to customer's expectation in Lego. Due to rapid shift in market, competition and sustainability becomes tougher. Hence, gathering data through focus-group is important as participants tends to provide real-time and deep insights in innovation ideas or product improvements which one could not have discussed elsewhere. This allows Lego to understand actual customers demand and gain competitive advantage.

4.5. Research Objectives 4

Multiple (3-group) Discriminant Analysis generated by SPSS software is proposed to examine significance factors and their respective classification accuracy of independent variables that influences between categories of dependent variables. The following variables are:

Independent Variables	Dependent Variables
• Attitude score towards current price of plastic bottles • Annual Household Income • Age • Household Size • Importance rating attached to current environmental-friendly campaigns	• Perception level towards recycling plastic bottles

Table 13: Independent and Dependent variables

Group Statistics

Perception level towards recycling plastic bottles		Mean	Std. Deviation
Less aware	Attitude score towards current price of plastic bottles		
	Annual Household Income		
	Age		
	Household Size		
	Importance rating attached to current environmental-friendly campaigns		
Somewhat aware	Attitude score towards current price of plastic bottles		
	Annual Household Income		
	Age		
	Household Size		
	Importance rating attached to current environmental-friendly campaigns		
Highly aware	Attitude score towards current price of plastic bottles		

	Annual Household Income		
	Age		
	Household Size		
	Importance rating attached to current environmental-friendly campaigns		
Total	Attitude score towards current price of plastic bottles		
	Annual Household Income		
	Age		
	Household Size		
	Importance rating attached to current environmental-friendly campaigns		

Table 14: Group Statistic of Multiple Discriminant Analysis

To analyse these statistical differences above, the following minimum (3-1 groups, 5 predicators) canonical discriminant function is proposed:

$$D = \beta_0 + \beta_1 AttitudeScore + \beta_2 AnnualHHIncome + \beta_3 Age + \beta_4 HHSize + \beta_5 ImptRating$$

- D = Discriminant score of Perception level towards recycling plastic bottles
- β_k = Discriminant coefficient for respective independent variable k where it ranges from 1 to k.

Often researchers use Structure Matrix in table 17 as the coefficient corresponds to correlation with each independent variables and discriminant function(s), resulting in producing more accurate model than unstandardized and standardized coefficients in table 15 and 16. Higher correlations indicates higher importance of independent variables, but less importance variables may be important in other functions, thus further analysis is proposed below.

Canonical Discriminant Function Coefficients

	Function	
	1	2
Attitude score towards current price of plastic bottles		
Annual Household Income		
Age		
Household Size		
Importance rating attached to current environmental-friendly campaigns		
(Constant)		

Unstandardized coefficients

Table 15: Canonical Discriminant Function Coefficients

Standardized Canonical Discriminant Function Coefficients

	Function	
	1	2
Attitude score towards current price of plastic bottles		
Annual Household Income		
Age		
Household Size		
Importance rating attached to current environmental-friendly campaigns		

Table 16: Standardized Canonical Discriminant Function Coefficients

Structure Matrix

	Function	
	1	2
Attitude score towards current price of plastic bottles		
Annual Household Income		
Age		
Household Size		
Importance rating attached to current environmental-friendly campaigns		
(Constant)		

Table 17: Structure Matrix

In table 18 and 19 below reflects the Eigenvalues and Wilk's Lambda to assess discriminatory ability of the discriminant function(s) proposed. In table 18, larger eigenvalues provide better discrimination, and this is supported by the % of variance explained by each function, and their large canonical correlation indicates higher discriminating power between the 3 groups. To measure if discriminant power is statistically significant, associated Chi-Square will be used in table 19 – Wilk's lambda and its following formulated hypothesis is:

- H_0 : The discriminant function does not discriminate between the groups.
- H_1: The discriminant function discriminates between the groups.

If the Wilk's lambda is close to 0, Chi-Square statistic is more than Chi-Square critical value, and P-value < 0.05, the null hypothesis will be rejected and thus, discriminant function will be considered statistically significant in discriminating between groups.

Eigenvalues

Function	Eigenvalues	% of Variance	Cumulative %	Canonical Correlation
1	a			
2	a			

a. First 2 canonical discriminant functions were used in the analysis.

Table 18: Eigenvalues

Wilk's Lambda

Test of Function(s)	Wilk's Lambda	Chi-Square	Df	Sig.
1 through 2				
2				

Table 19: Wilk's lambda

Table 20 assess the significant differences of individual independent variables that is potential useful for the discriminant model. The lower the coefficient of Wilks lambda, the closer to total discrimination, and this is supported by the test of significance where F-test is less than F-critical value and P-value <0.05.

Tests of Equality of Group Means

	Wilk's Lambda	F	df1	df2	Sig.
Attitude score towards current price of plastic bottles					
Annual Household Income					
Age					
Household Size					
Importance rating attached to current environmental-friendly campaigns					

Table 20: Tests of Equality of Group Means

In table 21 below summarised the Classification Matrix used to determine the accuracy of classification where percentage of cases are correctly classified. The HIT ratio is based on validation sample of cases not selected which are correctly classified over total cases not selected. Based on researchers, the analysis is significant if HIT ratio improvement over chances is at least greater than 25% than that obtained by chance. This allows Lego to understand customer's perception towards environmental issue involving use of plastic bottles.

Classification Results[a,b]

			Perception level towards recycling plastic bottles	Not aware	Somewhat Aware	Highly Aware	Total
Cases Selected	Original	Count	Not aware				
			Somewhat Aware				
			Highly Aware				
		%	Not aware				
			Somewhat Aware				
			Highly Aware				
Cases Not Selected	Original	Count	Not aware				
			Somewhat Aware				
			Highly Aware				
		%	Not aware				
			Somewhat Aware				
			Highly Aware				

Predicted Group Membership

a. ___ of selected original grouped cases correctly classified.

b. ___ of unselected original grouped cases correctly classified.

Table 21: Classification Results

4.6. Research Objectives 5.1

Cross-tabulation (Contingency table) is proposed to examine if types of advertisement encountered (Independent variable) differs by frequency of monthly encounters of advertisement (Dependent variable). The formulated hypothesis is:

- H_0 : There is no association between Type of advertisement encountered and Frequency of monthly encounters of advertisement.
- H_1: There is an association between Type of advertisement encountered and Frequency of monthly encounters of advertisement.

Type of advertisement encountered * Frequency of monthly encounters of advertisement Crosstabulation

Count		Frequency of monthly encounters of advertisement					
		Rarely	Seldom	Neutral	Often	Frequent	Total
Types of advertisement encountered	Social media						
	Newspaper						
	Online news site						
	TV						
	Radio						
	Email						
	Telephone						

*Table 22: Crosstabulation table of Type of advertisement encountered * Frequency of monthly encounters of advertisement*

In table 23 below shows the Chi-Square tests of association between independent and dependent variables. If Chi-Square test statistic is greater than Chi-Square critical value and P-value <0.05, null hypothesis will be rejected which indicates that there is a significant association between types of advertisement encountered and its frequencies. The statical test result is reliable if Superscript a. has cells with expected count less than 5 and meets the minimum requirement of expected count.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	a		
Likelihood Ratio			
Linear-by-Linear Association			
N of Valid Cases			

a. __ cells have expected count less than 5. The minimum expected count is __.

Table 23: Chi-Square Tests of Type of advertisement encountered * Frequency of monthly encounters

In table 24 below reflects the on Proportionate Reduction in Error Measures (PRE) to examine overall improvement when prediction is done in both direction (Dependent variable modelled using both variables) by using symmetry lambda. If one of the variables has higher overall improvement and is statistically significant where P-value <0.05, it will be used as dependent variables. This allows Lego to implement effective marketing strategy to attract customer's awareness.

Directional Measures

			Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Nominal by Nominal	Lambda	Symmetric				
		Type of advertisement encountered				
		Frequency of monthly encounters of advertisement				

a. Not assuming the null hypothesis

b. Using the asymptotic standard error assuming the null hypothesis

Table 24: Directional Measures

4.7. Research Objectives 5.2

Paired t-test is proposed to examine the effect of new advertising campaign on customer's concern towards recycling plastic bottles. Experimental Static Group design will be conducted first by allocating pre-test on new advertisement effect to experimental group and post-test on control group without new advertisement effect to compare its differences.

<div style="text-align:center">

Static group design

EG: X O_1

CG: O_2

</div>

O_1: Pre-test of experiment group

O_2: Post-test control group

X: Exposure to new advertisement

Table 25: Experimental Static group design

Table 26 and 27 reflects the Paired Sample Statistic and Tests on customer's level of concern before and after new advertising campaign. As Lego's required sample size is large, assumption on normality can be ignored. The following formulated hypothesis is:

- $H_0 : \mu_d = 0$
- $H_1 : \mu_d \neq 0$

If the t-test value is more than t-critical value, P-value <0.05 and value 0 not within 95% Confidence Interval, null hypothesis will be rejected and conclude that the customer's level of concern before and after new advertising campaign are significantly different. This allows Lego to understand the effectiveness of marketing strategy.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	BEFORE				
	AFTER				

Table 26: Paired Samples Statistics

Paired Samples Test

Paired Differences

		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Differences		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	BEFORE - AFTER								

Table 27: Paired Samples Test

4.8. Research Objectives 6

Similar to section 4.4 where Online Focus-Group will be conducted to identify customer's value proposition that are important in contributing to decision on expansion of retail store. This allows the company to collect relevant information and enhance the accuracy of decision-making on retail expansion that would reduce long-term business losses.

4.9. Research Objectives 7

Independent t-test is proposed to examine the extent to which customer satisfaction rating differs across different types of shopping experience where independent variable (uncorrelated): Types of shopping experience, and dependent variable: Customer satisfaction rating. As sample size required is large, assumption on normality is ignored. In table 29 below, Levene's Test will be conducted first to determine if independent variables have equal variances to meet the assumption of homogeneity before analysing t-test equality of means. The following hypothesis for Levene's Test is:

- $H_0 : \sigma_1^2 - \sigma_2^2 = 0$
- $H_1 : \sigma_1^2 - \sigma_2^2 \neq 0$

σ_1^2: Population variance of customer's satisfaction rating on online shopping experience.
σ_2^2: Population variance of customer's satisfaction rating on offline shopping experience.

If the F-test is more than the F-critical value (rejection region) and P-value (Sig.) <0.05, the null hypothesis will be rejected and conclude that population variances of customer's satisfaction rating on both online offline shopping experience is not equal.

The following formulated hypothesis for t-test is:

- $H_0 : \mu_1 - \mu_2 = 0$
- $H_1 : \mu_1 - \mu_2 \neq 0$

μ_1: Population mean of customer's satisfaction rating on online shopping experience.

μ_2: Population mean of customer's satisfaction rating on offline shopping experience.

If t-test value is more than t-critical value of rejection region, P-value (Sig. 2tailed) <0.05 and value 0 not within 95% Confidence Interval, the null hypothesis will be rejected and conclude that population mean of customer's satisfaction rating on both online offline shopping experience is not equal. This allows Lego to accurately decide on its store expansion and its location.

Group Statistics

		N	Mean	Std. Deviation	Std. Error Mean
Customer satisfaction rating	Online				
	Offline				

Table 28: Independent Group Statistics

Independent Samples Test

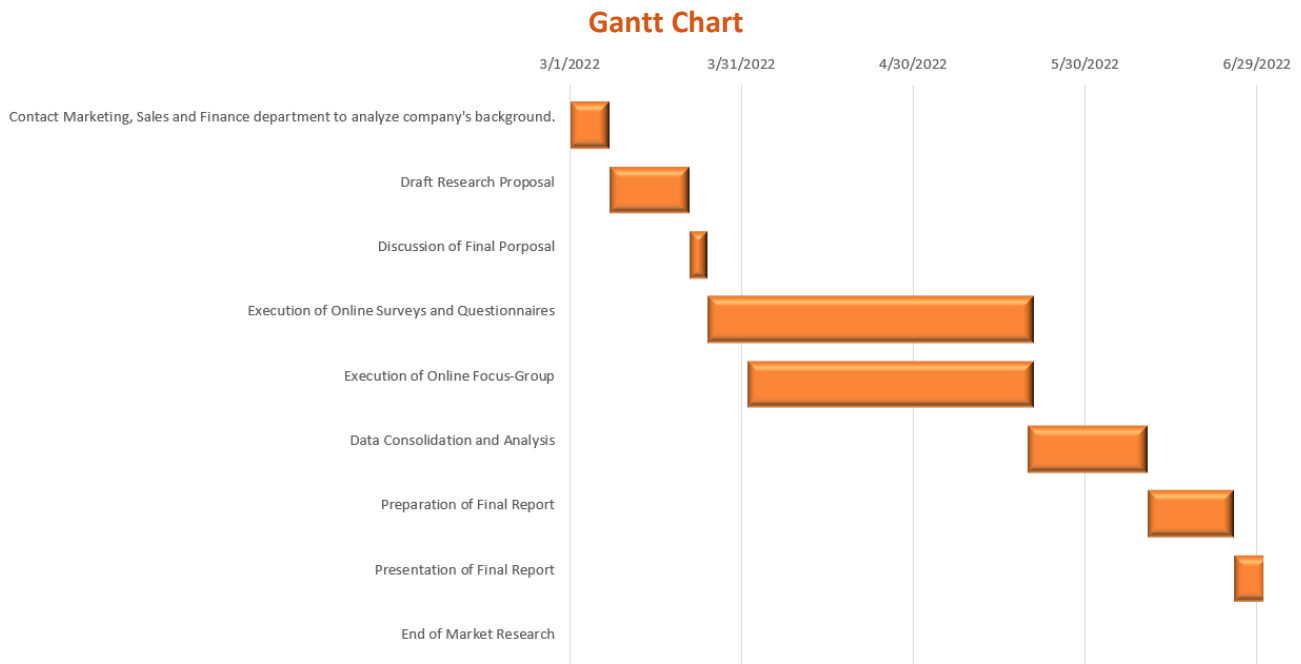|		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Customer Satisfaction Rating	Equal variances assumed									
	Equal variances not assumed									

Table 29: Independent Samples Test

5. Timescale

Gantt Chart

Figure 1: Timescale

6. Budget Table

| Activity	Estimated cost (£)	
Online Survey Incentives	20,000	10£ per participants
Online Focus-Group Incentives	20,000	20£ per participants
Cost of Outsourcing market research agent	100,000	Inclusive of contract salary for work done and agent fee
Data collection, preparation, analysis and report	40,000	Software cost such as SPSS premium, cloud storage and techonology required.
Administrative Cost	30,000	Legal proceedings/ expedited paperwork for collection of data (consumer privacy, security etc.)
Miscellaneous expenses	10,000	travel, allowance, etc
20% VAT/GST Tax	44,000	Additional taxes for services used.
Total	264,000	
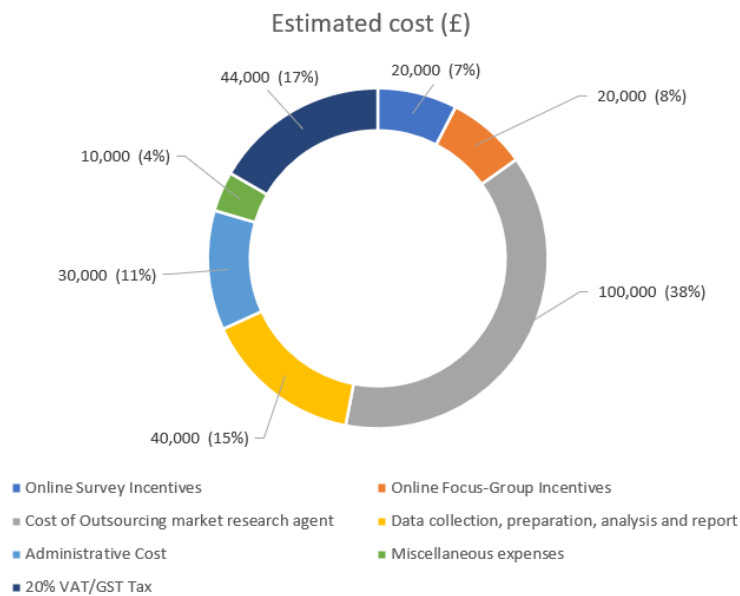
Figure 2: Budget breakdown

7. Recommendations

7.1. Data Intelligence

Due to rise in IoT, big data phenomenon may occur where there are excessive data exchanged online. Lego can consider using data intelligence tool such machine learning algorithms to analyse relationship and predict customer's demand due to its speed of processing data. This will improve quality and accuracy of data collection and enhances in decision-making, resulting increasing market growth and sustainability in the long run.

7.2. Improve Customer Relationship

It is important to improve customer relationship through online event such as product/service demonstration and feedbacks as it will retain existing or potential customer, increase participation rate, quality of data gathered and reduce response error from both online surveys and discussion.

-End of Report-

*Word Count **2997** –*

Excluding titles, tables, tables of content, equations, executive summary, and references.

8. References for Research:

En.wikipedia.org. 2022. *The Lego Group - Wikipedia*. [online] Available at: <https://en.wikipedia.org/wiki/The_Lego_Group> [Accessed 12 February 2022].

Statista Infographics. 2022. *Infographic: Brick by Brick: Lego's Growing Toy Empire*. [online] Available at: <https://www.statista.com/chart/24021/revenue-and-net-profit-of-the-lego-group/> [Accessed 12 February 2022].

Statista Infographics. 2022. *Infographic: Brick by Brick: Lego's Growing Toy Empire*. [online] Available at: <https://www.statista.com/chart/24021/revenue-and-net-profit-of-the-lego-group/> [Accessed 12 February 2022].

Statista. 2022. *Global IoT and non-IoT connections 2010-2025 | Statista*. [online] Available at: <https://www.statista.com/statistics/1101442/iot-number-of-connected-devices-worldwide/> [Accessed 12 February 2022].

Statista. 2022. *Internet of Things spending worldwide 2023 | Statista*. [online] Available at: <https://www.statista.com/statistics/668996/worldwide-expenditures-for-the-internet-of-things/> [Accessed 12 February 2022].

Climate Change: Vital Signs of the Planet. 2022. *Climate Change: Vital Signs of the Planet*. [online] Available at: <https://climate.nasa.gov/> [Accessed 12 February 2022].

World Economic Forum. 2022. *We know plastic pollution is bad – but how exactly is it linked to climate change?*. [online] Available at: <https://www.weforum.org/agenda/2022/01/plastic-pollution-climate-change-solution/> [Accessed 12 February 2022].

Base, K. and causation, C., 2022. *Correlation vs causation*. [online] Scribbr. Available at: <https://www.scribbr.com/methodology/correlation-vs-causation/#causal-research> [Accessed 12 February 2022].

Us.sagepub.com. 2022. [online] Available at: <https://us.sagepub.com/sites/default/files/upm-binaries/55588_Chapter_1_Sample_Creswell_Research_Design_4e.pdf> [Accessed 12 February 2022].

Base, K. and study?, W., 2022. *What is a cross-sectional study?*. [online] Scribbr. Available at: <https://www.scribbr.com/methodology/cross-sectional-study/> [Accessed 12 February 2022].

Stats.oecd.org. 2022. *Level of GDP per capita and productivity*. [online] Available at: <https://stats.oecd.org/index.aspx?DataSetCode=PDB_LV> [Accessed 12 February 2022].

QuestionPro. 2022. *Convenience Sampling: Definition, Advantages and Examples | QuestionPro*. [online] Available at: <https://www.questionpro.com/blog/convenience-sampling/> [Accessed 12 February 2022].

Satrix Solutions. 2022. *Top Five Ways to Distribute Your Online Survey - Satrix Solutions*. [online] Available at: <https://www.satrixsolutions.com/blog/top-five-ways-distribute-online-survey/> [Accessed 12 February 2022].

Small Business - Chron.com. 2022. *Important Skills That a Focus Group Moderator Should Possess*. [online] Available at: <https://smallbusiness.chron.com/important-skills-focus-group-moderator-should-possess-69910.html> [Accessed 12 February 2022].

Statistics LibreTexts. 2022. *7.4: Finite Population Correction Factor*. [online] Available at: <https://stats.libretexts.org/Bookshelves/Applied_Statistics/Book%3A_Business_Statistics_(OpenStax)/07%3A_The_Central_Limit_Theorem/7.04%3A_Finite_Population_Correction_Factor> [Accessed 12 February 2022].

Perzynska, K., 2022. [online] Available at: <https://survicate.com/surveys/survey-completion-rate/> [Accessed 12 February 2022].

Yu, S., Alper, H.E., Nguyen, AM. *et al.* The effectiveness of a monetary incentive offer on survey response rates and response completeness in a longitudinal study. *BMC Med Res Methodol* **17,** 77 (2017). [online] Available at: < https://doi.org/10.1186/s12874-017-0353-1 > [Accessed 12 February 2022].

O'Neill, A., 2022. *U.S.: Age distribution | Statista*. [online] Statista. Available at: <https://www.statista.com/statistics/270000/age-distribution-in-the-united-states/> [Accessed 12 February 2022].

Statistics How To. 2022. *Adjusted R2 / Adjusted R-Squared: What is it used for?*. [online] Available at: <https://www.statisticshowto.com/probability-and-statistics/statistics-definitions/adjusted-r2/> [Accessed 12 February 2022].

Youtube.com. 2022. [online] Available at: <https://www.youtube.com/watch?v=QkH13mO0uog> [Accessed 19 February 2022].

Ltd, A., 2022. *Educational Toy Market CAGR 8.2% by 2022-2028 | Industry Size, Shares, Revenue, Key Players, Regions | Report by Absolute Reports*. [online] GlobeNewswire News Room. Available at: <https://www.globenewswire.com/news-release/2022/01/20/2370495/0/en/Educational-Toy-Market-CAGR-8-2-by-2022-2028-Industry-Size-Shares-Revenue-Key-Players-Regions-Report-by-Absolute-Reports.html> [Accessed 19 February 2022].

Staff, R., 2022. *Prices for post-consumer plastic bales continue to rise - Plastics Recycling Update*. [online] Plastics Recycling Update. Available at: <https://resource-recycling.com/plastics/2021/02/17/prices-for-post-consumer-plastic-bales-continue-to-rise/> [Accessed 19 February 2022].

UNEP - UN Environment Programme. 2022. *100 Days to #BeatPlasticPollution for #CleanSeas*. [online] Available at: <https://www.unep.org/events/campaign/100-days-beatplasticpollution-cleanseas#:~:text=100%20Days%20to%20%23BeatPlasticPollution%20is,support%20for%20policies%20to%20change> [Accessed 19 February 2022].

9. Surveys Template

Surveys and Questionnaires

The Lego Group is a top private company which specialize in producing and selling plastic brick toys to wide range of customers as well as offering variety of collections and collaborations with popular movies such as Marvel, DC comics and Harry Potter. It's brand valuation is approximately 5.4 billion U.S. dollars and thus, labelled as one of the top 10 toy brands worldwide. Due to rapid changes in natural environment, technology and customer's demand in a dynamic market, the company would like to understand the customer's expectation and demand, as well as concern towards impact of production of plastics on natural environment.

Monetary incentive of $5 will be awarded upon completion of survey and will be mailed to you 3-5 working days after validating your answers. Thank you for your time and participation.

1. What is your age? *

Short answer text

2. What is your gender? *

○ Male

○ Female

○ Other...

3. Which country are you from? (Current residency) *

Short answer text

4. What is your household size? *

Short answer text

5. What is your annual household income? *

Short answer text

6. Have you purchased LEGO products before? *

◯ Yes

◯ No (Skip question 7)

7. What is your annual household expenditure on LEGO? *

Short answer text

8. Please rate the following brands in terms of appealingness on their respective physical products. *

	1	2	3	4	5	6	7	8	9	10
LEGO	○	○	○	○	○	○	○	○	○	○
Bandai…	○	○	○	○	○	○	○	○	○	○
HasBro	○	○	○	○	○	○	○	○	○	○
Mattel	○	○	○	○	○	○	○	○	○	○

9. How often do you use the internet? *

○ Never

○ Rarely

○ Sometimes

○ Often

○ All the time

10. How likely would you purchase digital goods with the given amount of internet usage that you have provided earlier? *

○ Definitely not

○ Less Likey

○ Neutral

○ Somewhat Likely

○ Most Likely

11. How well-aware are you when comes to recycling plastic bottles? *

	1	2	3	
Less aware	○	○	○	Highly aware

12. Based on research, the current market price of a PET plastic bottles is 8.63 cents/pound(450grams). In your opinion, please rate the score based on how you feel towards the price stated above. *

	1	2	3	4	5	
Cheap	○	○	○	○	○	Expensive

13. With the current environmental friendly campaigns on single-used plastic addressed by United Nations Environment Programme (UNEP), please rate the importance that you feel on current campaigns. *

	1	2	3	4	5	
Not important	○	○	○	○	○	Very important

14. What types of advertisement do you most often encounter and how frequent on a monthly basis? (You can pick more than 1) *

	Rarely	Seldom	Neutral	Often	Frequent
Social media - ...	○	○	○	○	○
Newspaper	○	○	○	○	○
Online news site	○	○	○	○	○
TV	○	○	○	○	○
Radio	○	○	○	○	○
Email	○	○	○	○	○
Telephone	○	○	○	○	○

15. What is your preferred type of shopping experience? *

○ Online

○ Offline

16. Please rate the satisfaction level of the chosen type of shopping experience. *

	1	2	3	4	5	
Least satisfied	○	○	○	○	○	Very satisfied

Surveys and Questionnaires

The Lego Group is a top private company which specialize in producing and selling plastic brick toys to wide range of customers as well as offering variety of collections and collaborations with popular movies such as Marvel, DC comics and Harry Potter. It's brand value is approximately 5.4 billion U.S. dollars and thus, labelled as one of the top 10 toy brands worldwide. Due to rapid changes in natural environment, technology and customer's demand in a dynamic market, the company would like to understand the customer's expectation and demand, as well as concern towards environmental friendly of Lego's production.

1. With the current advertising campaign on recycling plastic bottles, please rate your concern *
level towards recycling of plastic bottles.

	1	2	3	4	5	
Not concern	○	○	○	○	○	Extremely concern

2. Please take a look at the advertisement below and provide a revised version of rating on *
concern level towards recycling plastics bottles.

	1	2	3	4	5	
Not concern	○	○	○	○	○	Extremely concern