



ANLY482 – ANALYTICS PRACTICUM

GROUP 21: SHIKSHA

BRAINSMITH

PROPOSAL REPORT

PREPARED FOR: PROF. KAM TIN SEONG

PREPARED BY: EUNICE WANG, POH WEN JIE, PRAJVAL JUHNJUHNWALA

# Project Overview

Brainsmith is a children’s education company based in India that designs and delivers premium educational, learning products and content for early childhood learning and healthy brain development. Their products are targeted at families of higher economic standing with young children and besides having their own website to sell their goods and display their catalogue, they sell their products on various online e-commerce merchants and engages in digital marketing.

Brainsmith has been operating for over a year but have conversion rates that are lower the industry average. With the data provided, we plan to conduct our analysis to identify traffic patterns and understand customers’ navigation patterns on the website that leads to either purchases or exiting the website.

# Project Motivation

The establishment of Brainsmith was done to enable parents, teachers, and educationists to empower children using the best and most advanced learning tools and methods.

Brainsmith has attached tracking modules to all their online advertisements. As a result, the company has collected variegated amounts of substantial data - but has not utilised this data for further analysis to improve the business.

With as many as 300 clicks per day, Brainsmith is enjoying increasing brand awareness in this niche market, in a geography with strong growth potential. However, they are unable to convert them into purchases, resulting in a low conversion rate.

They also have a relatively large ad spend, and a thriving potential market with excellent quality products – all of which are goldmines waiting for the correct information and analysis so that better policies can be implemented for increased sales and profits, as well as created most amount of buzz.

Hence, our client and the team believe that early optimisation and quick outreach to tap into the market could help with a long-term steady growth in sales. Identification of the potential in the markets from this point forth, will help to achieve the goals of the retention and conversion rates that the company wishes to achieve.

# Project Objectives

Our sponsor, as a high-level summary, wishes to understand their difficulties in terms of online visibility, customer retention and conversion rates on the website.

By utilising the data that we are expecting from:

* The traffic on their website, through all of their placed trackers at each tab and intra-website hyperlinks within their website,
* Possible text-mining on related other e-commerce sites they use for sales,
* And the relevant data from all of their marketing endeavours on social media,

This project hence, has set out the following objectives:

1. In relation to retention and conversion rates and their direct website sales:
2. To understand customers’ browsing patterns upon entering the website and the duration they spend navigating the pages
3. To analyse customers’ behaviour on the site and explain the reason for leaving the site
4. To understand how to improve retention rates on the site
5. With respect to their online visibility:
6. To understand the key terminologies and buzzwords that attract the target demographic to perform the call of action of either opening the website, or to click and buy
7. To be informed about the avenues most popular, and the kind of marketing collaterals that get most traction and churn
8. A summary of costing and demand analysis:
9. To get a fair understanding of what products are moving the fastest with which demographic
10. To analyse and minimise the merchant and holding time and costs, as well as find the key product or product category to lay more emphasis on, based on the trackers on the website and their expected sales

# Methodology

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Tool Used | Tool Utility | Expected Usage |
| 1. | R | Language and environment commonly used for statistical use with computing and graphics. It is often utilized to visualise graphs and table outputs. | Can be used to effectively clean column-wise data and find correlations  Logistic regressions can be used to clean the data and analyse the different independent variables of each identified visitor and customer to the website and regress it against each individual binary occurrence of customers buying or not buying. |
| 2. | SAS JMP Pro | Analytical software that can handle large volumes of data efficiently. JMP Pro can also be used to visualise numbers in multiple formats, and will enable us to easily isolate specific data sets. | The program is also equipped with text-explorer analytics which will be useful to identify trends and the major routes that customers take when visiting the website. K-Means Clustering and similar analyses could help with the profiling of customers based on buying, visiting patterns and their demographic. |
| 3. | Tableau | Renowned for their beautiful visual formatting and colours will be a secondary tool to our project | Will use to create presentable visualisations for the client, and for easily consumable results of the analyses and the data collected and cleaned. |

# Scope of Work

The scope of our project is as follows:

* Discovery
  + Understanding of the business and its current problems
  + Additional research on the industry
* Data Preparation
  + Data Collection
  + Data Preparation
  + Initial Data Exploration
  + Data Cleaning
* Data Modelling
  + Data Exploration
  + Data Analysis and Visualisations
  + Formulating insights and recommendations
* Communication of Results
  + Presentation of results to client and Professor
  + Conference Day and final paper

# Work Plan

Please refer to the wiki page for a higher quality image of the Gantt Chart.