TUNIS BUSINESS SCHOOL
UNIVERSITY OF TUNIS

Internship Report in Business Analytics

as a partial fulfillment of the degree of

Bachelor in Business Administration

Implementing a Marketing Mix Modeling Analysis And Budget Optimization Using CRISP-DM Methodology

---

By:

Mohamed Majd Ayadi

Academic Supervisor:

Dr. Lassaad Moubarki

Professional Supervisors:

Mr. Fares Ben Mahmoud

Mr. Mens Ayed

MASS ANALYTICS

Tunis, 2023 - 2024

Abstract

This work was aimed to provide MASS Analytics, a marketing mix modeling company, with insights about the drivers of website traffic for an IT solution provider company called company X. MASS Analytics is a software company that provides MMM solutions and consultancy services to its clients. The project was intended to determine the effect of company X marketing mix on the number of visitors to their website using three years worth of data provided at the national level. In this project, we used two models to run a marketing mix modeling(MMM) project: the first was a log linear model using the company's software, MassTer, and the second was a multiple linear regression model using python. Additionally, we were able to measure the percentage contribution of various factors including media channels, events and others to determine the ones that significantly drive growth for website visitors. After comparing the outputs of each model and selecting the one with a better fit for our project, we run optimization and prediction for year 2024. We created two optimal scenarios aiming to improve website traffic and media budget allocation after receiving the media budget plan for the current year.

Results from this MMM project proved that the log linear model is more suited for this analysis than the multiple linear regression model. Moreover, marketing variables such as : Informative and Educational Books Submissions, Website publications, Newsletter and Youtube outperform other channels in driving website visitors. For the optimal scenarios 1 & 2 built within MassTer with different budget constraints settings, they yield respectively 63.96% and 15.13% increase in traffic growth with the same budget planned for 2024.

Keywords

Marketing Mix Modeling, Website Visitors, KPI, Log Linear Model, Multiple Linear Regression

Contents

List of Abbreviations

- TBS: Tunis Business School
- MMM: Marketing Mix Modeling
- CRM: Customer Relationship Management
- ROI: Return on Investment
- CPC: Cost per click
- CPM: Cost per Thousand Impressions
- CPA: Cost per Acquisition
- DR: Diminishing Returns
- KPI: Key Performance Indicator
- FB: Facebook
- CTR: Click Through Rate

List of Tables

List of Figures

Introduction

This study, framed within the CRISP-DM methodology, aims to evaluate and compare two different regression techniques for a marketing mix modeling project: a log linear model developed using the company's software and a multiple linear regression model built in python. We will evaluate both models from statistical and business perspectives to determine which technique provides the most accurate and robust results. Upon identifying the superior model, we will perform optimization and prediction analysis for year 2024 with two different scenarios(constrained and non-constrained scenario). The results showed that the log linear model is the most appropriate model in our situation. Moreover, outputs from the two optimal scenarios revealed that the number of visitors will rise by 63.96% in the non limited scenario and by 15.13% in the constrained one with the same budget spend planned for 2024.

Our analysis also explores the impact of various internal and external factors including paid and organic online media, events, newsletter, website publications and other variables on the number of website visitors. By using a three years of data at the national level, we identified the most effective marketing channels. The results showed that the contribution of Informative and Educational Books, Website publications, Newsletter and Youtube outperform other channels in driving website visitors. Through assessing the impact of these channels via marketing mix modeling, we aim to significantly improve the company X marketing effectiveness in order to boost their website traffic.

General Conclusion

During this MMM project, we compared the performance of a multiple linear regression model and a log linear model at explaining the variations in website traffic. Both models are statistically robust. However, from a business standpoint, the log linear model proved to be superior due to its ability to account for complex interactions between variables and provide more accurate insights into the factors driving website visits.

To accomplish this, we followed a CRISP-DM methodology that allowed us to systematically analyze the impact of various internal and external factors on website traffic. By applying the two models, we gained a comprehensive understanding of the contribution of various variables to our KPI, enabling us to identify the ones with the greatest influence on website visitors variation.

The insights and recommendations provided will help in making better data driven decisions to enhance overall website performance, ensuring that marketing efforts are aligned with strategic goals. By taking advantage of the superior capabilities of the log linear model, we are well prepared to execute prediction and optimization for year 2024, thus significantly enhancing the long-term growth of the company's website visitors. Offering strategic advice on the best way to allocate marketing budget across various media channels can effectively increase website visits without raising the media budget spend.

While the selection of the log linear model is needed to achieve results that are both statistically and commercially significant, it's important to note that no model is free of limitations. Some issues faced when using this model were related to reporting the individual contribution of each channel. The multiplicative form of the log linear model presents some challenges in separating contributions. This is when approximation methods are used to convert the multiplicative model into an additive one allowing for a clear distinction between the individual contributions of each variable. Obviously, this approximation will yield a decomposition error that will be added to the standard OLS error. Our role as analysts is to make sure that the decomposition error is as small as possible.