SCALETRIX.AI

DATA ANALYTICS CASE

Technology Students' Gymkhana

IIT Kharagpur

ANALYTIXLABS

www.analytixlabs.co.in

Transforming Careers Since 2011

2500+ Professionals Trained Every Year

13+ Years of Global Experience

4.5 + Average Feedback Consistently

- Premier analytics training institute founded in 2011 by alumni from McKinsey, IIM, ISB, and IIT
- Consistently ranked among Top 3 analytics institutes by Analytics India Magazine, currently holding #1 position
- Expertise in corporate and individual training, with 15,000+ professionals trained and 500+ corporate programs delivered
- Specialized in Data Science & AI education through experiential learning and blended formats
- Offers full-time postgraduate boot camps across major metropolitan cities

Leadership

Sumeet Bansal, CEO

- PEC University of Technology
- 18 Years of global experience

McKinsey & Company ZS

Ankita Gupta, COO

- St. Stephens, ISB Hyderabad
- 19 Years of global experience

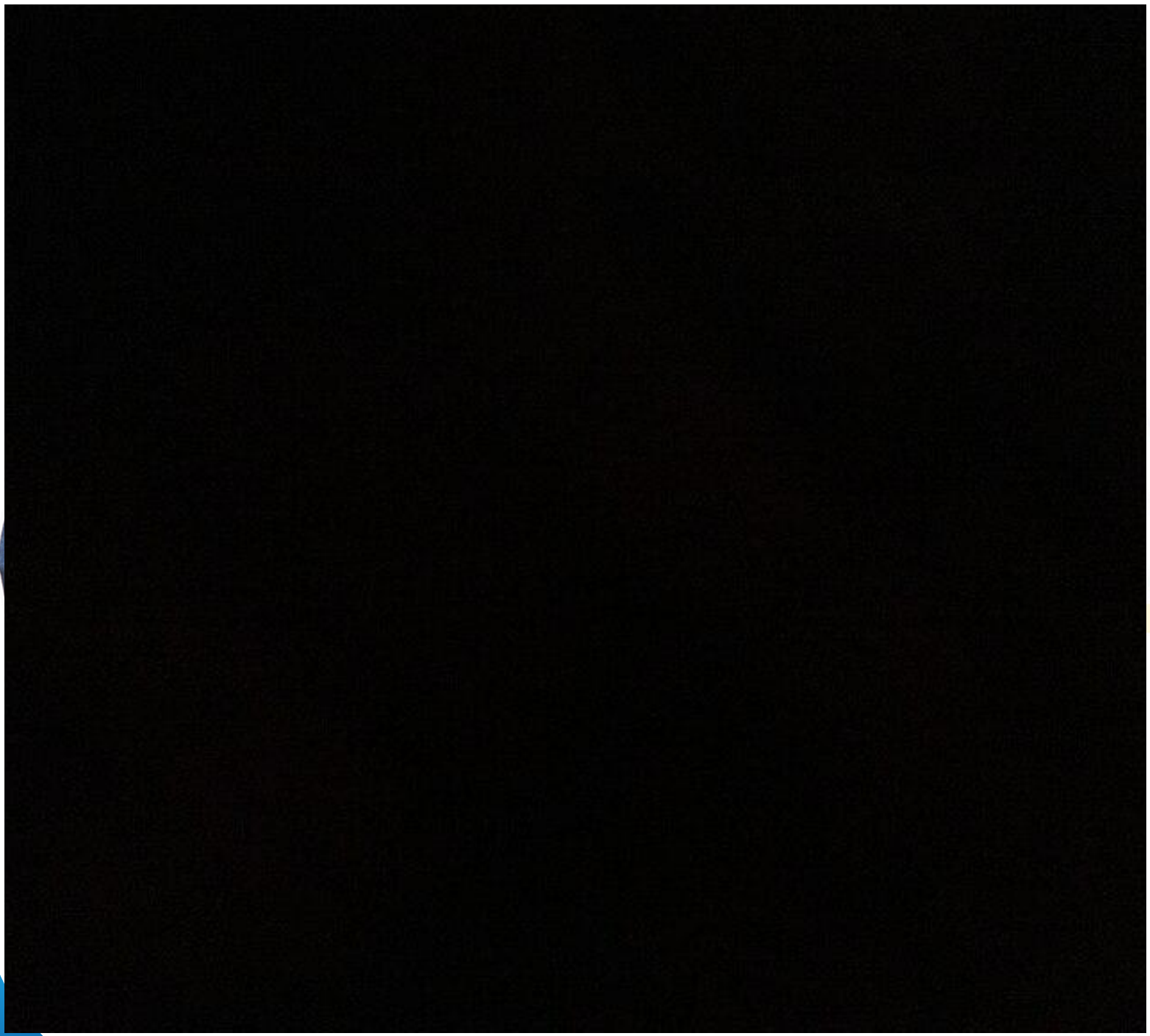
McKinsey & Company Fidelity INTERNATIONAL

Chandra Mouli K, CDS

- IIT Madras
- 18 Years of global experience

McKinsey & Company citibank

Judge and Mentor



Business Context

ElectroMart is an e-commerce firm based out of Ontario, Canada specializing in electronic products. Over the last year, they have spent a significant amount of money on marketing. Occasionally, they also offered big-ticket promotions (similar to the Big Billion Day offering by Flipkart in India). They are about to create a marketing budget for the next year, which includes spending on commercials, online campaigns, and pricing & promotion strategies.

Stake Holders:
Rajesh – Director of ElectroMart
Sanya – Chief Marketing Officer
Pandey – Marketing Operations Team Lead

Business Problem

The CFO feels that the money spent over the last 12 months on marketing was not sufficiently impactful, and, that they can either cut on the budget or reallocate it optimally across marketing levers to improve the revenue response

Imagine that you are a data analyst and supporting a marketing team and the marketing team is working on budget optimization for the next year. As part of this exercise, you are required to support the team by observing the actual impact of different marketing variables over the last year. Using your understanding of the data, you have to recommend the optimal budget allocation for different marketing levers for the next year. (Note: This required to be done using data analytics)

The objective is classified into the following sub-goals:

- Performance driver analysis: Which KPIs drive the top-line (revenue) performance?
- Impact analysis on marketing ROI: What is the quantitative impact of each commercial lever on revenue?
- Optimizing marketing spending: How to best allocate the marketing budget to gain the highest outcome?

Data Availability

The following data files were available to us for analysis:

- Main Consumer file with customer order details on a daily basis
- Media Investment file with the amount invested in each advertising medium for the past year
- Sale Calendar file showing dates from the past year when there was a promotional offer
- NPS file showing net promotion score and company stock value for last year
- Weather file having detailed weather reports from last year in the state of Ontario, Canada
- SKU Mapping with product hierarchy details to understand the product hierarchy
- Holidays list in Canada for the period of available data.

Data Description

Column Description for Customers_Orders.csv:

This table contains all the information about customer orders. You have to use the data from July 2023 to June 2024. The data consists of the following types of information:

- FSN ID: The unique identification of each SKU (SKU => Stock Keeping Unit)
- Order Date: The date on which the order was placed
- Order ID: The unique identification number of each order
- Order item ID: Suppose you order 2 different products under the same order, It generates 2 different order Item IDs under the same order ID; orders are tracked by the Order Item ID.
- Deliverybdays: days to get item or order from warehouse for shipping
- Deliverycdays: days to deliver item to customer
- GMV: Gross Merchandise Value or Revenue
- Units: Number of units of the specific product sold
- Order payment type: How the order was paid – prepaid or cash on delivery
- SLA: Number of days it typically takes to deliver the product
- Cust id: Unique identification of a customer
- Product MRP: Maximum retail price of the product
- Product procurement SLA: Time is typically taken to procure the product

Note: You can assume "\N" value in deliverybdays & deliverycdays is equal to 0

Data Description

Description for “Media data-Sale Calendar-NPS Scores_Data.xlsx“

- Media Spend: Monthly spending on various advertising channels
- Sale Calendar: Days when there was any special sale
- Monthly NPS score – this may work as a proxy to the ‘voice of the customer’
- Stock Index of the company on a monthly basis
- Canada_holiday.xlsx: This helps to understand the holiday list for the time period of data available

Description for “Product Hierarchy Details.docx”

- FSN ID: The unique identification of each SKU
- Super_Category: Product Super Category Information
- Category: Product Category Information
- Sub_Category: Product Sub-Category Information
- Vertical: Product that belongs to which vertical

As the E-commerce company is situated in the Ontario region. Here, we will include its climate data (Weather Data ONTARIO-2023.csv, Weather Data ONTARIO-2024.csv) to analyze if it has any effect on the revenue.

Business Objective

- Detailed Exploratory Data Analysis
- Insights Generation & Detailed Presentation
- Dashboards or Visualisation Creation

Expectations From Team

- The team should Understand the below and document:
  - Business Context
  - Technology Stack
  - Available data
  - Problem Statement
  - List of Outputs to Generate
  - Assumptions Etc.
- What technology stack do you want to use and why?
- Detailed list of data checks to perform or can perform on this data.
- Detailed list of data processing tasks to perform or can perform on this data.
- Detailed list of exploratory data analysis to perform or can perform on this data.
  - Perform Univariate Analysis, Bivariate Analysis, Correlation, Cross-tabs, and visualizations on the datasets. Obtain insights to solve the business problem.
- Detailed dashboards
  - KPI's/KRA's/KRI's (Key Performance Indicators, Key Resultant Areas, Key Risk Indicators)
  - Number of Dashboards
  - Reports in Each Dashboard
  - Filters/Slicers to use Etc
- You should leverage statistical concepts (Hypothesis Testing) to solve the problems
- You should also leverage AI to solve these business problems

Expectations From Team

- Identify what additional information that can help to solve the problem. Also, additional analysis can be done with the help of new data.
- Any alternative approaches can be leveraged to solve these problems (in addition to what you worked)

Deliverables

- Detailed Presentation that includes solutions for the below business objectives
  - Performance driver analysis: Which KPIs drive the top-line (revenue) performance?
  - Impact analysis on marketing ROI: What is the quantitative impact of each commercial lever on revenue? 💬
  - Optimizing marketing spending: How to best allocate the marketing budget to gain the highest outcome for the next year?
  - Which of the products/product categories to target for the upcoming campaigns? What is the rationality?
  - Which marketing channels to be leveraged for each of the product categories? What is the rationality?
- Code Files with Detailed comments
- Github Repository (you should not upload the data)

Glossary

- List Price: List Price = GMV * Units
- Payday Week: If Payday falls within the week, then payday week = 1, else 0
- Holiday Week: If Holiday falls within the week, then payday week = 1, else 0
- Product Type - Luxury /Mass-market: If GMV value is greater than 80 percentile, then luxury, else mass-market
- Discount%: Discount% = 100*(product_mrp – list price) / product_mrp

Sample Exploratory Data Analysis

- Relationship between revenue & advertisement spends
- How are the discounts driving revenue?
- How the customer satisfaction scores driving future revenues?
- Is there any relationship between revenue & payment types?
- Are holidays or paydays impacting sales?
- Is there any relationship between product categories and revenue
- Is there any trend & seasonality exists?
- Are there any trends that exists in Advertisement Investments on marketing channels?
- Etc…

Note: This is not an exhaustive list of analyses. For your understanding, they provided a few samples of analysis. As part of this exercise, you are required to work on as many analyses that help to understand the data & problems and provide insights & recommendations to solve the business problem.