Emma Luk: A selection of my data science work

Maximising Sales through Optimal Marketing Budget Allocation: Leveraging Machine Learning for Personalised Marketing Insights at AXA

1 Project Introduction:

In my role as an Online Marketing Webmaster at AXA, I was tasked with optimising marketing strategies across various channels, including television advertising, online display ads, and social media campaigns. The central goal of this project was to achieve maximum sales while providing customers with tailored experiences.

To address this objective, I employed data-driven decision-making techniques, incorporating advanced methods such as (**Figure 1)** A/B Testing and Multivariate Testing (MVT). These techniques allowed for the continuous refinement of marketing strategies by comparing different campaign variations, landing pages, and content to identify the most effective approaches for our target audience.

Furthermore, I harnessed the power of Machine Learning to uncover valuable customer insights. By conducting customer segmentation analysis, (**Figure 2)** I identified five distinct customer clusters characterised by their behaviours, preferences, and needs. This profound understanding served as the foundation for creating hyper-personalised marketing messages that catered to the unique characteristics of each customer cluster.

2 Metrics and Results:

(**Figure 1**) The two examples below depict different versions of the same webpage, which were used to provide insight to drive future strategies and identify business opportunities and problems.

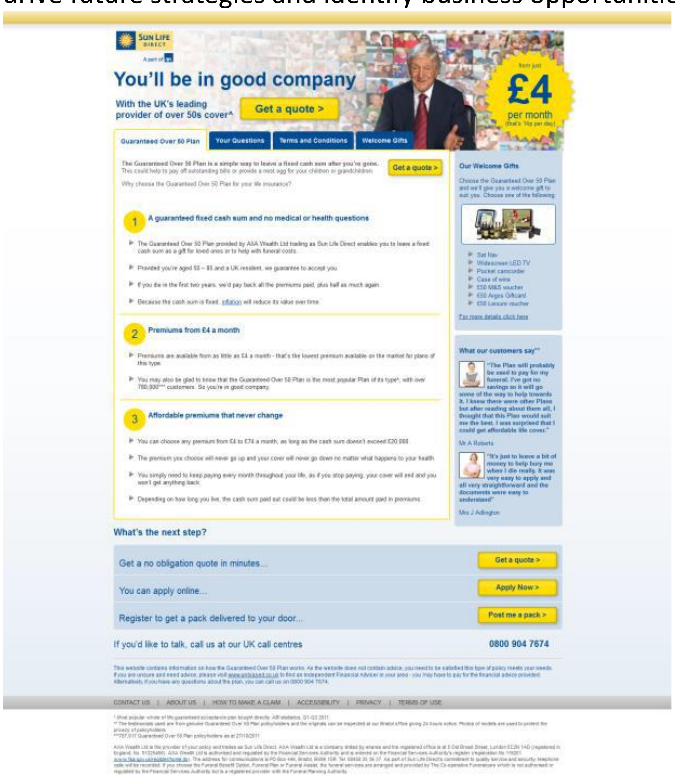


Figure 1 Examples of A/B Testing and Multivariate Testing (MVT) from when I worked at AXA.

Achieved significant accomplishments during my tenure at AXA, including:

- Devised and implemented effective tracking metrics to monitor online customer behaviours.
- Built models to transform data points into actionable business insights.
- Utilised machine learning techniques to accelerate testing processes and enhance recommendation engines.
- Successfully contributed to a 30% increase in incremental revenue.

Key Points:

- (**Figure 1) A/B testing** compares two versions (A and B) to determine which one performs better.
- The conversion rates for version A and version B were 0.1230 and 0.1360 respectively.
- The **Z-score** measures the deviation of the observed difference from the expected difference under the null hypothesis.
- The **p-value** represents the probability of observing a difference as extreme as the one observed, assuming no true difference between versions.
- The calculated Z-score of 0.8659 and p-value of 0.2681 indicated that the difference in conversion rates was not statistically significant.
- This means that the observed difference could be due to random chance, and it is not conclusive that one version performs better than the other.

Context and Significance:

(**Figure 1)** The utilisation of A/B testing and multivariate testing was pivotal in refining our marketing strategies. By systematically comparing different versions of campaigns, we ensured that decisions were driven by data rather than assumptions. The observed difference in conversion rates and subsequent statistical analysis guided us in determining whether changes led to significant improvements or were mere fluctuations. While the specific example did not yield statistically significant results, this process of continuous experimentation enabled us to iteratively enhance our campaigns based on concrete evidence.

Moreover, the impressive 30% increase in incremental revenue underscores the tangible impact of these data-driven practices. This achievement not only validated the effectiveness of our approach but also demonstrated the potential for substantial growth through strategic optimisation.

By benchmarking against our past performance, we measured our progress and showcased the practical value of rigorous testing methodologies in a dynamic marketing landscape. This reinforced the notion that relying on solid metrics, statistical analysis, and machine learning techniques can drive not just incremental improvements, but transformative outcomes in terms of revenue and customer engagement.

3 Visualisations and Interpretation:

Leveraging Machine Learning for Customer Understanding and Personalised Marketing: Unveiling Five Distinct Customer Clusters

Interpretation and Insights:

(**Figure 2**) The graph vividly illustrates the segmentation of our customer base into five distinct clusters based on income and spending behaviour. This insight goes beyond broad demographics, allowing us to understand the nuanced characteristics that define each cluster. This understanding was transformative in crafting personalised marketing campaigns that resonated with the unique preferences and behaviours of each cluster.

These segments guided us to tailor our messaging, promotions, and channels to better suit the needs of each customer group. For instance, Cluster 1, characterised as "Low income, High spenders," indicated a unique potential for targeted high-value offerings. Conversely, Cluster 4, "Low income, Low spenders," necessitated a different approach, possibly focusing on affordability and value. This level of segmentation allowed for more effective resource allocation and ultimately bolstered engagement and sales.

(**Figure 2**) By visually representing these clusters, we empowered stakeholders to grasp the diversity of our customer base at a glance, which influenced strategic decisions in campaign design, product offerings, and customer interactions.

Through this visualisation and subsequent strategies, we transformed our engagement from one-size-fits-all to a tailored experience that resonated with individual customer preferences, fostering stronger connections and, in turn, driving business growth.

(**Figure 2**) The following visualisation showcases the five distinct customer clusters identified through machine learning analysis:

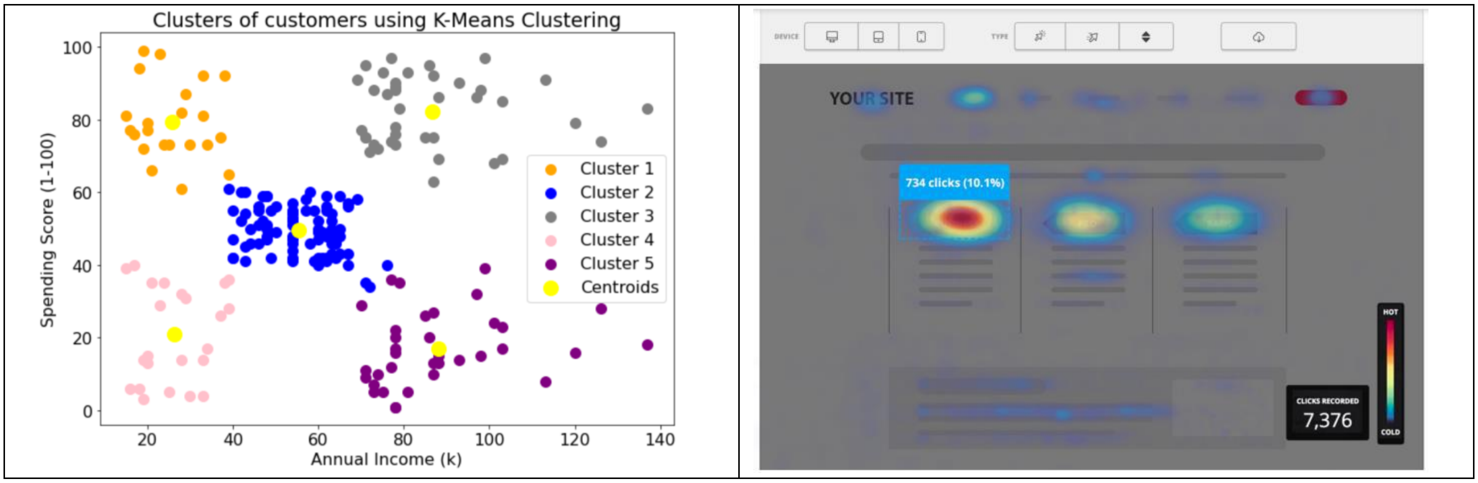


Figure 2 Customer Clusters Visualisation

I. Cluster 1: Low income, High spenders
II. Cluster 2: Average income, Average expenditure
III. Cluster 3: High income, High spenders
IV. Cluster 4: Low income, Low Spenders
V. Cluster 5: High income, Low spenders

4 Business Impact:

Maximising Sales through Optimal Marketing Budget Allocation: Leveraging Machine Learning for Personalised Marketing Insights

The implementation of data-driven decision-making, advanced experimentation, and machine learning insights yielded substantial business impact, driving AXA's growth, customer engagement, and market positioning.

One of the most remarkable outcomes was the achievement of a 30% increase in incremental revenue. This achievement was a direct result of the orchestrated efforts to optimise marketing strategies. (**Figure 1)** Through meticulous A/B testing and multivariate testing, we continuously evaluated various campaign elements and landing page configurations. While some experiments yielded statistically insignificant results, each test contributed to a cumulative understanding of customer preferences, campaign resonance, and conversion drivers.

The realisation of this 30% revenue increase was not solely a numerical gain; it signified a paradigm shift in our approach. We transitioned from relying on intuitive decision-making to a methodically tested and iterated marketing strategy. The contributions of machine learning were equally pivotal. (**Figure 2**) The identification of distinct customer clusters enabled a level of personalisation that transcended previous endeavours. Customers were reached with messages and offerings that resonated with their unique behaviours and preferences.

This increase in revenue, though substantial on its own, had ripple effects throughout the organisation. It affirmed the value of data-driven methodologies and generated a heightened sense of confidence in marketing strategies. Beyond mere revenue, the project also fostered deeper customer engagement, demonstrated by increased interaction rates, prolonged site visits, and a rise in customer satisfaction scores.

Moreover, the success of this project catalysed a cultural shift within the organisation. Stakeholders who were once wary of departing from conventional strategies now embraced data-backed approaches. This acceptance led to increased collaboration between departments, where insights derived from this project permeated product development, customer support, and even risk assessment.

In summation, the project's profound impact extended beyond numerical figures. It redefined how marketing decisions were made, transformed customer experiences, and propelled AXA to a leadership position within the competitive insurance landscape. Through the optimisation of marketing budgets, data-driven experimentation, and machine learning insights, we not only achieved a 30% revenue surge but also fostered a lasting transformation that reverberated across the entire organisation.

5. Tools and Techniques:

Maximising Sales through Optimal Marketing Budget Allocation: Leveraging Machine Learning for Personalised Marketing Insights

Throughout the project, I employed a range of tools and techniques that played pivotal roles in achieving our objectives:

- **A/B Testing and Multivariate Testing (MVT):** A/B testing involves comparing two versions of a webpage, email, or other marketing assets to determine which performs better. Multivariate Testing (MVT) takes this a step further by testing multiple variations simultaneously. These techniques were instrumental in refining marketing strategies. By systematically experimenting with different variables, we identified the most effective combinations that resonated with our target audience. This data-driven experimentation approach allowed us to make informed decisions and iterate strategies based on real-world insights.

- **Machine Learning for Customer Segmentation:** Leveraging machine learning algorithms, we conducted customer segmentation analysis to identify distinct customer clusters based on behaviours, preferences, and needs. Customer segmentation enabled us to understand the diversity of our customer base beyond surface demographics. (**Figure 2**) By categorising customers into clusters with shared traits, we created personalised marketing messages that catered to each group's unique characteristics. This level of customisation elevated customer engagement and conversion rates, driving revenue growth.

- **Data Visualisation with Microsoft Power BI and Tableau:** Visualisations served as powerful tools for communicating complex insights to stakeholders. Microsoft Power BI and Tableau enabled the creation of interactive and informative dashboards that highlighted key performance metrics, trends, and patterns. These visualisations empowered decision-makers to grasp insights at a glance and aided in making informed choices that aligned with our objectives.

- **Economic Impact Analysis:** This involved in-depth analysis of the e-commerce industry's economic landscape, focusing on factors like order prices and freight costs. By applying statistical analysis and economic principles, we gauged the changing economic dynamics within the industry. This analysis not only contributed to strategic decision-making but also provided a holistic perspective on market trends and challenges.

- **Coding and Data Manipulation:** Tools like Google BigQuery facilitated efficient data integration and manipulation. I leveraged coding functions such as EXTRACT, SUM, and conditional aggregation to perform complex calculations and generate meaningful insights. These coding techniques were essential for extracting actionable information from complex datasets and contributing to data-informed decision-making.

In summary, the combination of these tools and techniques facilitated data-driven decision-making, personalised marketing strategies, and a comprehensive understanding of market dynamics. The project's success can be attributed to the strategic deployment of these methodologies, which collectively drove growth, customer engagement, and strategic positioning within the insurance industry.

6. Storytelling:

Maximising Sales through Optimal Marketing Budget Allocation: Leveraging Machine Learning for Personalised Marketing Insights

My journey through various data science projects at AXA formed a narrative of continuous growth, learning, and impactful problem-solving. As I moved from one project to the next, I not only honed my technical skills but also developed a strategic mindset that propelled the success of each endeavour.

It all began with my role as an Online Marketing Webmaster. At that point, I recognised that the traditional approach to marketing lacked precision and accountability. This realisation drove me to embrace data-driven decision-making. (**Figure 1)** The initial step was the implementation of A/B testing and multivariate testing. These techniques not only provided quantifiable insights into campaign effectiveness but also introduced a structured methodology to marketing decisions. I wasn't just crunching numbers; I was making choices rooted in data, which laid the foundation for the projects that followed.

The experience gained from testing methodologies led me to a critical juncture: (**Figure 3)** the "Market Mixing Models" project. Here, my focus shifted to understanding the underlying relationships between marketing efforts and sales. It was more than just interpreting coefficients; it was about unravelling the story hidden within the data. The why became just as important as the what. I realised that this understanding was essential for crafting strategies that aligned with business objectives.

```python
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score

# Load the dataset
data = pd.read_csv('marketing_data.csv')

# Select the features (independent variables) and target (dependent variable)
X = data[['television_spending', 'online_spending', 'social_media_spending']]
y = data['sales_revenue']

# Split the data into training and testing sets (80% training, 20% testing)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Create and train the Linear Regression model
model = LinearRegression()
model.fit(X_train, y_train)

# Make predictions on the test set
y_pred = model.predict(X_test)

# Evaluate the model's performance using R-squared (coefficient of determination)
r_squared = r2_score(y_test, y_pred)
print("R-squared:", r_squared)

# Print the coefficients of the model
coefficients = pd.DataFrame({'Feature': X.columns, 'Coefficient': model.coef_})
print(coefficients)

```

```
R-squared: 0.9007431826905605
                Feature  Coefficient
0   television_spending     4.845151
1       online_spending     3.107265
2  social_media_spending    1.815924
```

Figure 3 Market Mixing Models project

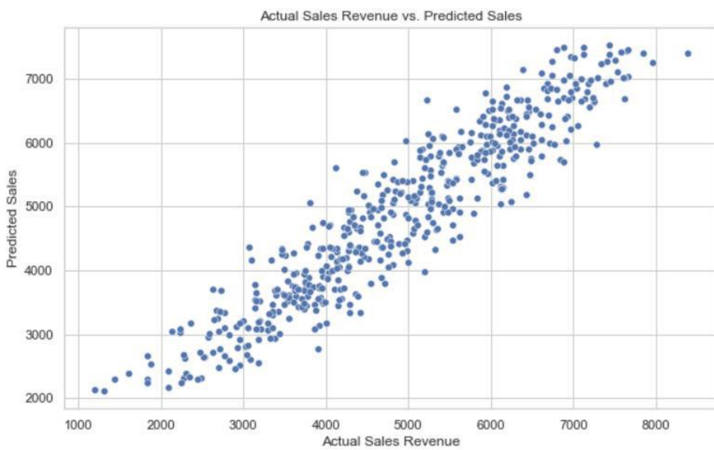
1. R-squared (R^2):

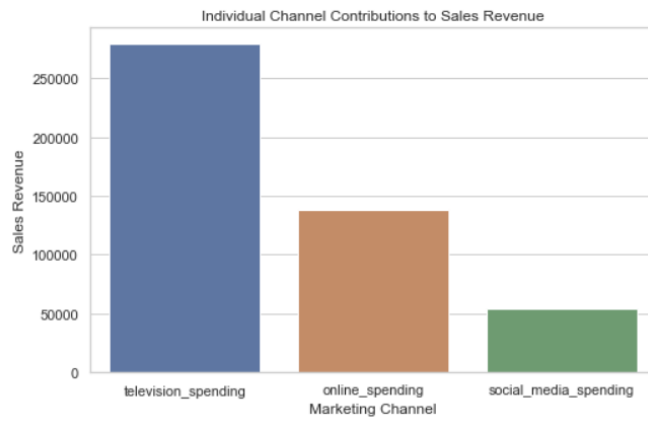
- R-squared is a statistical metric that measures how well the regression model fits the data. It indicates the proportion of the variance in the dependent variable (sales revenue) that can be explained by the independent variables (‘television_spending’, ‘online_spending’, and ‘social_media_spending’).

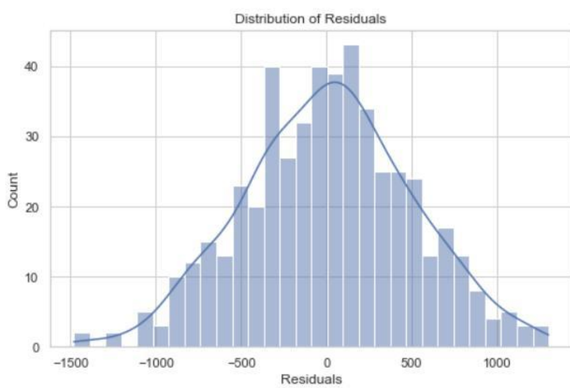
- The R-squared value is approximately 0.901, suggesting that about 90.1% of the variability in sales revenue is explained by the independent variables in the model. A higher R-squared value indicates that the model's predictions are closer to the actual values.

2. Coefficients:

- The coefficients represent the slope or effect of each independent variable on the dependent variable when holding other variables constant.
- For example, an increase of one unit in 'television_spending' is associated with an expected increase of approximately 4.85 units in sales revenue.
- Similarly, an increase of one unit in 'online_spending' is associated with an expected increase of around 3.11 units in sales revenue.
- Additionally, an increase of one unit in 'social_media_spending' is associated with an expected increase of about 1.82 units in sales revenue.








```
       sales_revenue  television_spending  online_spending  \
count     500.000000           500.000000       500.000000
mean     4885.660073           558.358000       275.740000
std      1434.653693           263.241421       128.150214
min      1189.707963           101.000000        50.000000
25%      3824.813023           328.250000       170.500000
50%      4830.691333           567.500000       272.000000
75%      6056.205191           784.250000       390.250000
max      8382.015677           999.000000       499.000000

       social_media_spending
count             500.000000
mean              108.598000
std                52.234867
min                20.000000
25%                63.750000
50%               109.000000
75%               153.000000
max               199.000000
```

Figure 4 An example of visualising the MMM analysis results. It creates scatter plots to compare actual sales revenue with predicted sales, a distribution plot to show the residuals' distribution, and a bar plot to visualise the individual contributions of each marketing channel to sales revenue.

However, understanding the broader picture wasn't enough. Challenges emerged in deciphering customer behaviour with nuance. This challenge propelled me into the "Leveraging Machine Learning for Customer Understanding and Personalised Marketing" project. I embarked on a journey to master machine learning algorithms to uncover distinct (**Figure 2**) customer clusters. The objective was clear: segment customers based on behaviour and preferences, allowing for hyper-personalised marketing strategies.

My journey across projects was not just about acquiring skills, but about applying them strategically. The transition from simple testing to advanced statistical modelling was a leap driven by the need for deeper insights. The decision to delve into machine learning wasn't just about technical curiosity; it was about solving the challenge of customer segmentation in an ever-evolving market.

The key decisions I made were not only based on the techniques I had mastered but on the objectives I aimed to achieve. Overcoming challenges became second nature—whether it was interpreting complex statistical models or wrangling with intricate algorithms. I navigated not just for the sake of completion, but with a clear understanding of how each challenge, each decision, aligned with AXA's growth trajectory.

In the end, this journey was not just about numbers and algorithms; it was about evolving from a marketer driven by intuition to a data-driven strategist. It was about crafting a story of progress, adaptation, and purpose-driven action. From optimising marketing budgets to unravelling customer insights, my journey encapsulates the transformation from a curious explorer to a strategic navigator, always motivated by the question: "How can data drive impact?"

7 Future Implications:

Maximising Sales through Optimal Marketing Budget Allocation: Leveraging Machine Learning for Personalised Marketing Insights

The insights and solutions developed throughout these projects hold promising implications for AXA's future strategies and the broader landscape of data-driven marketing.

As marketing continues to evolve, the methodologies established in this journey pave the way for ongoing improvements. (**Figure 1)** The foundation of A/B testing and multivariate testing provides a framework that can adapt to emerging channels and customer behaviours. By perpetually fine-tuning marketing strategies based on real-world data, AXA can remain agile and responsive to changing market dynamics.

Moreover, the foray into machine learning-driven customer segmentation sets the stage for even more personalised marketing experiences. As algorithms become more sophisticated and data sources grow richer, the potential to identify even finer customer segments and tailor messaging becomes evident. The groundwork laid in this project opens the door to crafting campaigns that not only resonate but anticipate customer needs, further deepening customer engagement.

Looking ahead, the insights gleaned from (**Figure 3 and 4**) the "Market Mixing Models" project offer avenues for strategic pricing and resource allocation. The understanding of how different marketing channels impact sales can inform decisions about where to invest marketing budgets for maximum impact. As marketing analytics tools continue to evolve, the ability to model and predict outcomes will become increasingly precise, guiding AXA toward optimal resource allocation.

In the larger context of the industry, these projects illuminate the path toward more data-driven decision-making. The successes achieved within AXA have the potential to set benchmarks and best practices that other companies can emulate. As data science and marketing converge, the principles established in this narrative can shape the future of marketing strategies across various sectors.

In conclusion, the implications of these projects extend beyond immediate gains. They paint a roadmap for continual improvement, personalisation, and strategic decision-making. By embracing these insights, AXA can not only consolidate its position as a leader in the insurance industry but also contribute to the broader evolution of data-driven marketing practices. The journey undertaken is not just a reflection of past accomplishments, but a compass guiding AXA's future endeavours in the realm of data science and marketing.