



OLLSCOIL NA
GAILLIMHE
UNIVERSITY
OF GALWAY

Scoil Ghnó agus
Eacnamaíochta J.E. Cairnes
J.E. Cairnes School of
Business and Economics

EXAMINATION SCRIPT COVER PAGE

Module Name and Code: Data Science and Big Data Analytics and MS5016

Assignment Title: Value in Analytics

Student ID: 22225743

In submitting this script, I am aware that it is my responsibility to adhere to the examination guidelines. Please tick (Yes/No) for the following:

	Yes	No
I have read the module owners' guidelines, description, and expectations for this exam and the submission process.	<input checked="" type="checkbox"/>	<input type="checkbox"/>
I have read the examination guidelines.	<input checked="" type="checkbox"/>	<input type="checkbox"/>
I am aware of the University Academic Integrity Policy https://www.universityofgalway.ie/media/registrar/docs/QA220-Academic-Integrity-Policy-Final.pdf and confirm the declaration below.	<input checked="" type="checkbox"/>	<input type="checkbox"/>
I have saved the files for submission (e.g., .docx, .xlsx) following strictly the format and naming required (e.g., 12345678_MS5106_MBY_Value_in_Analytics.pdf).	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Declaration for this Exam Submission:

I hereby declare that the work submitted is entirely my own work. It has not been taken from the work of others, except to the extent that such work has been cited and acknowledged within the text of my work. This work is not done in whole or in part by a machine or through Artificial Intelligence, such as ChatGPT. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as their own.

Introduction

A well-known Irish apparel brand called PriMarket is seeking to modernize with IoT activation and other technology. They are doing this by covertly monitoring customer movement using Wi-Fi devices placed in specific store areas. Customers can earn loyalty card points for discounts by using the mobile application. A customer's location can be tracked and a Wi-Fi dataset can be gathered by PriMarket when they log into the mobile application. Additionally, each item of clothing is equipped with RFID tags from the business, which can be used to track garment movement but not the customer holding the item.

(A) One Data-driven question solved using Analytics Lifecycle

Data-Driven question: Determine in which area the customer will spend more time in the upcoming months.

The process of generating, collecting, and processing data into insights that can be used to make informed business decisions is called the analytics life cycle.



Figure 1: Analytics Life Cycle

Step1: Discovery

Identifying the business problem that needs to be solved is the first step. It is employed to formulate a strategy and compile the data needed to address the issue. Throughout this phase, the entire team must plan and gather the required equipment, labor, resources, and technological advancements. Once the data's quality and availability have been determined, we can proceed with their preparation, modeling, and visualization. The amount of time users spend in each area or market segment is provided by the Wi-Fi information, which may be utilized to determine which product is more popular.

Step2: Data Preparation

The second stage typically involves cleaning, transforming, conditioning, and merging the data into a format that can be analyzed, as well as eliminating any NULL, duplicates, negative values, and outliers. The Wi-Fi dataset needs to be cleaned before any operations or model creation is done. The dataset also includes a date and time column, which we can be utilized to add a weekly and monthly column to check the sales in accordance. Pre-processing can be done using a variety of programs, including Python and Microsoft Excel.

Step3: Model Planning

The analyst chooses which tool, methods, methodology, and workflow to utilize to develop a model in the third step, called "Model Planning". In order to pick the features for the model and to understand the many factors and their relationships, the previously generated data is further studied. The objective is to estimate where the consumers will purchase in the upcoming months, so we may use crucial data like the date and time of the visit, the location, the name of the region, and the duration to construct a prediction model using Python or R-Studio.

Step4: Model Building

The dataset is divided into training and testing halves following the exploratory data analysis, which is the next phase of the life cycle. It's also critical to consider whether developing the models requires a steady environment. Classification and prediction can benefit from a variety of techniques, including Regression, Gradient boosting, Random Forest, and others. The best model is chosen based on the error numbers and accuracy ratings.

Step5: Communicate Results

To determine whether the business issue mentioned in phase 1 has been resolved, this phase mostly comprises comparing and choosing the best prediction model. Stakeholders are grouped and clearly informed of the results after analysis. Key discoveries are also discovered in this process, and the findings are reported together with the business value.

Step6: Operationalize

The last stage is to deliver a thorough report to the stakeholders that include significant results, source code, briefings, and relevant papers. To assess performance and ascertain whether model predictions are occurring in accordance with business requirements, the model is also deployed to a real environment. However, if any deviation from the intended output is discovered, the analytics life cycle can also be moved backward. The prediction model will display customer behavior in each location over the coming months, and the following are two takeaways:

- With the help of the prediction model PriMarket can predict which garments will be most sold and which will be least sold in the upcoming months.
- With the help of the prediction model PriMarket can fix prices and provide offers based on sales and customer preference.

(B) Leveraging RFID data to extract insights and create value

Wireless technology called RFID (Radio Frequency Identification) is used to gather information about inventory and client preferences. PriMarket has embedded RFID into each piece of clothing and put antennas in every gate to track in-store activities. RFID data can be utilized for a number of things, such as inventory management, product movement tracking, and consumer behavior research.

Based on the distinction between products with high and low sales made using the RFID data, PriMarket can manage replenishment. The retailer may enhance operations by using real-time tracking to learn which products customers prefer and which products are constantly moving at what time and where. PriMarket gains from this in the following ways:

- **Boost store sales:** RFID data may be utilized to track consumer traffic and product demand, enabling products to be displayed in the most efficient and pertinent locations to improve client shopping experiences.
- **Demand forecasting and restocking:** PriMarket may use the RFID data to pinpoint the merchandise that is frequently bought and moves a lot inside the store, which can assist with production scheduling and inventory replenishment. This raises client satisfaction.

Clothing can be placed in specific locations using the information gathered, such as near the entrance so that customers can see the most popular items as soon as they enter, close to the cashier where the busiest crowd may congregate, and also close to the replenishment area so that the restocking of the out-of-stock items can be done quickly. The information gathered is crucial because it reveals consumer behavior, which may help to explain whether a product was bought or not. The tags can occasionally aid in reducing customer shoplifting and things being transferred without permission.

Example 1: In supermarkets like Dunnes and Aldi, products like chewing gum and chocolate are always placed near the till controlled by RFID antennas, and also vegetables, fruits, and refreshing beverages are always placed at the entrance because they are the most popular items with lots of sections and can also help in easy customer accessibility

Example 2: For better restocking and customer support, businesses like Adidas and Zara also employed RFID tags affixed to each of their products to track sales and inventory levels in real time.

(C) Challenges arising when trying to leverage the actionable insights as described in questions A and B

PriMarket's goal is to transform their stores and make them IoT-enabled using data gathered from the Wi-Fi dataset using routers and the RFID dataset, which provides client information and behavior accordingly. We were able to gain several useful insights by using this data, including the use of real-time RFID tracking to forecast customer preferences and behavior and the building of models as part of the analytics life cycle to predict garment sales and price fixing. Nevertheless, there are some challenges to putting these useful insights into action,

1. Throughout the analytics life cycle, the quantity and caliber of data acquired are constantly vital because flaws in the dataset could impede the development of models and the precision of forecasts. Although the steps in data preparation are essential to the cleansing procedure, the prediction model may occasionally perform poorly if the data it receives is inaccurate and meaningless.
2. The prediction model that is used to forecast sales, set prices, and make offers is solely dependent on client behavior, which is constantly dynamic because it can alter sometimes depending on various factors.
3. It's critical to have a solid firewall, security, and regulations to safeguard the confidentiality and avoid non-compliance because monitoring customer behavior and collecting data could lead to privacy issues, breaches, and cyberattacks.
4. RFID technology can provide a lot of continuous, streaming data that is challenging to manage without the proper tools and approaches because all clothing has RFID tags to track movement. Additionally, occasionally there might be signal loss or interference that produces bad or inaccurate data and affects business decisions. The use of the proper instruments and antennas is crucial for these reasons.

Analytics Life Cycle challenge/failure example: Due to inaccurate market data and a failure to anticipate Apple's claim that smartphones will become a future customer preference, Nokia, a maker of mobile devices, saw a decline in the market. The failure of model creation and trend prediction was eventually caused by Nokia's study's complete reliance on internal surveys and debates, where the data was very limited and distorted. Outside data sources were also neglected. Nokia's failure of the future prediction model led to wrong business judgments and the company's demise because of its biased, distorted, and insufficient data from external as well as internal sources.

RFID data overload example: One of the first businesses to adopt the RFID idea was Walmart, which used the technology to track stock levels, movements, and locations. However, it was not anticipated that Walmart's use of the technology would immediately produce a large amount of data, and Walmart lacked the necessary infrastructure to support the vision. Additionally, the data wasn't always accurate, which made collecting, maintaining, and analyzing it very difficult and resulted in inaccurate business decisions.

References:

- [1] 6 Phases of Data Analytics Lifecycle Every Data Analyst Should Know - BPB Online (Oct 4, 2021) <https://bpbonline.medium.com/6-phases-of-data-analytics-lifecycle-every-data-analyst-should-know-1a9ed65a18a4#:~:text=Data%20analytics%20involves%20mainly%20six,communication%20of%20results%2C%20and%20operationalization.> [Accessed 04/28/2023]
- [2] 6 Phases of Data Analytics Lifecycle Every Data Analyst Should Know About – Rohit Sharma (AUG 31, 2022) <https://www.upgrad.com/blog/data-analytics-lifecycle/> [Accessed 04/28/2023]
- [3] Data Analytics/Analysis Life Cycle (December 22, 2022) <https://begincodingnow.com/data-analytics-analysis-life-cycle/> [Accessed 05/01/2023]
- [4] Where Nokia Went Wrong - James Surowiecki (September 3, 2013) <https://www.newyorker.com/business/currency/where-nokia-went-wrong> [Accessed 05/01/2023]
- [5] Some suppliers gain from failed Wal-Mart RFID edict - Sharon Gaudin (APR 28, 2008) <https://www.computerworld.com/article/2551910/some-suppliers-gain-from-failed-wal-mart-rfid-edict.html?page=2> [Accessed 05/01/2023]
- [6] Supply Chain Comment: Did Walmart's Failed Case Tagging Program Set RFID Back or Move it Forward? (April 19, 2017) https://www.scdigest.com/experts/gilmore_impact_walmart_rfid_program.php?cid=12273 [Accessed 05/01/2023]