

[Coles Organisation]

# REPORT

Prepared by: [Rokshana Pervin]

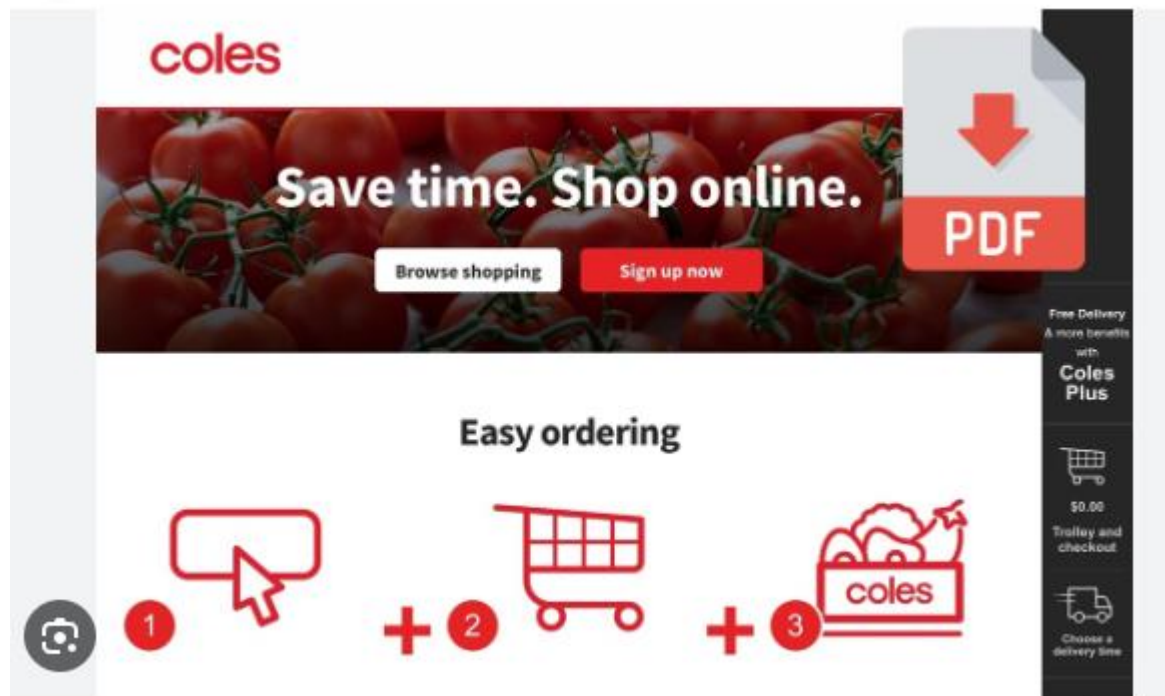
## Contents

Introduction .....	3
Literature review of DSA methods/ techniques and tools .....	4
Possible solution and recommendation .....	5
Conclusion .....	6
References .....	7

## Introduction:

### • Brief profile of the Coles organisation:

Coles supermarket is a national largest supermarket of Australia with more than 800 stores across Australia and employ more than 120,000 team members. With the age of 100 years old, Coles has been dedicated to delivering quality food, great customer value service to Aussie families and communities. Coles started its journey from 1914 with earning 935 pounds and adding with several services such as online, liquor, financial services.



**Figure 1:** Screenshot of the Coles online

Here, Coles online service is the biggest part of the Coles group or Coles supermarket which provides customers with an 'anytime, anywhere' shopping proposition, offering a choice of home delivery, including same-day and overnight drop and go-services' (Colesgroup n.d.).

### • Selected consumer friction points or problem:

As we are confronting in uncertain environments in this revolutionary age, people are searching alternative ways to live their lives in an easy way. Through these ways, leads are facing many different issues such as late online delivery, unloading the webpage or technical issues, as well as product unavailability in the webpage which can make business down and create friction. Here, it is researched that product unavailability in the webpage is the core issue that forms friction between consumer and webpage (ADRENALIN 2023) which results consumers leaving the webpage in a short time. In this case, consumers lose their trust and confident from this webpage and search for another webpage known as competitor from where they can get their desired items or products. Here, Coles supermarket is struggling with consumer friction point or consumer churn with their online service with product unavailability that are minimising their ROI and needs to be fixed by applying different approaches such as RFM (recency, frequency, monetary), CLV (customer lifetime value), ML or machine learning approach, supervised machine learning approach and so on.

## Literature review of DSA methods/techniques and tools:

### • The scope of search and selection of the resources for the literature review:

As consumer friction is reducing leads or consumers that impacts on organisations revenues, organisations need to focus on preventing consumer friction or consumer churn to increase organisational revenues and reputation. Study represents that Coles is confronting with various business problems such as price optimisation, customer shopping experiences in virtually and in person, product unavailability and so on. In this case, Coles can apply several approaches such as marketing funnel, RFM (recency, frequency, monetary) and CLV (consumer lifetime value) models, as well as PV (Perceived value), PR (Perceived risk), RI (Repurchase intention) metrics and personas to retain their prospects as a forever customer (Wong and Wei 2018).

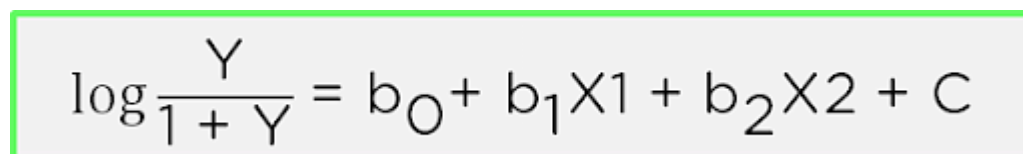
Here, marketing funnel is a visual tool. According to the marketing funnel, Coles organisation can understand the customers and their behaviours and maximise their marketing strategy effectively (Indeed 2022).

To understand customer behaviours, Coles organisation need to collect accurate data such as qualitative and quantitative data to test the accurate model and gain insights and informed decisions through the consumer demographics (Khanna 2017). In this case, Coles can use consumer personas profile using census data which will help to gain the insights of the qualitative and quantitative data.

In this digital era, organisational problem is converted into business or marketing problem and analytical problem. Therefore, it is observed that Coles organisation need to focus on their analytical problem to minimize their consumer friction effectively. In this case, Coles can use DSA or Data Science Analytics tools such as ML or machine learning tools to mitigate consumer friction point or consumer churn appropriately (Osmanbegović et al. 2022).

### DSA methods or techniques:

To prevent consumer churn, we can apply DSA or Data Science Analytics tools or machine learning tools such as supervised machine learning model where logistic regression model is one of the supervised machine learning models (GeeksforGeeks n.d.). Here, logistic regression is the statistical method of binary classification where output displayed with two classes such as churn or not churn (Osmanbegović et al. 2022). According to this model, we need to focus on target variable or dependent variable as well as predictive variables or independent variables where target variable is known as churn and predictive variables would be consumers' age, gender, postcode, location, and other demographic variables (Sharma 2021). As an outcome, Coles organisation or retail industry can reduce consumer churn at least (5-7) % using the classification model or logistic regression model that 'describes the relationship between one binary variable and independent variables' (Datacamp 2019).

A screenshot of a mathematical equation for the Sigmoid Function, enclosed in a green rectangular border. The equation is: 
$$\log \frac{Y}{1 + Y} = b_0 + b_1 X_1 + b_2 X_2 + C$$

**Figure 2:** Screenshot of the Sigmoid Function

The above equation represents  $Y$  as a probability of outcome,  $b_0$ ,  $b_1$ ,  $b_2$  is the co-efficients,  $X_1$ ,  $X_2$  is the independent variables and  $C$  is the constant.

## Possible solution and recommendations:

To achieve the possible solution, we need to prepare the data and check the data either it has null values (inaccurate data) or multicollinearity (variables that are exist with the same characteristics) before training the dataset. During the logistic regression process, the log odds of the probability of customer or dependent variable will be related to the independent variables such as age, gender, postcode and so on.

### Data Preparation:

Data availability ensures that data is prepared for analysis. To prepare the data, categorical data such as marital, gender and numerical data such as age, tenure will be collected from the census data (Australian Bureau of Statistics n.d.). Here, Categorical data will be converted into two discrete data such as 0 and 1 though churn feature's value will remain as yes or no. Prior to using the data for analysis, we need to normalise the data by measuring mean and standard deviation (Sharma 2021).

### Generating Initial and Reduced model:

After preparing data, an initial model will be created with respect to churn analysis to check the significance of all features. Apart from this, predictive analysis will be completed using various information through the initial model (Osmanbegović et al. 2022). Some features will be eliminated to create reduced model as well as accuracy score, confusion matrix and AUC or area under curve will be compared with the initial and reduced models. In the large dataset, to predict the outcome of the consumer churn or consumer friction point, we need to split the data 70% and 30% and parallelly train the data 70% and test the data 30%.

### Analysing Initial model and Reduced model:

During analysing the Initial model, it is observed that probability value or p-value of various independent variables are significant for predictive modelling (Sharma 2021). Here, p-value (0.05) will help to calculate the accuracy and confusion matrix of the model and will reduce the list of variables by showing the percentage of specificity (Service continuation) and Sensitivity (service discontinuation).

Using lower p-values with higher significance, Reduced model will be created with the list of independent variables. After acquiring the accuracy and confusion matrix with lower p-values of the reduced model, it can further reduce the list of variables considering the important points as well as forms another reduced model applying above process (Simplilearn 2023). When the present features represent the higher significancy than the previous features, it will indicate to predict the consumer churn or friction point accurately.

### Model comparison and Outcome:

After comparing both Initial and Final Reduced model, there is a big difference can be anticipated between accuracy, specificity, sensitivity where these terms are higher in the final reduced model than the Initial model. With this outcome, the final reduced model will be highlighted as 'lean, faster and less-resource intensive' (Sharma 2021). Apart from this, the AUC or area under curve of the final reduced model would be nearly 1 which will provide the great insight into which customers may discontinue with the product or organisation.

Moreover, co-efficient of the lr or logistic regression model will help to find the customer who will be intent to leave the online platform for forever (Wahba 2016).

Through the above observations, it is hypothesized that in this way both model (logistic regression model) and metrics (Specificity, Sensitivity, AUC) will help Coles organisation to predict the potential customers analytically who will remain with the webpage or platform for forever.

### **Scope of covered:**

Using the logistic regression model, it is analysed that customers monthly interest or daily necessity of the items can be predicted which will help Coles to calculate the number of the customers who are actually interested to check-out the product such as milk powder from the online platform.

### **Scope of not covered:**

In this report, one model analysed to predict and reduce the customer friction point. It is anticipated that if we compare more than one model, it would be better to gain the more real-insights and informed decisions regarding the consumer friction point of the unavailable product of the Coles Online platform.

### **Recommendations and justifications:**

Here, it is observed that the logistic regression model boosts the results of the consumer friction point and able to perform effectively to reduce the consumer churn (Osmanbegović et al. 2022).

It is also recommended that Coles need to monitor online platform as well as algorithms' operations continuously through business and analytical team (Stanford University School of Engineering 2016). Study represents that Coles can accommodate children approach from the census data as they are major consumer in the retail industry. Apart from this, Coles can increase their collaborative approach to the parent companies and suppliers using the metrics and co-efficient of the model (Palos-Sanchez1 et al. 2018). Hence, it is recommended and justified that if the model or algorithm is not provide accurate results, it will not display the consumer friction point appropriately which can cause the loss of revenue for the organisation (Analytics Vidhya 2023).

### **Conclusion:**

Nowadays, every industry is trying to adopt advanced technology to enhance their business effectively. In this case, retail industry is not behind across the world. Here, it is analysed from the above report that Coles organisation can use logistics regression to amplify their business from any sector specifically in online sector. As Coles is confronting with consumer friction point in online platform by highlighting unavailable product, it is observed that an accurate method such as logistic regression method or Data Science analytics method will help Coles to manipulate the consumer friction point as well as to retain the consumer for forever.

### **Limitations and future plan:**

As we are getting a lot of tons of data in the upcoming days, we need to be aware specifically in retail industry. From the study, it is observed that Coles is struggling with many consumers friction point which makes Coles revenue down than its competitor. Although in

this report one consumer friction point has been discussed, there are many friction points need to be addressed in order to acquire their goal.

Therefore, it is found from the above report that Coles can make strong foundation both their analytical and business strategy which will extend Coles consumers' lifetime as well as their revenues enormously.

## References:

ADRENALIN (2023) *Why friction is killing your online customer journey*, adrenalinmedia website, accessed 01 November 2023. <https://www.adrenalinmedia.com.au/why-friction-is-killing-your-online-customer-journey/>

Analytics Vidhya (2023) *Regression Analysis | Beginners Comprehensive Guide (Updated 2023)*, analyticsvidhya website, accessed 24 November 2023. <https://www.analyticsvidhya.com/blog/2020/12/beginners-take-how-logistic-regression-is-related-to-linear-regression/>

Australian Bureau of Statistics (n.d.) *Census*, abs website, accessed 06 November 2023. <https://www.abs.gov.au/census>

Colesgroup (n.d.) *About us*, colesgroup website, accessed 12 November 2023. <https://www.colesgroup.com.au/about-us/?page=about-us>

DataCamp (2019) *Understanding Logistic Regression in Python Tutorial*, datacamp website, accessed 17 November 2023. [Python Logistic Regression Tutorial with Sklearn & Scikit | DataCamp](#)

GeeksforGeeks (n.d.) *Supervised and Unsupervised learning*, geeksforgeeks website, accessed 20 November 2023. [Supervised and Unsupervised learning - GeeksforGeeks](#)

IBISWorld (2023) *Coles Group Limited – Australia Company Profile*, ibisworld website, accessed 06 November 2023. [Coles Group Limited - Company Profile Report | IBISWorld](#)

Indeed (2022) *How a Marketing Funnel Works (With Stages and Examples)*, indeed website, accessed 15 November 2023. [How a Marketing Funnel Works \(With Stages and Examples\) | Indeed.com](#)

Khanna V (2017) 'A Study on Factors Affecting Subscription Rates of Netflix in India: An Empirical Approach', *Research Paper*, 18(1), 83-100, <https://doi.org/10.51768/dbr.v18i1.181201717>

Osmanbegović E, Suljic M and Džinić A (2022) 'Prediction of telecom services consumers churn by using machine learning algorithms', *Economic Review – Journal of Economics and Business*, XX (2), DOI: 10.51558/2303-680X.2022.20.2.53

Palos-Sanchez<sup>1</sup> PR and Correia<sup>2</sup> MB (2018) 'The Collaborative Economy Based Analysis of Demand: Study of Airbnb Case in Spain and Portugal', *Journal of Theoretical and Applied Electronic Commerce Research*, 13(3), 85-98, DOI: 10.4067/S0718-18762018000300105

Sharma A (2021) *Customer Churn Data Analysis using Logistic Regression*, medium website, accessed 22 November 2023. [Customer Churn Data Analysis using Logistic Regression | by Akhil Sharma | Data Science on Customer Churn Data | Medium](#)

Simplilearn (2023) *Understanding The Difference Between Linear vs Logistic Regression*, simplilearn website, accessed 23 November 2023. <https://www.simplilearn.com/tutorials/machine-learning-tutorial/linear-regression-vs-logistic-regression>

Stanford University School of Engineering (2016), Data science and entrepreneurship panel: Personalized data is the key opportunity [YouTube Video], *YouTube*, accessed 22 November 2023. <https://youtu.be/zlq-GfPU3OQ>

Wahba A (2016) *How to conduct Aha-moment (aka Magic moment) analysis without knowledge of statistics or data science!*, medium website, accessed 11 November 2023. [How to conduct Aha-moment \(aka Magic moment\) analysis without knowledge of statistics or data science! | by Abdelrahman Wahba | Egyptian Startup Manual | Medium](#)

Wong E and Wei Y (2018) 'Customer online shopping experience data analytics: Integrated customer segmentation and customised services prediction model', *International Journal of Retail & Distribution Management*, 46(4): 406-420, DOI 10.1108/IJRDM-06-2017-0130.