RESEARCH PROPOSAL FORM

(also referred to as the 'Statement of Intent Form', or SOI)

To be submitted by the researcher to the Institute Research Sub-Committee (IRC)

Research Title: Market Basket Analysis in correlation with Product Assortment Strategies in Supermarkets
Institute name: Institute of Information and Communication Technology
Course / Programme: Bachelor of Science (Honours) in Business Analytics
Level and year of study Level 6 Year 3
Main area of study being proposed: Understanding consumer behaviour and refining product assortment strategies are essential for sustaining competitiveness and enhancing profitability in the retail industry. As key participants in the retail industry, supermarkets are always looking for new and innovative ways to improve the shopping experience and increase sales. In this effort, market basket analysis proves to be a useful tool since it provides insights into the preferences and buying habits of consumers. In order to clarify their relationship and implications for supermarket management, this research investigates the relationship between market basket analysis and product assortment techniques. I explore market basket analysis using Python based research approaches, making use online data sets and performing an extensive survey to gain a deeper understanding of customer behaviour. This study intends to give supermarket managers useful insights by clarifying the connection between market basket analysis and product assortment methods. These insights will help supermarket managers make well-informed decisions about product placement, inventory control, and marketing tactics. In the end, this study adds to the discussion on retail management practices by emphasizing the role that data driven strategies play in promoting operational brilliance and customer satisfaction.

Name of Researcher: Casey Portelli	**Researcher's I.D. Number:** 148704L
Signature of Researcher CPortelli	**Date of submission of Form** 31/10/2024
Name of Tutor (or Recommended Tutor): Frankie Inguanez	

Personal Motivation for the Choice of Research Theme.

My significant interest in marketing, and specifically customer behaviour, is the main reason I chose the research issue of market basket analysis to evaluate and improve assortment tactics within supermarkets. The complexities involved in consumer decision-making and the tactics used to sway these decisions intrigue me. It is intriguing to be able to integrate market basket analysis and focus my research on these components. As a customer, I've always been interested in product selection strategies, including how promotions are put together, why particular products are positioned in particular areas of supermarkets, and cross-selling techniques. My curiosity has led me to investigate the reasoning behind these choices and how they affect sales and customer satisfaction. With this research I can combine the study of customer behaviour with algorithms in supermarkets. Through the creation of dependable and successful product placement and promotion techniques, my goal is to offer supermarkets insightful information that will improve their organisational tactics. In addition to pursuing my academic and professional interests, my research has the potential to significantly influence the retail industry and, eventually, enhance the shopping experience for customers.

Outline of Key Literature and Theoretical Framework or Propositions.

In order to strike a balance between revenue generating and customer satisfaction, researchers have looked into a variety of strategies for optimising assortment procedures in the retail sector. Their insights into the developing topic of assortment optimisation are compiled in this review.

The Apriori algorithm is used in the UIN Malang study [1] to analyse customer purchasing habits. This study demonstrated how data mining methods may produce association rules and identify frequently occurring item sets, providing insightful information about consumer behaviour. These results demonstrated how well MBA programmes work to improve marketing and inventory management by anticipating consumer wants and placing products in the best possible locations, thereby utilising transactional data for strategic decision-making.

Similarly, the Apriori and FP-Growth algorithms are used in the work by Manpreet Kaur and Shivani Kang [2] to investigate supermarket assortment optimisation. This study showed that frequent item set analysis could enhance advertising and product positioning tactics. MBA has the potential to improve decision-making, client satisfaction, and revenue through data-driven marketing initiatives. Notably, the FP-Growth algorithm surpassed Apriori in terms of execution time and scalability with large datasets.

Expanding upon previous research, Santos's study [3] on rice subcategory online grocery retail assortment optimisation highlights the significance of SKU selection for optimising sales without compromising revenue. This study emphasises the importance of decision-makers' involvement and data-driven decision-making through transactional logs.

The influence of big data analytics on assortment strategy in Pakistan's organised retail industry is further examined by Shaikh et. al. [4]. Despite obstacles like small sample sizes and scarce data, they stress the importance of highly qualified data scientists and advanced algorithms.

Karki's approach [5] to managing retail assortments involves assigning products to customer segments based on their behaviour, which is a hybrid technique. Nonetheless, this method recognises that customer preferences are dynamic and that data biases may exist, implying that ongoing improvement is necessary.

In their exploration of the use of analytics tools in specialised retail, Kanagaraj and Venkatesh stress the use of data in focused marketing initiatives and enhancing sales results. Given the advantages, they point

out privacy and data quality concerns that need further investigation [6].

Finally, in order to improve assortment planning's sales prediction accuracy, Tamhankar et al. look into clustering techniques [7]. Their study demonstrates how combining clustering algorithms with regression and classification models can considerably enhance prediction accuracy, which in turn contributes to effective assortment optimisation. Their research establishes a connection between conceptual models and real applications.

Combined, these studies highlight the important developments and continuous difficulties in the field of retail assortment optimisation, highlighting the crucial role that advanced analytics and data-driven strategies play in achieving both corporate efficiency and customer satisfaction.

Significance of the Study.

This study is significant because it could provide useful information and workable solutions for the retail sector, particularly supermarkets. To maximise product placements and promotions, this research uses market basket analysis to evaluate and improve assortment methods. Supermarkets may create more successful cross-selling techniques and more precisely customise their assortments to match client preferences by having a better understanding of consumer purchase habits and the relationships between products. Through the provision of practical evidence and operational breakthroughs in the application of big data analytics and machine learning techniques, this study also makes a contribution to the academic field of marketing and consumer behaviour. In the end, the research's conclusions can help supermarkets make data-driven choices that will strengthen their competitive advantage and provide customers with a more tailored and fulfilling shopping experience.

Hypotheses and/or Research Question/s

The research hypothesis:
Effective product associations and improved product placement strategies can be found by using market basket analysis to stimulate consumer behaviour in product purchases, supported with a supermarket mangers' insights and customer surveys. In supermarkets, this strategy seeks to improve consumer satisfaction, sales and the shopping experience.

The research questions:
1. In what ways does market basket analysis help supermarket assortment plans be improved and optimized to better serve customers and increase sales?
2. How can cross-selling opportunities in supermarkets be improved and complimentary products found using market basket analysis?
3. How do external factors like pricing strategies of competitors, consumer preferences, and economic situations effect the understanding of market basket analysis in supermarkets?

Target Participants and Research Methods for Data Collection and Analysis

The target participants for this study will be supermarket owners and managers as well as customers. A total of two owners and/or managers will be interviewed to learn more about their assortment strategy and obstacles, and customers will be surveyed to learn more about their preferences and purchase behaviour. A target of 100 survey respondents representing the public that purchase items from supermarkets and/or grocery stores. Customer surveys, in-depth interviews with owners and managers, and (publicly available, not local private) transaction data from supermarkets will all be used in the data

collection process. Market basket analysis will be performed to find common item sets and product associations using Python's Apriori technique. To provide a thorough understanding of consumer behaviour and to guide supermarket assortment and product placement strategies, this study will use a triangulation of research methods: quantitative surveys, qualitative interviews, and computational analysis.

Anticipated Contributions of the Study.

It is anticipated that this study would significantly advance the understanding of consumer behaviour and retail management. Firstly, it will give businesses important information about supermarket customers buying habits, enabling them to determine which product combinations are most frequently purchased together. By optimising product placements and promotions, this data may raise sales and improve consumer satisfaction. Secondly, the study will provide useful suggestions for enhancing the assortment tactics that are now in use by evaluating their efficacy. Supermarkets may be able to better serve customers and improve organisational efficiency as a result. Finally, by showcasing the potential of big data analytics in the retail sector through the application of sophisticated data analysis techniques like the Apriori algorithm, the industry will be encouraged to embrace these strategies going forward.

Dissertation Project Plan.

		2024		2025					
WP	**Description**	**Nov**	**Dec**	**Jan**	**Feb**	**March**	**April**	**May**	**June**
Project Commencement									
1.	**Idea Formulation and Proposal**								
	1.1 Discussion with Lecturers								
	1.2 Mentor Allocation								
	1.3 SOI								
2.	**Literature Review**								
	2.1 Background Information								
	2.2 Datasets, Solutions								
	2.3 Metrics								
3.	**Research Methodology**								
	3.1 Hypothesis and Questions								
	3.2 Pipeline								
	3.3 Data collection Evaluation								
4.	**Development**								
	4.1 Conduct Survey and Interview								
	4.2 Market Basket Analyses Setup								
	4.2 Data Cleaning and Pre-processing								
	4.4 Algorithm Implementation								
	4.5 Model Training and Validation								
	4.6 Results Analysis								
5.	**Deliverables**								
	5.1 Dissertation Writing								
	5.2 VIVA Preparation								
	Project Closure								

Commented [FI1]: My only comment is that you do not mention the interviews and surveys. This would fall under the development section.

Ethical Considerations.

Refer to *guidance points below. You are also additionally required to read MCAST Document 074 'Research Ethics Policy and Procedure' that is available on the College website*

Research shall be conducted in such a manner so as to avoid any psychological and physical harm to humans and animals and financial damage to organizations:

1. *Only the supervisor and examiners will have access to any data gathered.*
2. *Participants will remain free to withdraw from the study at any time without having to provide any reason. In the case of withdrawal, all the records and information collection will be deleted.*
3. *The participant, who is the sole proprietor of the data provided, is granting that such data would be processed for this study purposes only.*
4. *The data collection process will be a transparent process.*
5. *All transcriptions and/or electronic recordings reflecting the data collected, once exhausted, are to be deleted*
6. *Confidentiality, anonymity and data protection procedures are to be ethically abided by.*
7. *The researcher would provide a soft copy of the study to the participant, if required.*

Enter details here regarding possibility of issues regarding confidential personal data:

Strict data management procedures, such as anonymization strategies, should be used to ensure participant anonymity. Informed consent should also be obtained to protect against the exposure of private or sensitive information.

When undertaking interviews, strict confidentiality procedures shall be undertaken. The target participants are informed about the research and its scope via an information participant letter outlining details about the research and how to contact the lead researcher and respective mentor for additional information. Should they accept to participate they need to sign a consent form which includes a declaration that they are 18 years or older. The interview shall be recorded then transcribed. The identity of the participant and the respective business shall be anonymised, and no identifiable information shall be retained. Upon the anonymization process the anonymised transcripts and audio/visual media are destroyed. Following this process, it would not be possible for the participant to withdraw consent since there would not be a manner to track the person. All information gathered through this research shall be stored solely on a password protected computer, will not be shared on any cloud service or any other device.

With regards to the survey, no personal identifiable information shall be gathered such as names, home/e-mail/IP addresses or phone numbers. Questions will be kept generic. Only adults are eligible for this survey and therefore a question asking the participants to confirm that they are 18 years of age or older. The text of the participation letter and consent form are included in the description of the survey.

Enter details here regarding possibility of physical harm:

The surveys shall be held online and therefore there is no risk of physical harm, whilst the interviews shall all be held in person. For this matter the participant is invited to select a location of their choosing, in which they feel safe and comfortable. Should they prefer to have the interview online, then this can be also accommodated.

Enter details here regarding possibility of moral harm:

With surveys and interviews conducted only to study consumer behaviour and organisational methods, this research ensures no possibility for moral harm and focuses purely on market basket analysis and organisational techniques.

Enter details here regarding possibility of business harm:

The transactional data used in this research shall be from public available datasets which enjoy a free to use and free to research license. The items in question are identified as generic products with no branding available.

The identities of the businesses in which the interview participants operate shall be anonymised. No transactional data of local businesses shall be utilised.

All data shall be retained on a password protected laptop and all safeguards possible by the lead researcher shall be taken to ensure that no data breaches or leaks occur.

*Please see **Annex 1** for a sample Participant Information Letter and **Annex 2** for a sample Participant Consent Form. Student is to submit a copy of the proposed Participant Information Letter and Participant Consent Form where applicable. Both documents should be attached to the end of the SOI that is being submitted by the student.*

List of Key References:

[1] F. Kurniawan, Binti Umayah, J. Hammad, Supeno Mardi Susiki Nugroho, and Mochammad Hariadi, "Market Basket Analysis to Identify Customer Behaviours by Way of Transaction Data," *Knowledge Engineering and Data Science*, vol. 1, no. 1, pp. 20–25, 2018, Accessed: May 21, 2024. [Online]. Available: https://journal2.um.ac.id/index.php/keds/article/view/1386/1377

[2] M. Kaur and S. Kang, "Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining," *Procedia computer science*, vol. 85, pp. 78–85, Jan. 2016, doi: https://doi.org/10.1016/j.procs.2016.05.180.

[3] H. Beatriz, O. Santos, D. Machado, and M. Fontes, "Assortment Optimization in Online Retail," 2019. Available: https://repositorio-aberto.up.pt/bitstream/10216/141248/2/368050.pdf

[4] S. Shaikh, F. Sultan, and M. Asim, "Assortment on the Bases of Big-Data Analytics: A Quantitative Analysis on Retail Industry," *JISR management and social sciences & economics*, vol. 19, no. 2, pp. 154–167, Dec. 2021, doi: https://doi.org/10.31384/jisrmsse/2021.19.2.9.

[5] D. Karki, "A hybrid approach for managing retail assortment by categorizing products based on consumer behavior MSc Research Project Data Analytics." Available: https://norma.ncirl.ie/3446/1/dhirajkarki.pdf

[6] K. Kanagaraj and R. Venkatesh, "Enhancing Business Performance: A Comprehensive Study of Sales and Distribution Analytics in Speciality Retail Sectors," Nov. 2023, doi: https://doi.org/10.1109/rmkmate59243.2023.10369436.

[7] R. Tamhankar, S. Khattar, X. Che, S. Zhu, and M. Lanham, "Optimal Clustering of Products for Regression-Type and Classification-Type Predictive Modeling for Assortment Planning." Available: https://matthewalanham.com/Students/2018_MWDSI_Optimal%20Clustering%20of%20Products.pdf

This section is to be filled in by the representative of the Institute Research Sub-Committee (IRC) prior to forwarding of this Form to the 'MCAST Research Ethics Committee' for final ethics approval:

Nature of Ethical Consideration	***Outcome (Tick)***	***Comments/Advice***
All ethical issues have been adequately tackled.		
Possibility of issues regarding misuse of data or some form of harm.		

Details of Representative to the Institute Research Sub-Committee.	
Name	Signature
Designation	Date

Annex 1: Participant Information Letter

MCAST

Sample:

Title of Research: _____

You are being invited to take part in a research study. Before you decide to participate, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

What is the purpose of the study?
This research is being undertaken on…

Why have I been chosen?
You have been chosen because…

Do I have to take part?
It is up to you to decide whether or not your take part. If you decide to take part you will be given this information sheet to keep and be asked to sign a corresponding consent form.

What will happen to me if I take part?
You will then be given a questionnaire on.../your data will be used…/your image will be used…

What are the possible disadvantages and risks of taking part?
There are no disadvantages or risks foreseen in taking part in the study.

What are the possible benefits of taking part?
By taking part you will be contributing to the development of a set of recommendations for…

What if something goes wrong?
If you wish to complain or have any concerns about any aspect of the way in which you have been approached or treated during the course of this study, please contact…(researcher is to give his/her MCAST email as a contact)

Will my details be kept confidential?
All information which is collected about you during the course of the research will be kept strictly confidential so that only the researcher carrying out the research will have access to such information and will not be shared with any other individuals. Participants should note that data/images collected from this project may be retained and published in an anonymized form. By agreeing to participate in this project, you are consenting to the retention and publication of data.

What will happen to the results of the research study?
The results will be written up into a dissertation for my final project of my Bachelor…

Who is organizing the research?
The research is conducted as part of a degree in …

Who may I contact for further information?
If you would like more information about the research before you decide…(researcher is to give his/her MCAST email as a contact)

Thank you for your interest in this research…

Annex 2: Participant (or Guardian) Consent Form

Sample:

Title of Research: _____

Name of Researcher: _____

Please initial box

1. I confirm that I have read and understand the Information Letter for the above study and have had the opportunity to ask questions. ☐
2. I understand that my/my charge's participation is voluntary and that I/my charge am/are free to withdraw at any time without giving any reason. ☐
3. I agree to allow my daughter/son/charge to take part in the above study. ☐

(Statement 3 is to be included only when guardians/parents are involved in giving consent)

_____	_____	_____
Name of Participant/ Guardian	Date	Signature
_____	_____	_____
Researcher	Date	Signature

1 for participant; 1 for researcher