

September 2024

Category review: Chips

Retail Analytics



Classification: Confidential



Our 17 year history assures best practice in privacy, security and the ethical use of data

We all have a responsibility to use data for good

Privacy

- We have built our business based on privacy by design principles for the past 17 years
- Quantum has strict protocols around the receipt and storage of personal information
- All information is de-identified using an irreversible tokenisation process with no ability to re-identify individuals.

Security

- We are ISO27001 certified - internationally recognised for our ability to uphold best practice standards across information security
- We use 'bank grade' security to store and process our data
- Comply with 200+ security requirements from NAB, Woolworths and other data partners
- All partner data is held in separate restricted environments
- All access to partner data is limited to essential staff only
- Security environment and processes regularly audited by our data partners.

Ethical use of data

Applies to all facets of our work, from the initiatives we take on, the information we use and how our solutions impact individuals, organisations and society.

Quantum believes in using data for progress, with great care and responsibility. As such please respect the commercial in confidence nature of this document.

Executive summary

01

Task 1

We need to present a strategic recommendation to Julia that is supported by data which she can then use for the upcoming category review. However, to do so, we need to analyse the data to understand the current purchasing trends and behaviours. The client is particularly interested in customer segments and their chip purchasing behaviour. Consider what metrics would help describe the customers' purchasing behaviour.

02

Task 2

The client has selected store numbers 77, 86 and 88 as trial stores and want control stores to be established stores that are operational for the entire observation period.

We would want to match trial stores to control stores that are similar to the trial store prior to the trial period of Feb 2019 in terms of :

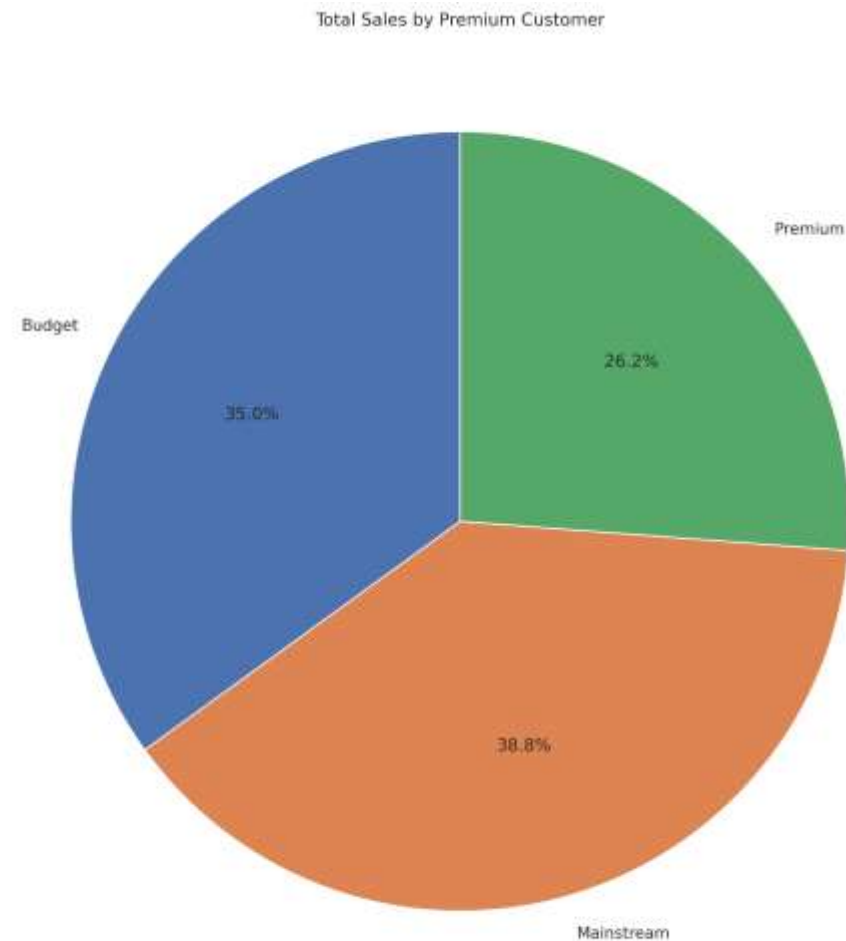
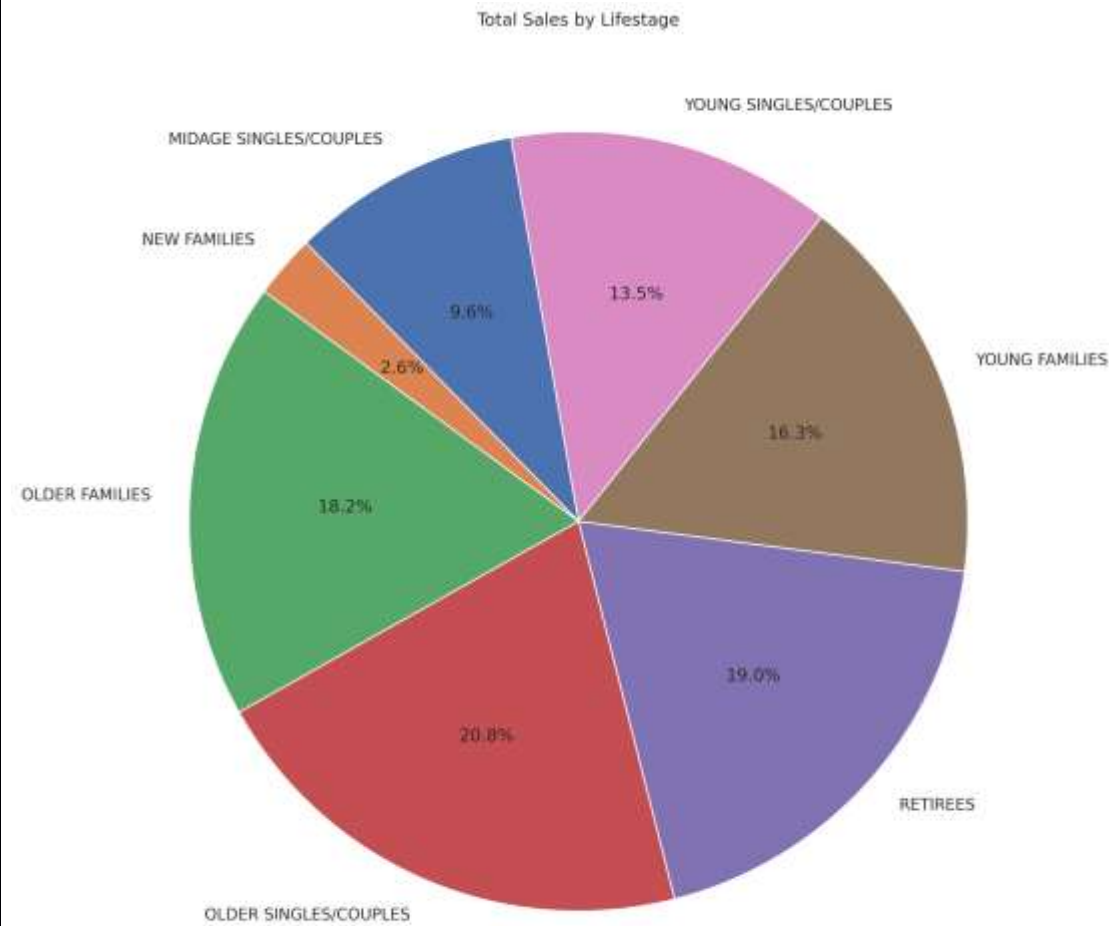
- Monthly overall sales revenue**
- Monthly number of customers**
- Monthly number of transactions per customer** Let's first create the metrics of interest and filter to stores that are present throughout the pre-trial period.

01

Category

```
Data columns (total 12 columns):
#      Column      Non-Null Count  Dtype
---  -
0     LYLTY_CARD_NBR  264834 non-null  int64
1     DATE            264834 non-null  object
2     STORE_NBR       264834 non-null  int64
3     TXN_ID          264834 non-null  int64
4     PROD_NBR        264834 non-null  int64
5     PROD_NAME       264834 non-null  object
6     PROD_QTY        264834 non-null  int64
7     TOT_SALES       264834 non-null  float64
8     PACK_SIZE       264834 non-null  int64
9     BRAND           264834 non-null  object
10    LIFESTAGE        264834 non-null  object
11    PREMIUM_CUSTOMER  264834 non-null  object
dtypes: float64(1), int64(6), object(5)
```

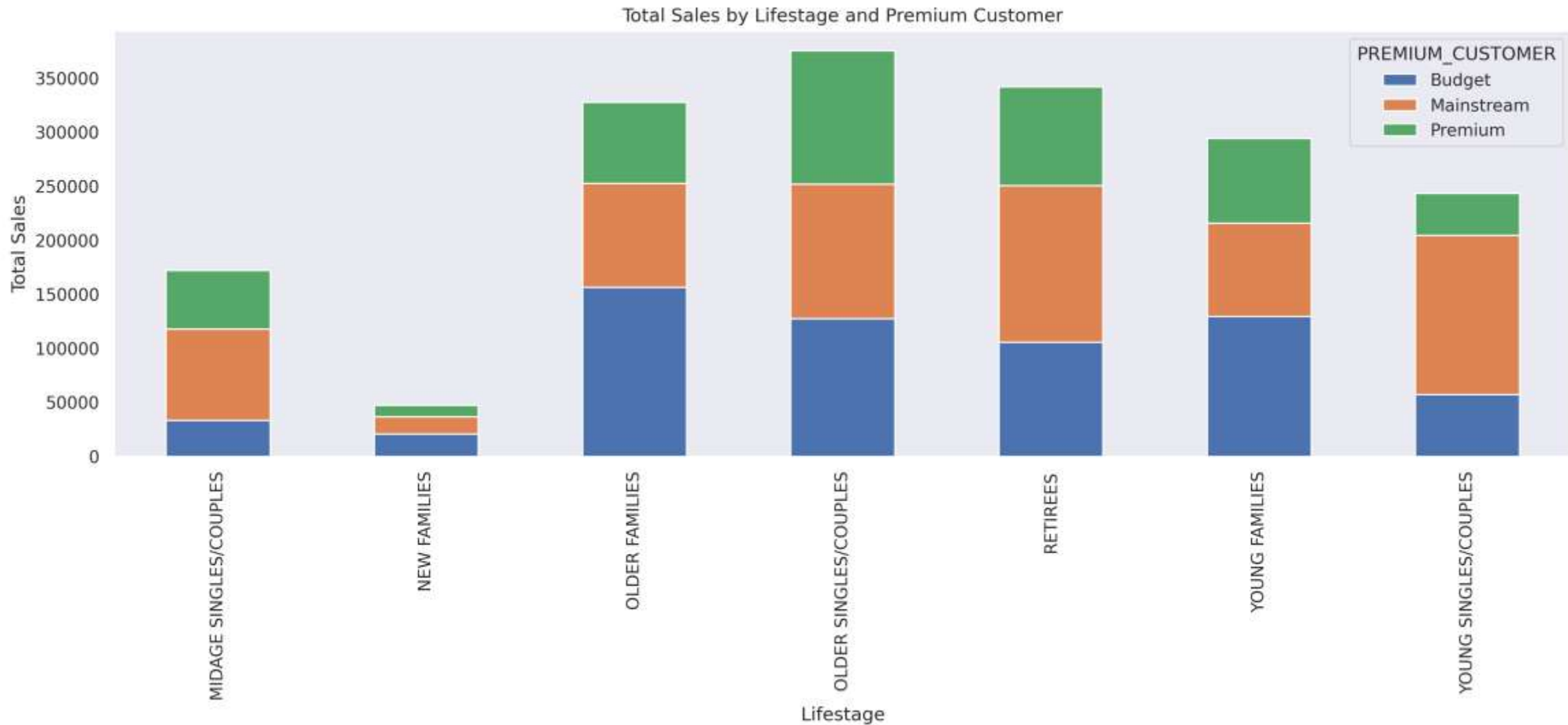
Overview:



Observations

- 1) Main outliers in this data usually reflect unexpected buying from customers.
- 2) Transactions of customers starkly increase in March and December month and August is likely just below them in the graph.
- 3) I searched December month to see those dates that created this difference and it was around Christmas which is common sense but we noticed the initial hit around 17-18 December which describes our range of action.
- 4) Maximum number of transactions by pack size is around 100k and its pack size is between 150-200.
- 5) Data analysis on customer segment reflects that customers who are in the category of older singles/couples & retirees with mainstream premium are likely to capture market dominance.
- 6) Maximum number of customers are also from mainstream premium ranging around 25-30k.
- 7) Older singles/couples and retirees are ranging maximum and numbered around 14k.
- 8) Total sales are also led by older singles/couples, contributing a whopping more than 350k in all premium sales with mainstream topping again with 700k.

The proportion of customers by affluence and life stage on this slide



02

Trial store performance

Control store vs other stores

- Control stores are similar to the trial stores in terms of these metrics, which will allow for a more accurate assessment of the impact of the trial.
- Trial stores are the stores where the new initiative or strategy is being implemented.
- Control stores are similar stores where the new initiative is not implemented.
- By comparing the trial stores to the control stores, we can isolate the impact of the new initiative.
- For example, if the trial stores show an increase in sales compared to the control stores, we can conclude that the new initiative is likely responsible for the increase.
- If the trial stores show no significant difference in sales compared to the control stores, we can conclude that the new initiative is not having a significant impact.
- In this case, the control stores are selected based on their similarity to the trial stores in terms of pre-trial metrics such as total sales, customer count, and transactions per customer.
- This ensures that the control stores are as similar as possible to the trial stores so that any differences in sales or customer behavior during the trial period can be attributed to the new initiative.

Performance in the trial store, determining if it was successful

The code analyzes the impact of a trial in selected stores (trial stores) compared to similar stores (control stores). It aims to determine whether the trial had a positive effect on sales and customer behavior.

Here's a theoretical explanation of how this analysis can be successful:

1. Identify Trial and Control Stores:

1. The analysis starts by selecting trial stores where a specific intervention (e.g., a new product placement, a promotional campaign) is implemented.
2. Control stores are chosen to be similar to the trial stores in terms of pre-trial sales, customer demographics, and other relevant factors. This ensures that any observed differences in performance during the trial period can be attributed to the intervention and not other external factors.

2. Pre-Trial Analysis:

1. The code calculates average monthly metrics (e.g., total sales, customer count, transactions per customer) for both trial and control stores during a period before the trial.
2. This pre-trial analysis establishes a baseline performance for each store, allowing us to compare the impact of the trial against the expected behavior.

3. Trial Period Analysis:

1. During the trial period, the code tracks the performance of both trial and control stores.
2. It calculates the percentage change in sales for the trial store compared to the control store.

4. Compare Trial and Control Store Performance:

1. If the trial store shows a significantly higher percentage change in sales compared to the control store, it suggests that the intervention was successful.
2. This indicates that the trial intervention positively impacted sales and customer behavior in the trial stores.

5. Conclusion:

1. If the trial store's performance is better than the control store's performance, it suggests that the implemented intervention was successful.
2. This analysis provides valuable insights into the effectiveness of the trial and can inform future decisions about implementing similar interventions in other stores or markets.

In essence, the code provides a framework for comparing the performance of stores with and without a specific intervention. By carefully selecting control stores and analyzing the pre-trial and trial period data, we can determine whether the intervention had a positive impact on sales and customer behavior.

Summary & Recommendation

Based on the analysis of the QVI data, we have identified several key insights regarding customer segments and their chip purchasing behavior, as well as the selection of control stores for the trial period. Customer Segmentation:

- **Lifestage and Premium Customer:** The analysis reveals that the 'Older Singles/Couples' and 'Young Singles/Couples' segments contribute significantly to total sales. Additionally, premium customers are also a key driving force for sales.

- **Pack Size:** The most popular pack sizes are 175g and 270g.

- **Brands:** 'Red' and 'Smith's' are the most popular brands.

Trial Store Analysis:

- We have identified control stores for each trial store (77, 86, 88) based on pre-trial metrics such as total sales, customer count, and transactions per customer.

- The control stores are similar to the trial stores in terms of these metrics, which will allow for a more accurate assessment of the impact of the trial.

Recommendations:

- **Target Customer Segments:** Focus marketing efforts on the 'Older Singles/Couples' and 'Young Singles/Couples' segments, as they represent the largest customer base and contribute significantly to sales.

- **Premium Customer Engagement:** Develop strategies to retain and engage premium customers, as they are a valuable customer segment.

- **Product Optimization:** Consider offering promotions or discounts on the most popular pack sizes (175g and 270g) and brands (Red and Smith's).

- **Trial Period Evaluation:** Monitor sales and customer behavior during the trial period in the trial stores and compare them to the control stores to assess the impact of the trial.

- **Data-Driven Decisions:** Continue to analyze data to identify new trends and insights that can inform business decisions.

By implementing these recommendations, Julia and the client can gain a deeper understanding of customer behavior and optimize their retail strategy for increased sales and profitability.



Disclaimer: This document comprises, and is the subject of intellectual property (including copyright) and confidentiality rights of one or multiple owners, including The Quantum Group Pty Limited and its affiliates (**Quantium**) and where applicable, its third-party data owners (**Data Providers**), together (**IP Owners**). The information contained in this document may have been prepared using raw data owned by the Data Providers. The Data Providers have not been involved in the analysis of the raw data, the preparation of, or the information contained in the document. The IP Owners do not make any representation (express or implied), nor give any guarantee or warranty in relation to the accuracy, completeness or appropriateness of the raw data, nor the analysis contained in this document. None of the IP Owners will have any liability for any use or disclosure by the recipient of any information contained in, or derived from this document. To the maximum extent permitted by law, the IP Owners expressly disclaim, take no responsibility for and have no liability for the preparation, contents, accuracy or completeness of this document, nor the analysis on which it is based. This document is provided in confidence, may only be used for the purpose provided, and may not be copied, reproduced, distributed, disclosed or made available to a third party in any way except strictly in accordance with the applicable written terms and conditions between you and Quantum, or otherwise with Quantum's prior written permission