



## **Quantium Virtual Internship**

**Data Analysis and Customer Insights**

## Executive Summary

This report presents a detailed analysis of chip sales data across Quantum's retail network for the 12-month period from July 2018 to June 2019. The analysis examines sales trends, product preferences, and customer purchasing behaviour across different life stages and premium-tier segments.

The goal is to equip the retail strategy team with clear, actionable insights into who buys chips, how much they spend and what drives their purchasing decisions, without requiring any technical background to interpret the findings.

### Key Findings:

- December is the biggest month for chip sales, driven by the lead-up to Christmas. Stores are closed on Christmas Day, resulting in zero recorded sales.
- Budget Older Families, Mainstream Young Singles/Couples and Mainstream Retirees are the three highest spending customer groups overall.
- Mainstream Young Singles/Couples represent the single largest group of customers buying chips.
- Older Families and Young Families tend to purchase more bags per shopping visit than other groups.
- Mainstream Young and Mid-Age Singles/Couples consistently pay a higher price per bag than Budget or Premium shoppers in the same age groups.
- Among Mainstream Young Singles/Couples, Tyrrells is disproportionately popular (purchased 23% more than average), while Burger brand is heavily avoided (56% below average).

## Data Preparation and Cleaning

### Data Quality Assessment

Before analysis, a comprehensive data cleaning process was performed in Excel to ensure data quality and consistency. The original dataset contained several issues that required attention:

### Cleaning Methodology

#### Product Name Standardisation

The PROD\_NAME column contained inconsistencies, including mixed case, extra whitespace, and merged pack-size information. Using Excel's TRIM and PROPER functions, text formatting was standardised across all product names.

## Pack Size Extraction

Pack sizes were embedded within product names and needed to be extracted into a separate column:

1. Product names were loaded into Power Query
2. Pack sizes (in grams) were extracted using text parsing functions
3. Inconsistent capitalisation of 'g' and 'G' was standardised using Find and Replace

## Brand Name Creation

A new BRAND column was created by extracting brand names from product names using TEXT and LEFT functions. Duplicate brand names (e.g., 'RED' and 'RRD' both referring to Red Rock Deli and 'Dorito' and 'Doritos') were identified and consolidated to ensure consistency.

## Outlier Removal

Data analysis revealed a customer who purchased 200 packets in a single transaction, an outlier indicating commercial rather than retail purchasing behaviour. This customer's transactions were removed from the dataset to prevent skewing of consumer insights.

## Product Category Filtering

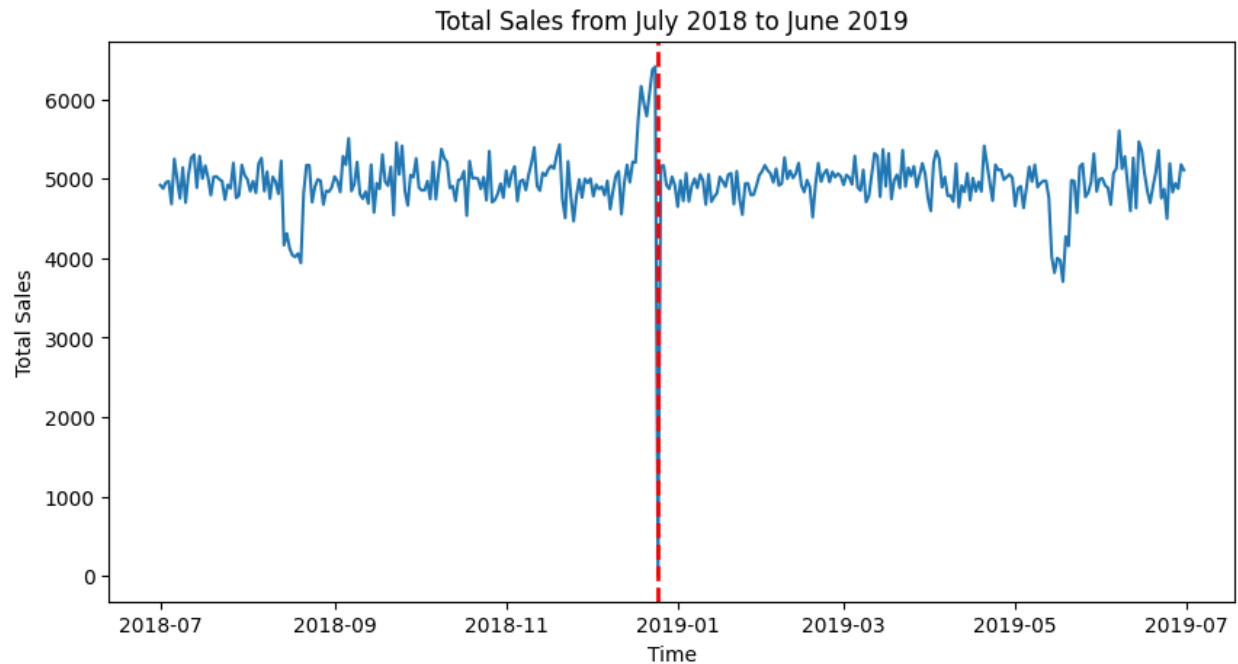
Non-chip products (such as salsa) were identified and removed from the dataset to ensure analysis focused exclusively on the chips category.

C	D	E	F	G	H	I	J	K	L	M
TY_CARD_NBR	TXN_ID	PROD_NBR	PROD_NAME	PROD_NAME_CLEANED	NEW_PROD_NAME_CLEANED	PROD_GRAMS	UNIT_PRICE	PROD_QTY	TOT_SALES	LIFESTAGE
1000	1	5	Natural Chip Compy SeaSalt 175g	Natural Chip Compy SeaSalt 175g	Natural Chip Compy SeaSalt	175g	3	2	6	YOUNG SIN
1307	348	66	CCs Nacho Cheese 175g	CCs Nacho Cheese 175g	CCs Nacho Cheese	175g	2.1	3	6.3	MIDAGE SI
1343	383	61	Smiths Crinkle Cut Chips Chicken 170g	Smiths Crinkle Cut Chips Chicken 170g	Smiths Crinkle Cut Chips Chicken	170g	1.45	2	2.9	MIDAGE SI
2373	974	69	Smiths Chip Thinly S/Cream&Onion 175g	Smiths Chip Thinly S/Cream&Onion 175g	Smiths Chip Thinly S/Cream&Onion	175g	3	5	15	MIDAGE SI
2426	1038	108	Kettle Tortilla ChpsHny&Jlono Chili 150g	Kettle Tortilla ChpsHny&Jlono Chili 150g	Kettle Tortilla ChpsHny&Jlono Chili	150g	4.6	3	13.8	MIDAGE SI
4074	2982	57	Old El Paso Salsa Dip Tomato Mild 300g	Old El Paso Salsa Dip Tomato Mild 300g	Old El Paso Salsa Dip Tomato Mild	300g	5.1	1	5.1	MIDAGE SI
4149	3333	16	Smiths Crinkle Chips Salt & Vinegar 330g	Smiths Crinkle Chips Salt & Vinegar 330g	Smiths Crinkle Chips Salt & Vinegar	330g	5.7	1	5.7	MIDAGE SI
4196	3539	24	Grain Waves Sweet Chili 210g	Grain Waves Sweet Chili 210g	Grain Waves Sweet Chili	210g	3.6	1	3.6	MIDAGE SI
5026	4525	42	Doritos Corn Chip Mexican Jalapeno 150g	Doritos Corn Chip Mexican Jalapeno 150g	Doritos Corn Chip Mexican Jalapeno	150g	3.9	1	3.9	MIDAGE SI
7150	6900	52	Grain Waves Sour Cream&Chives 210g	Grain Waves Sour Cream&Chives 210g	Grain Waves Sour Cream&Chives	210g	3.6	2	7.2	MIDAGE SI
7215	7176	16	Smiths Crinkle Chips Salt & Vinegar 330g	Smiths Crinkle Chips Salt & Vinegar 330g	Smiths Crinkle Chips Salt & Vinegar	330g	5.7	1	5.7	MIDAGE SI
8294	8221	114	Kettle Sensations Siracha Lime 150g	Kettle Sensations Siracha Lime 150g	Kettle Sensations Siracha Lime	150g	4.6	5	23	MIDAGE SI
9208	8634	15	Twisties Cheese 270g	Twisties Cheese 270g	Twisties Cheese	270g	4.6	2	9.2	MIDAGE SI
13213	12447	92	WW Crinkle Cut Chicken 175g	WW Crinkle Cut Chicken 175g	WW Crinkle Cut Chicken	175g	1.7	1	1.7	MIDAGE SI
19272	16686	44	Thins Chips Light& Tangy 175g	Thins Chips Light& Tangy 175g	Thins Chips Light& Tangy	175g	3.3	1	3.3	MIDAGE SI
20164	17136	54	CCs Original 175g	CCs Original 175g	CCs Original	175g	2.1	1	2.1	MIDAGE SI
20418	17413	94	Burger Rings 220g	Burger Rings 220g	Burger Rings	220g	2.3	4	9.2	MIDAGE SI
22411	18646	98	NCC Sour Cream & Garden Chives 175g	NCC Sour Cream & Garden Chives 175g	NCC Sour Cream & Garden Chives	175g	3	1	3	MIDAGE SI
22456	18696	93	Doritos Corn Chip Southern Chicken 150g	Doritos Corn Chip Southern Chicken 150g	Doritos Corn Chip Southern Chicken	150g	3.9	1	3.9	MIDAGE SI
23067	19162	56	Cheezels Cheese Box 125g	Cheezels Cheese Box 125g	Cheezels Cheese Box	125g	2.1	1	2.1	MIDAGE SI
25105	21815	7	Smiths Crinkle Original 330g	Smiths Crinkle Original 330g	Smiths Crinkle Original	330g	5.7	1	5.7	MIDAGE SI
33081	29949	98	NCC Sour Cream & Garden Chives 175g	NCC Sour Cream & Garden Chives 175g	NCC Sour Cream & Garden Chives	175g	3	1	3	MIDAGE SI
36012	32077	31	Infzns Crn Crnchers Tangy Gcmole 110g	Infzns Crn Crnchers Tangy Gcmole 110g	Infzns Crn Crnchers Tangy Gcmole	110g	3.8	1	3.8	MIDAGE SI
36302	33188	32	Kettle Sea Salt And Vinegar 175g	Kettle Sea Salt And Vinegar 175g	Kettle Sea Salt And Vinegar	175g	5.4	1	5.4	MIDAGE SI
38142	34181	108	Kettle Tortilla ChpsHny&Jlono Chili 150g	Kettle Tortilla ChpsHny&Jlono Chili 150g	Kettle Tortilla ChpsHny&Jlono Chili	150g	4.6	2	9.2	MIDAGE SI
39144	35506	57	Old El Paso Salsa Dip Tomato Mild 300g	Old El Paso Salsa Dip Tomato Mild 300g	Old El Paso Salsa Dip Tomato Mild	300g	5.1	1	5.1	MIDAGE SI
39167	35638	111	Smiths Chip Thinly Cut Original 175g	Smiths Chip Thinly Cut Original 175g	Smiths Chip Thinly Cut Original	175g	3	2	6	MIDAGE SI
41423	38393	46	Kettle Original 175g	Kettle Original 175g	Kettle Original	175g	5.4	1	5.4	MIDAGE SI
41486	38472	13	Red Rock Deli Thai Chili&Lime 150g	Red Rock Deli Thai Chili&Lime 150g	Red Rock Deli Thai Chili&Lime	150g	2.7	1	2.7	MIDAGE SI

## Sales Trends

### Full Year Overview (July 2018 – June 2019)

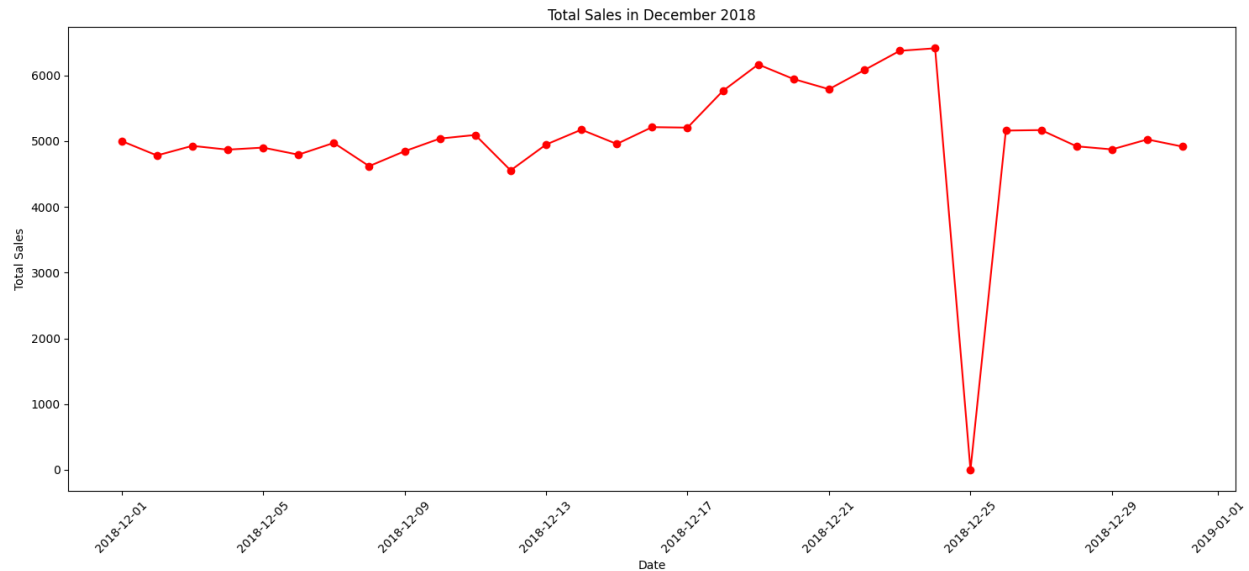
The chart below shows total chip sales for every day of the year. Sales remain relatively steady throughout the year, with a notable uplift as December approaches. The most striking feature is a complete drop to zero on 25 December 2018, Christmas Day, when all stores were closed.



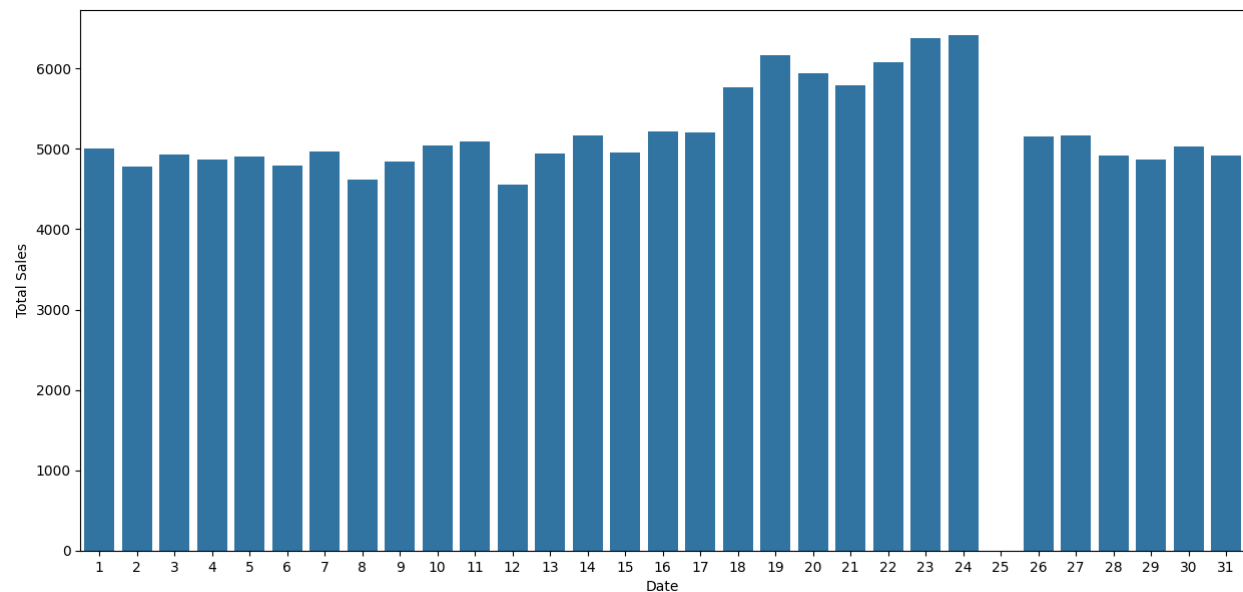
Outside of the holiday period, daily sales show a gentle seasonal rhythm that peaks slightly in the colder months. No other dates showed a complete sales dropout, confirming that Christmas Day is a unique outlier in the data.

## December 2018 in Detail

Because December has the highest concentration of elevated sales, it is worth examining it separately. The two charts below show daily December sales first as a line to highlight the trend, then as bars to show each individual day.



*Daily chip sales throughout December 2018. The sharp drop to zero on the 25th is clearly visible.*



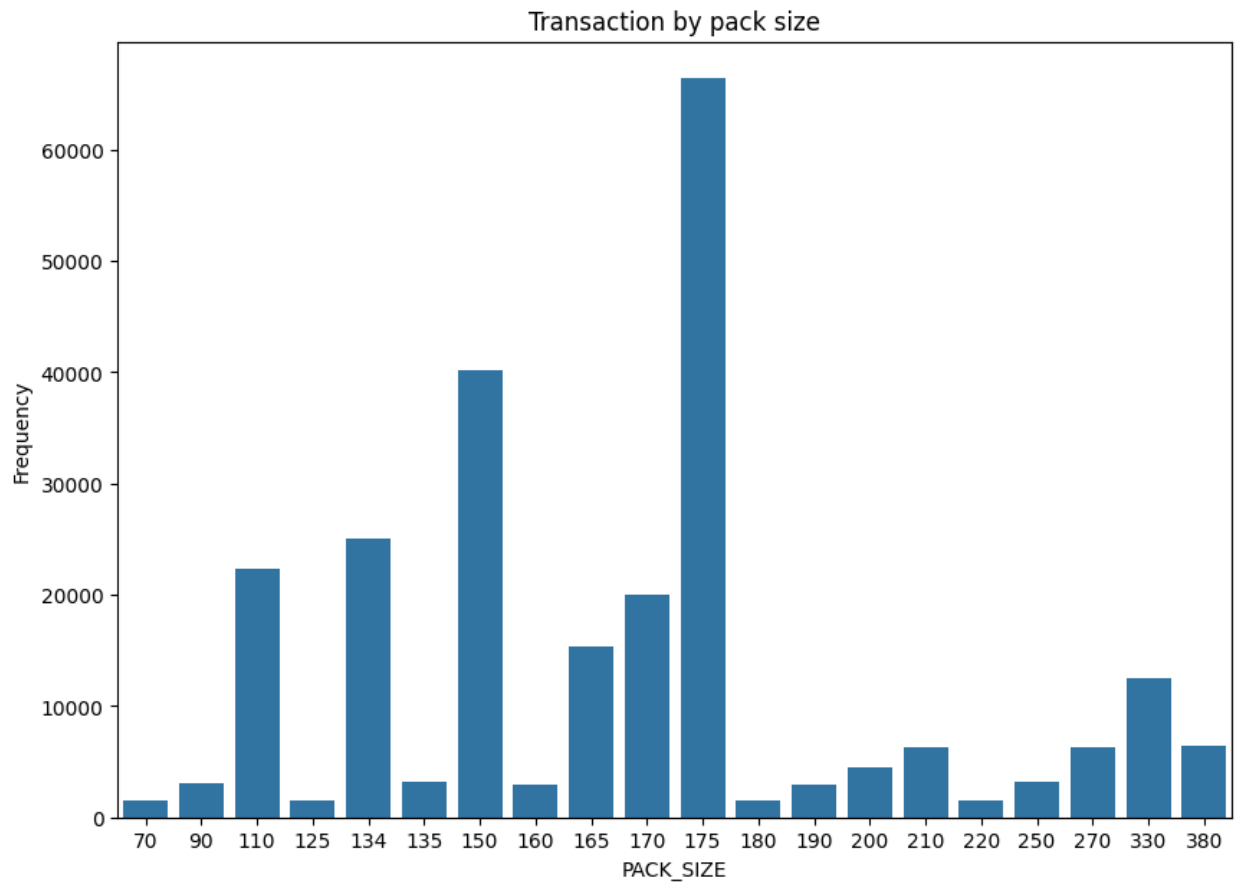
*Sales by day in December 2018. The bar chart shows no sales on Christmas Day (zero sales).*

Sales build consistently through the first three weeks of December as customers stock up ahead of Christmas gatherings and parties. After Christmas Day, sales recover quickly, though they do not reach the same pre-Christmas peak again before year's end. This pattern suggests customers are primarily buying chips for the Christmas period rather than for New Year celebrations.

## Product Preferences

### Most Popular Pack Sizes

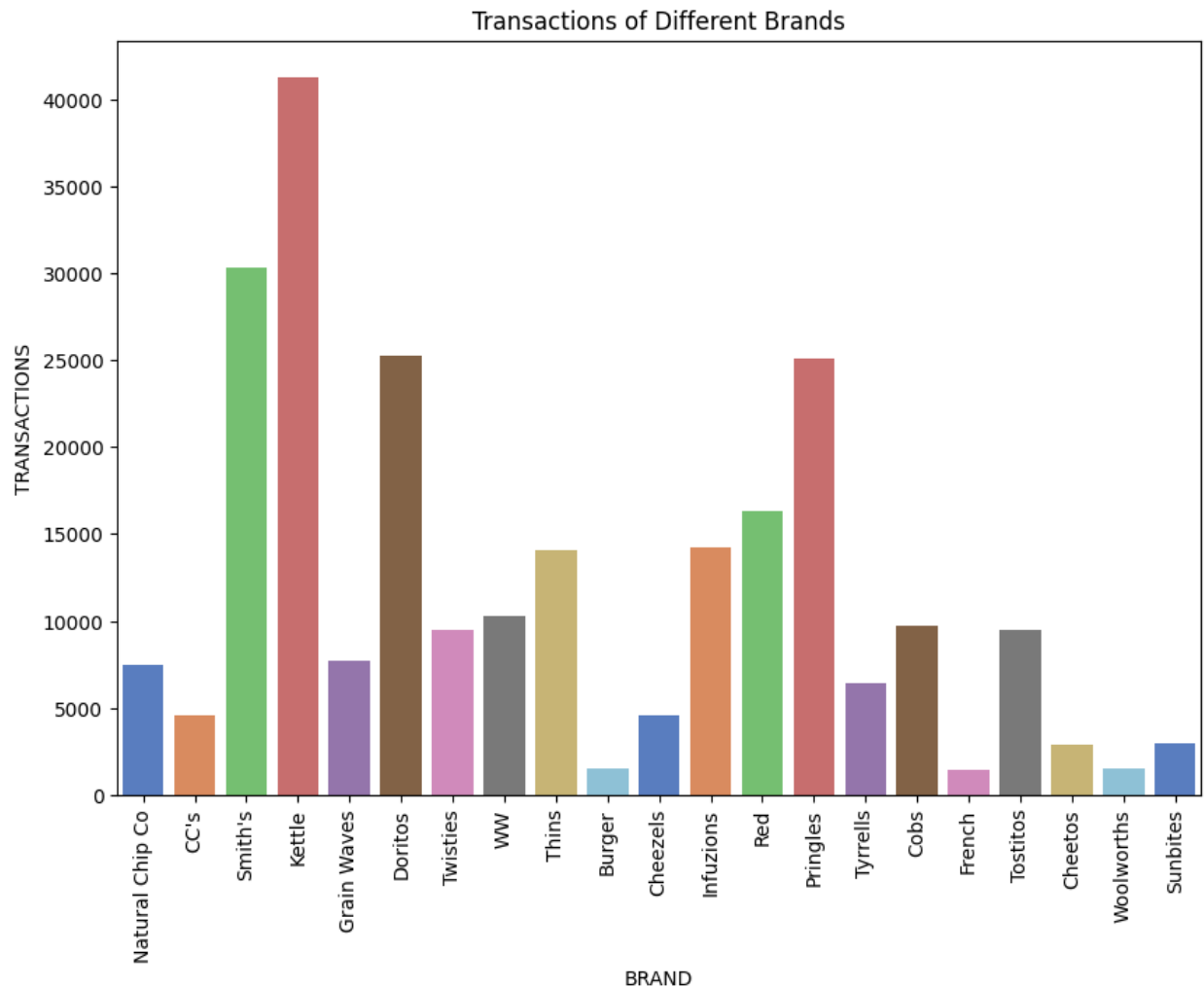
Chips are sold in a wide range of pack sizes. The chart below shows how often each pack size is purchased across all transactions. Pack size gives us a useful indicator of whether customers are buying for personal snacking, sharing, or stocking up in bulk.



The 175g pack is the dominant choice by a significant margin; it accounts for far more transactions than any other size. This suggests it hits the sweet spot between price, value and convenience. Smaller sizes (under 100g) and very large sizes (over 300g) are comparatively unpopular, suggesting that customers prefer mid-range bags for sharing or at-home consumption.

## Most Popular Brands

Beyond pack size, brand preference tells us which products customers actively seek out. The chart below ranks all chip brands by total number of transactions during the year.



Kettle is the clear market leader, outperforming the second-placed brand (Smiths) by a comfortable margin. The top five brands, Kettle, Smiths, Pringles, Doritos and RRD, together account for a disproportionate share of all chip transactions, following a typical power-law distribution where a small number of brands generate most of the sales activity.

## Understanding Our Customers

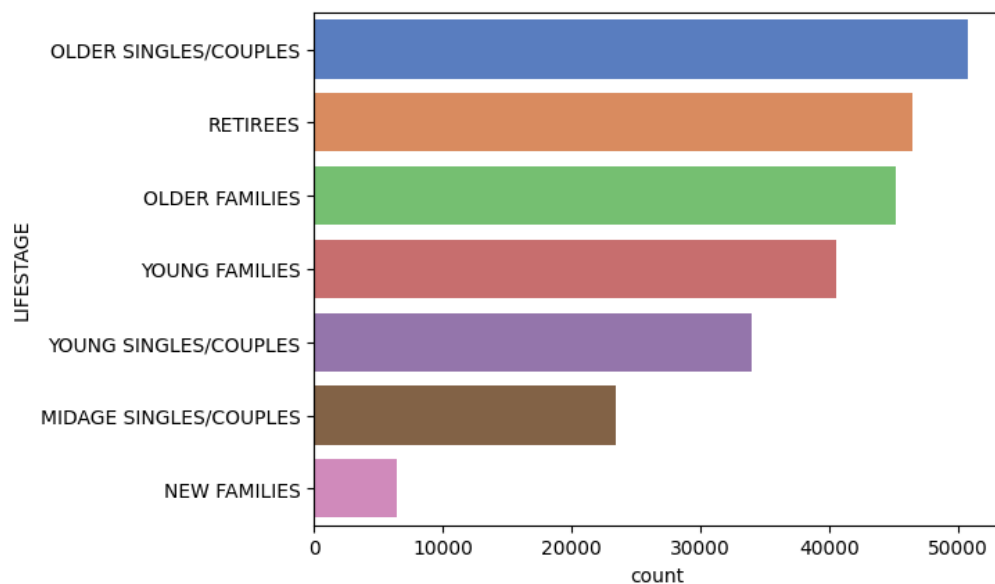
To develop effective sales strategies, it is not enough to know what is being sold; we need to understand who is buying it. Customers in this dataset are described along two dimensions:

- Lifestage
  - How customers are categorised by their life situation, for example, whether they are young singles, families with young children, or retirees.
- Premium Tier
  - Overall, whether customers shop in a Budget, Mainstream, or Premium style.

## How Many Customers Are in Each Group

### By Lifestage

The chart below shows the number of unique customers within each lifestage category.

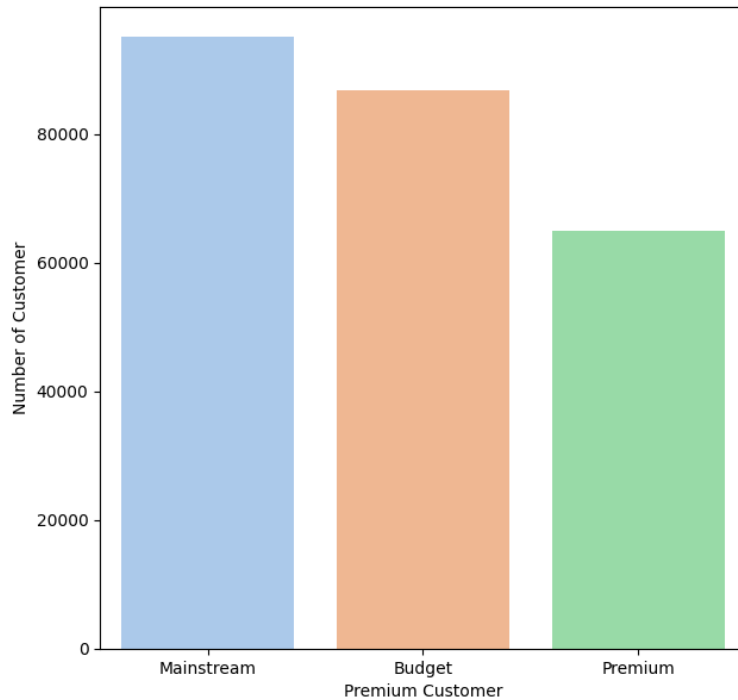


Older Singles/Couples is the largest lifestage group, followed by Retirees and Older Families. New Families represent a very small slice of the chip-buying customer base, likely because this group is still establishing household routines and may prioritise other food categories.



### By Premium Tier

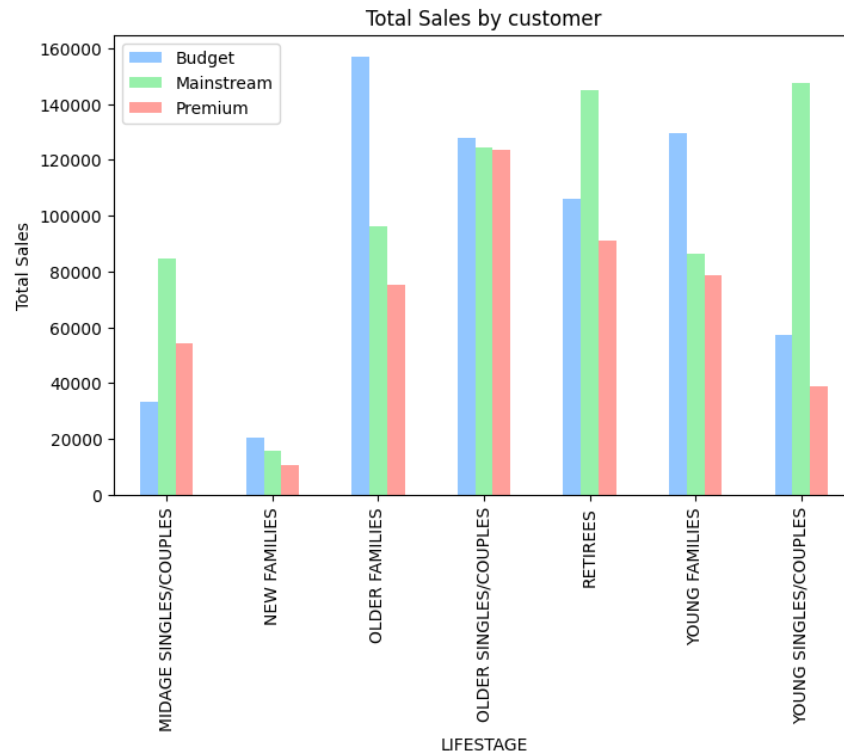
The chart below breaks down the same customer population by their shopping tier.



Mainstream is the largest lifestage group, followed by Budget and Premium. Premium represents a very small slice of the chip-buying customer base, likely suggesting this segment either prefers other snack categories or buys chips less frequently.

## Who spends the Most on Chips?

Now that we know how customers are distributed across segments, we can examine how much each group spends in total. This identifies which combinations of lifestage and Premium tier drive the most revenue for the chip category.

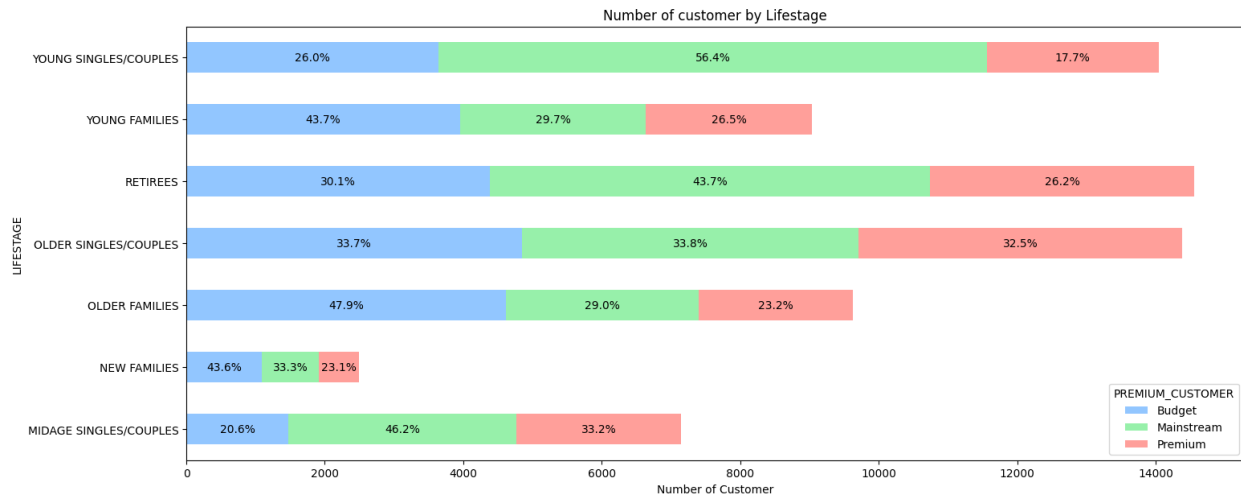


Three groups stand out as the highest contributors to total chip sales:

1. Budget Older Families
  - Despite not being the largest group by customer count, Budget Older Families generate significant total sales, indicating they buy in larger quantities per visit.
2. Mainstream Young Singles/Couples
  - Despite not being the largest group by customer count, Budget Older Families generate significant total sales, indicating they buy in larger quantities per visit.
3. Mainstream Retirees
  - Retirees with a Mainstream spending style represent a loyal and consistent customer segment with strong total sales.

## Sales Mix by Lifestage

The chart below shows, for each lifestage, what proportion of their total chip spending comes from Budget, Mainstream, and Premium shoppers. This reveals how the three tiers compete within each life group.



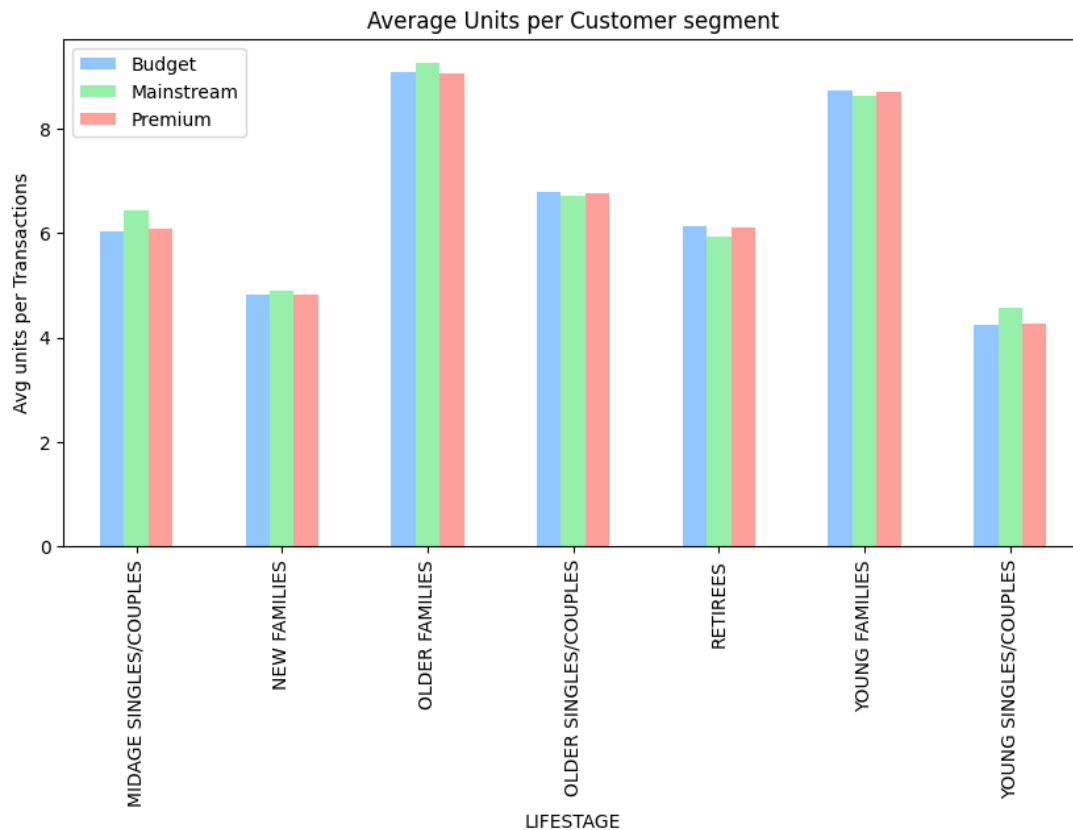
Mainstream spending dominates across most lifestage groups. However, in the Older Families category, Budget shoppers punch above their weight; they contribute a larger share of sales relative to their size than other groups. Premium shoppers consistently represent the smallest share of chip sales across all lifestages, suggesting this segment either prefers other snack categories or buys chips less frequently.

## How Do Different Groups Buy?

Beyond the headline sales numbers, two key behavioural dimensions help explain why some groups spend more than others: how many bags they buy per visit and how much they pay per bag.

### Average Number of Bags Purchased per Customer

The chart below shows the average number of chip bags purchased per customer within each segment. A higher number indicates that customers in that group consistently buy more bags when shopping for chips.



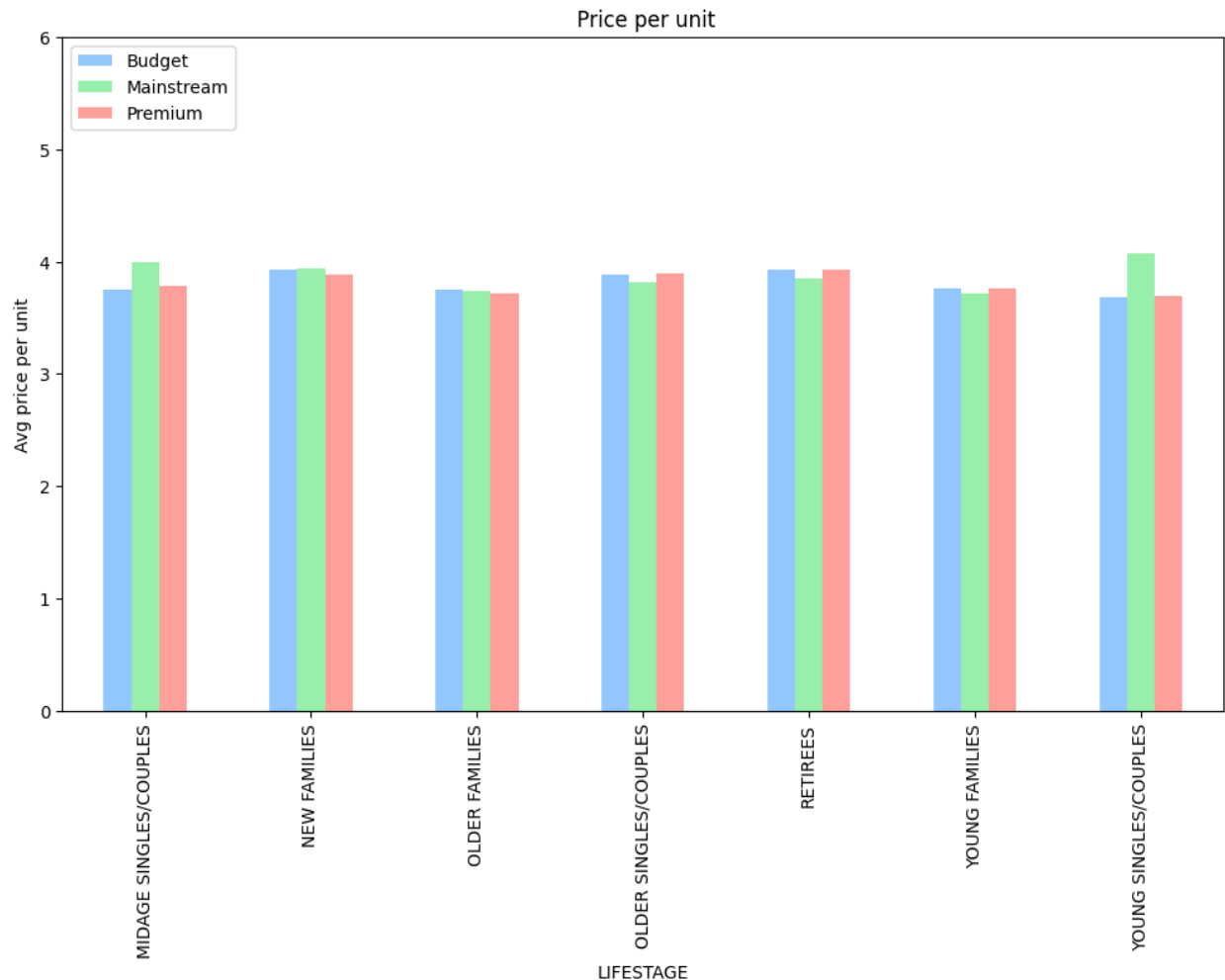
Older Families and Young Families buy significantly more bags per customer than any other group. For Older Families, this helps explain why their total sales are so high despite a relatively modest customer base; each customer simply buys more. Budget-tier families are particularly prolific buyers, likely stocking up for multiple household members in a single shopping trip.

### Insight: Family Buying Behaviour

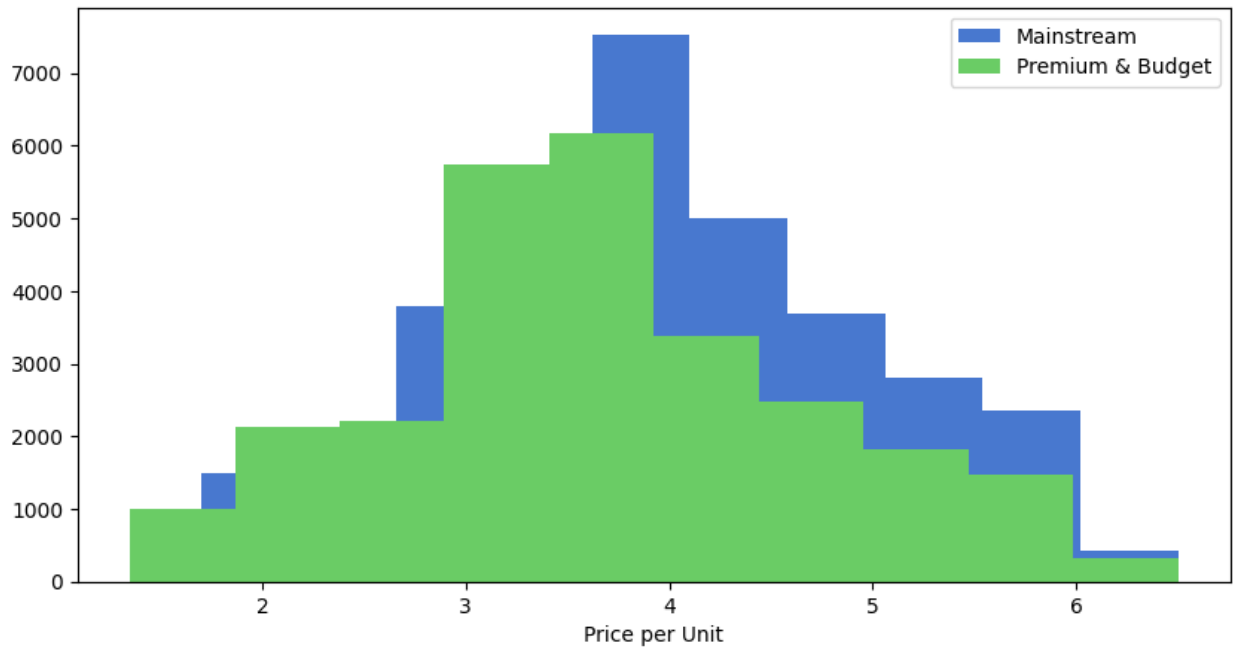
Family segments are bulk buyers. Multi-pack promotions, bundle deals and larger pack sizes are likely to resonate strongly with both Older and Young Families.

## Average Price Paid per Bag

While some groups buy more bags, others may spend more per bag. The chart below examines the average price per bag across all segments.



Mainstream shoppers consistently pay a slightly higher price per bag than Budget or Premium shoppers. This is particularly evident among Mainstream Mid-Age and Young Singles/Couples. Counterintuitively, Premium-tier shoppers do not pay more for chips; this likely reflects that Premium shoppers are more selective about their snacking choices and tend to buy chips less frequently, often for entertaining rather than as a personal staple.



The histogram confirms the pricing difference is genuine and statistically significant. Mainstream Young and Mid-Age Singles/Couples pay, on average, approximately \$0.30 more per bag than their Budget and Premium counterparts. A statistical t-test confirms that this difference is not due to chance ( $p < 0.05$ ), indicating it is a reliable pattern in customer behaviour.

### **Insight: Mainstream Price Sensitivity**

Mainstream Young and Mid-Age Singles/Couples show willingness to pay a modest premium for chips. This segment is well-suited to mid-range premium product launches or single-serve premium formats

## Deep Dive: Mainstream Young Singles & Couples

Given that Mainstream Young Singles/Couples are the single largest chip-buying customer group, it is valuable to understand their product preferences in more detail. Do they gravitate towards certain brands or pack sizes that other customers do not?

### Brand Affinity

The analysis compares the share of chips purchased by brand for Mainstream Young Singles/Couples versus all other customers. A score above 1.0 means this group buys that brand more than average; below 1.0 means they buy it less.

Brand	Affinity Score	What This Means
Tyrrells	1.23 ↑	23% more likely to purchase than other groups
Twisties (270g)	1.18 ↑	18% more likely strongly prefer this large pack format
Kettle	1.05 ↑	Slight preference, consistent with the overall market
Burger Brand	0.44 ↓	56% less likely to purchase than other groups
Pringles	0.92 →	Slight under-index, near average

### Pack Size Affinity

When it comes to pack size, Mainstream Young Singles/Couples show a clear preference for the 270g size, which is significantly higher than for other customer groups. This size is sold exclusively by Twisties, aligning with their strong brand affinity, as shown above. It suggests that for this customer group, the appeal of Twisties and the convenience of a larger share-size bag are closely linked.

### Targeting Recommendations

Mainstream Young Singles/Couples are a high-value, brand-aware segment. In-store placements of Tyrrells and Twisties (270g) near areas frequented by this demographic, alongside targeted promotions, are likely to drive incremental sales.

## Conclusions & Recommendations

The analysis of the 2018–2019 chip sales data reveals a clear picture of who is driving the chip category and how. Below are the key recommendations for the retail strategy team.

Priority Action	Rationale
Maximise Christmas stocking (1–24 Dec)	Sales peak sharply in December; a single day of stockout in this period represents significant lost revenue
Prioritise the 175g pack size on shelves	It is the most purchased size across all customer types; insufficient availability risks losing sales to competitors
Protect Kettle brand availability	Kettle leads all brands by transaction count; it drives store visits and basket attachment
Launch bundle promotions for Family segments	Older and Young Families buy the most bags per visit; multi-pack deals will increase basket size and loyalty
Target Mainstream Young Singles with Tyrrells & Twisties 270g	These are the brands this high-volume segment actively seeks; targeted placement and promotions will drive spend
Consider mid-tier premium products for Mainstream Singles	This group pays above average per bag; there is headroom to introduce premium single-serve formats at a higher price point

These recommendations are grounded in observed purchasing behaviour across the full year of sales data. Implementation should be prioritised ahead of the next trading period in December to maximise impact.