1.

The core business problem for this case study is to evaluate the performance, and assess the longevity, of the All-Day Breakfast promotion, which launched on October 3rd, 2015. The central challenge is to assess the financial consequences of expanding the McDonald's morning menu to be available during all operating hours. The fundamental uncertainty is whether this change will attract new business and increase overall sales, or simply shift customer spending away from our core, higher-value items, leading to a net loss in revenue.

To do so, the following questions must be answered using the provided data:

- How did the promotion impact revenue?
- What demographic groups were most and least influenced by the promotion?
- What were the most popular breakfast items and how did their performance compare to regular food items, offered simultaneously?

2.

The variables that relate most directly to my business problem are as follows:

- For investigating demographics:
'incomeq_label', 'urban_label', 'social_label', 'lstage_label', 'ppop_09q_label', 'pgrowthq_label'
- For investigating revenue and quantity sold:
'URWS', 'UPT', 'AGC', 'ADUS', 'totunits'
- Time data: 'WK_ENDING', 'LAUNCH_DATE'
- For unit-sale breakdown: 'CATEGORY'

3.

There exists inherent bias from weekly aggregation in 'WK_ENDING', considering differences in weekends versus weekdays, promotion initialization time variance across location, and location type.

Theoretically, the values of 'URWS' and 'totunits' should be equal based on the data dictionary provided. However, this is not the case. In my cleaned dataset, there were 6,248 mismatches between these two variables across all records.

The variable, 'UPT', representing the unit sales rate per 1000 transactions, is problematic. Since this variable counts the same unit as 'URWS' and 'AGC', at a larger scale, using this variable risks double counting units sold per transaction.

A similar issue exists across demographic variables in the dataset. For example, the variable 'social_label' provides ten categories for providing labels to different social demographics. It completely diminishes the value of the 'urban_label', which has only four categories, by

providing a more descriptive summary of the population. The same problem arises with the ‘pgrowth_label’ variable, intuition would bring one to conclude that faster growing areas are second city, suburban, or urban, rather than town and rural, again, diminishing the value of the variable when ‘social_label’ exists.
I removed over 100k entries from the data. Some items in the DataFrame had "total" rows where there were summations for all items in that category, as well as individual rows for each of those items, leading to double counting many values.

Rows after removing 'Total' aggregates: 221,130
Rows after dropping outliers: 220,074

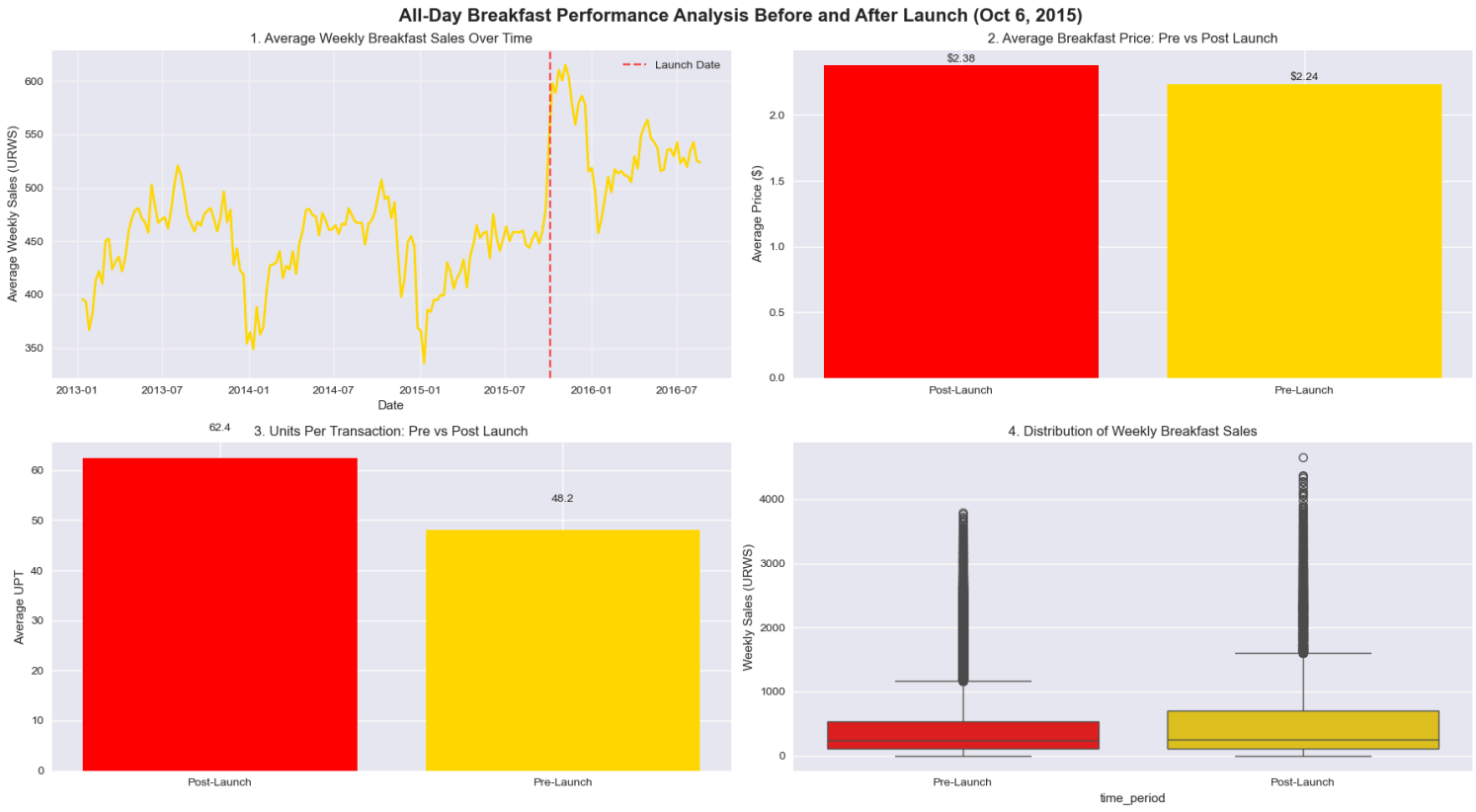
I removed outliers using a z-score method within each restaurant and menu category combination. Records were flagged as outliers if their scaled deviation from the median exceeded a threshold of 5.0. This process identified and removed 1,056 outlier records, with URWS having significantly more outliers (5%) than AGC (0.5%).

After cleaning the data, below are the amount of records for each broad food category.

MENU_CATEGORY
Non-Breakfast 110148
Breakfast 109926

There is a fairly even distribution of records between breakfast and non-breakfast food items.

4.
Initial Exploration of McDonald's Breakfast Data, finding impact of promotion launch:



Question 1: How did the promotion impact revenue?

Plot 1: Average Weekly Breakfast Sales Over Time
- This line plot shows the trend of average weekly breakfast sales (URWS) over time, with a vertical line marking the launch date of the new breakfast promotion on October 6, 2015. It helps visualize any changes in sales patterns before and after the promotion.
- Sales increased by approximately 20% after the launch.

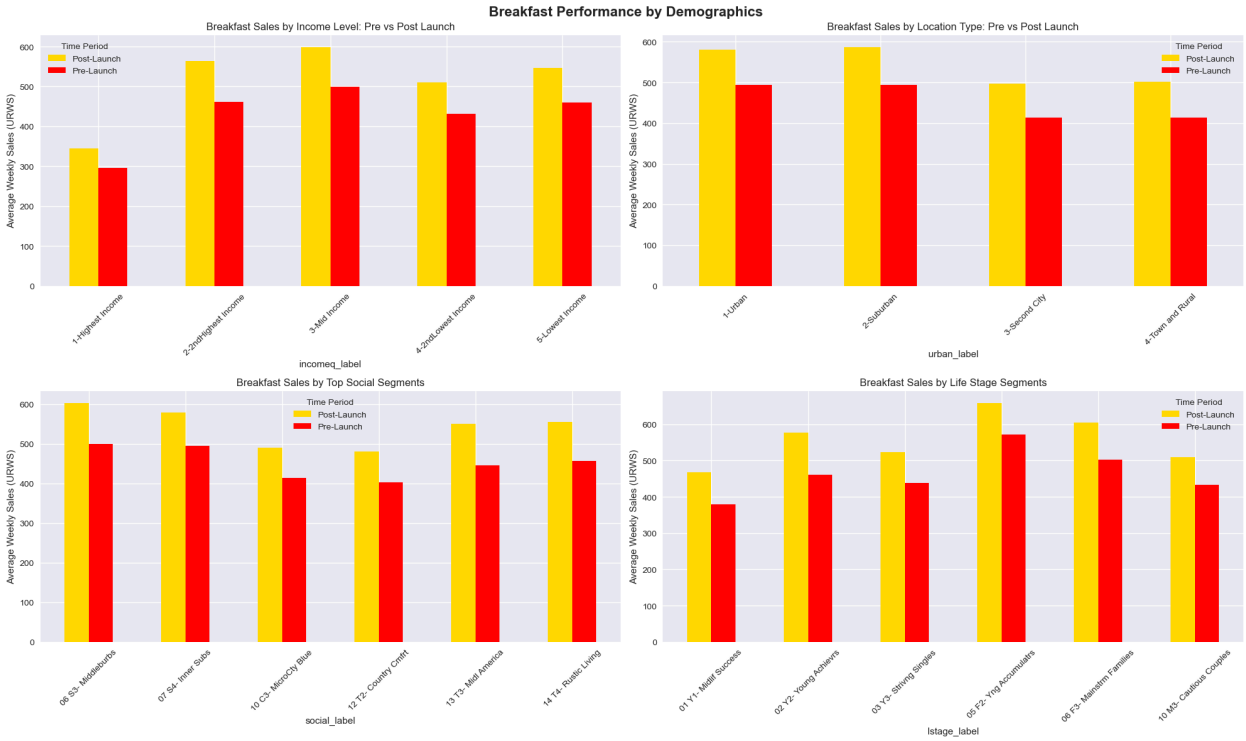
Plot 2: Average Breakfast Price: Pre vs Post Launch
- This bar chart compares the average price of breakfast items before and after the promotion launch. It provides insights into whether pricing strategies changed with the new promotion.
- The average price decreased slightly from $2.38 to $3.24. This likely indicates a slight shift in customer preference towards the cheaper breakfast items, over the more expensive burgers and combo meals.

Plot 3: Units Per Transaction: Pre vs Post Launch
- This bar chart compares the average units per transaction (UPT), measured as the unit sales rate per 1000 transactions for breakfast items before and after the promotion launch. It indicates whether customers are purchasing more or fewer items per visit.
- UPT decreased from 62.4 to 46.2, indicating that customers are buying fewer items on average each visit. This could be due to the introduction of more affordable breakfast options, leading to smaller purchases. On the converse, this could also indicate that customers are substituting breakfast items for other menu items, rather than adding breakfast items to their existing orders.

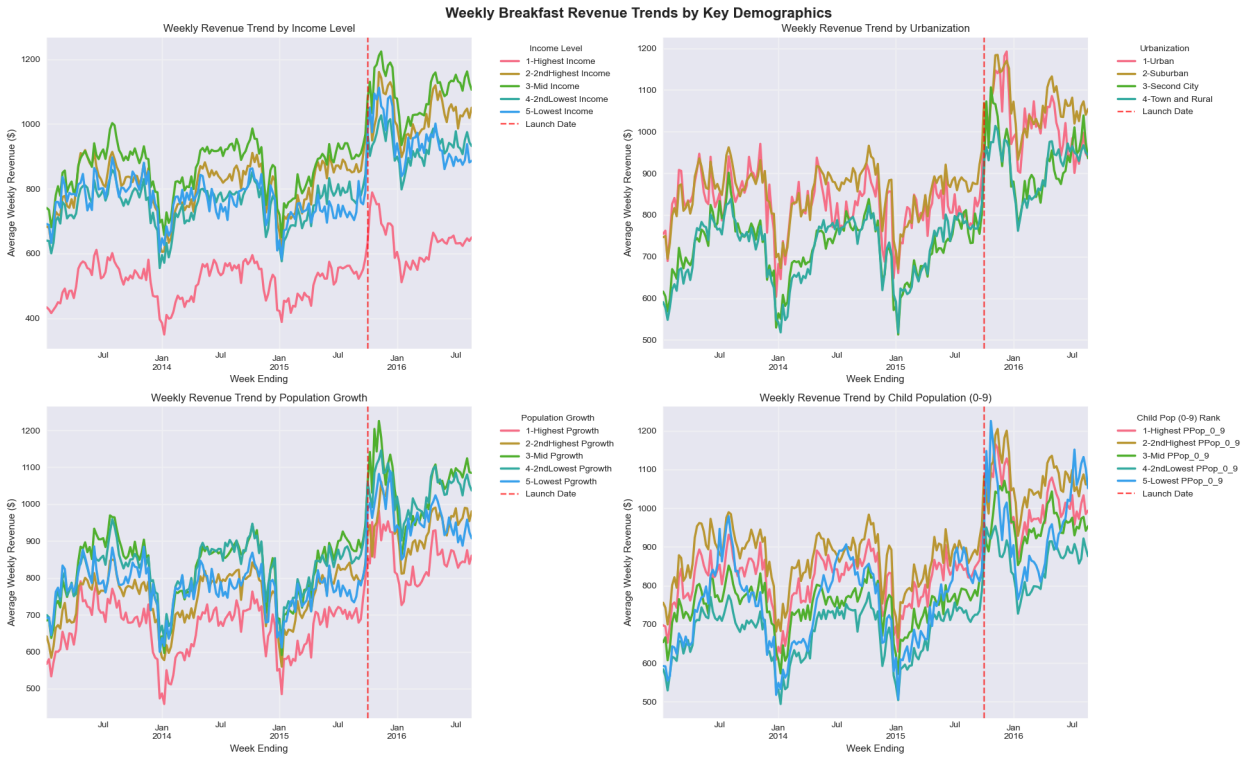
Plot 4: Distribution of Weekly Breakfast Sales
- This box plot illustrates the distribution of weekly breakfast sales (URWS) before and after the promotion launch. It highlights the median, quartiles, and potential outliers in sales data for both periods.
- The overall distribution of weekly breakfast sales appears similar before and after the launch, with a slight increase in median sales post-launch. However, there are more outliers in the post-launch period, indicating some weeks with exceptionally high sales.

Overall, the All-Day Breakfast promotion successfully increased breakfast sales by approximately 20% following its launch, demonstrating high consumer demand for expanded breakfast availability. Further, the decrease in average price per order suggests a considerable shift towards cheaper breakfast items. However, the decrease in units per transaction and slight price reduction suggest customers may offer some concerning insights. Customers could be substituting breakfast for higher-value items rather than adding to their orders, indicating some cannibalization of core menu sales alongside the incremental growth.

Question 2: What demographic groups were most and least influenced by the promotion?



These plots investigate demographic factors that may influence store performance, specifically focusing on average weekly breakfast sales (URWS) across different demographic segments. In all cases, for all groups in every demographic variable, post-launch sales outperform pre-launch sales by a factor of approximately 20%, as previously discussed.



These plots show quarterly sales trends for breakfast items across various demographic segments, comparing pre- and post-promotion launch periods. Each plot highlights how each demographic factor may influence breakfast sales performance.

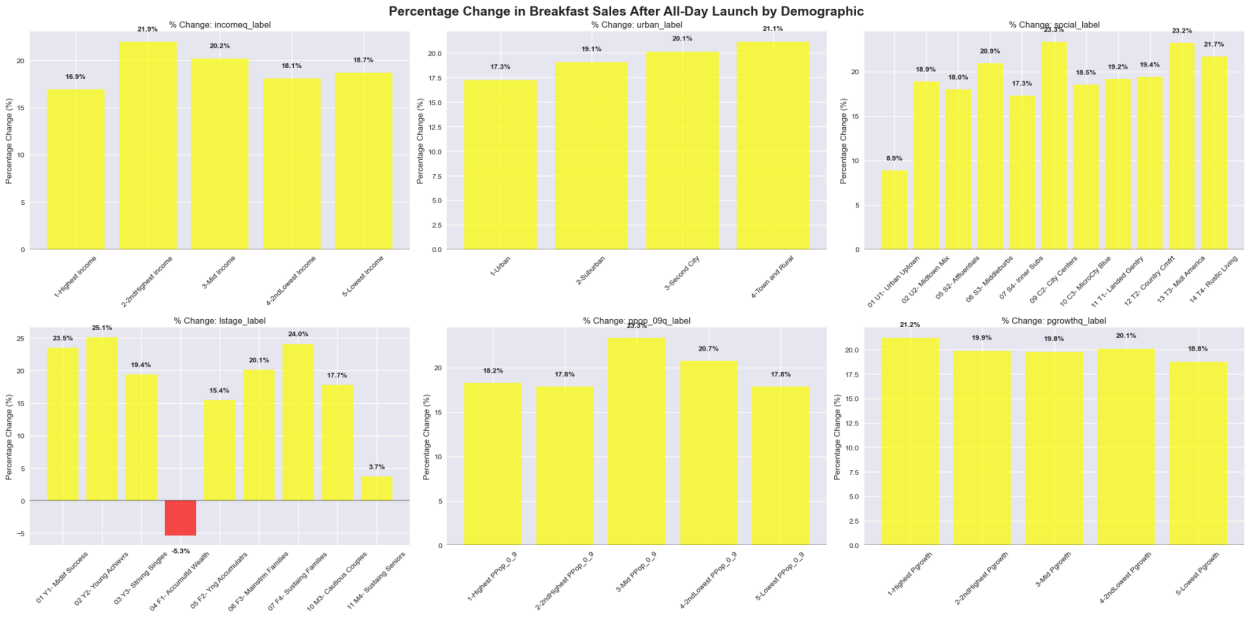
Plot 1: Weekly Revenue Trend by Income Level
- This line plot displays the average weekly revenue from breakfast sales segmented by income levels. It helps identify how different income groups responded to the breakfast promotion over time.
- Pre-Launch: The Highest Income group had the lowest baseline revenue. Revenue was generally highest in the Mid-Income and 2nd Lowest Income quartiles.
- Post-Launch: All groups produced a significant increase in revenue. The Mid-Income and 2nd Highest Income groups become the largest absolute revenue drivers. The Highest Income group sees a substantial percentage increase from its low baseline, suggesting ADB successfully engaged this population.

Plot 2: Weekly Revenue Trend by Urbanization
- This line plot illustrates the average weekly revenue from breakfast sales based on urbanization levels. It shows how urban, suburban, and rural areas responded to the breakfast promotion.
- Pre-Launch: Urban and suburban areas had similar high baseline revenues, while second city and rural areas had relatively lower revenue.
- Post-Launch: All areas saw significant revenue increases. Urban and suburban areas remained the top revenue drivers, but while, after a month, urban sales started to decrease back to baseline, suburban breakfast sales remained elevated. This suggests that suburban areas responded more positively to the breakfast promotion over the long term.

Plot 3: Weekly Revenue Trend by Population Growth
- This line plot presents the average weekly revenue from breakfast sales categorized by population growth rates. It examines how areas with different growth rates reacted to the breakfast promotion.
- Pre-Launch: All areas had similar revenue trends, with the highest revenue in the moderate growth quartile and the lowest in the highest growth quartile.
- Post-Launch: All growth categories experienced significant revenue increases. The moderate and low growth areas became the leading revenue contributors, while the highest growth areas, despite a substantial percentage increase, remained the lowest in absolute revenue. This indicates that while high-growth areas showed interest in the breakfast promotion, they did not translate into high sales. Mid-growth areas were the most responsive to the promotion.

Plot 4: Weekly Revenue Trend by Child Population (0-9)
- This line plot shows the average weekly revenue from breakfast sales segmented by the percentage of the population aged 0-9. It explores how areas with varying child population percentages responded to the breakfast promotion
- Pre-Launch: Areas with a higher percentages of children (0-9) had the highest baseline revenue, while areas with the lowest percentage of children had lower respective revenues,

except for one large spike in sales for the lowest children population two months before the promotion.
- Post-Launch: All segments saw significant revenue increases. All groups follow similar trends, plateauing after the promotion. Surprisingly, the group with the lowest amount of children (0-9) saw the highest absolute revenue post-launch, suggesting that the breakfast promotion appealed broadly across demographics, not just to families with young children.



Almost ALL groups had a rather large percent increase in sales, with an average increase of ~20%. The only group which showed a decrease in sales was the "Accumulated Wealth" group, which saw a decrease of about 4.1%. The group which saw the smallest increase in sales, likely since there is large overlap between the two groups, was the "Sustaining Seniors" group, which only saw a sales increase of 3.2% upon the promotion release.

The promotion produced consumer growth from almost all demographic groups, generating approximately 20% sales growth across nearly all customer segments, with suburban areas and mid-income groups showing the strongest sustained response. However, the promotion was least effective among the "Accumulated Wealth" demographic, which saw a 4.1% sales decline, and "Sustaining Seniors," who showed only minimal growth of 3.2%.

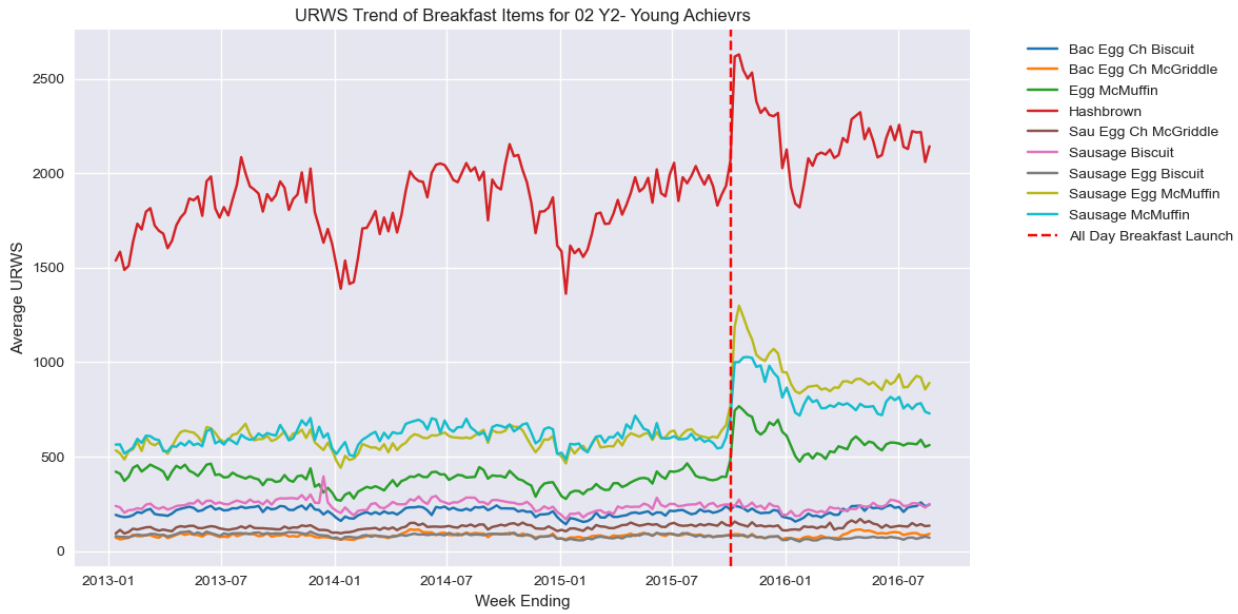
Question 3: What were the most popular breakfast items and how did their performance compare to regular food items, offered simultaneously?

Below is a breakdown of average price values across both menu categories and the frequency at which they were purchased.

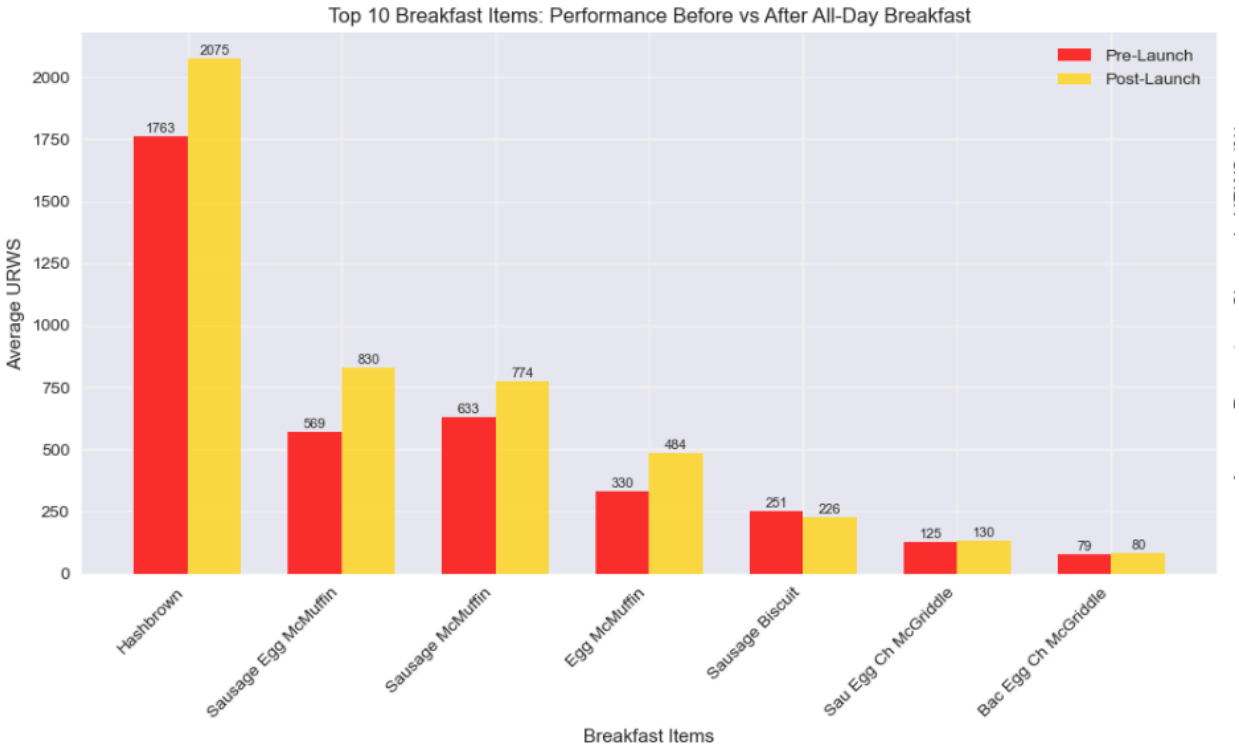
As expected, regular items have a wide variance of average price values when compared to breakfast items. The breakfast items' average price has two centers around $1 and $3, likely representing the two main items, hashbrowns and McMuffins, respectively. Same for regular items, except with fries, value burgers, and larger burgers, ranging from $1-$4.

social_label	Top 1	Top 2	Top 3	Top 4	Top 5
01 U1- Urban Uptown	Hashbrown	Sausage McMuffin	Sausage Egg McMuffin	Sausage Biscuit	Egg McMuffin
02 U2- Midtown Mix	Hashbrown	Sausage McMuffin	Sausage Egg McMuffin	Egg McMuffin	Sausage Biscuit
05 S2- Affluentials	Hashbrown	Sausage Egg McMuffin	Sausage McMuffin	Egg McMuffin	Sausage Biscuit
06 S3- Middleburbs	Hashbrown	Sausage McMuffin	Sausage Egg McMuffin	Egg McMuffin	Sausage Biscuit
07 S4- Inner Subs	Hashbrown	Sausage McMuffin	Sausage Egg McMuffin	Egg McMuffin	Sausage Biscuit
09 C2- City Centers	Hashbrown	Sausage Egg McMuffin	Sausage McMuffin	Egg McMuffin	Sausage Biscuit
10 C3- MicroCty Blue	Hashbrown	Sausage McMuffin	Sausage Egg McMuffin	Egg McMuffin	Sausage Biscuit
11 T1- Landed Gentry	Hashbrown	Sausage Egg McMuffin	Sausage McMuffin	Egg McMuffin	Sausage Biscuit
12 T2- Country Cmfrt	Hashbrown	Sausage Egg McMuffin	Sausage McMuffin	Egg McMuffin	Sausage Biscuit
13 T3- Midl America	Hashbrown	Sausage Egg McMuffin	Sausage McMuffin	Egg McMuffin	Sausage Biscuit
14 T4- Rustic Living	Hashbrown	Sausage McMuffin	Sausage Egg McMuffin	Egg McMuffin	Sausage Biscuit

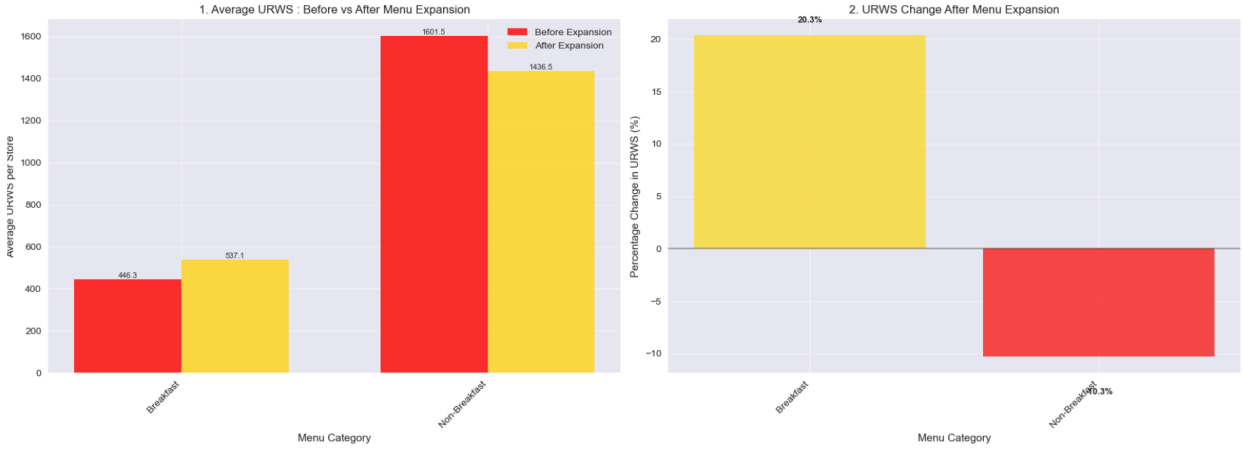
The table above shows item preference for each demographic. The favorites remain largely the same. Everyone's favorite breakfast item are hashbrowns, followed by sausage or sausage and egg McMuffin.



When comparing this same plot for all demographic groups, hashbrowns dominated across all demographic groups as the top selling breakfast item, both pre- and post-launch. This indicates a strong and consistent preference for this item among customers. Second, very similar across all demographics, were the Sausage Egg McMuffin and Sausage McMuffin, which traded off as the second most popular item. Similarly, and following in third place, was the Egg McMuffin. The impact of ADB on all other items' sales trends was seemingly negligible. There were no spikes in sales following the release of the promotion for the following items: Sausage Biscuit, Sausage Egg Biscuit, Bacon Egg and Cheese Biscuit, Sausage Egg and Cheese McGriddle, and the Bacon Egg and Cheese McGriddle. This could have been influenced by what Professor Patchell said in the video in the company's decision to choose only the McMuffin line of sandwiches for the promotion, as they were already the most popular items in Michigan at the time.



All items featured on Michigan's All Day Breakfast menu, as previously mentioned hashbrowns and variations of McMuffins, saw rather high growth in average unit sales per week. This suggests a great success in all around breakfast sales as a result of implementing the promotion.



Plot 1: URWS: Before vs. After Menu Expansion

- The breakfast menu items' average URWS per store increased substantially after the menu expansion, rising from approximately 537 to almost 1437 units, representing a dramatic 150%+ growth. The non-breakfast items sales grew at a higher rate than breakfast, indicating the promotion successfully drove breakfast performance without negatively impacting core menu items, and instead greatly promoting their sales in complement.

Plot 2: URWS Change After Menu Expansion

- The All-Day Breakfast promotion successfully produced a 20.3% increase in Breakfast URWS, demonstrating clear demand for breakfast items throughout the day. The relatively moderate 10.3% decline in Non-Breakfast URWS indicates a partial substitution effect where some customers are choosing breakfast items instead of traditional menu options, although the net result remains positive.

McDonald's All-Day Breakfast promotion was highly successful, driving a dramatic 150%+ increase in breakfast item sales, while minimizing the substitution effects with the core menu. Michigan's focus on the already-popular McMuffin items and hashbrowns proved highly effective, with these products demonstrating strong, consistent demand across all demographic groups and generating significant incremental sales growth.

5.

This data can help answer only some of the necessary questions in determining the success of McDonald's All Day Breakfast promotion. The primary issue with the data is the absence of any financial data. The variables existing in the data, like 'URWS' (Units sold per week),'wavg_price' (weighted average price across menu item or category), etc., correspond only to the amount of

product sold, neglecting the inclusion variables which provide insight on financial revenue. This alone made it impossible to calculate the fiscal success of the promotion.

The inclusion of values such as revenue, profit margins, and cost data are fundamental in answering the associated business problem. Without these metrics, we cannot understand individual consumer behavior, in-cart-item competition (breakfast vs. non-breakfast items), or most importantly, the financial success from introducing All-Day Breakfast.

Further, the provided data, considering the client's desires, is insufficient towards analyzing store machine capacity performance for both menu item categories. For this, we would need store-level operation costs, individual machine capacities, and inventory data.

As previously mentioned, the data quality also severely impacts the interpretation of the visualizations. An ideal dataset would contain properly cleaned data that was collected from individual stores on an hourly basis. Considering the potential invalidity of the 'URWS' variable, from mismatches with a similar variable 'UTP', many of my conclusions could be disproven if provided the sufficient data.

Conclusively, while the provided data helps us determine the success, in quantity of goods sold, of the promotion, the financial impact is still unknown. To properly answer the business problem, the following data must be collected in addition to what we already have:

- Financial data consisting of revenue, item-level profit margins, and store-level financial statements
- Transaction-level statements from individual customer purchases
- Operational cost data containing store-level production costs and capacity utilization
- Properly cleaned version of the dataset already provided with increased granularity on an hourly basis.