Anusha Bhat

31-016

October 17, 2024

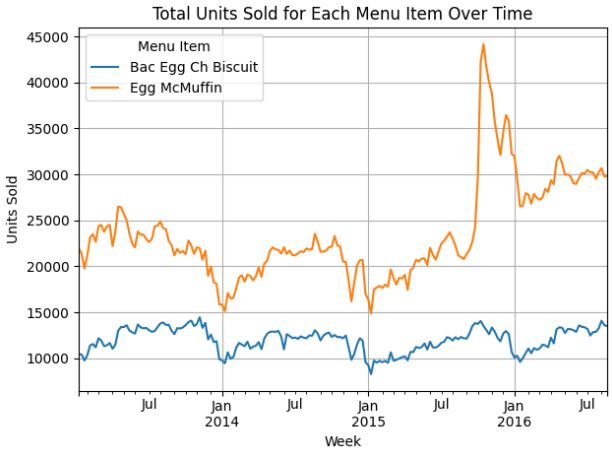
McDonald's Case Exploration

1) Business problem

Restaurants face many issues when it comes to sales, with one of the biggest issues being whether they should host promotions or not. While promotions can potentially bring in higher sales, they can lower profit margins for the restaurant due to operational costs and pricing. This is exactly the problem that the McDonald's Regional Manager is facing.

Over the past few years, the business has seen declining transactions and sales. However, they have noticed that when they do certain promotions like "All Day Breakfast", there is an increase in sales transactions. Additionally, they noticed that some customers like breakfast items but can't come early enough to buy them in the morning, so "All Day Breakfast" is beneficial for drawing in that customer base. In order to support "All Day Breakfast", though, the business must have enough equipment to make the breakfast items the entire day in addition to the regular day menu items since they typically only make breakfast items in the morning and then use the same equipment for the day time menu items after breakfast is over. Additionally, customers may come in and only buy breakfast items rather than more expensive items. These issues combined can lower profit margins even if transactions increase due to the promotion. So, the manager would like to investigate if there truly is a benefit to profit margins if the company hosts "All Day Breakfast". An alternative solution the business tried out was offering one breakfast menu item all day as opposed to the entire menu item. Restaurants had to choose between selling a Biscuit or McMuffin and the Michigan stores chose the McMuffin. The manager also wants to know if there was any difference in doing this promotion compared to the "All Day Breakfast".

We can begin to see that this concern is warranted when we observe the total weekly units sold between the Egg McMuffin and the Bacon and Egg Cheese Biscuit over the course of a few years.



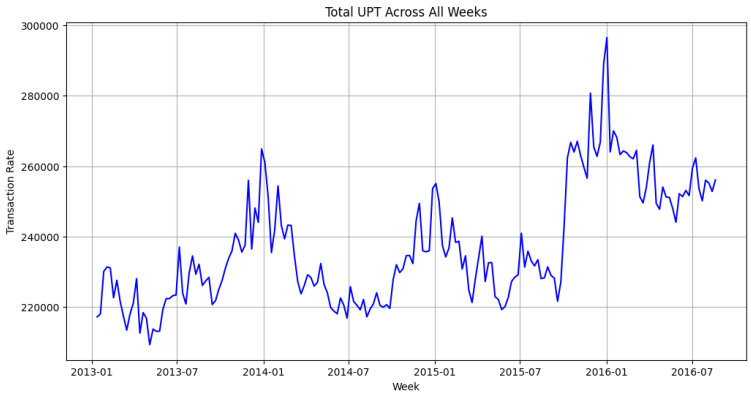
We can observe that the McMuffin is sold at higher rates than the Biscuit for the Michigan stores, so there is an issue regarding how to decide which items and promotions to offer.

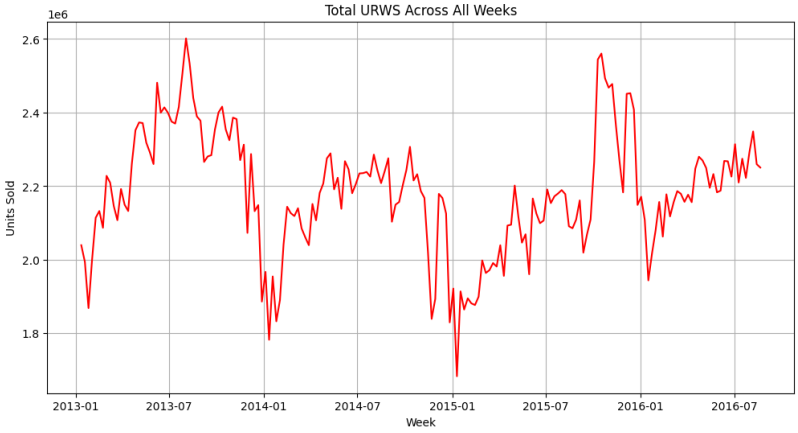
2) Important Variables

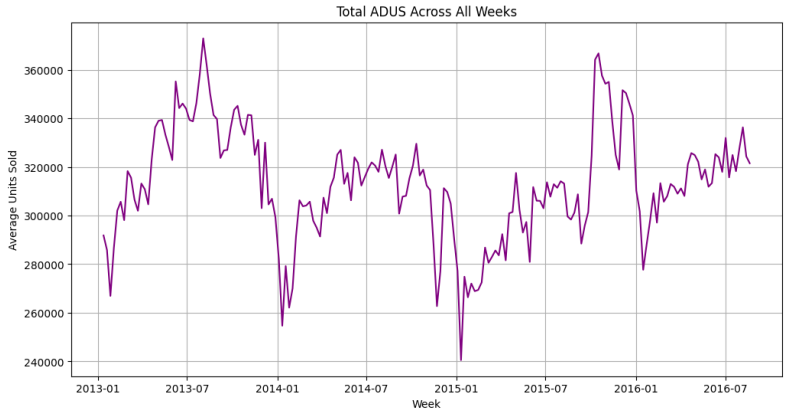
The most important variables are the restaurant key, city location, if the restaurant is traditional, item description, week of sales tracking period, units sold for the week, weighted average price, unit sales rate per 1000 transactions, average transactions per day, and average daily units sold. The item description, units sold, and transactional data is helpful to observe whether some items are sold more frequently than others, if some items are only bought at higher or lower prices, and making comparisons in sales across seasons/time. The restaurant key and city location is relevant as different locations have different customer bases, and each customer base has its own taste that influences which items are bought more in that location. Additionally, whether the restaurant is a traditional restaurant or not is interesting to the problem since location and style of restaurant may also impact the rates at which customers visit the shop and whether they will choose another restaurant over a specific McDonald's.

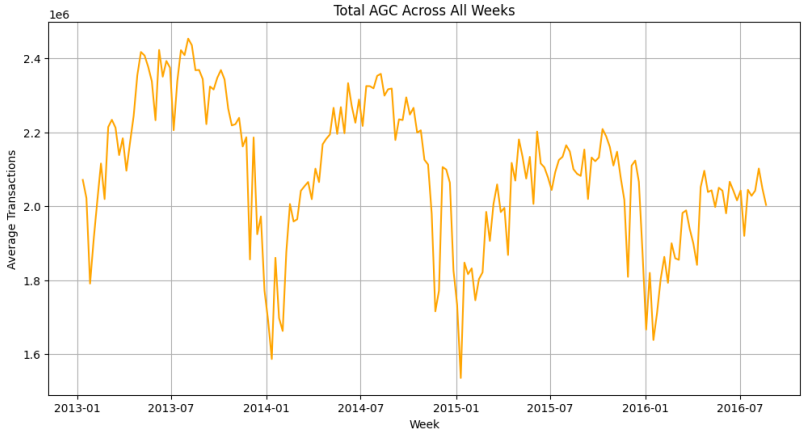
The longitude and latitude of each restaurant are not important as it's too granular of information. The county and state are also not important since we have the city information. Owner label is not important since McDonald's is a franchise, and local owners are held to the same standard as every other location so more than likely customers can't tell the difference. Item numbers do not need to be used since we are using item description. Lastly, the total units variable is the same as the units sold for the week variable so we don't need to use that variable either.

We can observe the different distributions of our important variables through elementary data analysis. We see that there are 28 unique cities in the dataset, 64 unique restaurants, there are 292,950 restaurants that are traditional (9,450 which are non-traditional), and 25 menu items (ranging from breakfast items to popular items such as burgers, fries, and chicken). Below we display time-series plots of aggregated data across all of the restaurants per week for the continuous variables.









It appears that the transaction rates and units sold respectively have increased over time as the end points for each of these graphs are all higher amounts than they are at the beginning—with the exception of the AGC chart. Furthermore, we observe a cyclical trend in each of these charts where there is a dip in the transactions or units sold respectively at the start of each year, with a gradual increase in the numbers as the year progresses.

3) Data Quality

There are several concerns with the quality of the data. To begin, the average prices, units sold per 1000 transactions, units sold for the week, average transactions per day, and average daily units sold, all have 4,762 rows of missing data (NaN values). This is an issue as missing data prevents us from seeing the complete picture. Additionally, as mentioned before, the total units sold variable is the same as the units sold for the week, which brings integrity concerns to the data—was the data entered incorrectly or is the redundancy purposeful? Lastly, there is an issue with how the variables are named in the dataset. Below is a table listing all of the food items.

Total All Day Breakfast Items	Total Breakfast Sandwich	Total McMuffins	Egg McMuffins	Sausage McMuffin
Total Biscuits	Bac Egg Ch Biscuit	Sausage Biscuit	Sausage Egg Biscuit	Hashbrown
Total McGriddles	Bac Egg Ch McGriddle	Sau Egg McGriddle	Total Bagels	Beef Category
Large Burgers	Big Mac	Core QP and QPC	Hamburger	Cheeseburger
Chicken Category	Total Fries	McChicken	McDouble	Sausage Egg McMuffin

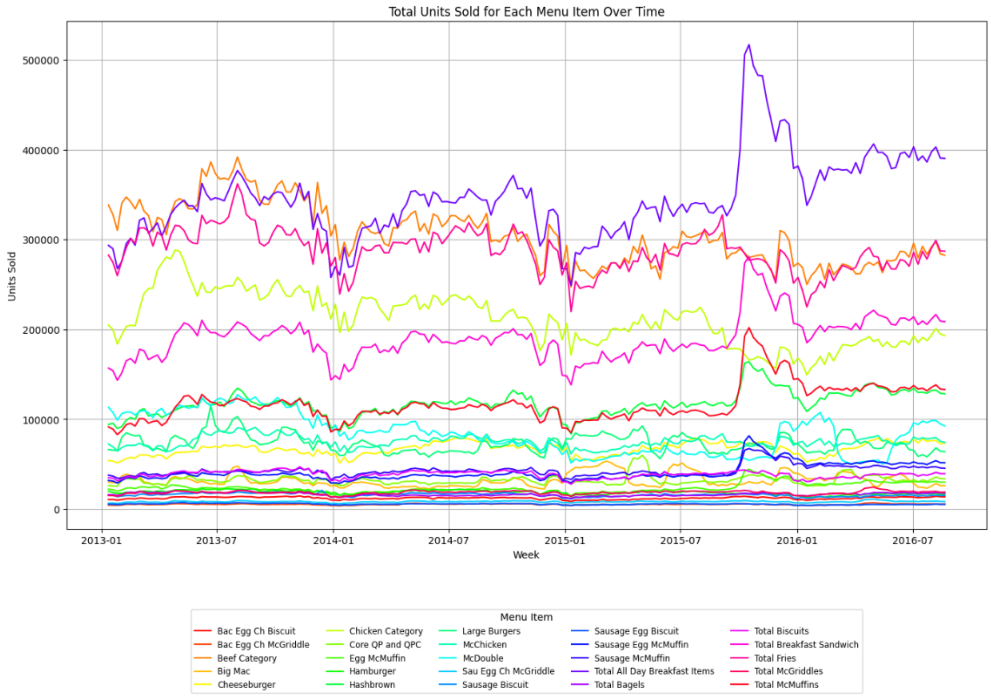
The food items are labelled in an unclear manner—some items are listed individually and within their overall groupings (e.g. different flavors of McMuffins along with Total McMuffin), while some items are only individual (e.g. no total Burger group). This makes it hard to compare between overall groups and it also creates an issue with double counting transactions and units sold. Moreover, the labeling can be confusing as some items have similar names (e.g. Total All Day Breakfast v.s. Total Breakfast Sandwich, and Chicken Category v.s. McChicken). This creates an issue with being able to identify which item you actually are analyzing.

4) Hypothesis

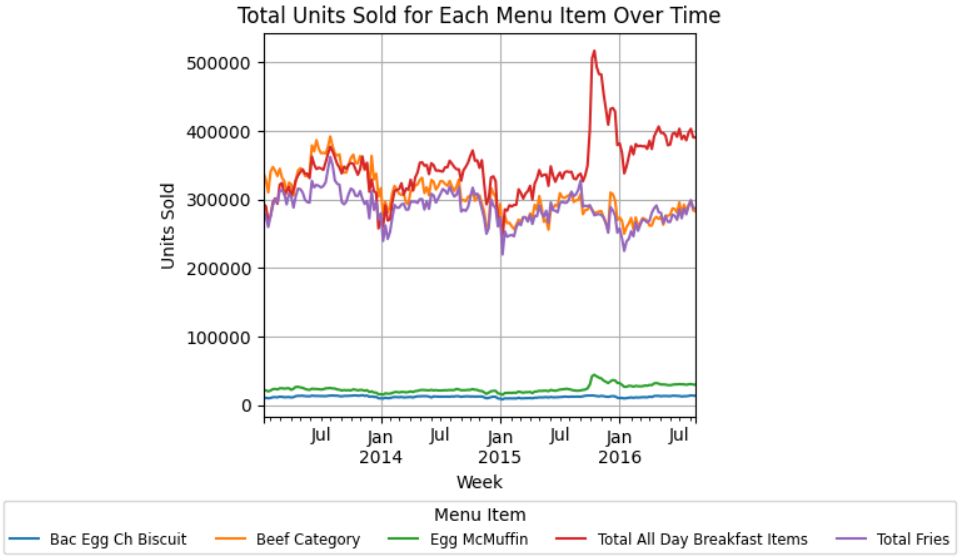
This specific data can be used in multiple interesting ways to investigate and answer business problems as there's a plethora of continuous and categorical variables. Some questions we could investigate are (but not limited to):

- Are different menu items bought more frequently than others?
- Do transactions and units sold go down at particular times of the year?
- Are more items sold during a promotion than individually?
- Does pricing impact the units sold of an item?

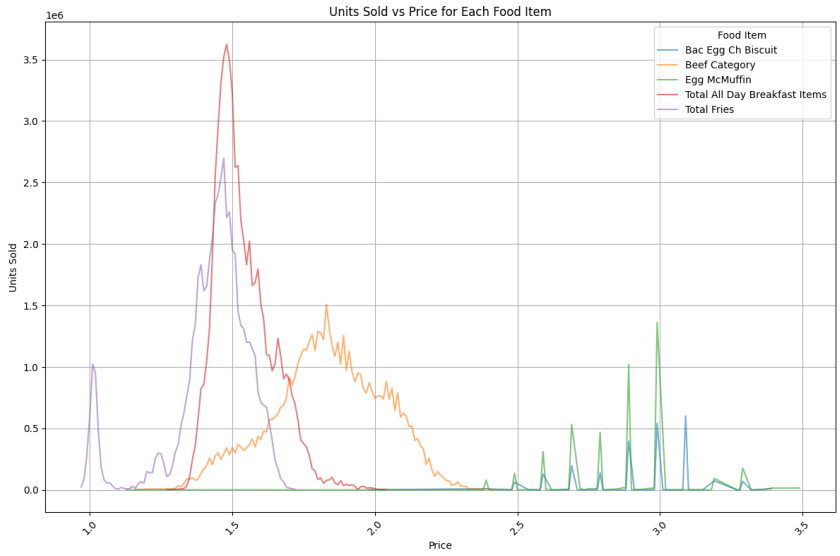
As we saw with the graph in the business problem section, we have already observed that Egg McMuffins are sold at higher rates than Bacon and Egg Biscuits in Michigan. Additionally, in the graph below we see that certain menu items are sold more frequently than other menu items.



In particular, the Total All Day Breakfast Items, Beef Category, and Total Fries appear to sell the most. On closer inspection, we see in the graph below that All Day Breakfast Items are sold more frequently than the individual McMuffin and Biscuit items, so the promotion may work better for sales.



Lastly, regarding prices, we observe below that All Day Breakfast items are sold more at lower prices than other popular menu items, and that more food items are sold at lower prices in general. More analysis into this could reveal how pricing impacts consumer spending for better identification of whether a promotion is worth the operational costs.



5) Confirming/Rejecting the Hypothesis

At this point, I do not think the data is enough to confirm the hypothesis. Although in the previous sections, we see through the EDA that pricing may have an impact on customer spending and there may be a benefit to selling promotional items, there needs to be more data collected and more analysis done to come to a proper conclusion. We saw in question 2 that there is a substantial amount of missing data, some data redundancy, and lack of clarity with menu items. If this data is fixed then the analysis done above will be more conclusive for answering the business problem. Moreover, I assumed that the McMuffins and Biscuits referred to by the regional manager were the egg varieties as he did not state a specific type, if this assumption is wrong then I can't use my EDA to immediately answer the business problem. Lastly, I think additional data should be collected regarding operational costs of making each menu item, revenue made by the stores each day/week v.s. profit margin, and information on the customer demographic to properly investigate whether a promotion will lead to a higher profit margin and if it will even work on the particular customer base for each store. As of now, we only have transactional and units sold which is only enough to see if there's an increase in sales during a promotion but not enough to decipher whether the increase in sales truly leads to an increase in profit.