RETAIL TRIAL EFFECTS: COMBINING MACHINE LEARNING AND QUASI EXPERIMENTAL APPROACHES

PostgreSQL

Power BI

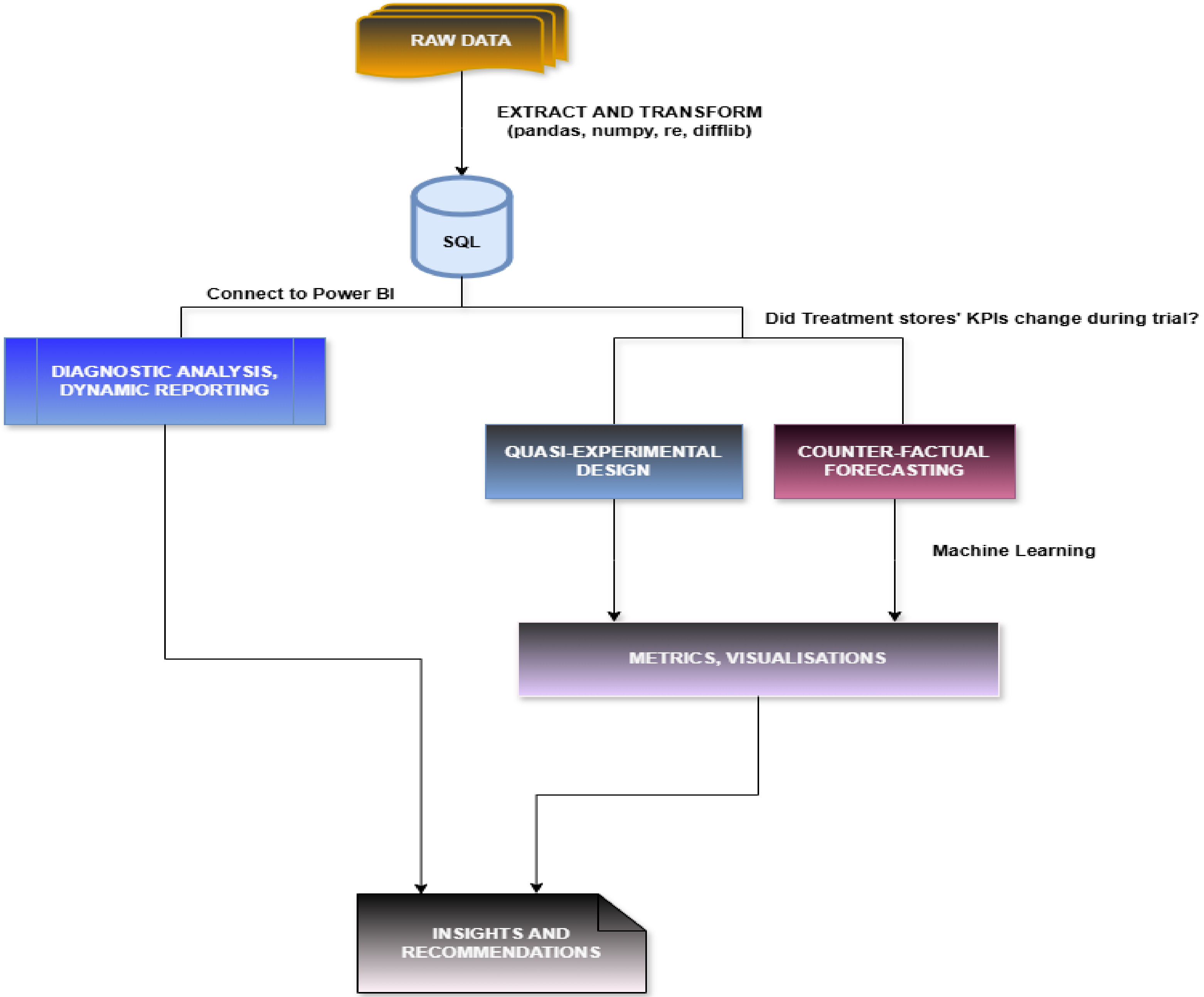
CASE

The client, who manages the Potato Chips category for a national supermarket chain, recently implemented a trial involving modified store layouts aimed at increasing sales revenue. They are seeking to evaluate whether the trial delivered the expected uplift in sales performance. Additionally, the client requires an interactive dashboard to monitor key performance indicators (KPIs) in real time and to support ongoing decision-making.

Approach

- Build an ETL pipeline, in order to automate reporting. The pipeline was built using Python (pandas, re, numpy, SQLAlchemy, psycopg2). The transformed raw data, along with new pivot tables and helper tables were loaded to a PostgreSQL database.
- Connect database to Power BI, build the data model, design the dashboard, depicting KPIs and interesting visuals.
- **Control store method:** Find control stores for the trial stores and measure uplift during trial period, ie, compare metrics between control and trial stores during the trial period.
- Build a **Machine Learning model** to predict sales during trial period for selected stores and compare the results with the Control store method
- Generate final insights and recommendations

PROJECT SCHEMA



Diagnostic Analysis

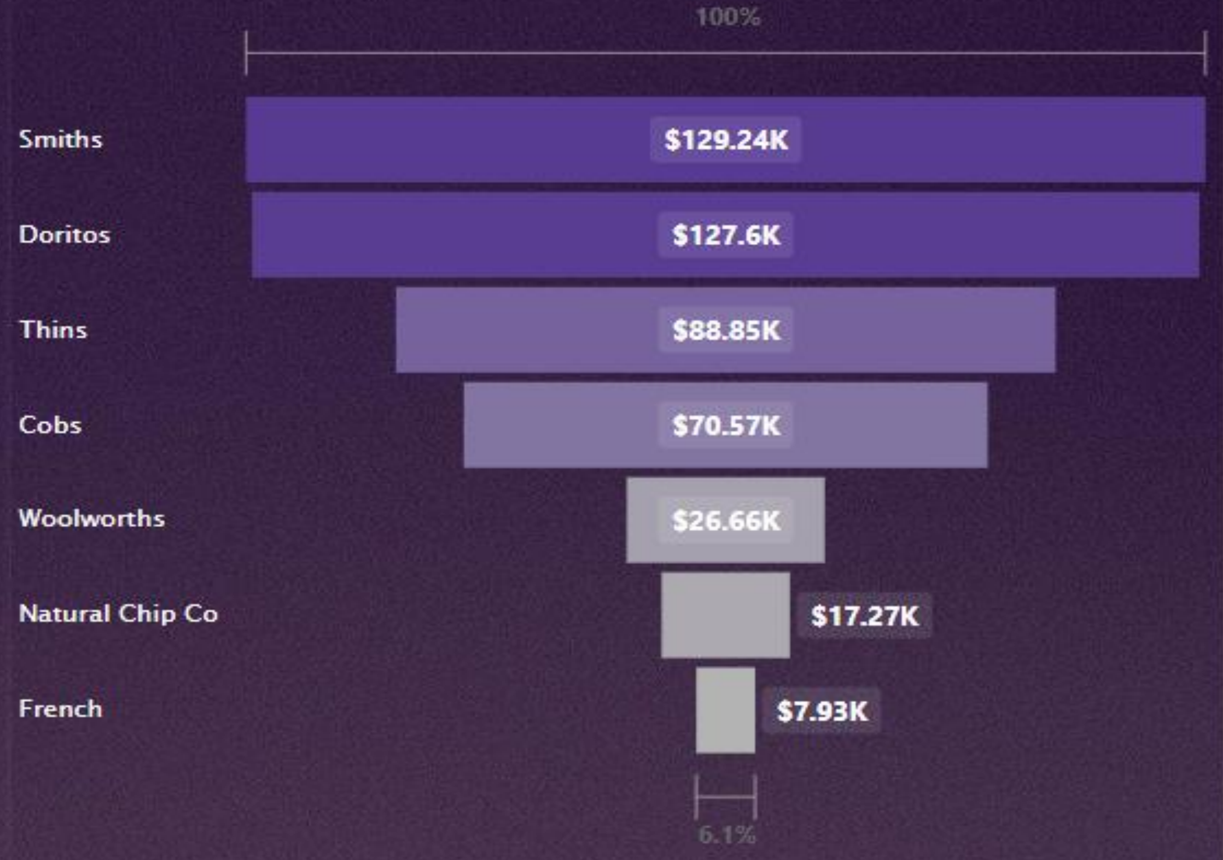
- **Approach**: Python to clean and transform raw data, load it to PostregeSQL database (ETL Pipeline), connect database to Power BI, leverage DAX measures to generate insights.
- **Product KPIs:** Total Sales, Average sales per product, Total quantity sold per product., Highest selling brand and product.
- **Customer Segment**: Age range of customers, membership status, segment of customers that drives highest percentage of sales.

DASHBOARD SNAPSHOTS

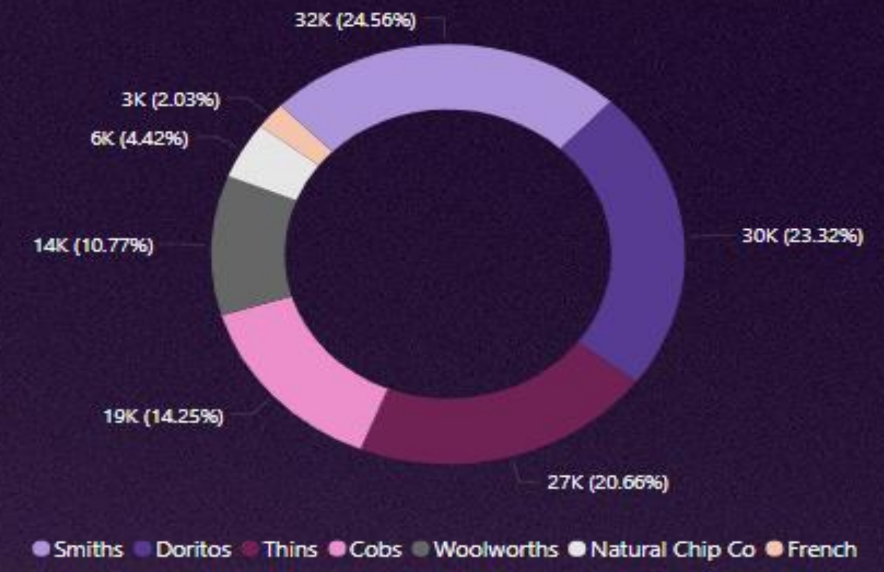
TOTAL REVENUE	TOTAL QUANTITY SOLD	AVERAGE REVENUE PER PRODUCT
$468.11K	130K	$15.6K

1	10	100	101	102	103	104	105	106	107	108	109	11	110	111	112	113	114	115	116	117	118	119	12	120	121	122	123	124	125	126

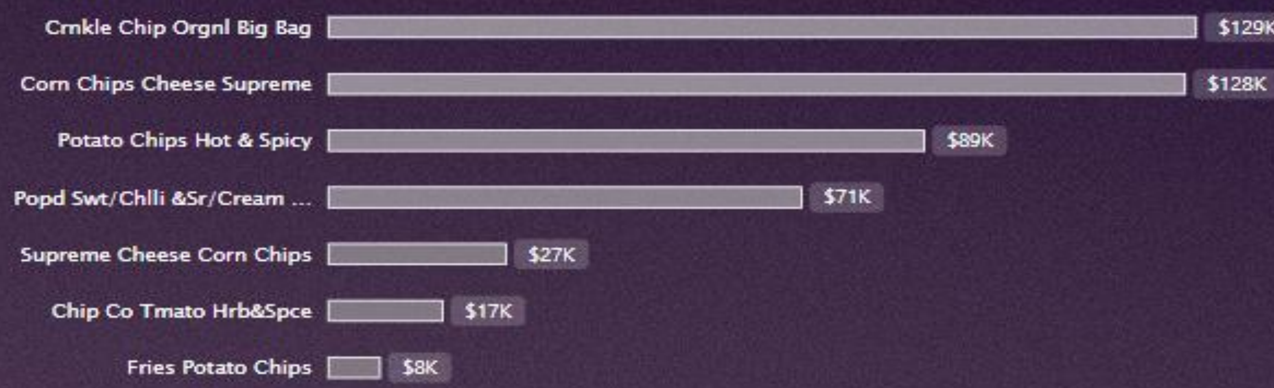
REVENUE PER BRAND



QUANTITY SOLD BY BRAND

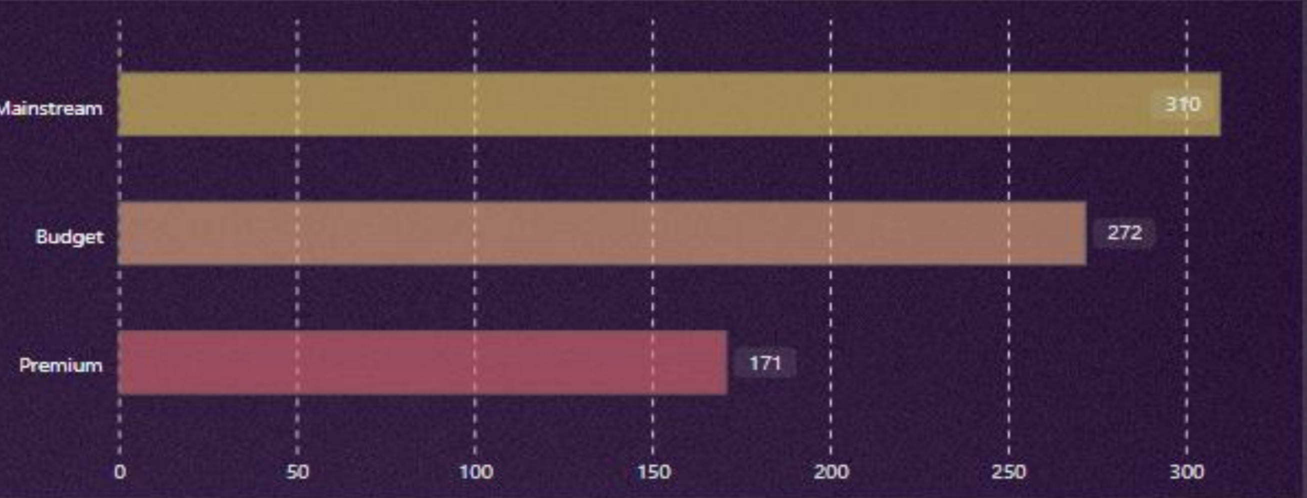


TOP SELLING PRODUCTS

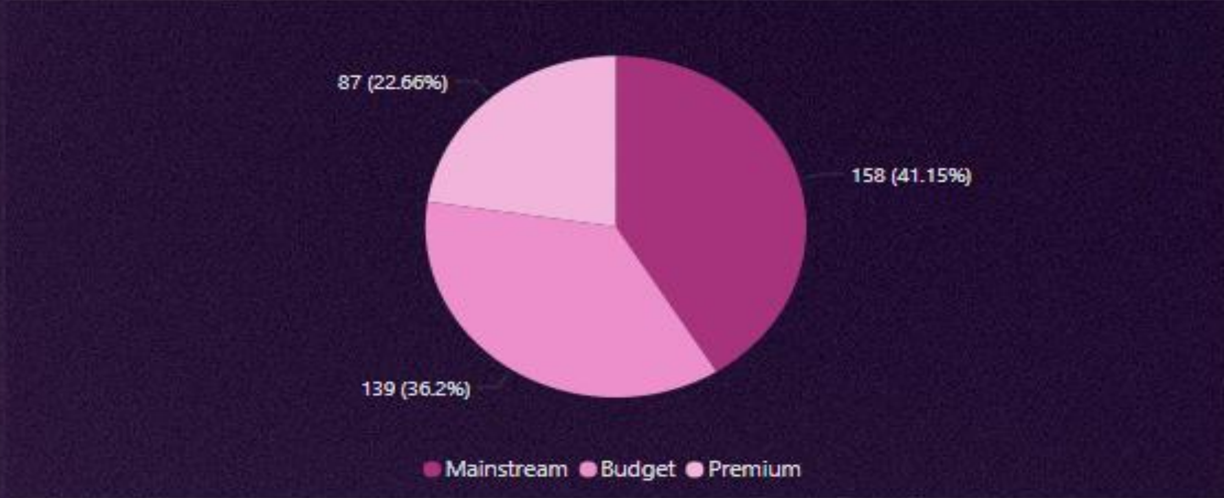


1 | 10 | 100 | 101 | 102 | 103 | 104 | 105 | 106 | 107 | 108 | 109 | 11 | 110 | 111 | 112 | 113 | 114 | 115 | 116 | 117 | 118 | 119 | 12 | 120 | 121

QUANTITY PURCHASED BY MEMBERSHIP

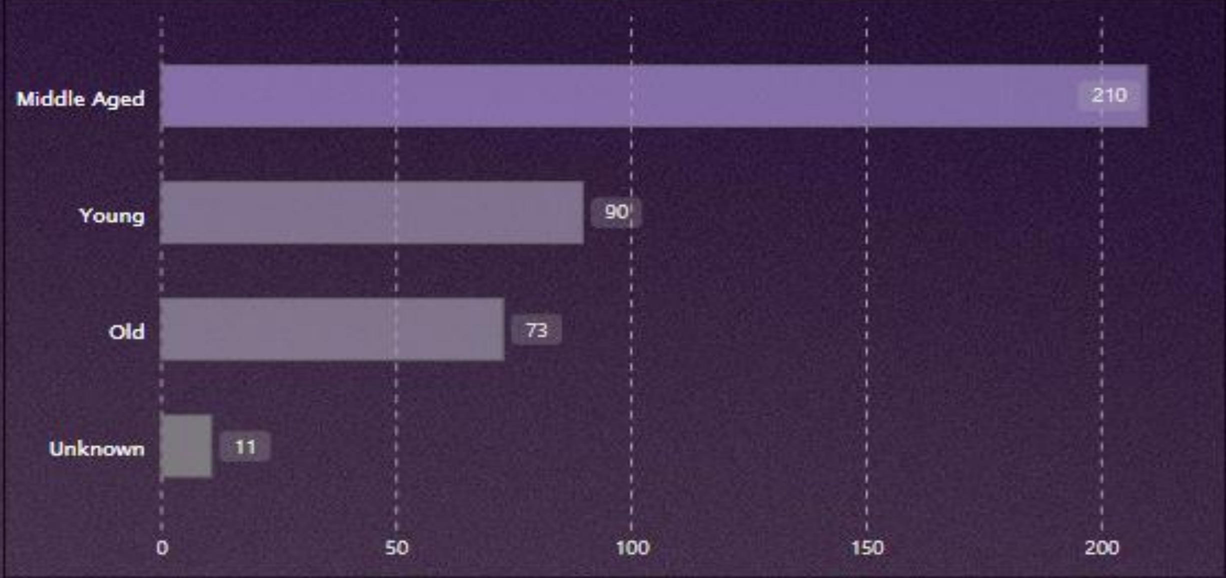


CUSTOMER DISTRIBUTION BY MEMBERSHIP STATUS



SALES CONTRIBUTION BY MEMBERSHIP STATUS

CUSTOMER DISTRIBUTION BY AGE GROUP



PRODUCT INSIGHTS

- **Total revenue** from chip sales across all stores in the given period is approximately **$468 K.**
- **Smith's Crinkle Cut Originals** had the highest proportion of contribution towards total revenue which was about **28%** of the total revenue generated.
- Doritos Corn chips had the highest contribution towards average revenue per product.
- 130 thousand units of product were sold, where Smith's and Doritos dominated, accounting for almost half of the total (48%).
- French Fries potato chips appeared to be the least popular and profitable product.

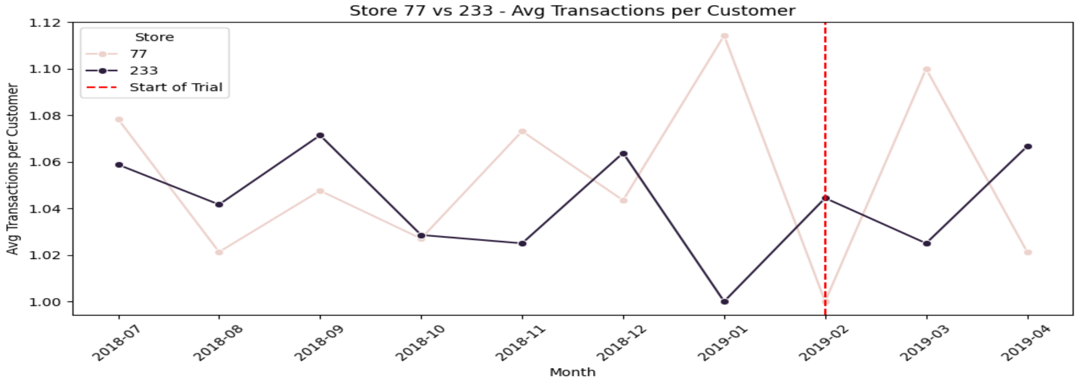
CUSTOMER BASED INSIGHTS

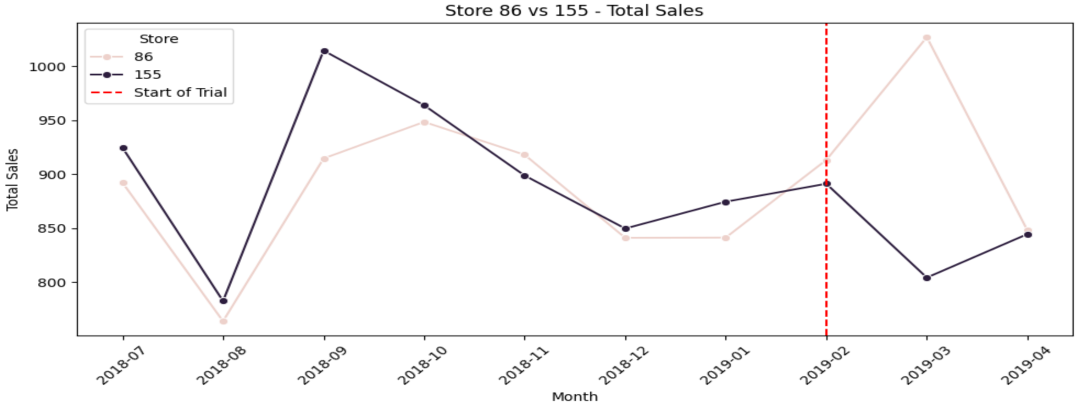
- Over 70% of customers are either mainstream or budget customers. Together, they contribute to approximately 73% of the revenue.
- The customer base is biased towards middle-aged customers, accounting for about 48% of total customers.
- The distribution of customer's membership did not appear to have any significant impact. The distribution matched sales contribution quite closely, i.e, Premium customers account for 26% of the total base and their contribution towards revenue was also calculated to be approximately 26%. It held true for all groups.

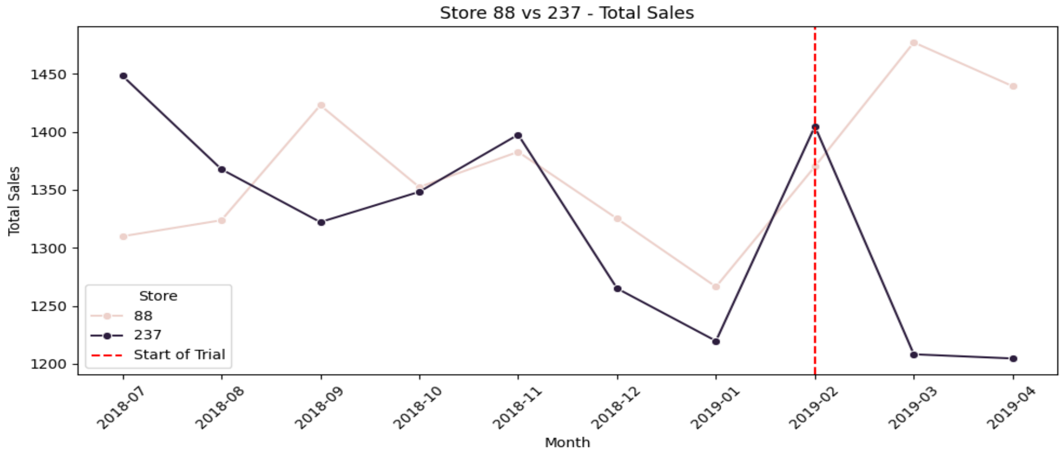
EFFECT OF TRIAL: STORES 77, 86 AND 88

- **Quasi-Experimental design approach**
  - Analyse all stores pre-trial to check which store compares most closely to the trial stores (77, 86, 88) in terms of trends and behaviour. Calculate revenue, average number of transactions, quantity of products sold.
  - Calculate the **Eucledian distance or Pearson Correlation coefficient** between metrics of selected stores and every other stores in the dataset.
  - Select the stores with lowest distance (most similar) as Control stores. For example Store 77 happened to be nearest to Store 233, 86 to 155 and 88 to 237 in terms of behaviour based on selected metrics.
  - Compare the metrics of Control stores with Trial stores during the trial period and visualise the uplift.

UPLIFT IN TOTAL SALES

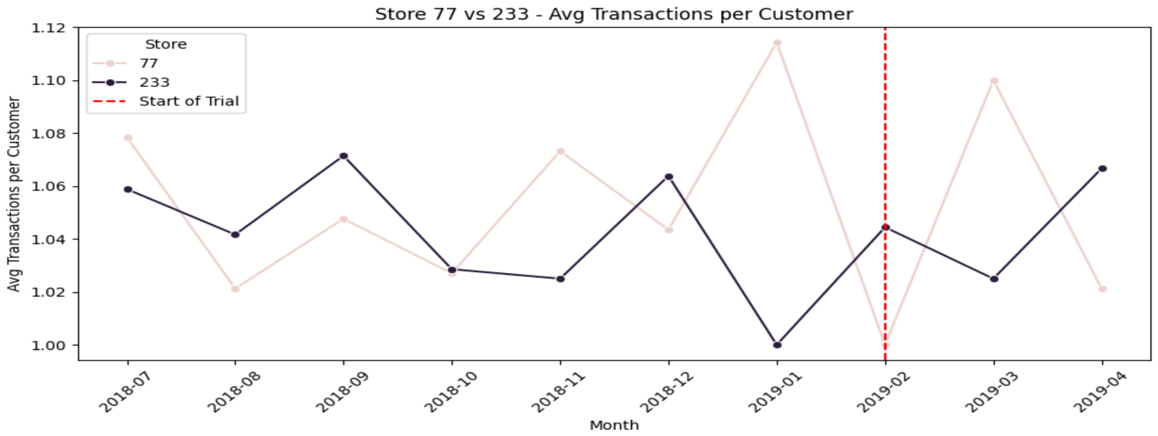


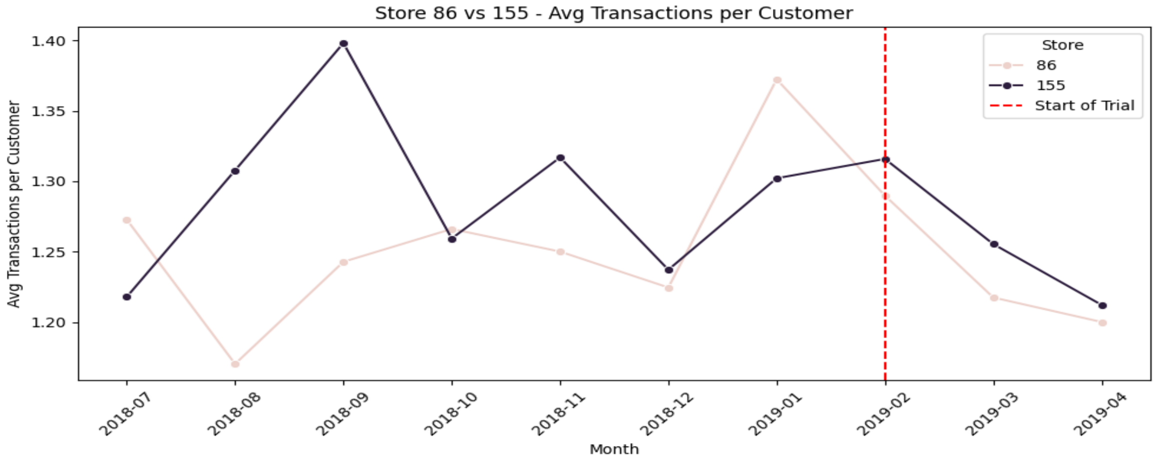




77	86	88
29.13%	9.76%	12.29%

UPLIFT IN AVERAGE TRANSACTIONS PER CUSTOMER





77	86	88
-0.12%	-1.98%	7.43%

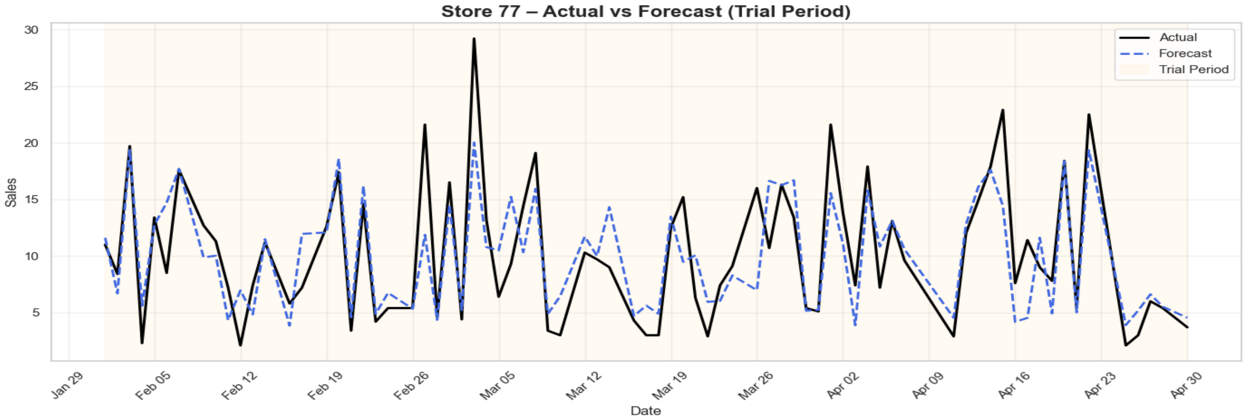
UPLIFT IN TRAFFIC

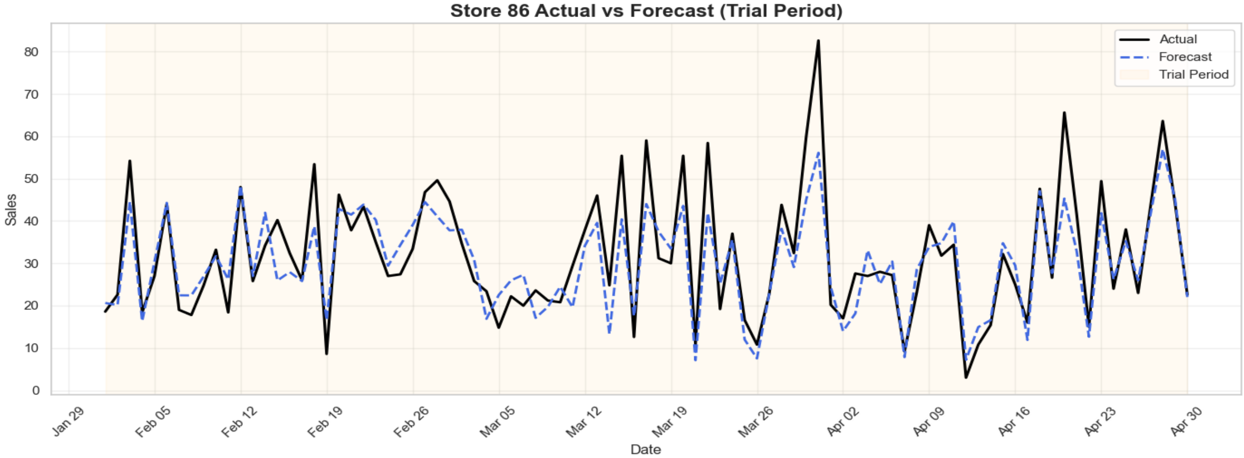
77	86	88
23.48%	13.54%	5.75%

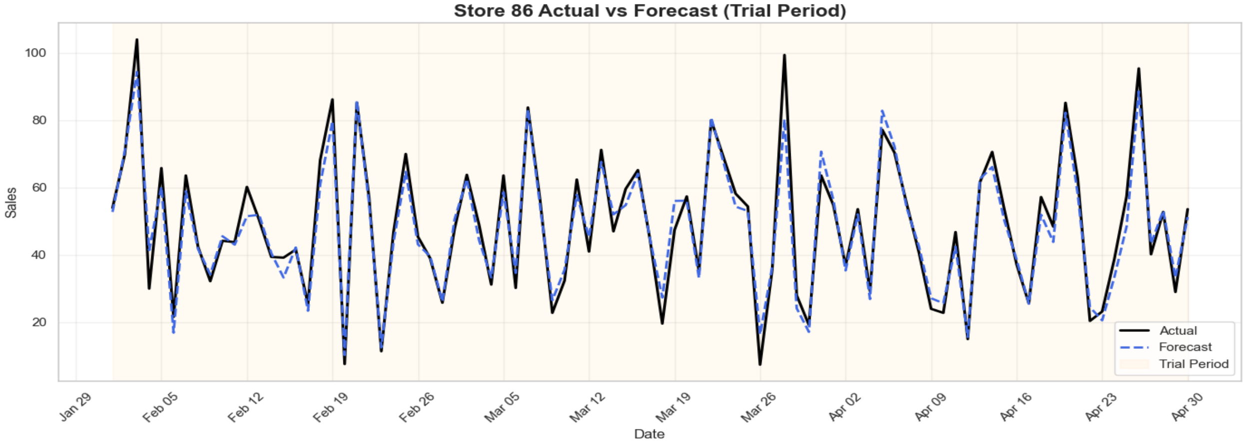
FORECASTING TRIAL PERIOD SALES

- **Approach:**
  - Use features such as average transaction, total sales, lag values of sales (to avoid leakage), age group, membership etc, to predict sales during trial period for stores 77, 86 and 88.
  - The algorithm of choice here was **XGBoost**, tuned with **Bayesian Serach** using **Optuna.**
  - **Solving the memory issue:** XGBoost doesn't have the capability to perform Autoregressive tasks like time-series models. To solve that, lags were calculated on a day, week, and fortnight level and were fed into the model as features.
  - **Does it leak data?** No, since only past values are used as features.

PREDICTION RESULTS







STORE 77:

Actual revenue: $777.00
Predicted revenue: $727.22
Uplift: 6.84%

STORE 86:

Actual revenue: $2788.20
Predicted revenue: $2642.84
Uplift: 5.50%

STORE 88:

Actual revenue: $4286.00
Predicted sales: $4150.00
Uplift: 3.27%

Insights

- The control store method demonstrated very significant uplifts in Store 77 and 88 metrics, which suggests that changing the store layouts (treatment), yielded significant positive effects.
- The forecasted sales with based on historical data suggested much more conservative uplifts in total revenue.
- The control stores seemed to be taking a dip in revenue whereas the trial stores saw a positive uplift especially during the month of March 2019.

Recommendations

- Why did sales drop during the trial for control stores? Was it expected? Was it an effect of Seasonality? Did the store see competitions in the same area? Did the store suffer qualitatively? Were there any issues in Logistics?
- Is the Control store method a true representation of the effect of treatment? Or is the uplift compared to forecasted sales a more accurate representation?

In order to make an accurate recommendation to drive business decisions, the set of questions in the first bullet point needs to be clarified. However, some important insights can be extracted with the given observation:

- No significant change was noticed in customer base and product trends.

.....Continued

Recommendations

- It is safer to assume that the actual effect of trial stores is somewhere in between the metrics quantified by the two methods.
- Since both methods show a positive correlation between Revenue and treatment, any loss in revenue would be very unlikely due to trial effects. In other words, the treatment does seem to have a positive impact on sales.
- The increase in sales but an overall decline in average number of transactions does suggest that customers are purchasing more at a time. It is most likely due to the layout change.
- The trial seems to be successful in increasing the revenue, however, the extent of the success needs to be tested more robustly, accounting for other factors mentioned previously. It is also advisable to increase the observation period, check for any drifts over time, before scaling it to other stores.

THANK YOU