

MARKETING & RETAIL ANALYTICS

**Presented by,
Renisha R**

Table of contents

- 01 Data understanding
- 02 Data pre-processing / objectives
- 03 EDA
- 04 Menu Analysis
- 05 Business insights & Recommendation

BUSINESS PROBLEM STATEMENT & UNDERSTANDING

PROBLEM STATEMENT:

The data set provided to you is the data set of a Café Chain for one of its restaurants. Do a thorough analysis of the data and come up with the following analysis.

The owner of the restaurant wants you to use this data to come up with a set of recommendations that can help his Café Chain increase its revenues. He is able to provide you with a data set for POS (point of sale data) for one of his chains.



PROBLEM UNDERSTANDING:

Restaurant (café chain) sells a usual mix of Coffee, pastas, beer, and sandwiches. Despite having an incredible spread of menu items, the Café' is still struggling to achieve the sales targets since last 2 Quarters.



Observational data is the Café's sales data collected for around 12 months. Dataset contains more than one lac of billing record for 5 quarters (from April 2010 to Mar 2011)

BUSINESS OBJECTIVES

- To understand the sale pattern across different categories with respect to different time frame.
- Conduct market basket analysis to find out the potential of cross selling and up selling.

Dataset:

Source: MRA

Data as of : 14/01/2023

Number of rows (customers) :145831

Features:

1. Date
2. Bill number
3. Item description
4. time
5. Quantity
6. Rate
7. Tax
8. Discount
9. Total
10. Category

Exploratory Data

We are exploring a dataset of transactions at a particular café – for a period of 1 year, from 1 April 2010 to 31 march 2011.

Each record has 9 features: which includes a Bill number, the item that was ordered, the category under which the said item is classified, data and time stamps of the order, along with the price, Quantity, Tax and the total amount charges for the item.

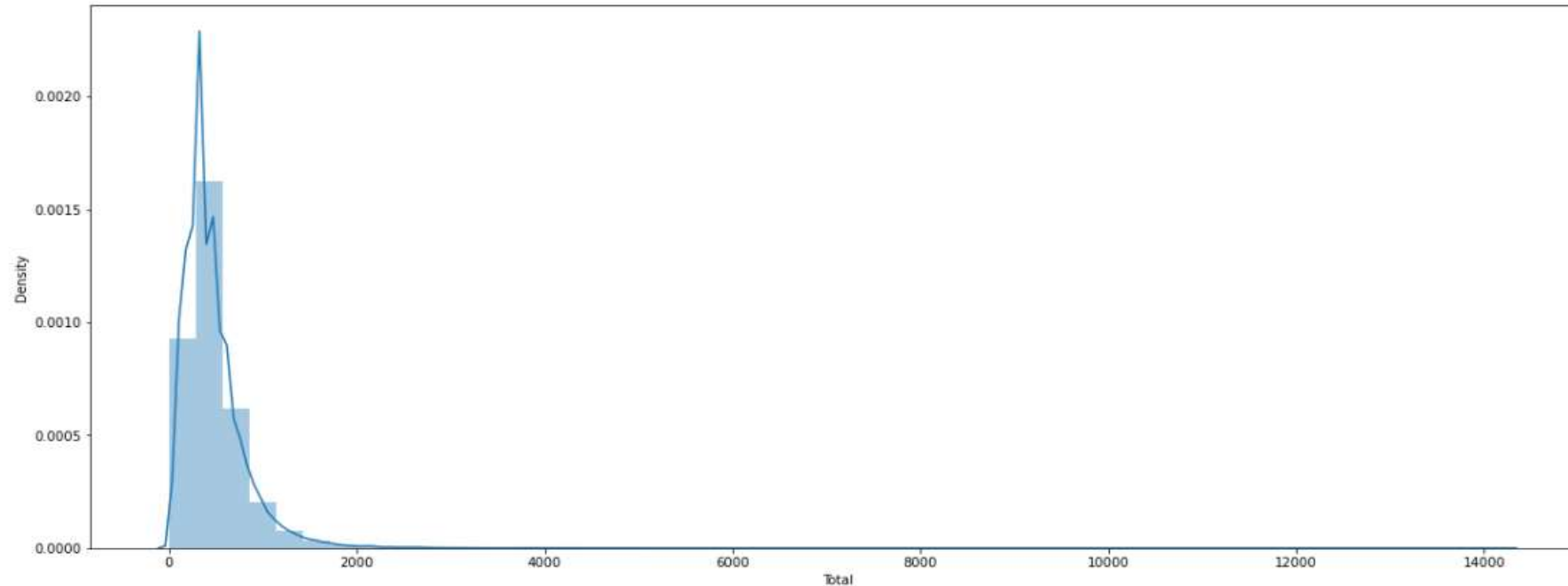
The dataset has 145830 rows, accounting for 69982 orders (identified by unique bill numbers)

The total Revenue from sales at this cafe, for the 1 year period, is a little over INR 3.28 Crore. We assume the currency is Indian rupees or INR for this exercise.

There are no missing values in the dataset



ORDERS AT THE CAFE



By grouping the data on the basis on unique bill numbers, we can study the orders placed at the café.

Number of orders placed at the café during the year- 69982.

Median order value(bill total) : INR 391

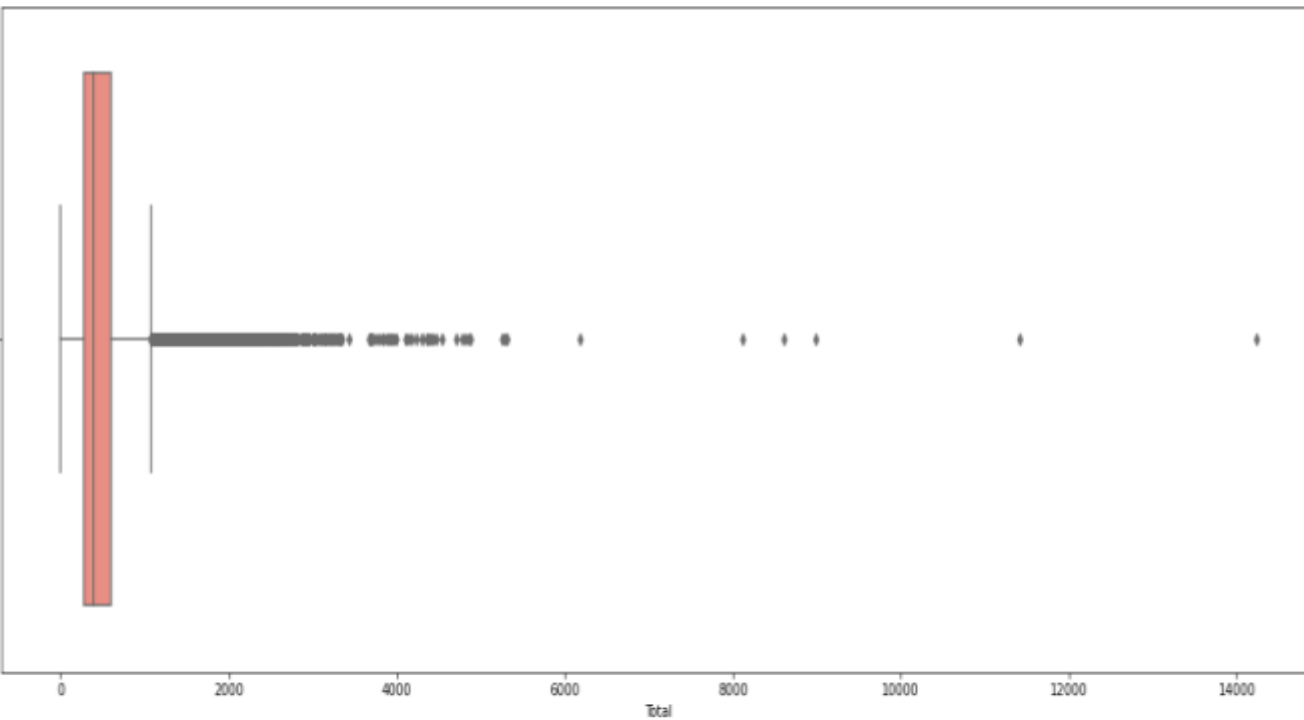
Mean order value (bill total) : INR 468.78

The notable difference in the mean and median values suggests a skewed distribution.

The featured histogram and box plot for order values (total per bill number), both corroborate the right skew.

Typically, order values are concentrated between INR 250 to INR 600. But there is a significant number of high value orders, which constitute outliers.

One can see the large range of values for order placed, from INR 1 to INR 14231.



It also observed that the typical order consists of 1 to 3 units, 2 being the median quantity per order.

In relatively rare instances do quantity of items in an order cross 5 units. Although in exceptional cases, outlier orders that have over 50/60 units of items.

OUTLIERS:

- We take a broad look at the high value orders or outliers.
- On the right are bar plots of Bills, arranged in descending order of value (Total Amount).
- The top plot features the constituent Categories of each Bill.
- Food, along with non-alcoholic beverage, appears to be the predominant combination of the outlier orders.
- The bottom plot colour codes Bills based on the quarter in which the order was placed.
- Q1 (Jan 2011 - Mar 2011) appears to have the orders with the highest values, followed by Q4 (Oct 2010 - Dec 2010).

Summary:

- We reviewed transaction data of a particular cafe for a period of 1 year, from 1 April 2010 to 31 March 2011. There were a net 69,982 orders placed, and total revenue earned during the period was around INR 3.28 Crore.
- The Median Order Value (Bill Total) is INR 391. And median quantity of item/s per order is 2. Improving on both of these counts can see significant growth in business.
- The Mean Order Value (Bill Total) is INR 468.78, suggesting a skewed distribution owing to a notable number of high value orders.
- There are 580 unique items in the dataset.
- But over 77% of the cafe's annual revenue earnings come from 10% of the menu, ie from 58 items. The Top 5 items bringing the highest revenues belong to the Tobacco category. Tobacco as a category accounts for over 44% of the cafe's annual revenues, even as it accounts for only 22% of the orders.
- Food is the most popular category in terms of quantity sold. Often ordered together, Food and Beverage account for over 68% of the ordered items. The average sales at the cafe, per day, is close to INR 90,000, with the IQR ranging roughly from INR 78,000 to INR 1,00,000.

Trends Across Months

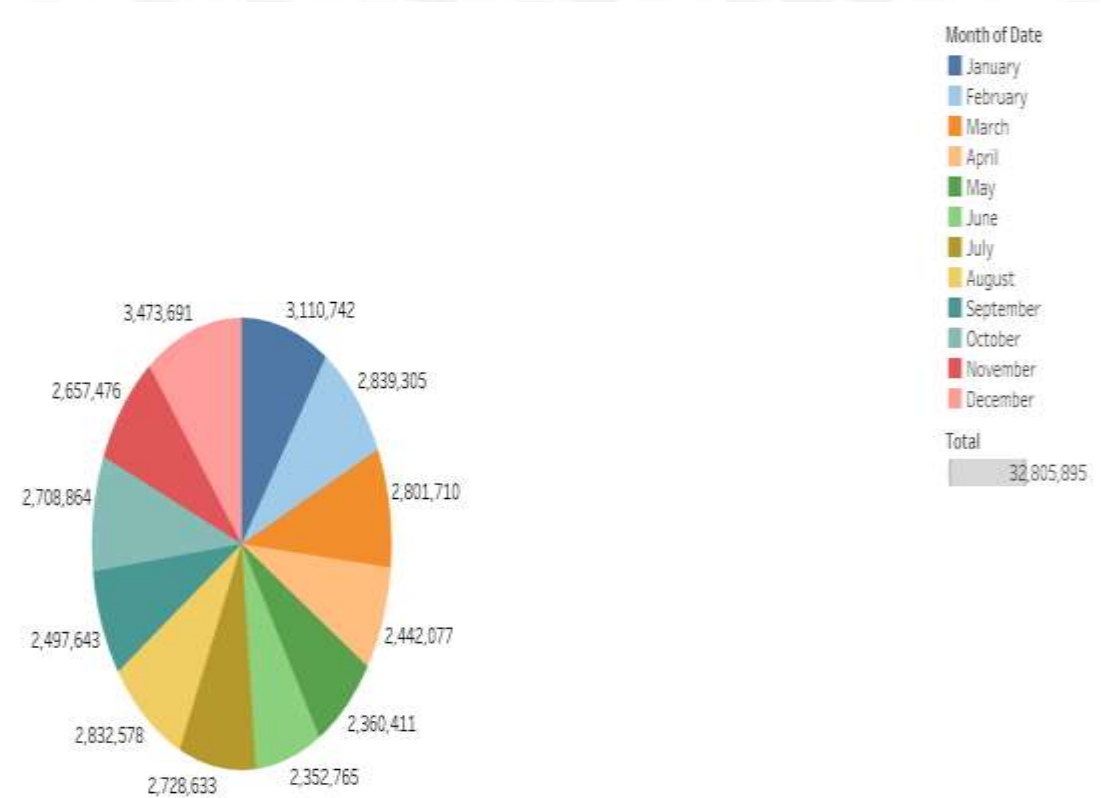
The 2 adjacent plots depict Monthly Sales, aggregated for Weekdays and Hours of the Day, respectively.

The Sales by Hours of the Day plot shows a fairly consistent pattern of monthly sales across all months, with Sales gradually building up with every hour since opening, with the peak at 11 PM.

March however shows peak sales at 10 PM, which is distinct from any other month

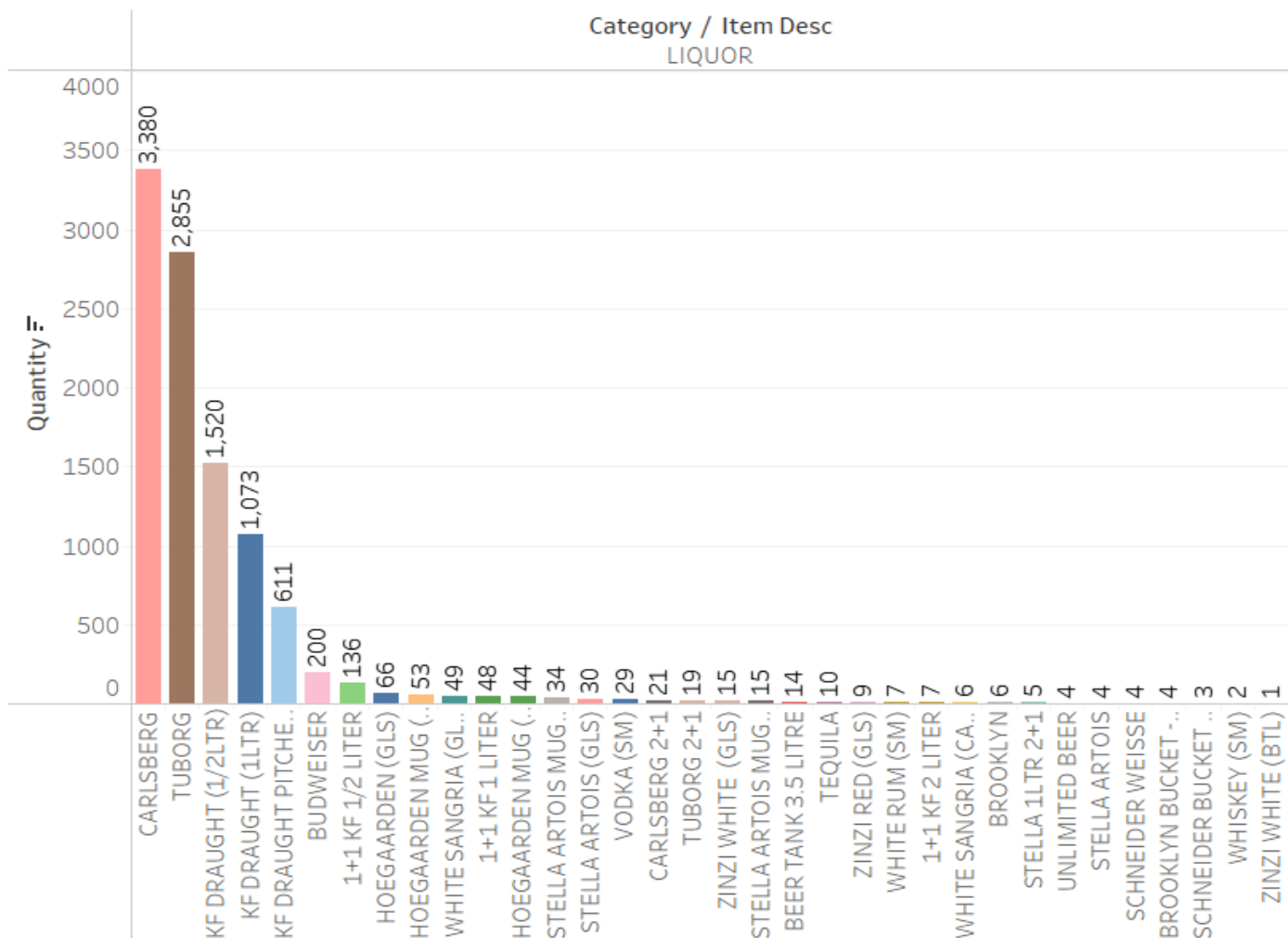
Also noticeable is that for the month of February, May and August, business starts at 9 AM, which is earlier than usual. December shows late hours business until 6 AM, which is probably accounted for by the New Year's Eve celebration.

month wise revenue



LEAST POPULAR ITEMS ON THE MENU

items less frequently sold



- We have earlier noted that the highest earning 55 items contribute to over 75% of the cafe's earnings.
- In contrast, there are 54 items that have sold only one unit each, throughout the year.
- The adjacent plot lists them in order of the revenue generated from the sales of these items.
- It is seen that over 40% of these are from the Merchandise category.
- And around 23% belong to Miscellaneous - which really is an assortment of items from all the other categories.
- So effectively, the 60% of the items can be uniformly attributed to the remaining categories.

TREND OVER MONTHS:

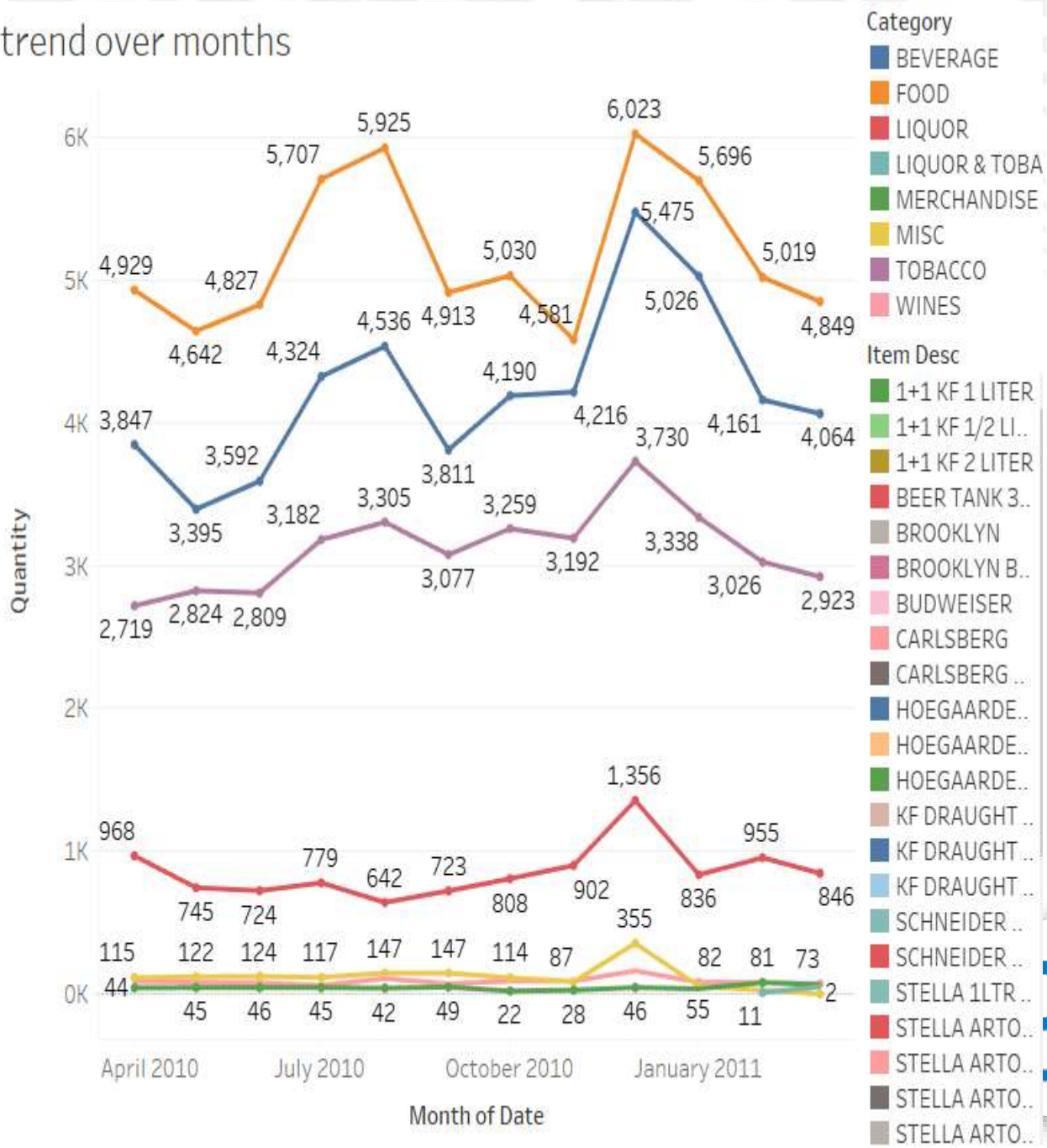
The plots on the right depict Sales across months, from April 2010 to March 2011. The top plot depicts quantities sold and net proceeds from Sales for the period. The plot below splits Monthly Sales by Category.

December 2010 is clearly the winner when it comes to Sales, with revenue nearing INR 35 lakhs. January 2011 is the next highest ranking month, with INR 31 lakhs in revenue.

June 2010 shows the lowest earnings at around INR 23 lakhs, and May 2010 has the lowest quantities sold.

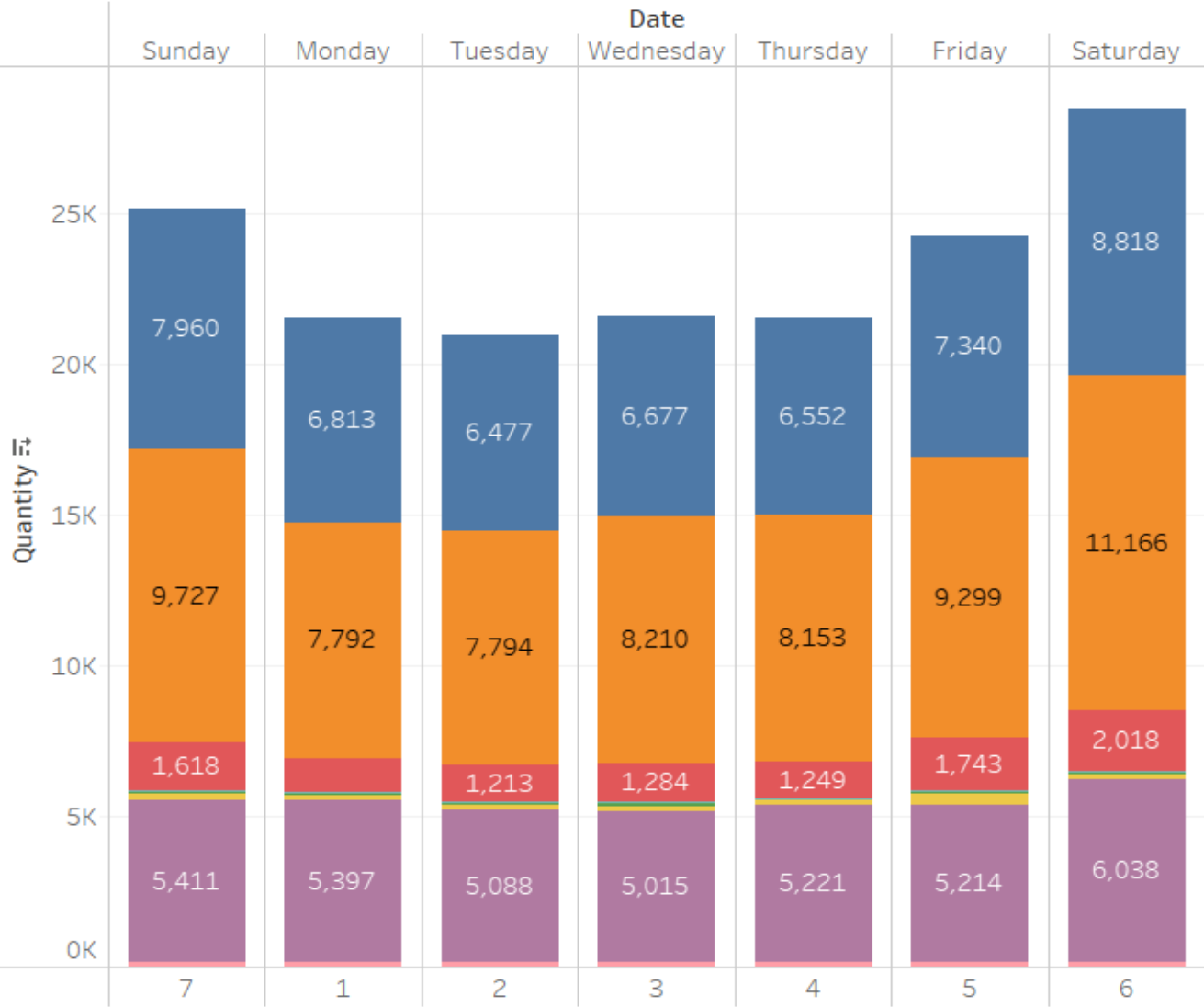
At a Category level, a similar trend is seen. A couple of interesting observations: Food records the lowest sales in the month of November 2011, where the general slump is seen in May and June 2010. Liquor Sales are seen picking up in February and March 2011, where other categories show a gradual decline after December 2010.

trend over months



TRENDS OVER WEEKDAYS:

customer choice of food over days of a week



The adjacent plots depict Annual Sales, consolidated and per category, aggregated by weekdays.

Saturday is typically the busiest day, and most lucrative from a revenue standpoint.

Sunday and Friday follow as the two next most lucrative days.

Basically indicating that Friday through Sunday, the weekend, is when the cafe conducts most business

Tuesday shows the lowest earnings, though Monday through Thursday show a generally uniform pattern



Recommendations:

We see a general concentration of business over weekends, and in the evenings. It does appear that the cafe and its resources aren't optimally engaged over weekdays, and in the morning and afternoon slots.

In order to promote more traffic during this period, the cafe can be made more amenable to working individuals and teams. Freelancers, Remote workers, Entrepreneurs, Professionals and students can then work out from this space in the 'off hours'. The cafe's menu and operations can be more tailored to the requirements of the working individual who treats the cafe as more of a workspace in these hours.

Different monetisation means can be explored. Aside from charges for food, there can even be a periodic subscription to use the cafe's premises and infrastructure on a regular basis.

MENU ANALYSIS

MARKET BASKET ANALYSIS

We know from the EDA exercise that a typical Bill / Order at the cafe has 2 units of an item or items, and has a net order value of INR 391.50 - going by the median Quantity and median Total for Bills

We've also observed that outlier orders have many more items, and some combinations of categories such as Food and Beverage were apparent during the course of our EDA

In order to increase Revenues and Profits, it will be important to induce people to buy more per order.

So we conduct a Market Basket Analysis to build Association rules, which will indicate which items are likely to be bought together

If an association is established between Items, the customer can be induced to spend more per visit in various ways.

For the purpose of this exercise we will focus on building Combo offers. That is, the aim will be to incentivise purchase of more items by offering the right combos at the right price.



Item Desc	1 axe twist	1+1 btl4 season white	1+1 gls 4season red	1+1 gls 4seasons white	1+1 kf 1 liter	1+1 kf 1/2 liter	1+1 kf 2 liter	1+1 vln cab sauv (btl)	1+1 vln cab sauv (gls)	1+1 vln cab sauv clasiq (btl)	...	white wine sheesha	wild mushroom linguini	yemeni great lakes (aulait)	yemeni great lakes (reg)	yin n yang fondue	zen rock square vase	zero size shot gls	zinzi red (gls)	zinzi white (gls)	zinzi white (btl)
Category																					
BEVERAGE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	50.0	67.0	0.0	0.0	0.0	0.0	0.0	0.0
FOOD	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	45.0	0.0	0.0	26.0	0.0	0.0	0.0	0.0	0.0
LIQUOR	0.0	0.0	0.0	0.0	42.0	111.0	7.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.0	9.0	1.0
LIQUOR	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LIQUOR & TOBACCO	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Configuration specifics:

- Minimum Support 0.001
- Minimum Confidence 0.05

Results - 313 rules were generated using the above thresholds.

The rules here have been sorted in descending order of Lift. We will focus principally on rules and combinations that have a lift of greater than 4. There are 33 such instances in our analysis.

As a secondary measure, we will also look at Combinations that have relatively High support and Lift between 2 and 3.

For this exercise we will ignore all associations with lift less than 2 - for that implies little or no association.



All combinations here have a support of 0.001, except for BMT Panini and Fries, which is slightly higher at 0.002.

Arabian Mist and Red Bull 2 + 1 are the most expensive combination. This also stands out as the only mixed-category combination. The rest of the combinations are arguably items in the same category.

Cappuccino with an added Hazelnut flavour is least expensive of combinations, but given that Cappuccino ranks 2 in terms of popularity, it will make up in volumes.

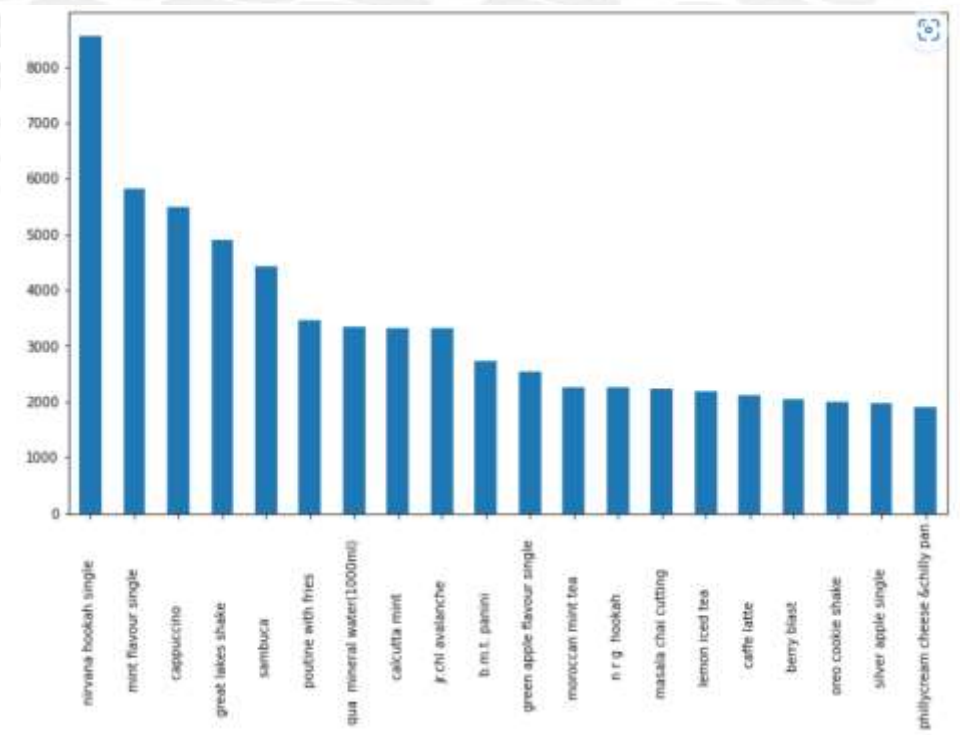
Interestingly, Fries / French Fries is an accompaniment that features in many Food item combinations.

We include combinations with Lift value 2 or 3, when the corresponding support is 0.003 or more.

Here we find the combination of Sambuca and Red Bull appearing in 2 sets.

Also, we have more of cross-category combinations in this case, unlike in the previous selection based on Lift alone.

So combining both approaches, we have shortlisted 24 associations which we will develop into combos presently.



POSSIBLE COMBOS AND OFFERS:

Combo 1. 'Add Fries to it!' : Paninis and French Fries

It can be noticed that out of the 20 associations shortlisted, 6 of them were paired with French Fries / Add Fries (which are essentially the same item).

The items are:

- SATAY CHICKEN PANINI
- B.M.T. PANINI
- COUNTRY ROAST CHICKEN PANINI
- COTTAGE CHEESE PANINI
- PHILLYCREAM CHEESE &CHILLY PAN



All these Food Items listed are essentially Paninis, price points between INR 105 to INR 115. Now this indicates Paninis and Fries have a strong fit, but the Confidence indicators show that only 6% to 12% of people opt for Fries. In the case of BMT Panini it is 20%.

We could provide Fries at an aggressively discounted price for more people to opt for it. If an additional INR 35 dissuades customers from buying Fries, perhaps a substantial discount would induce a larger number of people to opt for Fries with their order. Fries / French Fries could be priced at INR 15 / 20 when purchased with a Panini.

Combo 2. 'Twinning!'

Applied to Cold Beverages such as Ice Teas, Non-Alcoholic Drinks and Shakes - to induce more people to opt for the more expensive non-alcoholic beverages, instead of the cheaper coffees and drinks.

The combinations featured here are:

- LEMON ICED TEA and BLACK CURRANT ICED TEA (both priced at INR 85)
- COOL CALIFORNICA paired with MIAMI MELONS, PINK LEMONADE, BERRY BLAST (INR 85 / 95)
- THE FERROR ROCHER SHAKE (INR 175) and KIT KAT SHAKE (INR 135)



The confidence measure for all these combinations vary. The Iced Teas show a 20% confidence but other drinks range from 6 to 14%. Nonetheless, there is scope to increase consumption of these drinks using combo offers.

One could offer a 25% to 50% off on the second drink in the cold beverage category or sub-category.

Combo 4: Premium Tobacco and Red Bull



Strong associations seen in the following:

- ARABIAN MIST (INR 350) and RED BULL 2+1 (INR 250)
- SAMBUCA (INR 350) and RED BULL 2+1 (INR 250)
- SAMBUCA (INR 350) and RED BULL ENERGY DRINK (INR 125)

Sambuca / AM Combo with Red Bull 2 + 1 could be priced at INR 500, and with Red Bull Energy Drink at INR 450.

APPENDIX:

The supporting files are,

Jupyter notebook – [Renisha R_28Jan23](#)

Tableau: <https://public.tableau.com/app/profile/renisha8818/viz/RenishaR-27jan2023/MRA?publish=yes>

PPT – [Converted as PDF](#).

THANK YOU