



Market Basket Analysis

By,
Anitha E

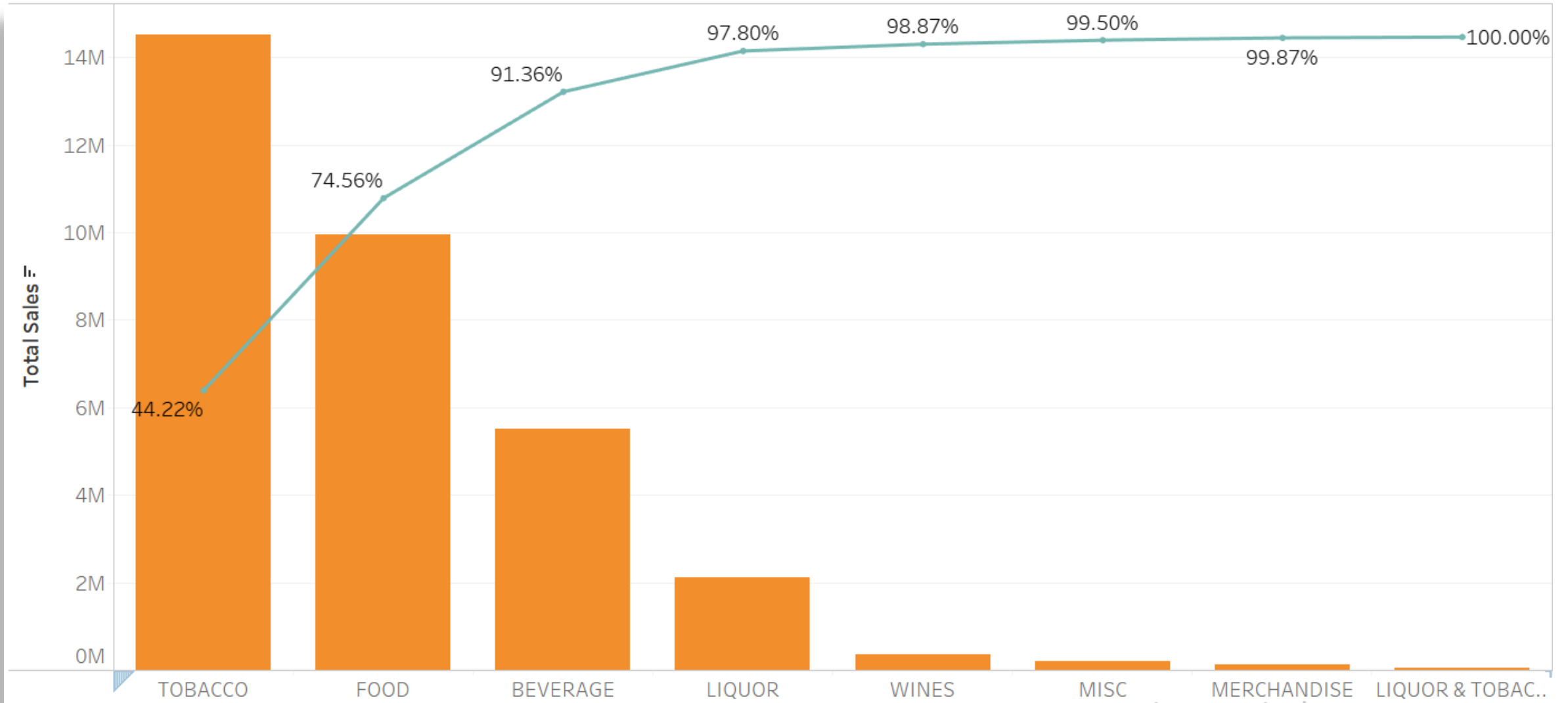
Executive Summary

- Categories such as “Tobacco”, “Beverage” and “Food” are the major contributors of Sales, as they contribute 90% of the total sales.
- Among various categories, “Tobacco” has the highest sales with 14,507,076 and lowest contributor to sales is “Liquor & Tobacco”
- Items categorized under ‘Food’ & ‘Beverage’ categories are sold most number of times.
- The average price of ‘Liquor & Tobacco’ is high and ‘Beverage’ is low compared to other categories.
- The sales is high during
- It is observed that sales high during the months of December, January, February and March in terms of total amount of sales and number of orders.
- It is also observed that average amount sales across months are almost similar and not widely distributed.
- It could be visualized that the 75% amount of sales on every day ranges within 1 and 1120. The highest orders in terms of amount are purchased during Friday, Saturday and Sunday

Exploratory Data Analysis

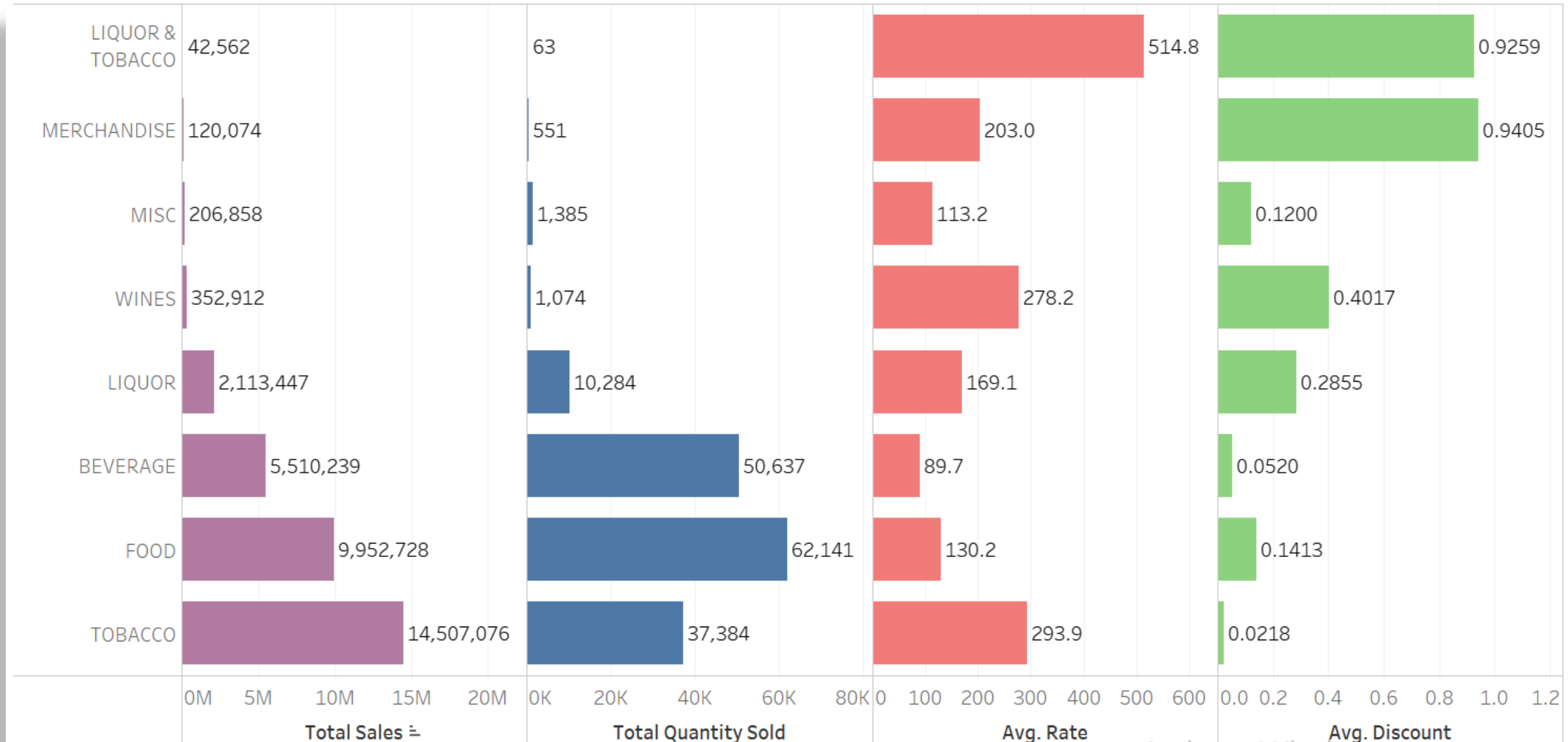
- There are 69982 orders in the dataset purchased for a period of one year, i.e from 01-04-2010 to 31-03-2011
- The total sales to the company is about 32805895.27 and total discount given is 13865
- On an average, customers buy 2 items per order.
- The average cost of single order is about 468.77
- Orders contain items from different categories such as Food, Beverage, Merchandise, MISC, Liquor, Tobacco and Liquor & Tobacco.
- Categories such as “Tobacco”, “Beverage” and “Food” are the major contributors of Sales, as they contribute 90% of the total sales.
- The following are the number of items in each category:
 - ❑ **Food – 208**
 - ❑ **Beverage – 118**
 - ❑ **Merchandise – 87**
 - ❑ **Tobacco – 52**
 - ❑ **Wines – 50**
 - ❑ **MISC - 49**
 - ❑ **Liquor – 34**
 - ❑ **Liquor & Tobacco - 4**

Contribution of each Category towards Total Sales



Categories such as “Tobacco”, “Food” and “Beverage” are the major contributors of Sales, as they contribute 90% of the total amount of sales.

Category Vs. Total Sales / Quantity / Average Rate / Discount



Frequently sold Items under each Category

		Product Rate	Quantity	FOOD
1	GREAT LAKES SHAKE	115	5,914	
2	POUTINE WITH FRIES	125	3,741	
3	JR.CHL AVALANCHE	167	3,446	
4	B.M.T. PANINI	105	3,010	
5	OREO COOKIE SHAKE	145	2,259	
6	PHILLYCREAM CHEESE &CHILLY PAN	105	2,173	
7	KIT KAT SHAKE	135	1,819	
8	COTTAGE CHEESE PANINI	105	1,780	
9	COUNTRY ROAST CHICKEN PANINI	115	1,436	
10	MAGGI NDLCREAM/ CHEE/GARLIC	125	1,373	

		Product Rate	Quantity	BEVERAGE
1	CAPPUCCINO	62	7,144	
2	QUA MINERAL WATER(1000ML)	50	3,633	
3	MASALA CHAI CUTTING	42	2,993	
4	MOROCCAN MINT TEA	46	2,812	
5	LEMON ICED TEA	85	2,687	
6	CAFFE LATTE	72	2,523	
7	RED BULL ENERGY DRINK	132	2,266	
8	BERRY BLAST	94	2,216	
9	COUNTRY LEMONADE	85	2,017	
10	PINK LEMONADE	85	1,554	

		Product Rate	Quantity	LIQUOR
1	CARLSBERG	121	3,380	
2	TUBORG	102	2,855	
3	KF DRAUGHT (1/2LTR)	126	1,520	
4	KF DRAUGHT (1LTR)	234	1,073	
5	KF DRAUGHT PITCHER (2LTR)	413	611	
6	BUDWEISER	108	200	
7	1+1 KF 1/2 LITER	125	136	
8	HOEGAARDEN (GLS)	243	66	
9	HOEGAARDEN MUG (1 LITRE)	617	53	
10	WHITE SANGRIA (GLS)áááá áááááá	175	49	

		Product Rate	Quantity	LIQUOR & TOBACCO
1	BEER HOOKAH	500.0	49.0	
2	2 DOM BEER + 1SPL SHEESHA	500.0	10.0	

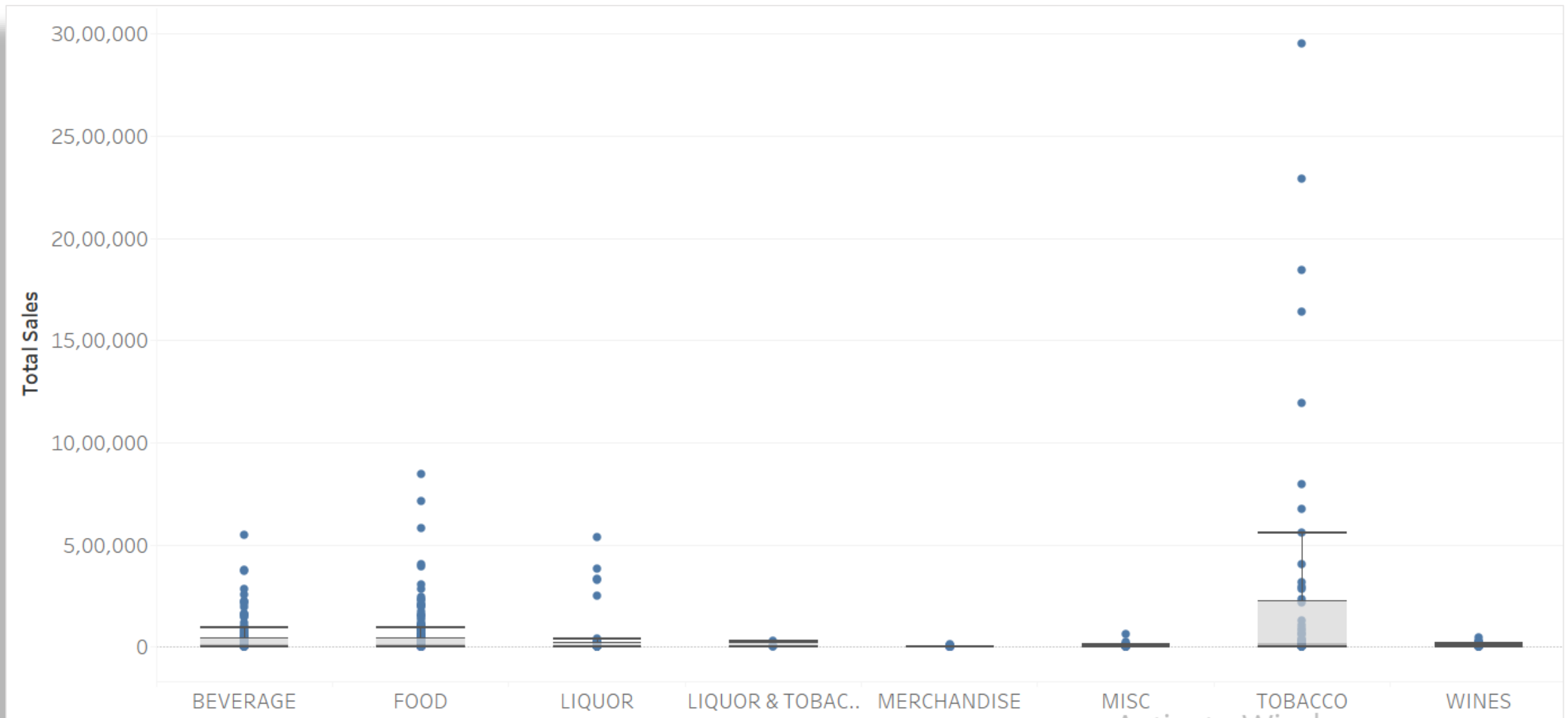
		Product Rate	Quantity	MERCHDIS
1	GREAT LAKES T-SHIRTS	345	37	
2	CH NOTE BOOKS	65	32	
3	MUGS - PATTERN	89	27	
4	CH COFFEE MUGS	200	27	
5	CH TINS	214	26	

		Product Rate	Quantity	MISC
1	ADD ON S	15	448	
2	PLAIN JANE (CHOCOLATE)	100	182	
3	ADD HERB ROAST CHICKEN	30	179	
4	RED BULL SHEESHA	450	103	
5	PLAIN JANE (STRAWBERRY)	100	100	

		Product Rate	Quantity	TOBACCO
1	NIRVANA HOOKAH SINGLE	258	8,686	
2	MINT FLAVOUR SINGLE	232	6,019	
3	SAMBUCA	365	4,765	
4	CALCUTTA MINT	372	3,339	
5	GREEN APPLE FLAVOUR SINGLE	232	2,590	
6	N R G HOOKAH	391	2,314	
7	SILVER APPLE SINGLE	257	1,985	
8	ARABIAN MIST	350	1,213	
9	MISCHIEF HOOKAH SINGLE	264	816	
10	RED WINE SHEESHA	392	778	

		Product Rate	Quantity	WINE
1	VLN CAB SAUV (GLS)	175	216	
2	RED SANGRIA (GLS)á áááááááááá	175	111	
3	SULA BLUSH ZINFANDEL(GLS)	200	107	
4	VLN CHENIN BLANC (GLS)	175	102	
5	VLN SAUV BLANC (GLS)	175	98	

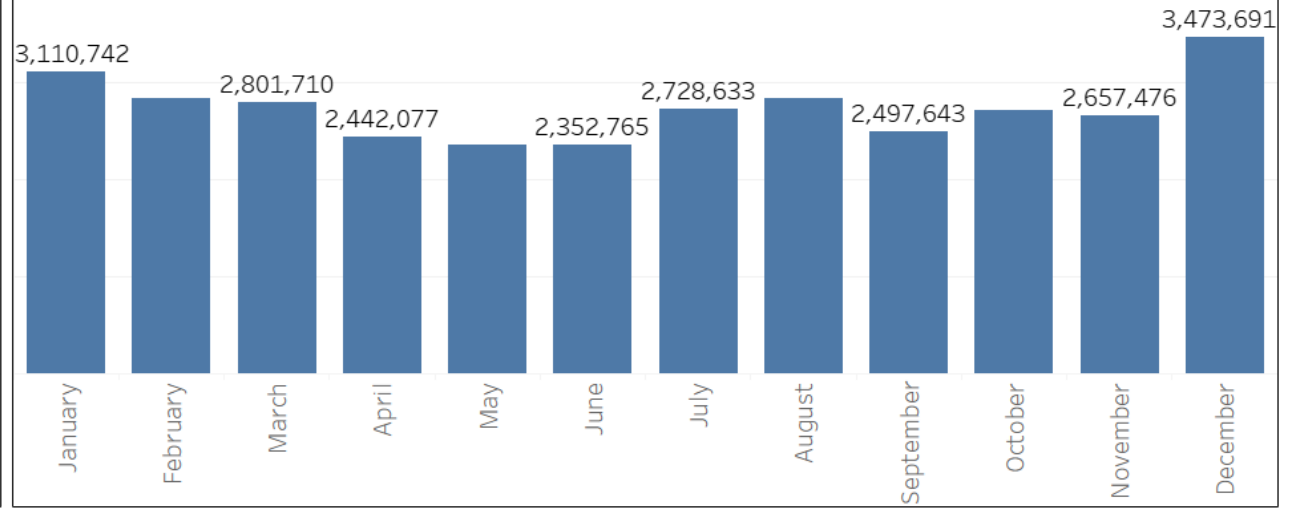
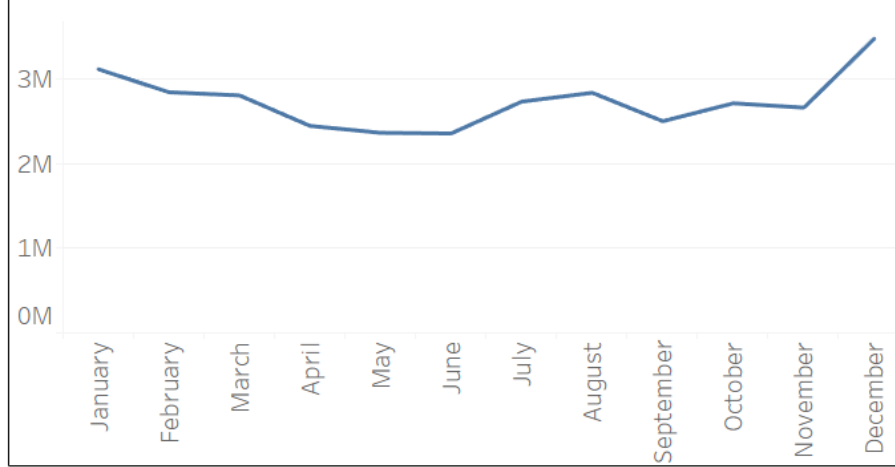
Sales across each Category



✓ It could be visualized that the sales of Tobacco are widely spread compared to other categories in the orders purchased.

Sales Trend across Months

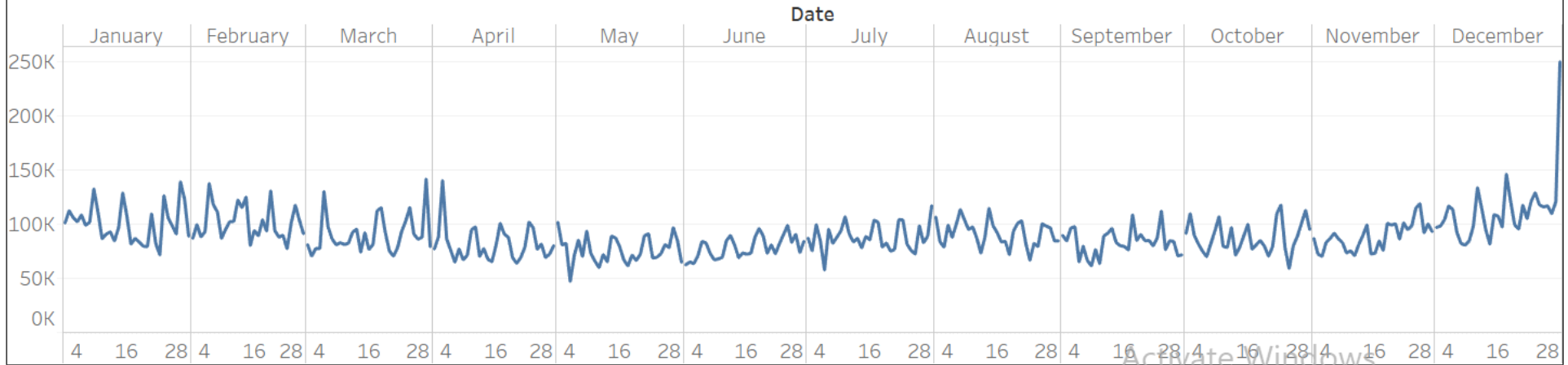
Total Sales cross Months



Mean Sales

January	233.70
February	239.60
March	241.42
April	218.63
May	219.51
June	216.76
July	213.59
August	213.07
September	217.19
October	224.24
November	228.03
December	231.83

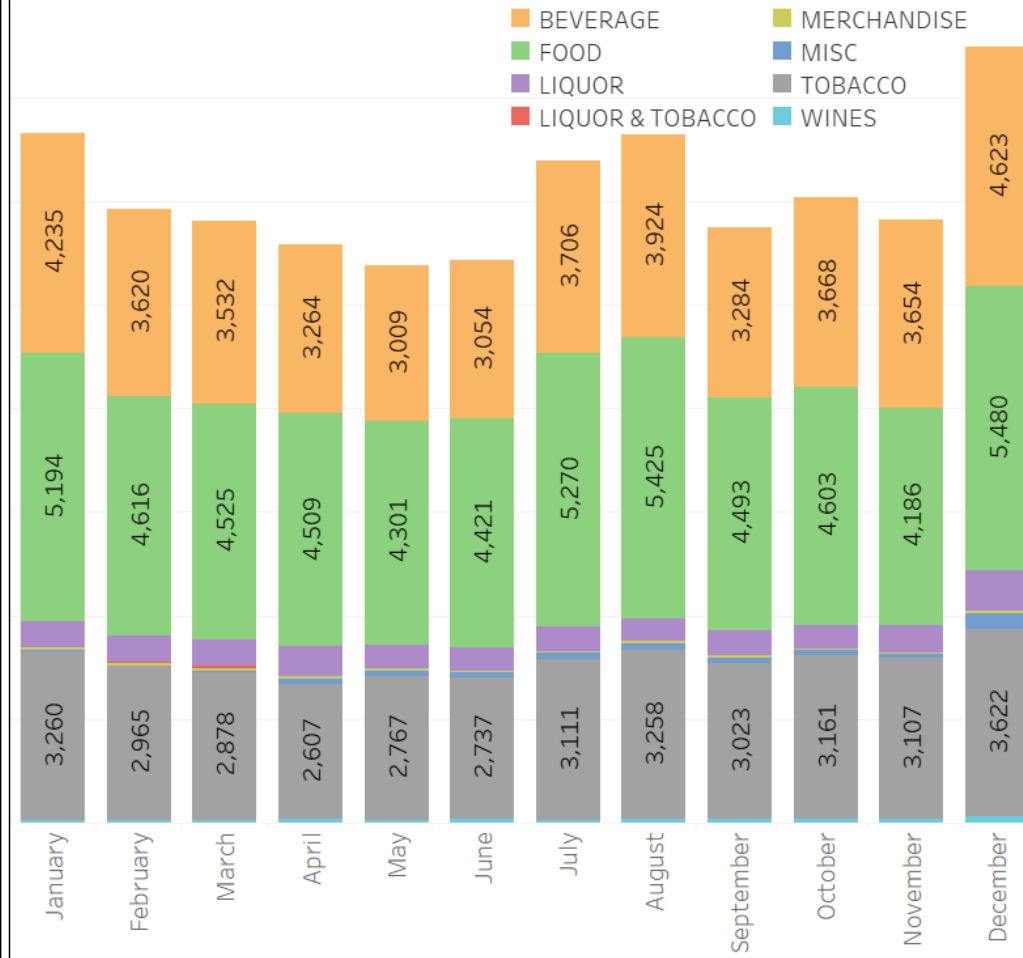
Total Sales across days/months



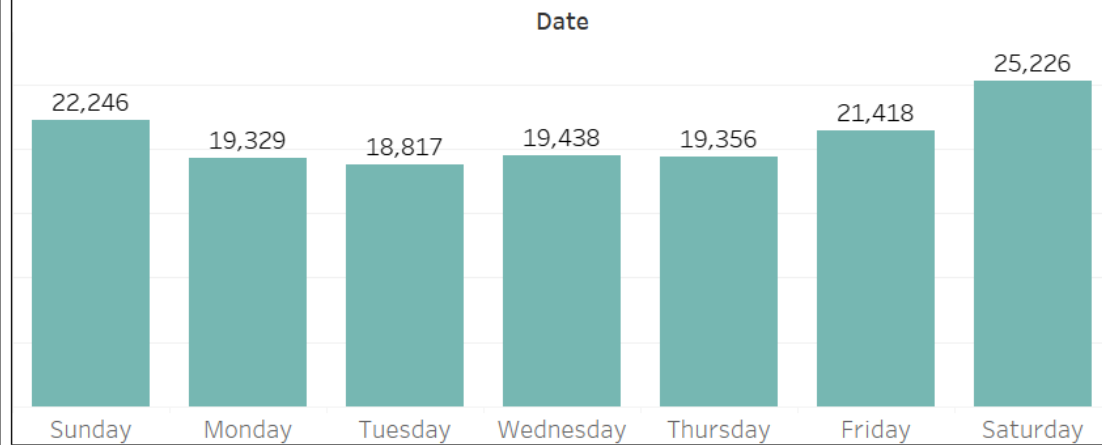
- ✓ It is observed that the total amount sales is higher in the months of December, January, February & March compared to other months.
- ✓ It is also observed that average amount sales across months are almost similar.
- ✓ The sales across days is tend to have a weekly rise in the amount of sales and recording the highest sales during the year end festive time.

Customer Behaviour across Months/Days of Week/Hour of Day

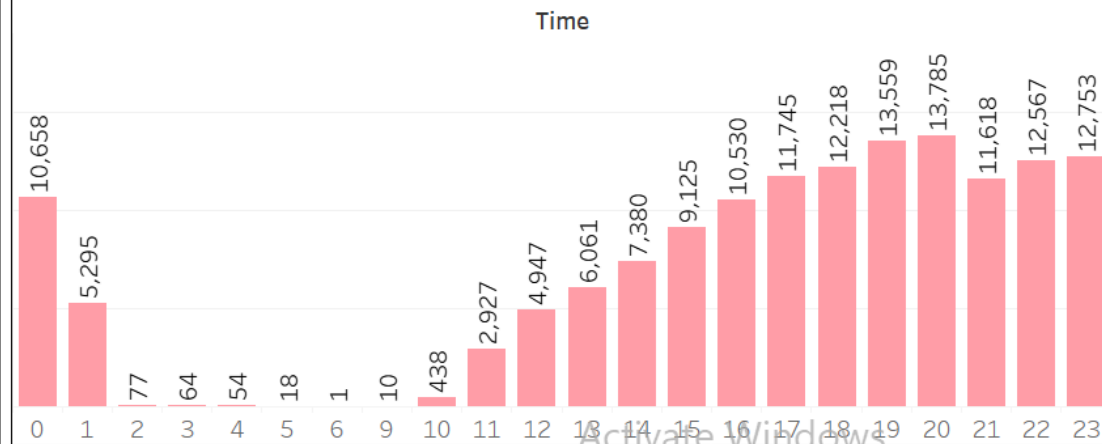
Order Counts of each Categories across Months



Order Counts over each Day of Week

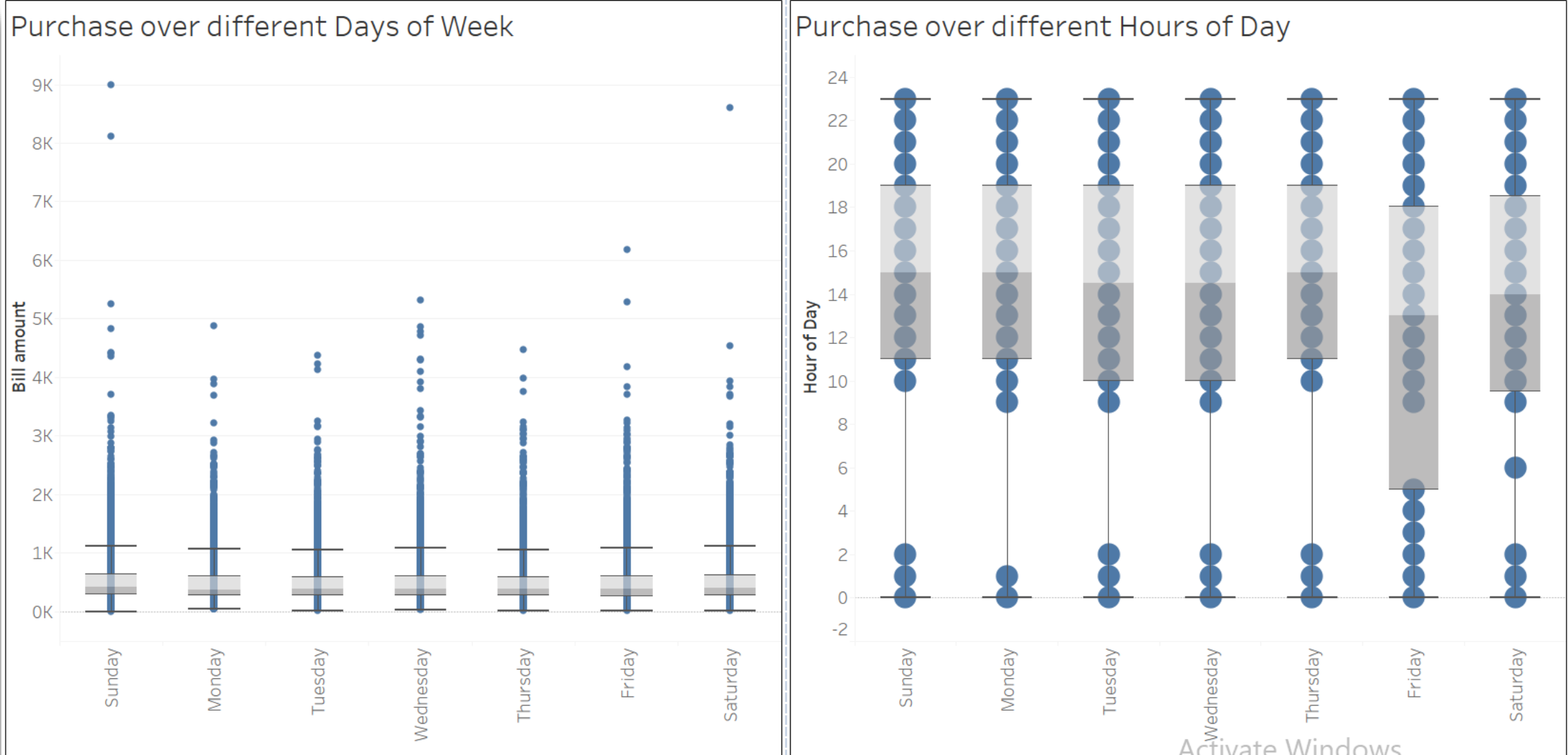


Order Counts over Different Hours of Day



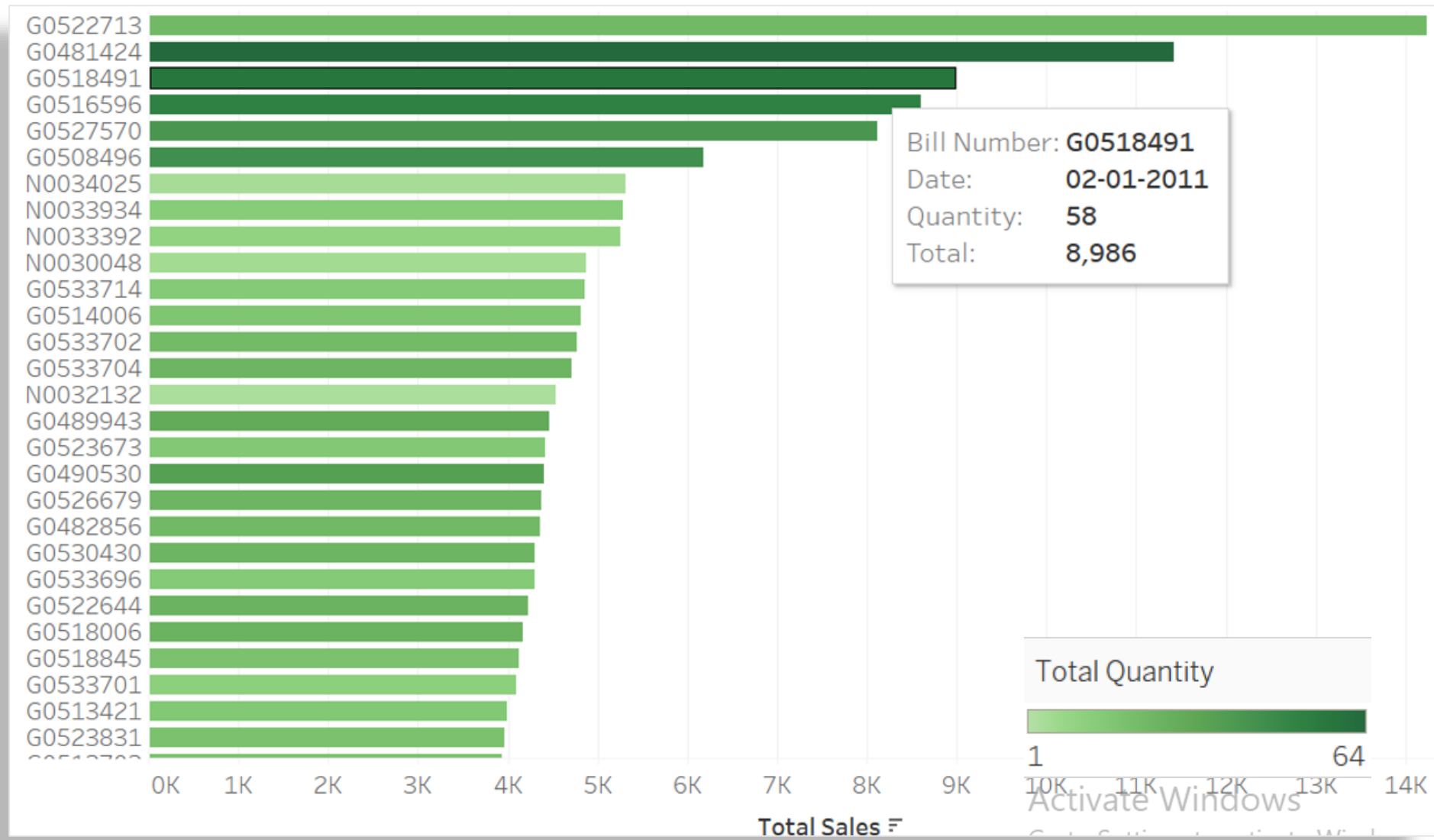
- ✓ It is observed that December, July & August records the highest number of orders with 'Food' & 'Beverage' being bought mostly.
- ✓ It could be identified that week ends (Friday, Saturday, Sunday) registers the highest number of orders.
- ✓ Night hours records the highest number of orders as compared to Day time, especially from 5 PM TO 12 PM. The early morning hours 2.00 AM to 9.00 AM hardly sees any orders and orders start to gradually increase in the noon to midnight.

Purchase across Days of Week & Hour of Day



- ✓ It could be noted that customers tend to buy a lot during Friday. Almost every hour of Friday, we could observe the order been placed and other days show similar pattern of purchase across hours.
- ✓ 75% of the bill amount ranges from 1 to 1120 and highest purchases of bill occurs on Saturday and Sunday.

Amount Vs. Quantity of Top Bills



- ✓ The above are the top Bills in terms of their amount of purchase.
- ✓ The most number of quantities sold in a single order is 64

Items sold less frequently

BEVERAGE

Item Desc	Product Rate	Quantity
2 AXE TWIST	140.0	1.0
DECAFFINATE COFFEE FRAPPE	75.0	1.0
MIXED FLAVOUR SINGLE	225.0	1.0
MOCAFE HOT CHOCOLATE(SF)	65.0	1.0
PEACH BULL	150.0	1.0

FOOD

Item Desc	Product Rate	Quantity
ADD BUTTERED TOAST	15	1
CAPONATA	150	1
NIRVANA HOOKAH DOUBLE	295	2
SCHNEIDER 2+1	600	2
STRAWBERRY MERINGUE	200	2
SUNNY SIDEUP + BEVERAGE	99	2

TOBACCO

Item Desc	Product Rate	Quantity
AL SIKANDARI HOOKAH DOUBLE	345.0	1.0
CLASSIC REGULAR	78.0	1.0
GOLD FLAKE ULTRA LIGHTS(20)	73.0	1.0
INDIA KINGS OCEAN BLUE	91.5	1.0
APPLE FLAVOUR DOUBLE	280.0	2.0
CLASSIC MENTHOL RUSH	83.3	2.0
GREAT LAKES HOOKAH SINGLE	295.0	2.0
ICE SPICE SHEESHA	200.0	2.0

MERCHANDISE

Item Desc	Product Rate	Quantity
ASH TRAYS	150	1
BEACH GREEN	262	1
BENARAS BLUE	351	1
CH TIN SMALL	171	1
CHAIRMAN COOL	262	1
COUNTRY LEMONADE GLASS(HANSA)	150	1
DHARMATEA LIGHT HOLDER	249	1
DIP BOWL	60	1
ETCHED LEAF TLIGHT HOLDER	249	1
FLAVOR 1000 GMS	1,470	1
GOLDEN DELIGHT 1.1	351	1
GREAT LAKES MUG SINGLE	142	1
GUERRILLA COOL	262	1
KITSCH BLUE	262	1
KITSCH PINK	262	1
KONKAN STRIPE	262	1
MODEL-P (IRON ASHTRAY)	444	1
MUGS - PLAIN COLOUR	67	1
ROCK THE BOAT TEA LIGHT HOLDER	222	1
SHAKE GLASS	250	1
SILVER STYLER 1.0	351	1
UDAIPUR LILY	351	1
ZEN ROCK SQUARE VASE	440	1

MISC

Item Desc	Product Rate	Quantity
GNOCCI CON POMMODORO	150	1
GRAPPO SHEESHA	450	1
HOEGAARDEN GLS (2+1)	700	1
HOEGAARDEN LTR MUGS (2+1)	1,300	1
JAPANESE YAKITORI WRAP	150	1
M & M SHAKE	180	1
MANGO FLAVOUR SINGLE	225	1
MEXICAN CHILLY CREPES	165	1
POLLO PELOTA CON TAMATE SALSA	190	1
RASPBERRY LITE SHAKE	115	1
RED BULL 2+1	250	1
RED SANGRIA (CARAFE) áááááááá	800	1
ROMA TOMATO & JALAPENO CROQUE	125	1

LIQUOR

Item Desc	Product Rate	Quantity
ZINZI WHITE (BTL)	700	1
WHISKEY (SM)	200	2

WINES

Item Desc	Product Rate	Quantity
1+1 VLN CAB SAUV (BTL)	800	1
1+1 VLN SAUV BLANC (BTL)	800	1
2 OCEAN PINOTAGE (BTL)	1,900	1
4 SEASONS CLAS SYRAH(BTL)	800	1
B1G1 4SEASON CLAS SAUV(BTL)	900	1
B1G1 4SEASON CLAS SAUV(GLS)	200	1
GOSSIPS CHARD AUS (BTL)	2,100	1
MANDALA VALLEY CHENIN BLA..	175	1
MATEUS ROSE PORTUGAL(BTL)	2,000	1
SULA CHENIN BLANC (BTL)	800	1

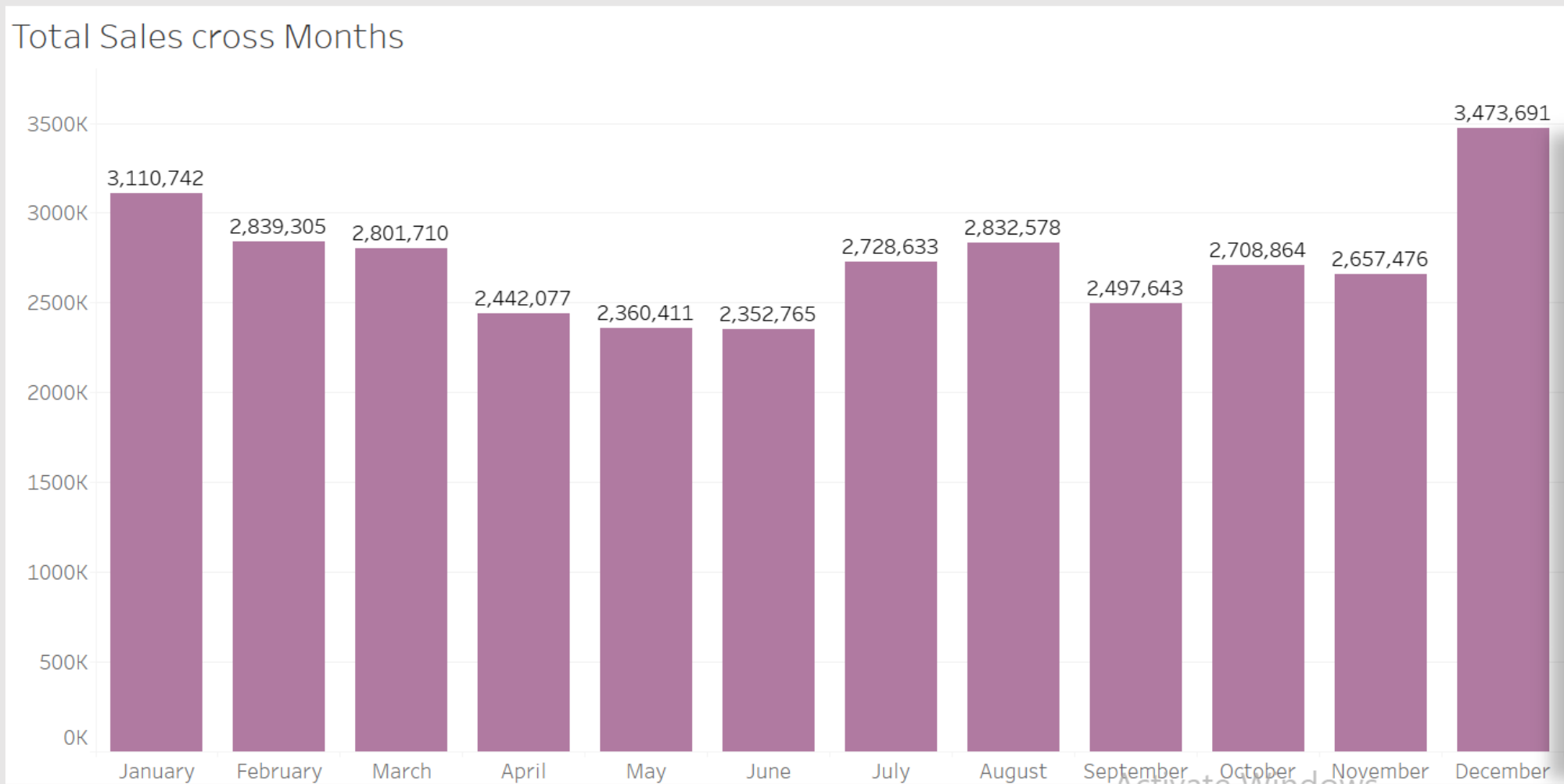
LIQUOR & TOBACCO

Item Desc	Product Rate	Quantity
2 DOM BEER + 1PREM SHEESHA	650.0	2.0
4 DOM BEER + 1SPL SHEESHA	750.0	2.0

- ✓ The above are the items under each category that are sold for once or twice in year. Hence it could be suggested that these products could be taken off from the menu.

Sales Trend across Months

- ❑ It is observed that sales high during the months of December, January, February and March in terms of total amount of sales and number of orders.
- ❑ It is also observed that average amount sales across months are almost similar and not widely distributed.

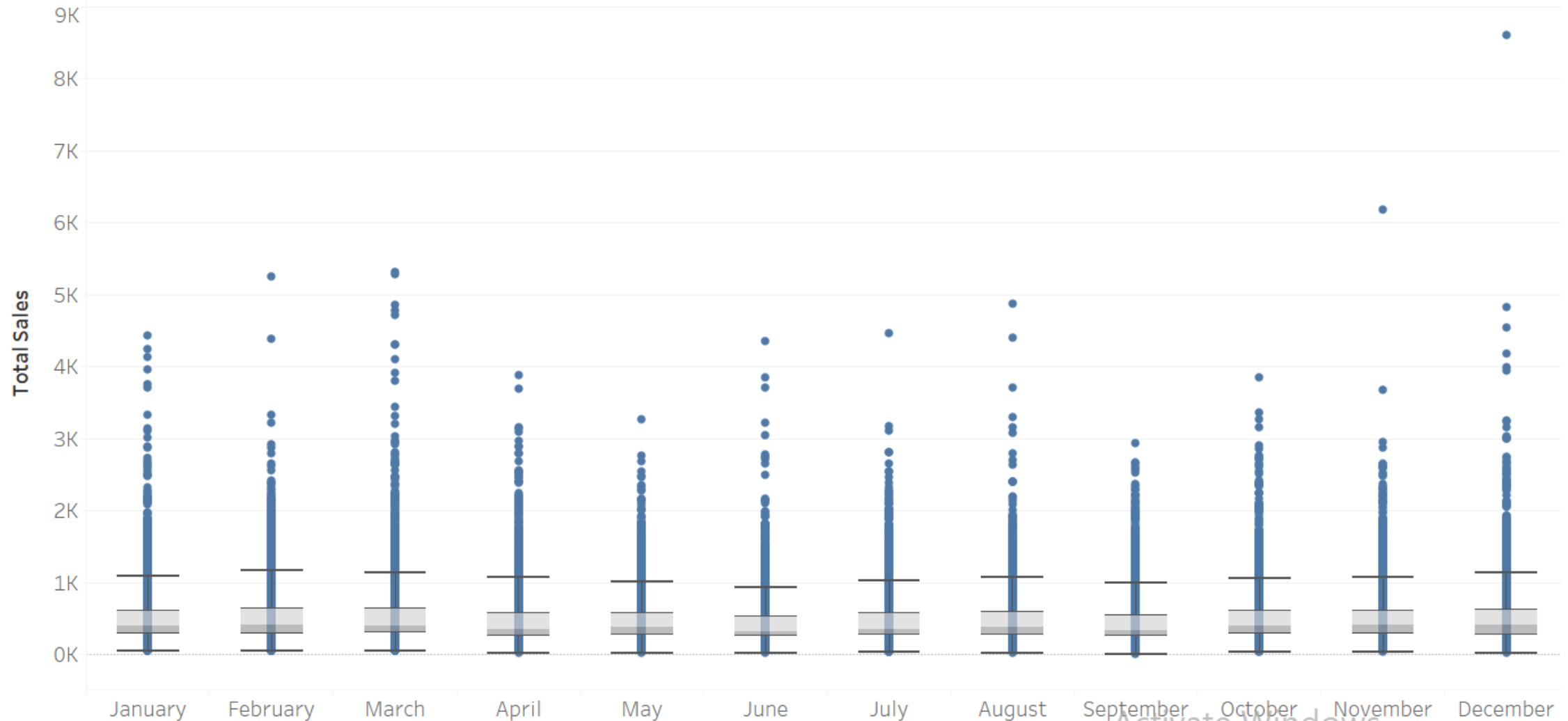


MEAN SALES

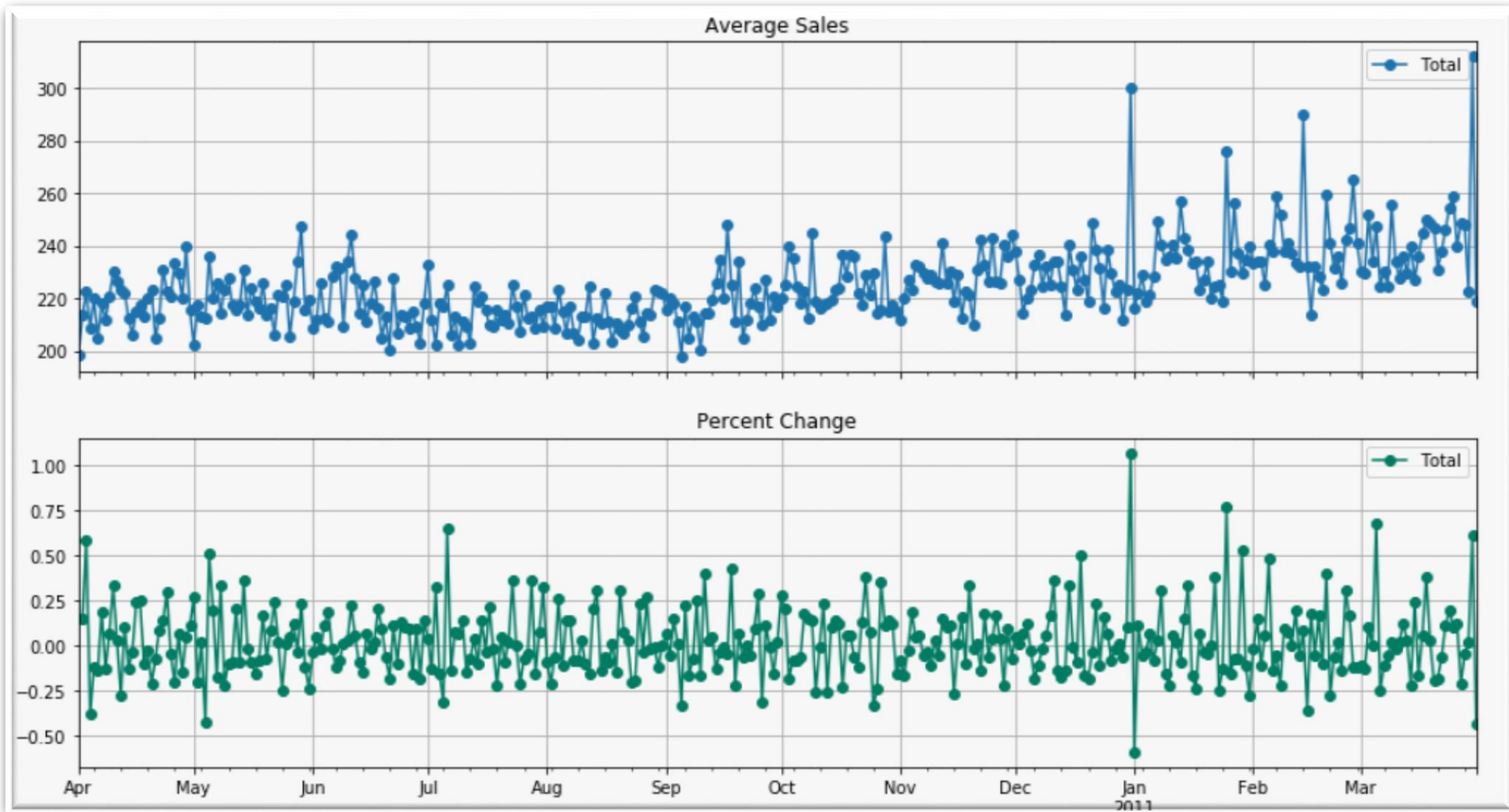
January	233.70
February	239.60
March	241.42
April	218.63
May	219.51
June	216.76
July	213.59
August	213.07
September	217.19
October	224.24
November	228.03
December	231.83

Sales Trend across Months – Box Plot

- ❑ It is observed that 75% of the total sales ranges from 1 to 1120 across various months.
- ❑ It is also observed that average amount sales across months are almost similar and not widely distributed.



Average Sales Trend & Percent Change across Months



Market Basket Analysis (MBA)

- In Retail, each customer purchase different set of products, different quantities and in different period of time.
- MBA uses the information to:
 - Understand the behaviour of customers.
 - Identify the pattern of customer purchases.
 - Gather insights about their product items.
 - ☐ Most & Least frequently bought
 - ☐ Products bought together
 - ☐ Profits gained from Promotions
 - Decide the combination of products to be sold.
 - Determine the store layout.

MBA - Association Rules

- Finding frequent patterns, associations, correlations, or causal structures among sets of items in transaction databases
- Understand customer buying habits by finding associations and correlations between the different items that customers place in their “shopping basket”

The diagram illustrates the relationship between an association rule and its associated metrics. A central rule, $Rule: X \Rightarrow Y$, is shown on the left. Three blue arrows originate from this rule and point to the right, each leading to a specific metric:

- The top arrow points to the formula for **Support**: $Support = \frac{freq(X, Y)}{N}$
- The middle arrow points to the formula for **Confidence**: $Confidence = \frac{freq(X, Y)}{freq(X)}$
- The bottom arrow points to the formula for **Lift**: $Lift = \frac{Support}{Supp(X) \times Supp(Y)}$

MBA - Association Identified

Table "default" - Rows: 1961 Spec - Columns: 6 Properties Flow Variables						
Row ID	D Support	D ▼ Confidence	D ▼ Lift	S Recommended Product	S implies	[...] Basket Items
rule231	0	1	3,332.476	2 RED BULL	<---	[N R G HOOKAH]
rule172	0	0.933	14.767	SAMBUCA	<---	[QUA MINERAL WATER(1000ML),RED BULL 2+1,GREAT LAKES SHAKE]
rule1427	0.001	0.8	26.723	CAFFE LATTE	<---	[ADD VANILLA FLAVOUR]
rule817	0	0.8	12.658	SAMBUCA	<---	[QUA MINERAL WATER(1000ML),MAGGI NDL ARRABIATA,B.M.T. PANINI]
rule232	0	0.714	3,332.476	N R G HOOKAH	<---	[2 RED BULL]
rule168	0	0.7	11.076	SAMBUCA	<---	[QUA MINERAL WATER(1000ML),RED BULL ENERGY DRINK,GREAT LAKES SHAKE]
rule382	0	0.64	21.379	CAFFE LATTE	<---	[ADD IRISH CREAM FLAVOUR]
rule1485	0.001	0.638	21.301	CAFFE LATTE	<---	[ADD CARAMEL FLAVOUR]
rule1068	0	0.625	20.878	CAFFE LATTE	<---	[ADD CINNAMON FLAVOUR]
rule816	0	0.625	16.79	B.M.T. PANINI	<---	[QUA MINERAL WATER(1000ML),MAGGI NDL ARRABIATA,SAMBUCA]
rule1299	0	0.561	4.683	NIRVANA HOOKAH SINGLE	<---	[QUA MINERAL WATER(500ML),MOROCCAN MINT TEA]
rule496	0	0.548	959.431	SAIGON NOODLES	<---	[ADD CHICKEN]
rule1795	0.002	0.546	18.235	CAFFE LATTE	<---	[ADD HAZELNUT FLAVOUR]
rule731	0	0.543	8.589	SAMBUCA	<---	[MAGGI NDL ARRABIATA,RED BULL 2+1]
rule1506	0.001	0.535	8.463	SAMBUCA	<---	[QUA MINERAL WATER(1000ML),RED BULL 2+1]
rule1391	0.001	0.507	8.02	SAMBUCA	<---	[RED BULL 2+1,GREAT LAKES SHAKE]
rule1154	0	0.482	4.022	NIRVANA HOOKAH SINGLE	<---	[RED BULL ENERGY DRINK,QUA MINERAL WATER(500ML)]
rule654	0	0.474	7.495	SAMBUCA	<---	[QUA MINERAL WATER(1000ML),B.M.T. PANINI,GREAT LAKES SHAKE]
rule1620	0.001	0.47	7.439	SAMBUCA	<---	[MAGGI NDL ARRABIATA,B.M.T. PANINI]
rule1670	0.001	0.465	79.364	LEMON INFUSED CHAR GRILLE...	<---	[ADD HERB ROAST CHICKEN]
rule1749	0.001	0.464	5.936	CAPPUCCINO	<---	[ADD HAZELNUT FLAVOUR]
rule1517	0.001	0.462	101.891	KHEEMA GHOTALA	<---	[BUTTERED TOASTS]
rule169	0	0.452	6.488	GREAT LAKES SHAKE	<---	[QUA MINERAL WATER(1000ML),RED BULL ENERGY DRINK,SAMBUCA]
rule955	0	0.449	7.104	SAMBUCA	<---	[PHILLYCREAM CHEESE &CHILLY PAN,RED BULL 2+1]
rule495	0	0.425	959.431	ADD CHICKEN	<---	[SAIGON NOODLES]
rule1455	0.001	0.414	6.553	SAMBUCA	<---	[RED BULL 2+1,B.M.T. PANINI]
rule1724	0.001	0.412	5.924	GREAT LAKES SHAKE	<---	[VANILLA ICECREAM]
rule590	0	0.409	8.386	POUTINE WITH FRIES	<---	[ADD SMOKED CHICKEN]
rule1175	0	0.406	5.194	CAPPUCCINO	<---	[ADD WHIPPED CREAM]
rule332	0	0.405	6.414	SAMBUCA	<---	[MAGGI NDL CREAM/ CHEE/GARLIC,RED BULL 2+1]

Rule - Creation, Summary, Interpretation

Rule Creation:

- We pass $\text{supp}=0.0002$ and $\text{conf}=0.05$ to return all the rules that have a support of at least 0.02% and confidence of at least 5%.
- We sort the rules by decreasing confidence.

Summary of the Rules:

- The number of rules: 1961
- The summary of quality measures: ranges of support, confidence, and lift.

Interpretation of Rules:

- 100% customers who bought "N R G HOOKAH" also bought "2 RED BULL".
- 93.3 % customers who bought "QUA Mineral Water (1000 ML) , RED BULL 2+1, GREAT LAKES SHAKE" also bought SAMBUCA
- 80% customers who bought "ADD VANILLA FLAVOUR" also bought "CAFFE LATTE"
- 80% customers who bought "QUA Mineral Water (1000 ML) , MAGGI NDL ARRABIATA , B.M.T. PANINI" also bought SAMBUCA

Suggestion of Combos

- “N R G HOOKAH” & “2 RED BULL” can be given as a combo, as 100% customers who bought them together.
- “QUA Mineral Water (1000 ML) , RED BULL 2+1, GREAT LAKES SHAKE” and SAMBUCA could be given as a combo, as there is a confidence of 93% that these items are bought together.
- “ADD VANILLA FLAVOUR” and “CAFFE LATTE” makes good combo as they are bought together with a confidence of 80%
- “QUA Mineral Water (1000 ML) , MAGGI NDL ARRABIATA , B.M.T. PANINI” and “SAMBUCA” could be the next combo, as they a confidence of 80% that they are bought together.
- “CAFFE LATTE” could be given as a combo with “ADD IRISH CREAM FLAVOUR” or “ADD CARAMEL FLAVOUR” or “ADD CINNAMON FLAVOUR”
- “QUA Mineral Water (500 ML) , MOROCCAN MINT TEA” and “NIRVANA HOOKAH SINGLE” could be given as a combo, as they are bought together with a confidence of 56%

THANK YOU

