Analysis of Cafe Great Transaction Dataset

by
Gaurang Swarge

Marketing & Retail Analytics

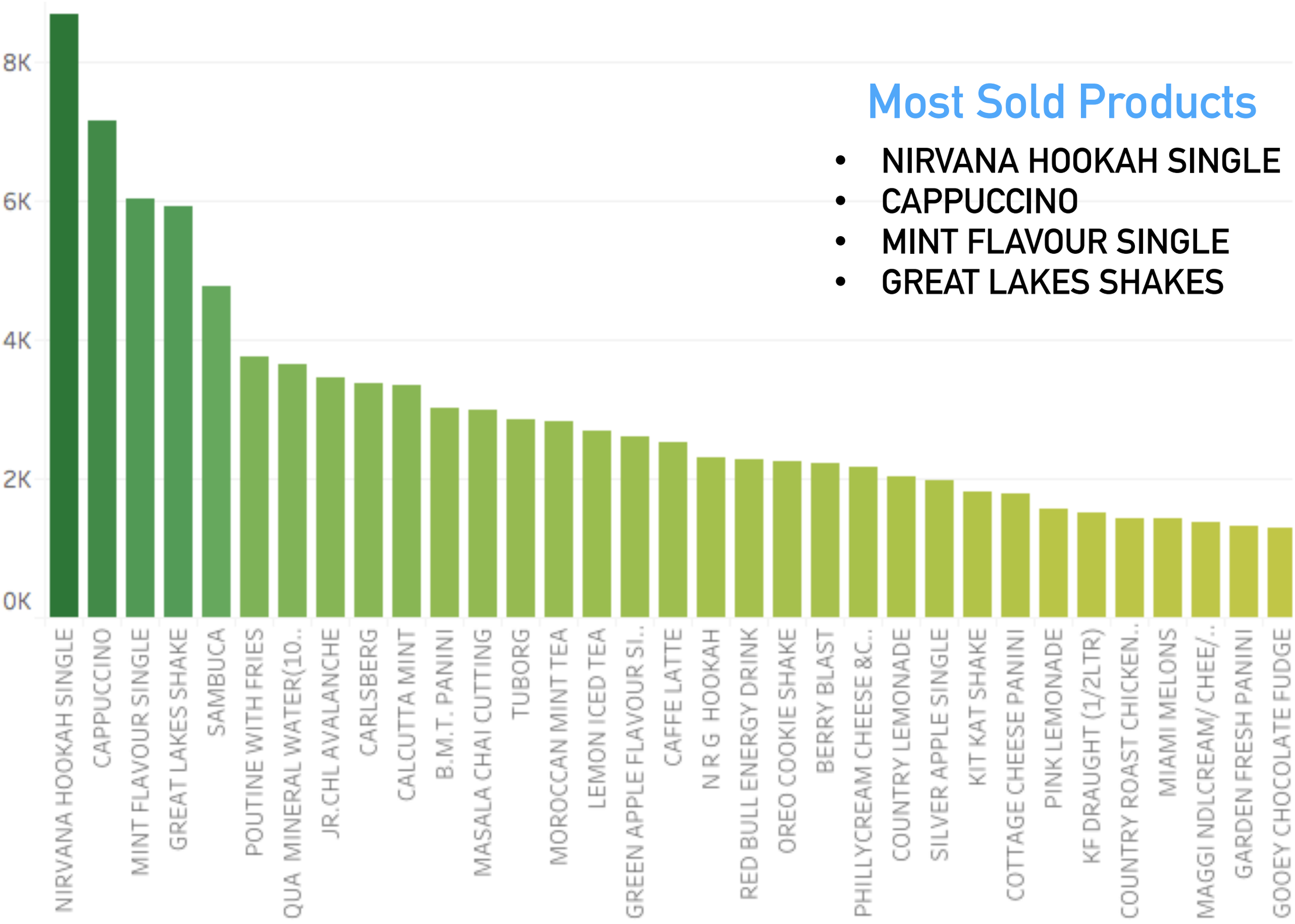
SNAPSHOT

Dataset: Cafe Great Transaction Data set
Time Frame: 1st April 2010 - 31st March 2011

Sample Data:

	A	B	C	D	E	F	G	H	I	J
1	Date	Bill Number	Item Desc	Time	Quantity	Rate	Tax	Discount	Total	Category
2	01-Apr-10	G0470115	QUA MINER	1:15:11 PM	1	50	11.88	0	61.88	BEVERAGE
3	01-Apr-10	G0470115	MONSOON M	1:15:11 PM	1	100	23.75	0	123.75	BEVERAGE
4	01-Apr-10	G0470116	MASALA CHA	1:17:35 PM	1	40	9.5	0	49.5	BEVERAGE
5	01-Apr-10	G0470117	QUA MINER	1:19:55 PM	1	50	11.88	0	61.88	BEVERAGE
6	01-Apr-10	G0470283	MOROCCAN	1:20:18 AM	1	45	10.69	0	55.69	BEVERAGE
7	01-Apr-10	G0470283	QUA MINER	1:20:18 AM	1	50	11.88	0	61.88	BEVERAGE
8	01-Apr-10	G0470118	MASALA CHA	1:21:34 PM	1	40	9.5	0	49.5	BEVERAGE
9	01-Apr-10	G0470118	MOROCCAN	1:21:34 PM	1	45	10.69	0	55.69	BEVERAGE
10	01-Apr-10	G0470284	CAPPUCCINO	1:29:38 AM	1	60	14.25	0	74.25	BEVERAGE

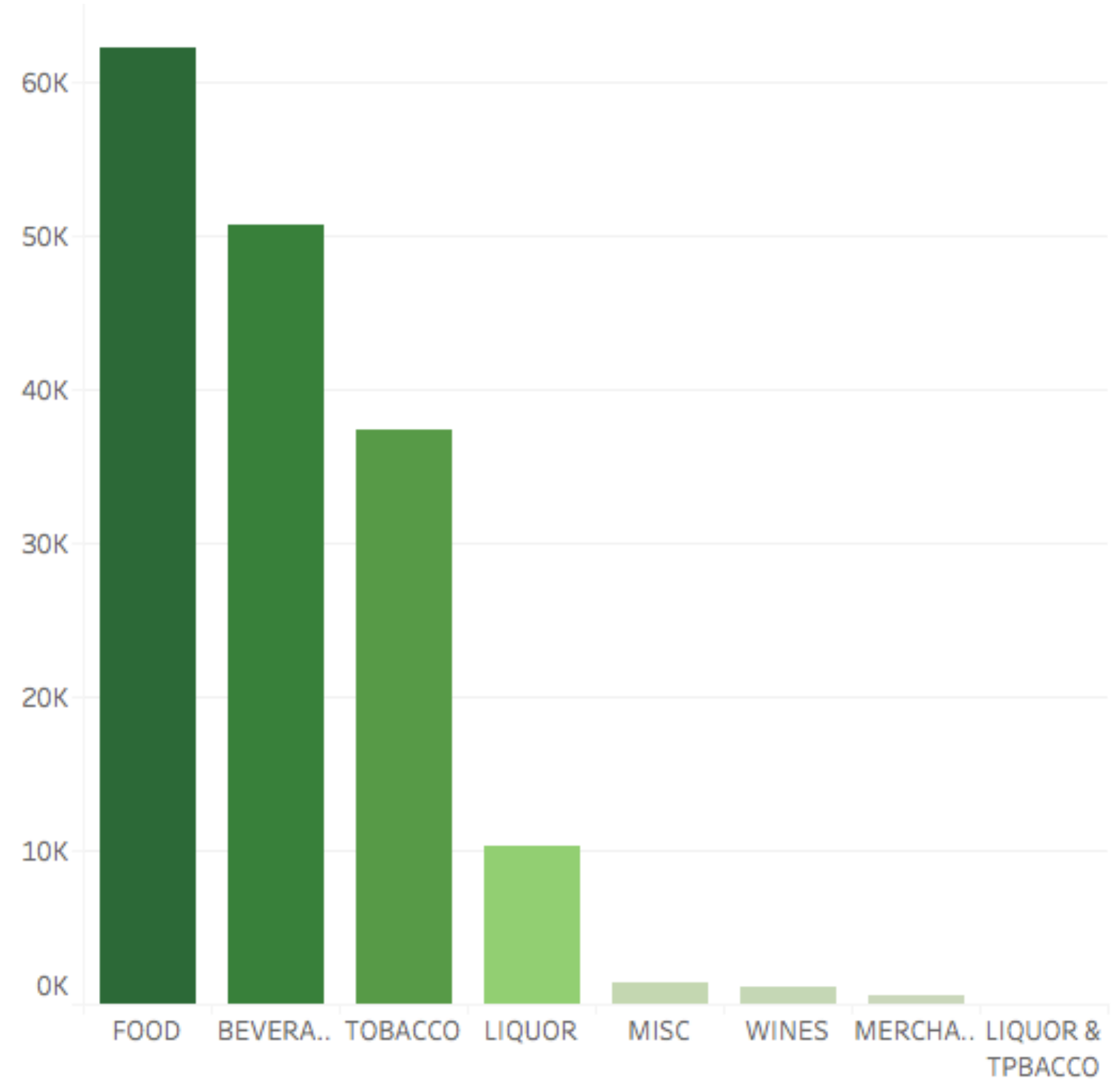
Exploratory Data Analysis



Most Sold Products

- NIRVANA HOOKAH SINGLE
- CAPPUCCINO
- MINT FLAVOUR SINGLE
- GREAT LAKES SHAKES

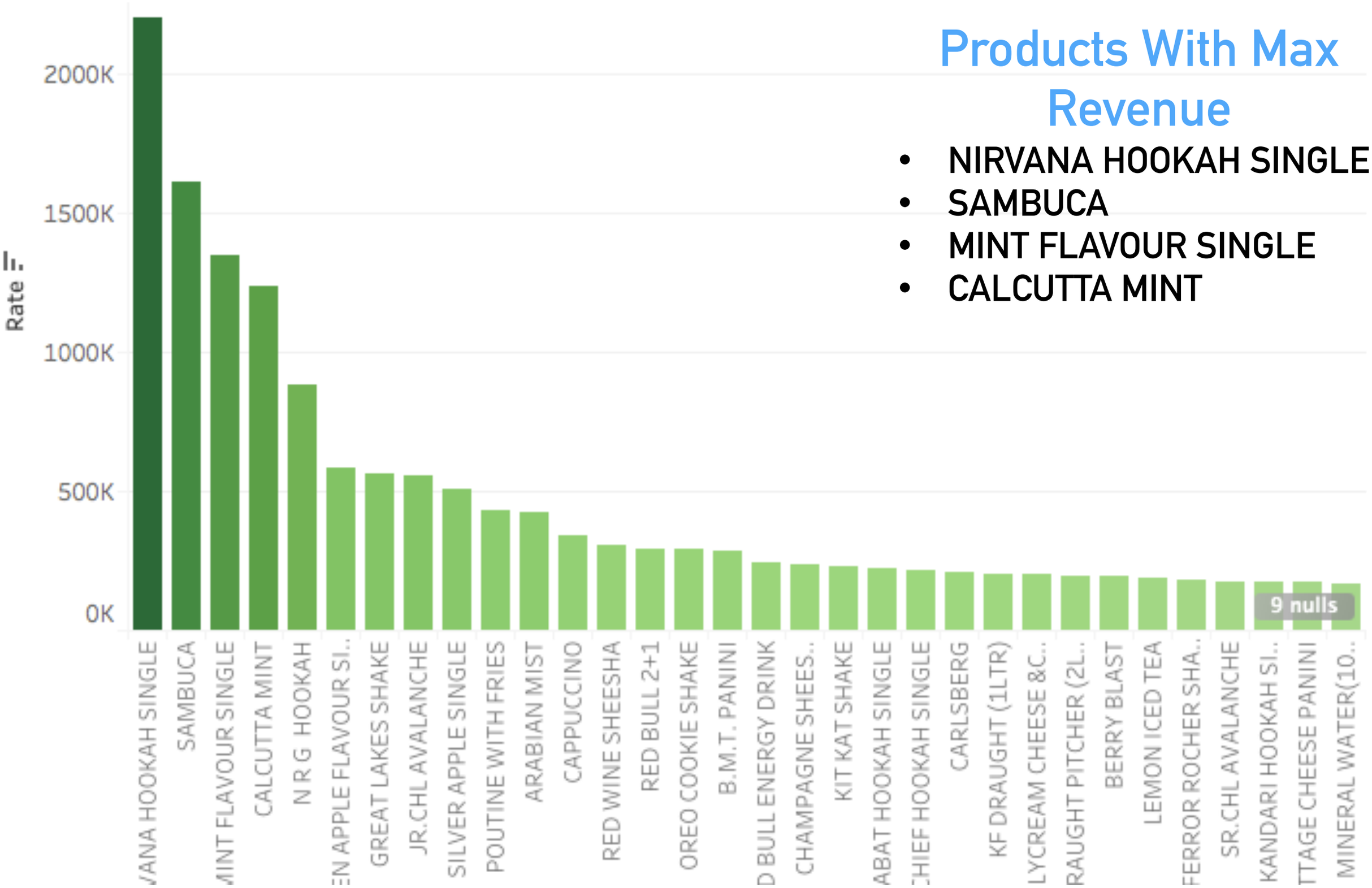
Exploratory Data Analysis



Most Sold Categories

- FOOD
- BEVERAGES
- TOBACCO
- LIQUOR

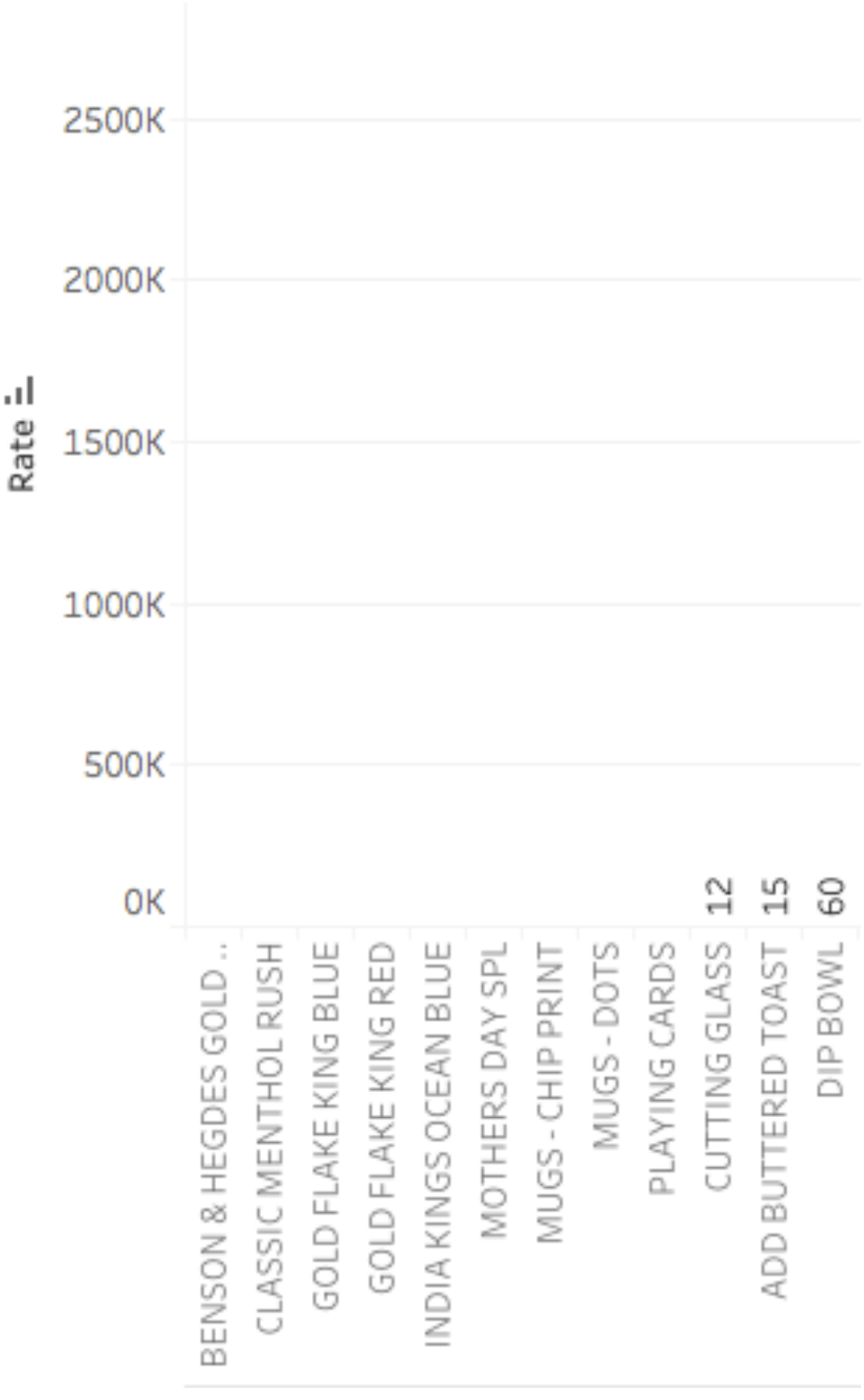
Exploratory Data Analysis



Products With Max Revenue

- NIRVANA HOOKAH SINGLE
- SAMBUCA
- MINT FLAVOUR SINGLE
- CALCUTTA MINT

Exploratory Data Analysis

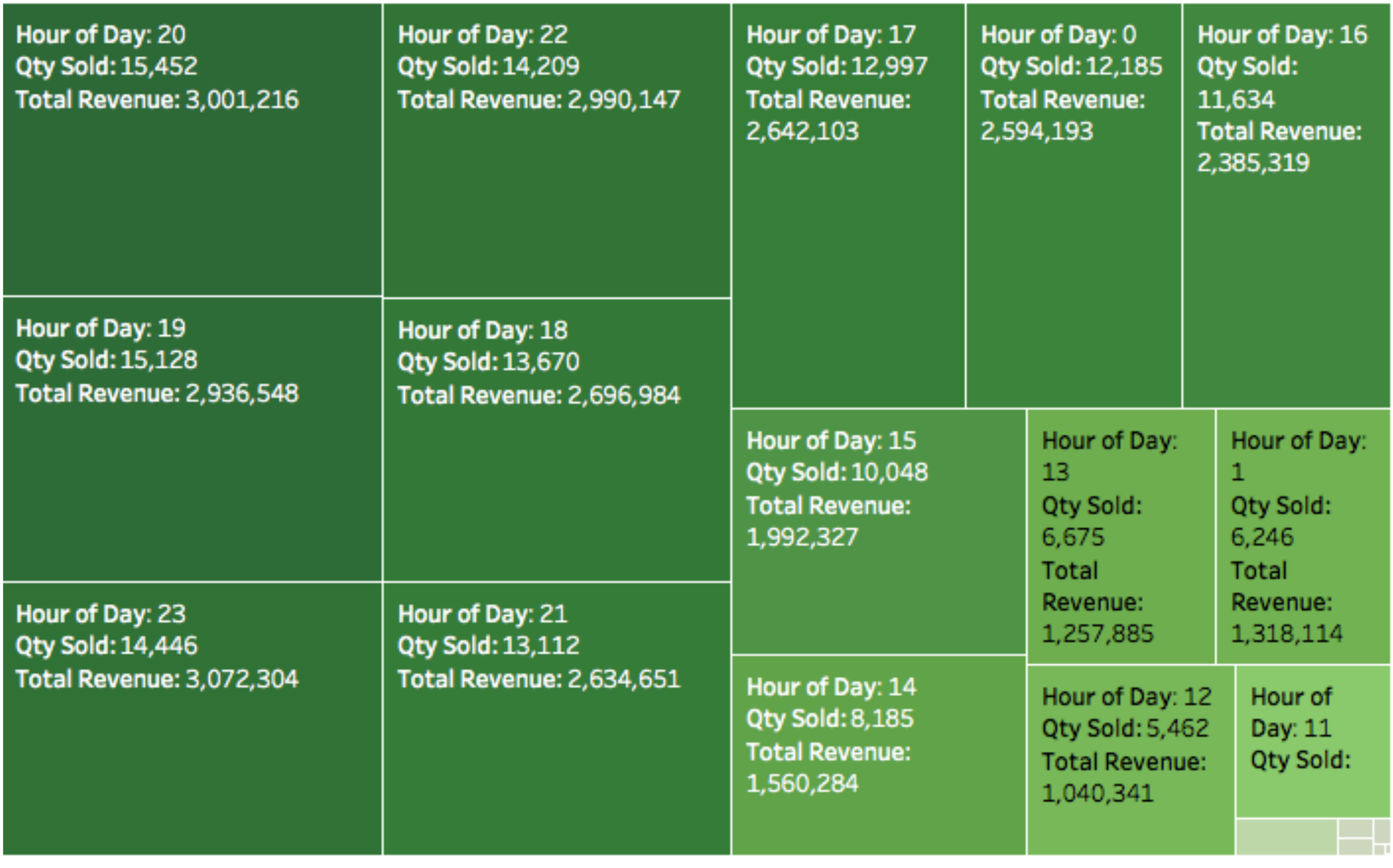


Products With NO Revenue

- BENSON & HEDGES GOLD
- CLASSIC MENTHOL RUSH
- GOLD FLAKE KING BLUE
- GOLD FLAKE KING RED
- INDIA KINGS OCEAN BLUE
- MOTHERS DAY SPL
- MUGS CHIP PRINT
- MUGS _ DOT
- PLAYING CARDS

Exploratory Data Analysis

BUSIEST HOURS : Sales wise

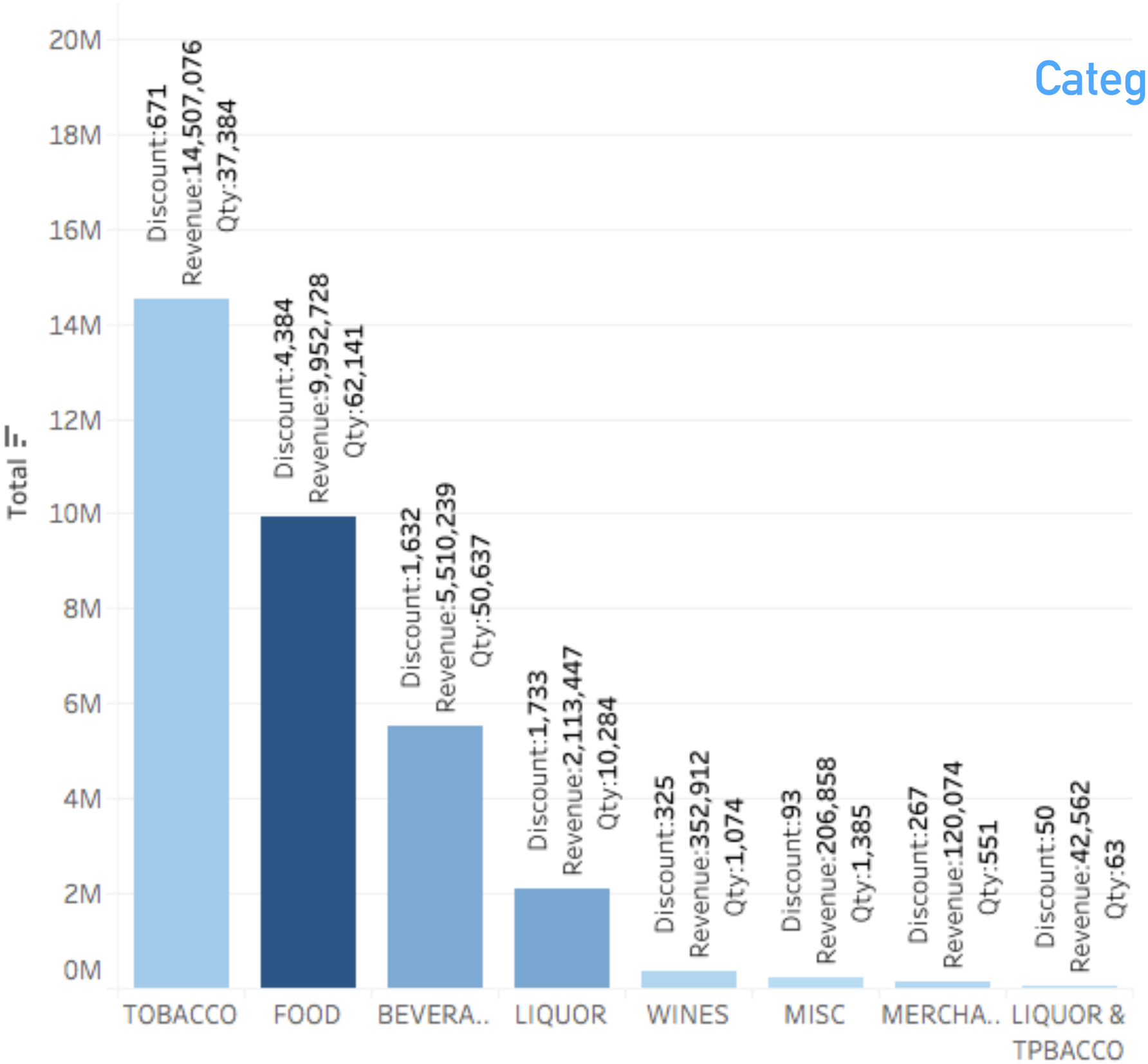


Exploratory Data Analysis

BUSIEST WEEKDAYS

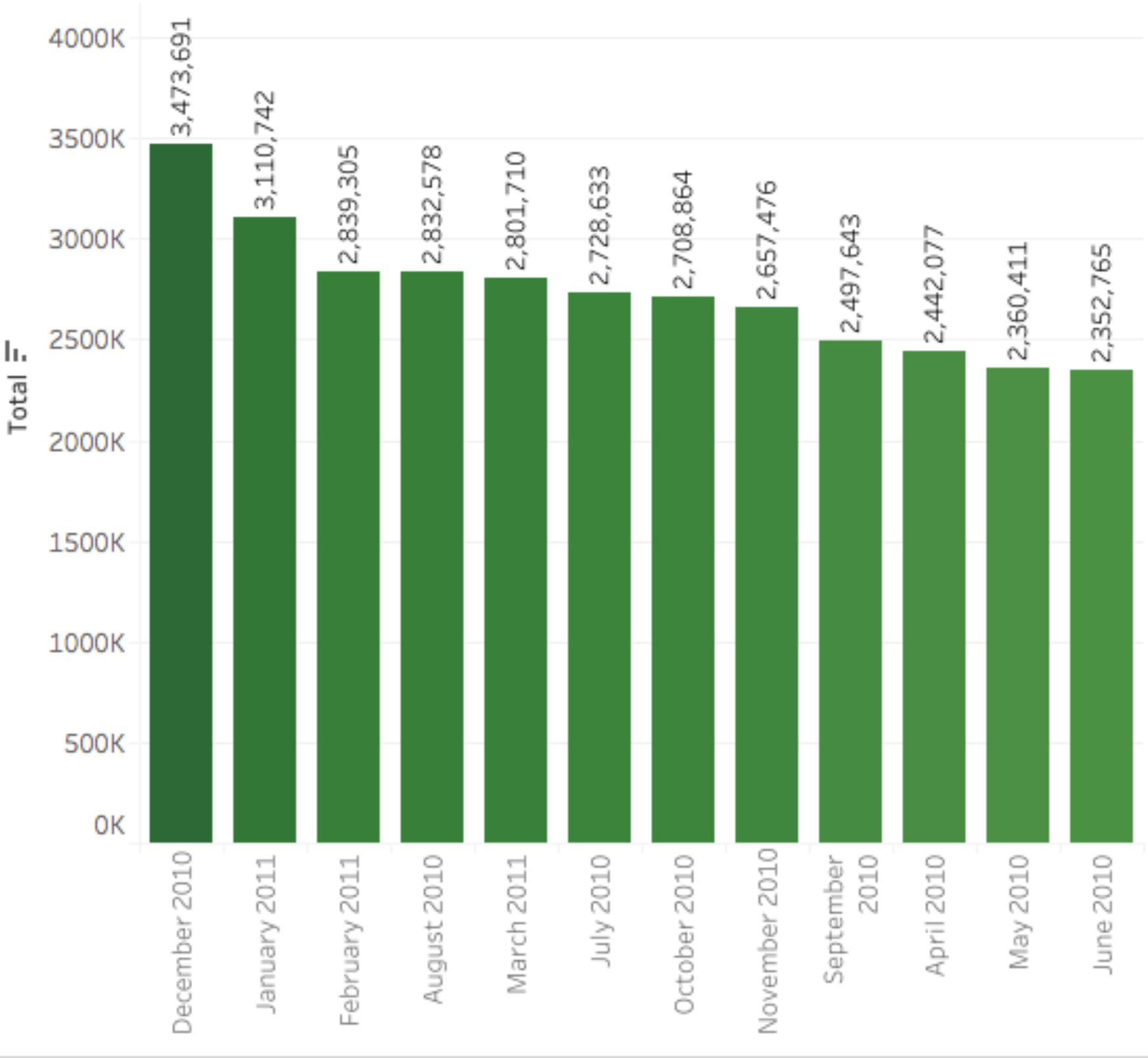
Exploratory Data Analysis

Category wise Discount Vs Total Revenue

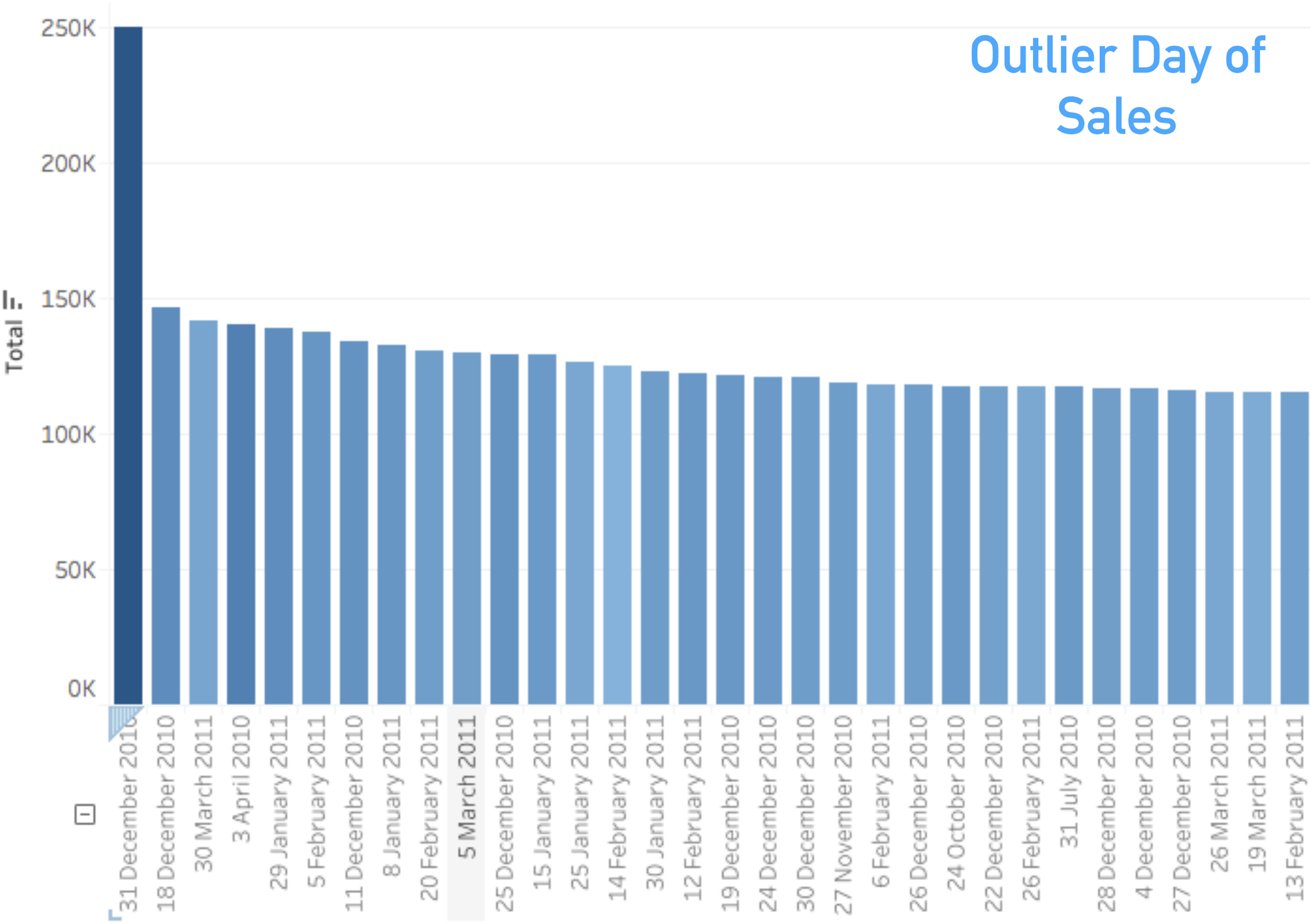


Exploratory Data Analysis

Months with Max Revenue

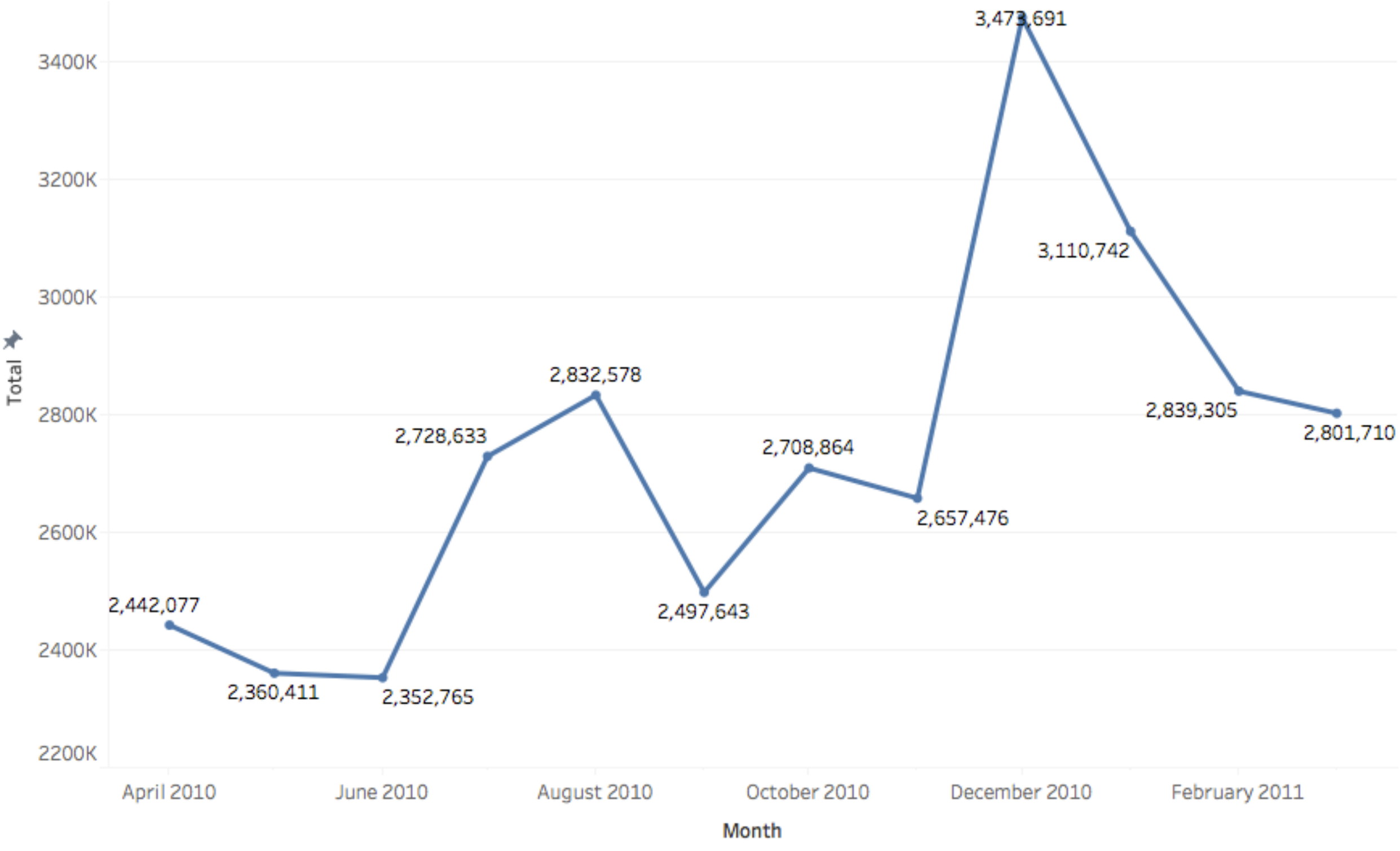


Exploratory Data Analysis



Exploratory Data Analysis

Sales Trend : Month-wise



Exploratory Data Analysis

Observations & Conclusions from the EDA:

- Customers visit the cafe most on Saturday followed by Sunday and Friday
- Customers visit mostly in the evening's specifically between 5:00 PM to Midnight, 8:00PM bringing the maximum sales
    - Interesting thing to note is inspite of 8:00PM and 7:00PM recording the most amount of sales, its the hours 11:00 PM and 12:00PM which brings equal or more revenue despite relatively lower sales
- There are 9 menu items which can immediately removed from the menu as they have recorded absolutely zero sales over the 10 month period.
- Most Sold categories in terms of quantity is FOOD, BEVERAGES, TOBACCO, LIQUOR
- Despite previous point, Tobacco category records the maximum amount of revenue, where as food category despite giving out maximum discounts records less revenue than Tobacco
- Sales trend show sales pick up in Summer i.e July-August and in Winter i.e. December & Jan
- December 2010 and Jan 2011 are the top 2 revenue grossing months
- 31st December 2010 is the single day which grossed almost 40% more revenue than the next highest revenue day

Menu Analysis

Association Rule Mining

- Association rules are written in “IF-THEN” format
- IF - LHS & THEN - RHS. Here, if a is purchased then b is also purchased
- Item sets illustrated in the following manner: a,b, count
- An Item set is a mathematical set of products in the basket.
- Apriori algorithm envisions an iterative approach where it uses k-Item sets to search for (k+1)-Item sets. The first 1-Item sets are found by gathering the count of each item in the set. Then the 1-Item sets are used to find 2-Item sets and so on until no more k-Item sets can be explored; when all our items land up in one final observation. One exploration takes one scan of the complete dataset.
- A rule with high lift but little support might be a coincidence or a flavor of some particular customer, and that a rule with high lift but little confidence means that the occurrence of Y is still too unlikely comparing to other items, therefore hard to create value in business.

Menu Analysis

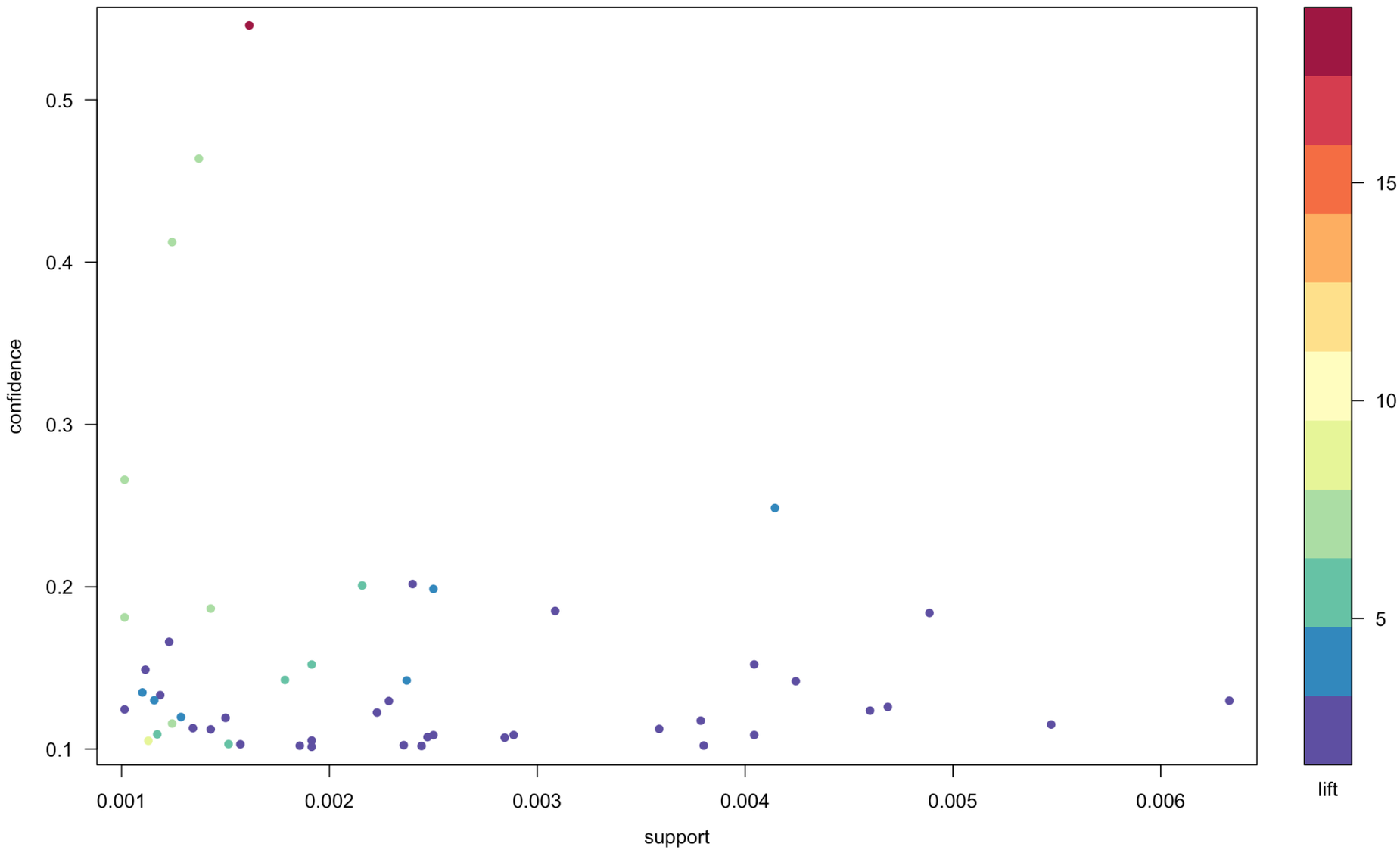
Association Rule Mining

- With these in mind, our way of creating rules would be setting a minimum support and confidence, then sorting by lift to find the most important rules among all.
- The rules with confidence of 1 imply that, whenever the LHS item was purchased, the RHS item was also purchased 100% of the time.
- A rule with a lift of 2 imply that, the items in LHS and RHS are 2 times more likely to be purchased together compared to the purchases when they are assumed to be unrelated.

- I ran the association Rule Mining algorithm on the transactional data and came up with 53 rules.
  - Parameters used were:
    - Support: 0.001 (1% of the transactions)
    - Confidence: 0.1

Menu Analysis

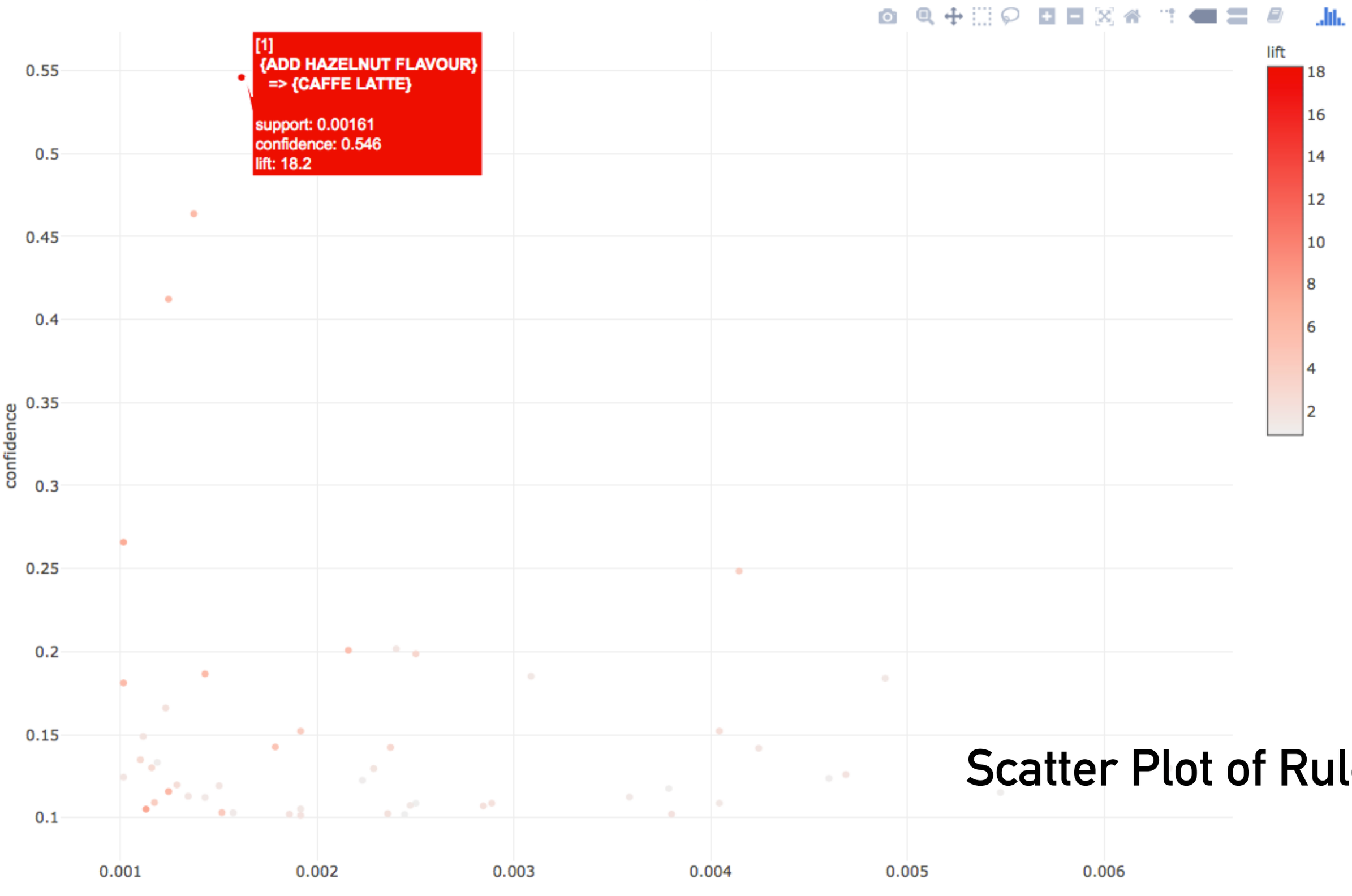
Results from Association Rule Mining



Scatter Plot of Rules

Menu Analysis

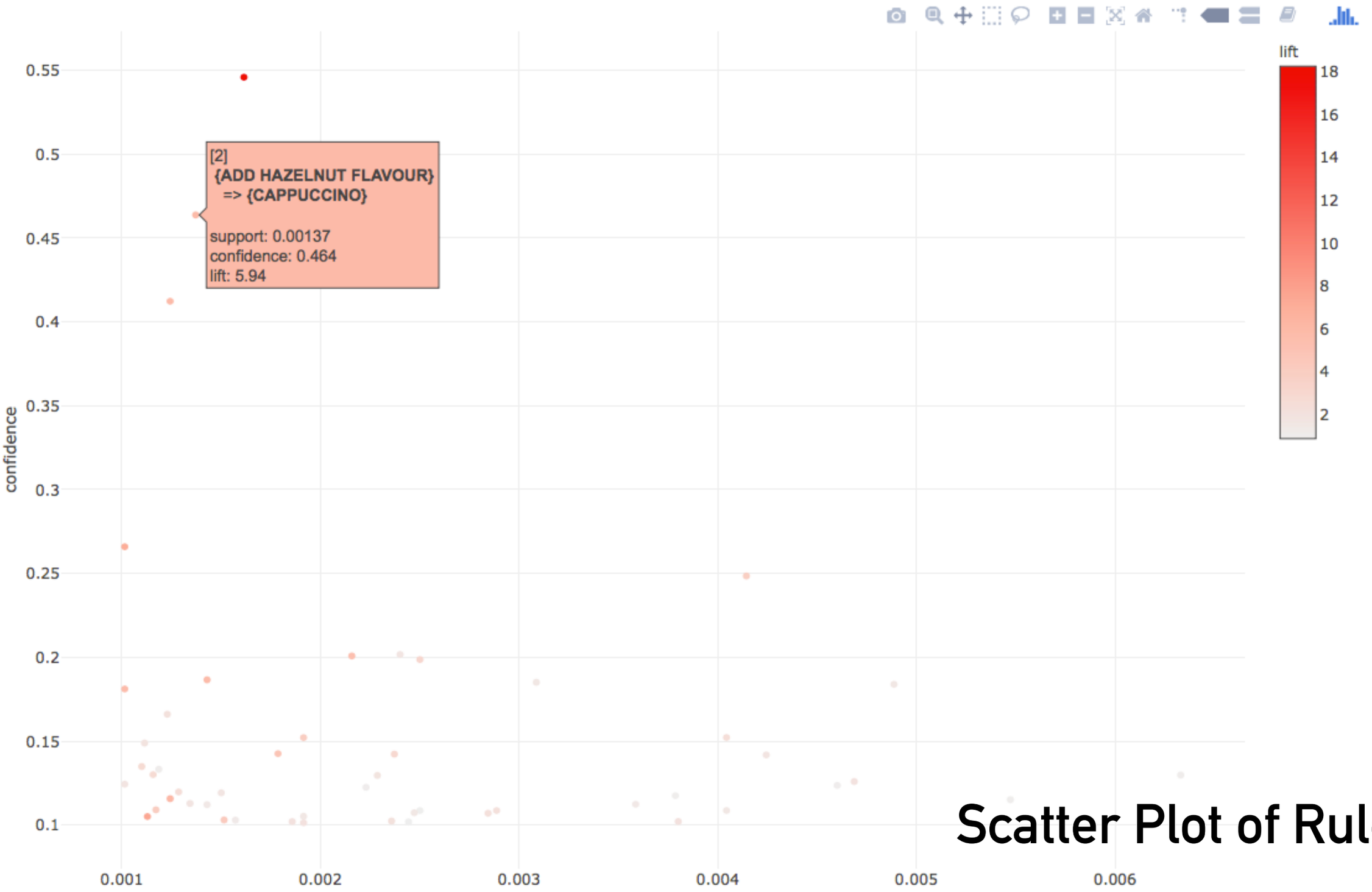
Results from Association Rule Mining



Scatter Plot of Rules

Menu Analysis

Results from Association Rule Mining



Scatter Plot of Rules

Menu Analysis

Results from Association Rule Mining

High Confidence Rules:

```
##     lhs                              rhs                     support     confidence lift      count
## [1] {ADD HAZELNUT FLAVOUR}        => {CAFFE LATTE}           0.001614701 0.5458937  18.235195 113
## [2] {ADD HAZELNUT FLAVOUR}        => {CAPPUCCINO}            0.001371781 0.4637681   5.935519  96
## [3] {VANILLA ICECREAM}            => {GREAT LAKES SHAKE}     0.001243177 0.4123223   5.923863  87
## [4] {FRENCH FRIES}                => {B.M.T. PANINI}         0.001014547 0.2659176   7.143741  71
## [5] {RED BULL 2+1}                => {SAMBUCA}               0.004143923 0.2485004   3.931846 290
## [6] {QUA  MINERAL WATER(500ML)} => {NIRVANA HOOKAH SINGLE} 0.002400617 0.2016807   1.682443 168
```

High Lift Rules:

```
##     lhs                          rhs                               support     confidence lift      count
## [1] {ADD HAZELNUT FLAVOUR} => {CAFFE LATTE}                    0.001614701 0.5458937  18.235195 113
## [2] {ADD FRIES}            => {SATAY CHICKEN PANINI}           0.001128862 0.1050532   7.478975  79
## [3] {FRENCH FRIES}         => {B.M.T. PANINI}                  0.001014547 0.2659176   7.143741  71
## [4] {ADD FRIES}            => {COUNTRY ROAST CHICKEN PANINI} 0.001243177 0.1156915   6.124298  87
## [5] {ADD HAZELNUT FLAVOUR} => {CAPPUCCINO}                     0.001371781 0.4637681   5.935519  96
## [6] {VANILLA ICECREAM}     => {GREAT LAKES SHAKE}              0.001243177 0.4123223   5.923863  87
```

Menu Analysis

Results from Association Rule Mining

Following Items are always purchased together

- Caffe Latte with Hazel Nut Flavour
- Cappuccino with Hazel Nut Flavour
- SAMBUCA with Red Bull
- Great Lakes Shakes with Vanilla ICE Cream

- B.M.T. Panini with French Fries
- SATAY Chicken Panini with French Fries
- Country Roast Chicken Panini with French Fries

- NIRVANA Hookah with Mineral Water Bottle

Menu Analysis

Results from Association Rule Mining

Combo Meal Suggestions:

- B.M.T. Panini with Caffe Latte or Cappuccino with French Fries
- B.M.T. Panini with Sambuca & with Poutine
- SATAY Chicken Panini with French Fries & with SAMBUCA with Red Bull
- Country Roast Chicken Panini with French Fries with Great Lakes Shakes with Vanilla ICE Cream
- SATAY Chicken Panini with Caffe Latte or Cappuccino with French Fries
- Nirvana Hookah with Poutine & B.M.T. Panini & with Mineral Water
- Nirvana Hookah with Poutine & B.M.T. Panini with Red Bull or Shake
- Bun Maska & Chai
- Veg Club Wrap with Poutine with Fries

Price Analysis

- For the sake of simplicity and narrow down products, we looked at only those menu items whose rates changed more than twice in the dataset

Positive Impact of Price Change

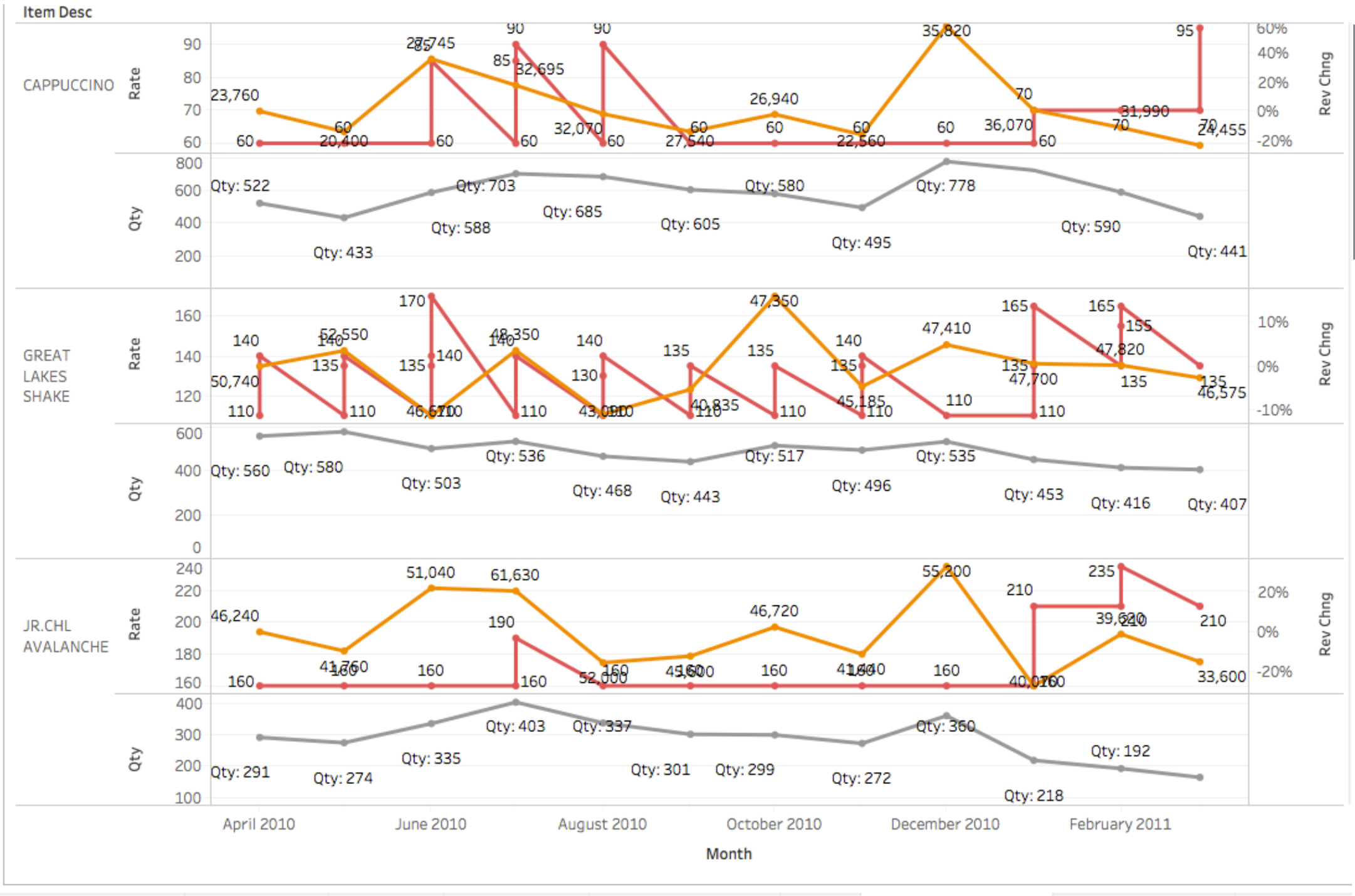
- Following Items increased in sales despite of rates increase:
  - Asian Wok Tossed Salad
  - MAGGIE NDL ARRABIATA
  - Lemon Infused Char Grilled Veg

Negative Impact of Price Change

- Almost all menu items saw a negative impact of the price change,
  Products like Cappuccino, Cafe Latte, Shakes, changed prices almost on a monthly basis, accordingly revenue was also directly affected.
  Though, maybe due the seasonality or some other factor, price change affected some items positively while other times it affected negatively

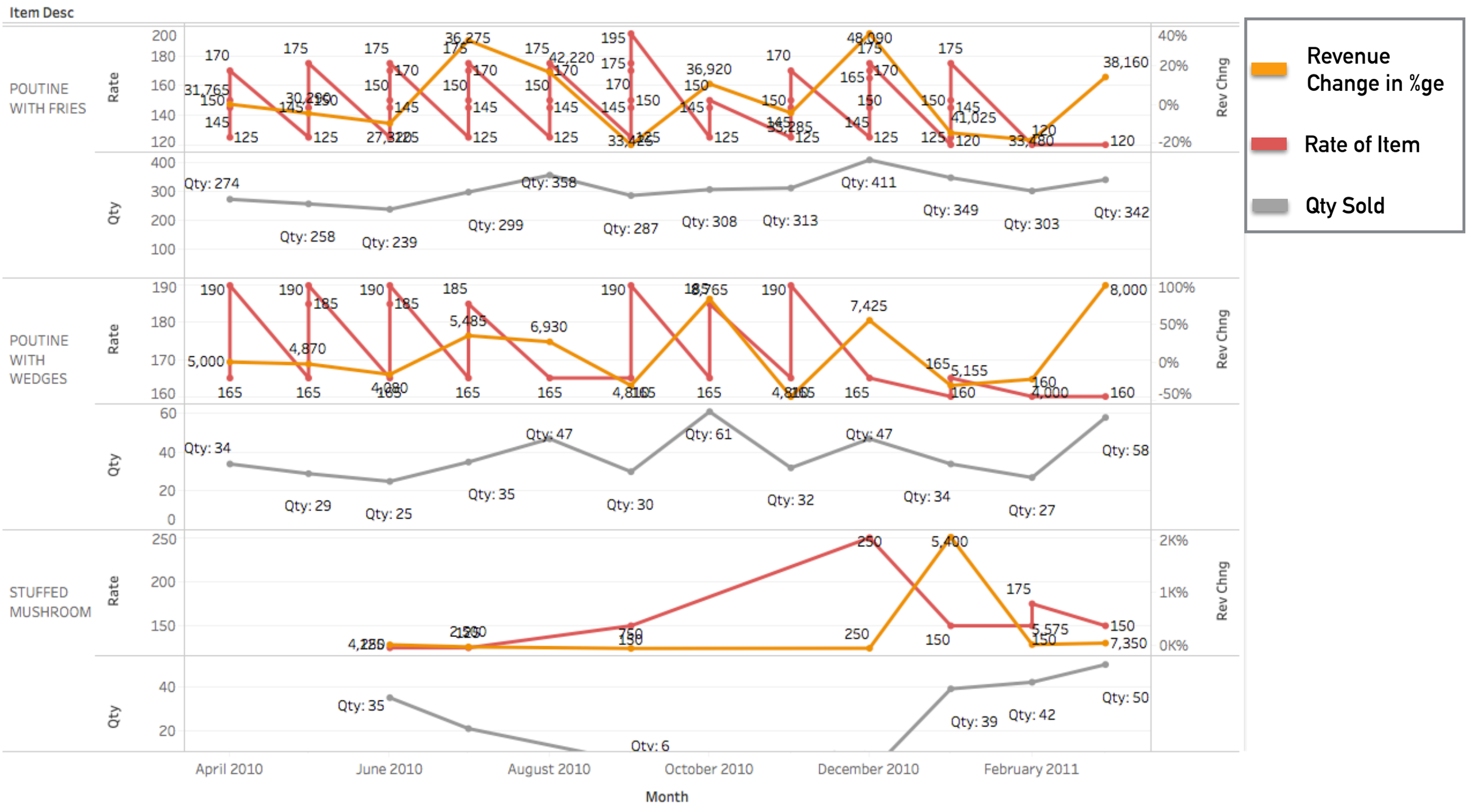
Price Analysis

Products with Rate Changes and Qty + Revenue over months



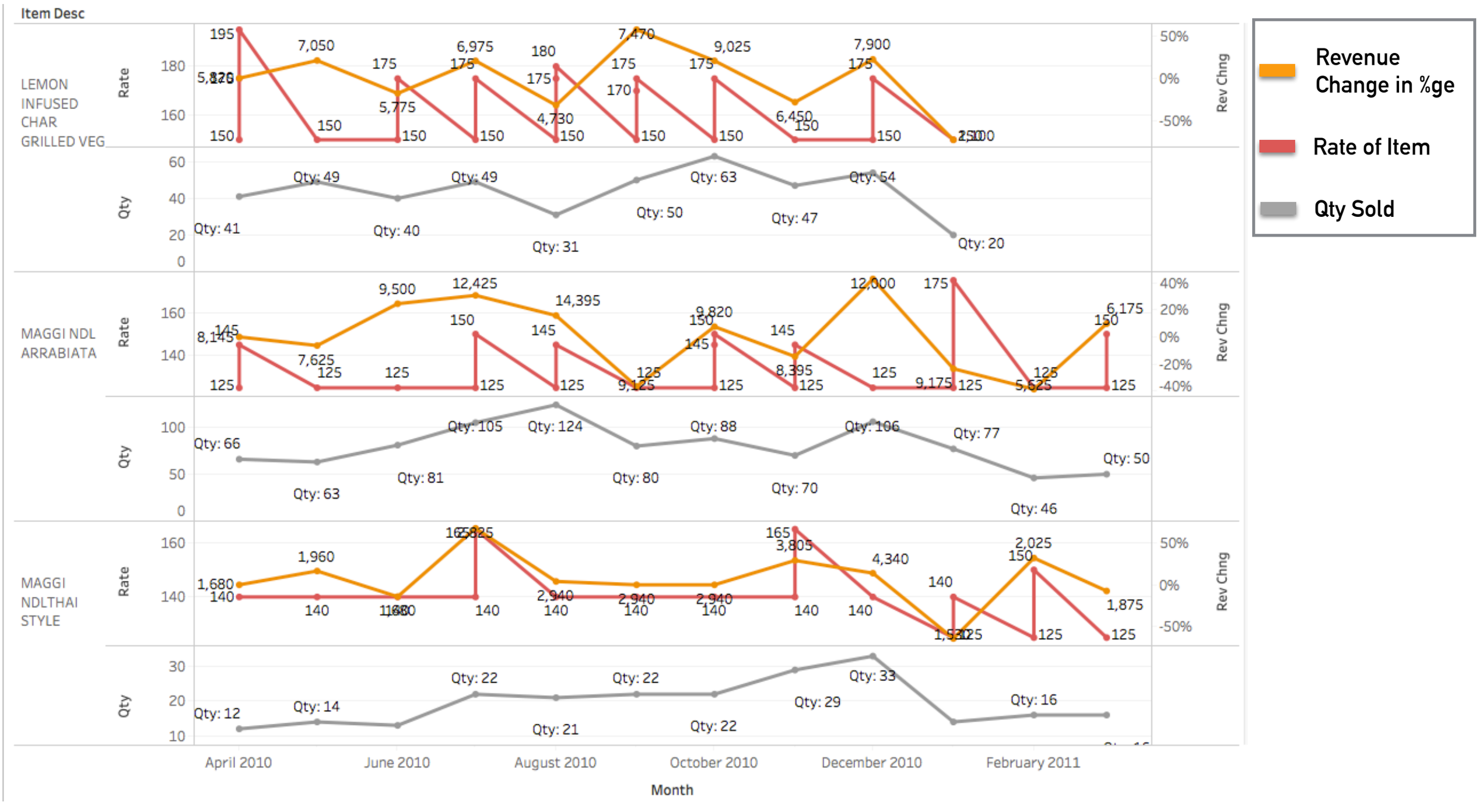
Price Analysis

Products with Rate Changes and Qty + Revenue over months



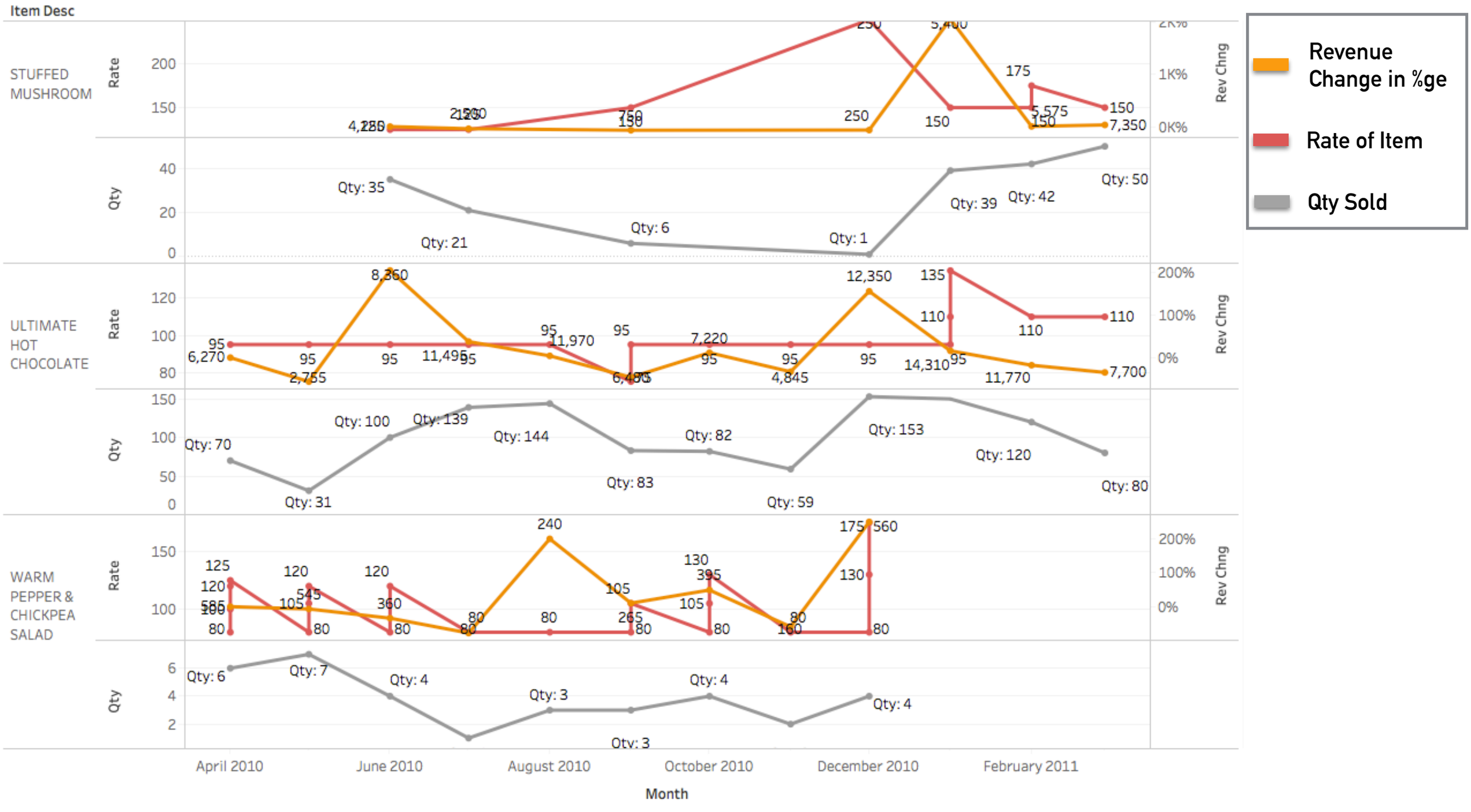
Price Analysis

Products with Rate Changes and Qty + Revenue over months



Price Analysis

Products with Rate Changes and Qty + Revenue over months



THANK YOU