

# *Café – Retail Analytics*

MRA Assignment

Submitted by

- Mohamed Yusuf



# Exploratory Data Analysis

## Executive Summary of data

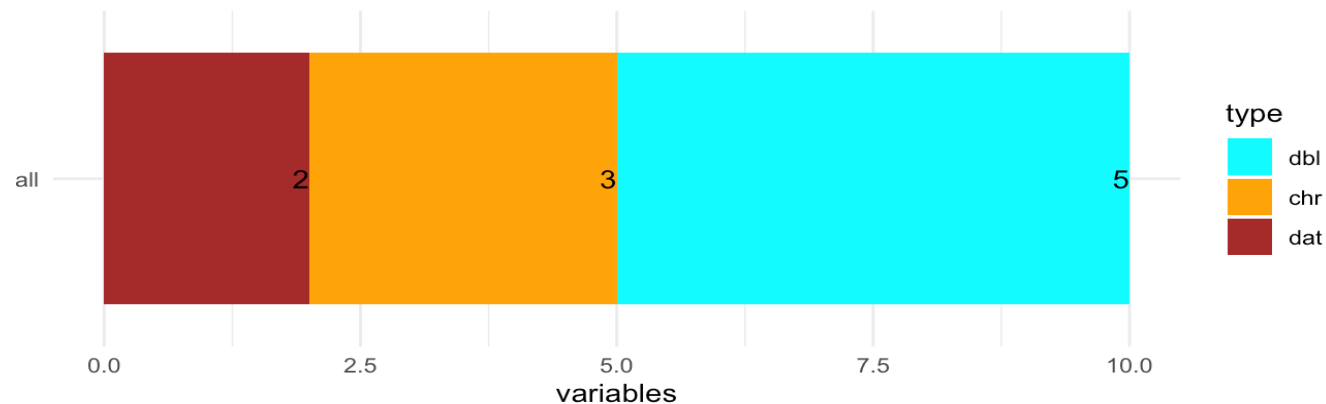
S.No	Description	Variable Name	Type
1	Date of Transaction	Date	Date
2	Bill reference number of transaction	Bill Number	Character
3	Description of Item	Item Desc	Character
4	Time the item is sold	Time	Time
5	No of Units the item is sold	Quantity	Numeric
6	Price of the item	Rate	Numeric
7	Tax charges for the item	Tax	Numeric
8	Discount offered for the item	Discount	Numeric
9	Total transaction value of the item	Total	Numeric
10	Category in which item is sold	Category	Character

With the given POS dataset we have 65535 records with 10 variables. No missing values found in the data

# An Overview of Data

## Describe Data

10 variables  
with 65.5k observations

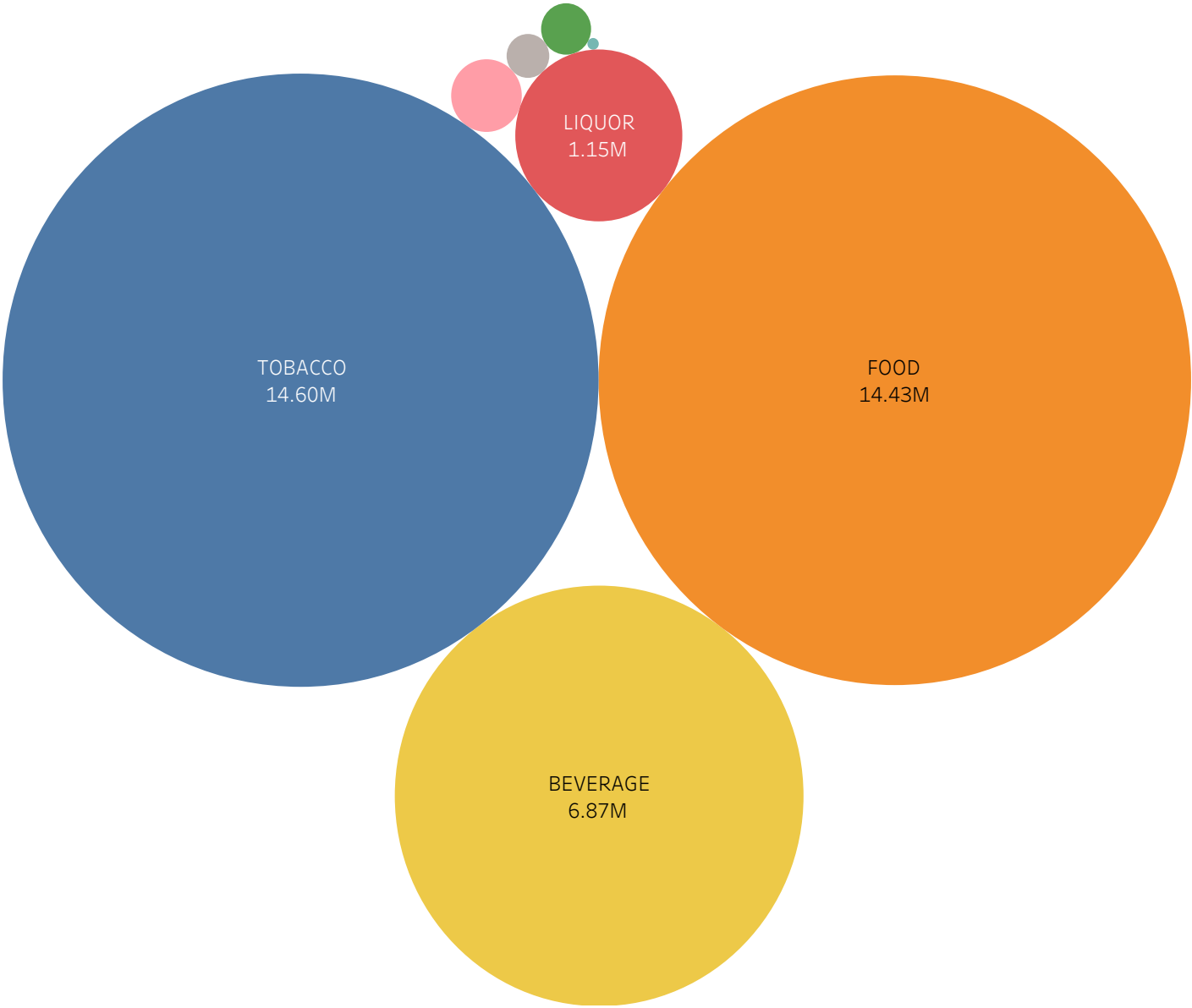


```
## 65 535 (65.5k) observations with 10 variables
## 0 variables containing missings (NA)
## 0 variables with no variance
```

variable	type	na	na_pct	unique	min	mean	max
Date	dat	0	0	318	NA	NA	NA
Bill Number	chr	0	0	34527	NA	NA	NA
Item Desc	chr	0	0	453	NA	NA	NA
Time	dat	0	0	23628	NA	NA	NA
Quantity	dbl	0	0	16	1.00	1.12	30.00
Rate	dbl	0	0	117	1.00	162.60	2100.00
Tax	dbl	0	0	320	0.00	49.42	914.06
Discount	dbl	0	0	62	0.00	0.10	396.00
Total	dbl	0	0	322	1.26	225.98	3839.06
Category	chr	0	0	8	NA	NA	NA

Category Sales

Tobacco and Food category contributes 77% of total sales

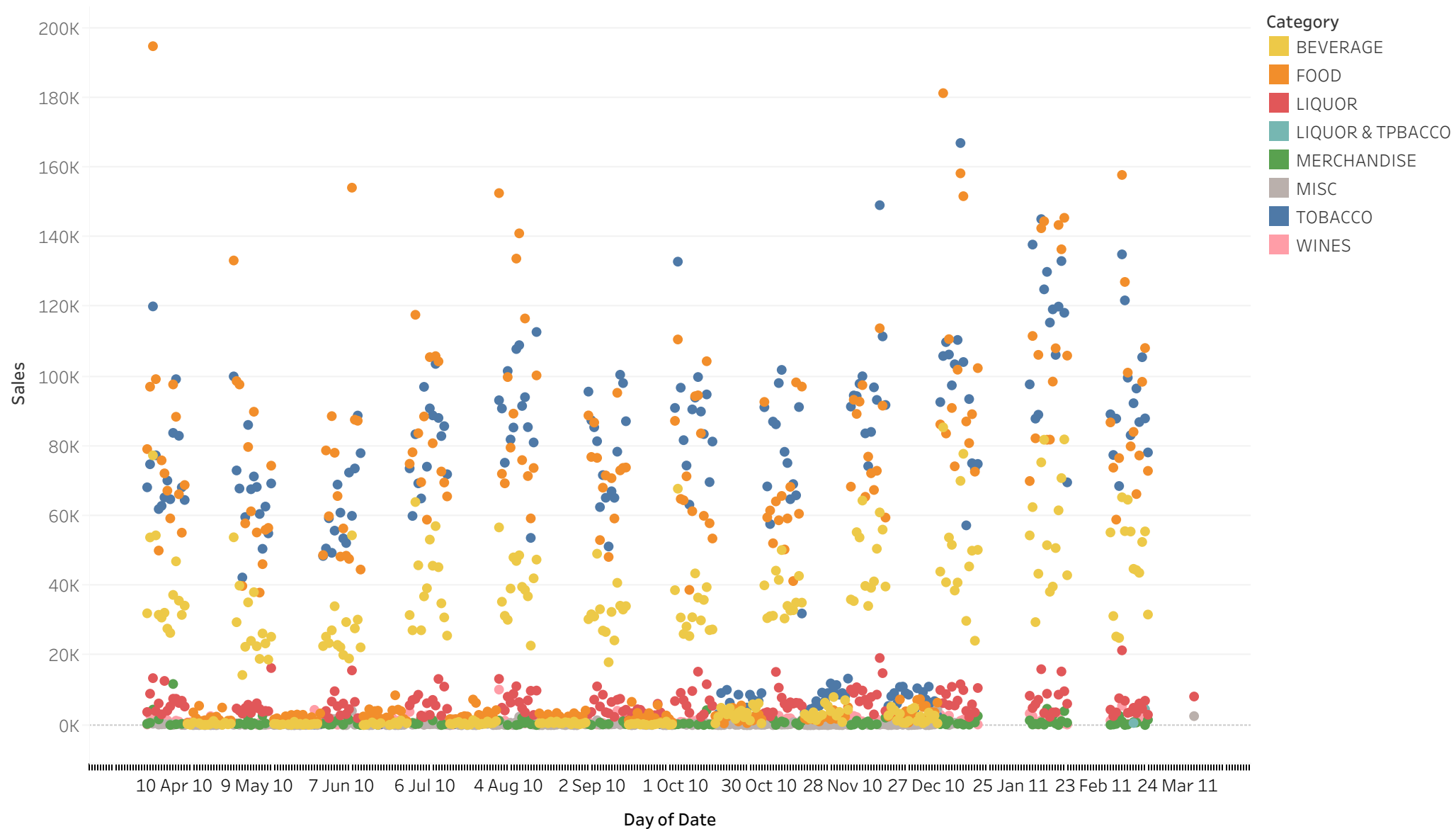


- Category
- TOBACCO
  - FOOD
  - BEVERAGE
  - LIQUOR
  - WINES
  - MISC
  - MERCHANDISE
  - LIQUOR & TPBACCO

Category and sum of Sales. Color shows details about Category. Size shows % of Total Sales. The marks are labeled by Category and sum of Sales.

## Sales Trend - Category

No sale transaction has been found during 3<sup>rd</sup> and 4<sup>th</sup> week of January, February and March of 2011.



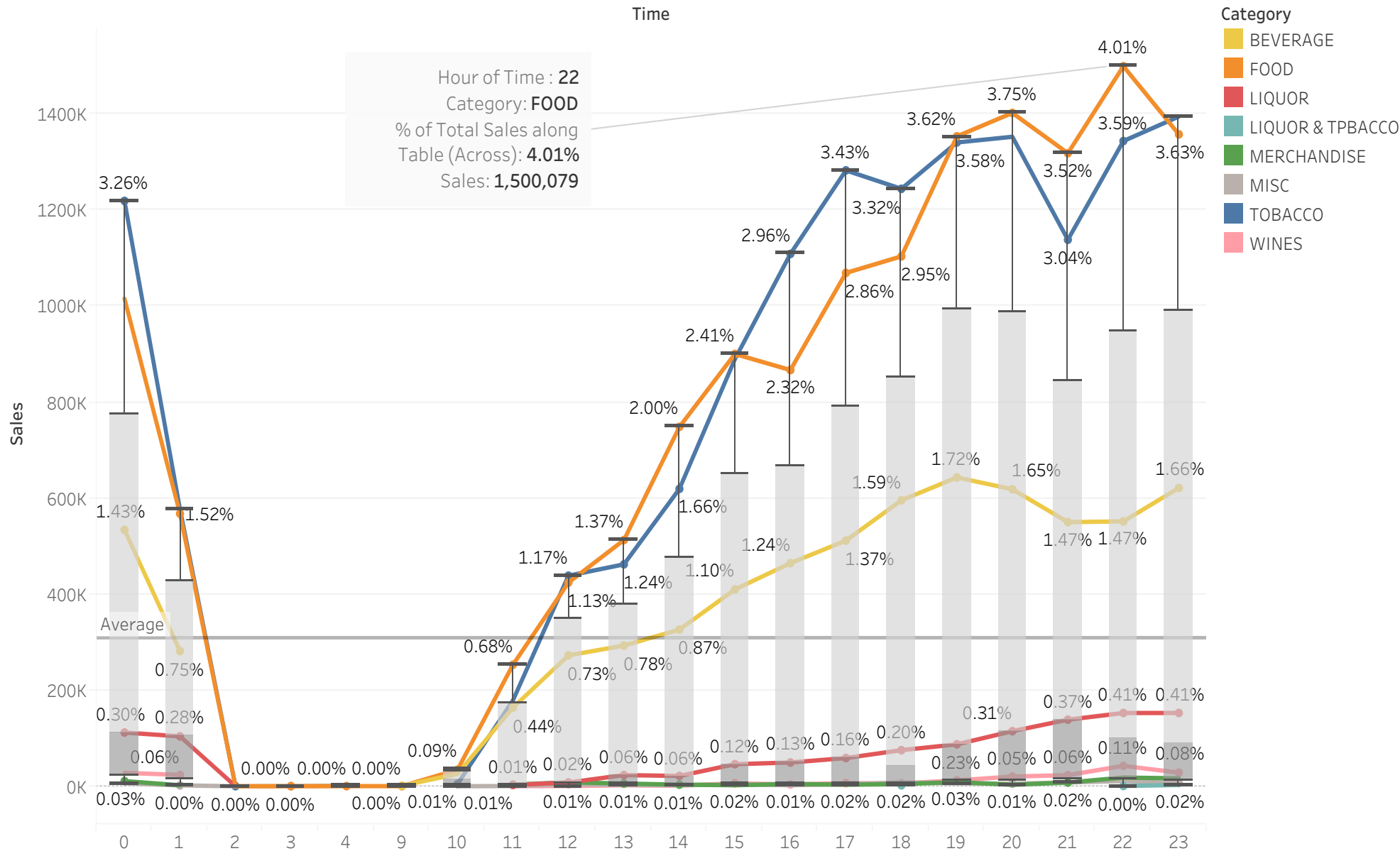
The plot of sum of Sales for Date Day. Color shows details about Category.

# Hourly Sales Pattern

## Hourly Sales pattern

- We could notice clearly that customer walk-ins start from 10 Am onwards. Sales peaks between late evening and night hours on most of days in a week.
- Sales drops and increase at 21 hours for all other category except for Alcohol.
- There is gradual increase in Alcohol sales till late night.
- Food sales overtakes the Tobacco sales during Lunch and Dinner time.

Hourly sales Trends



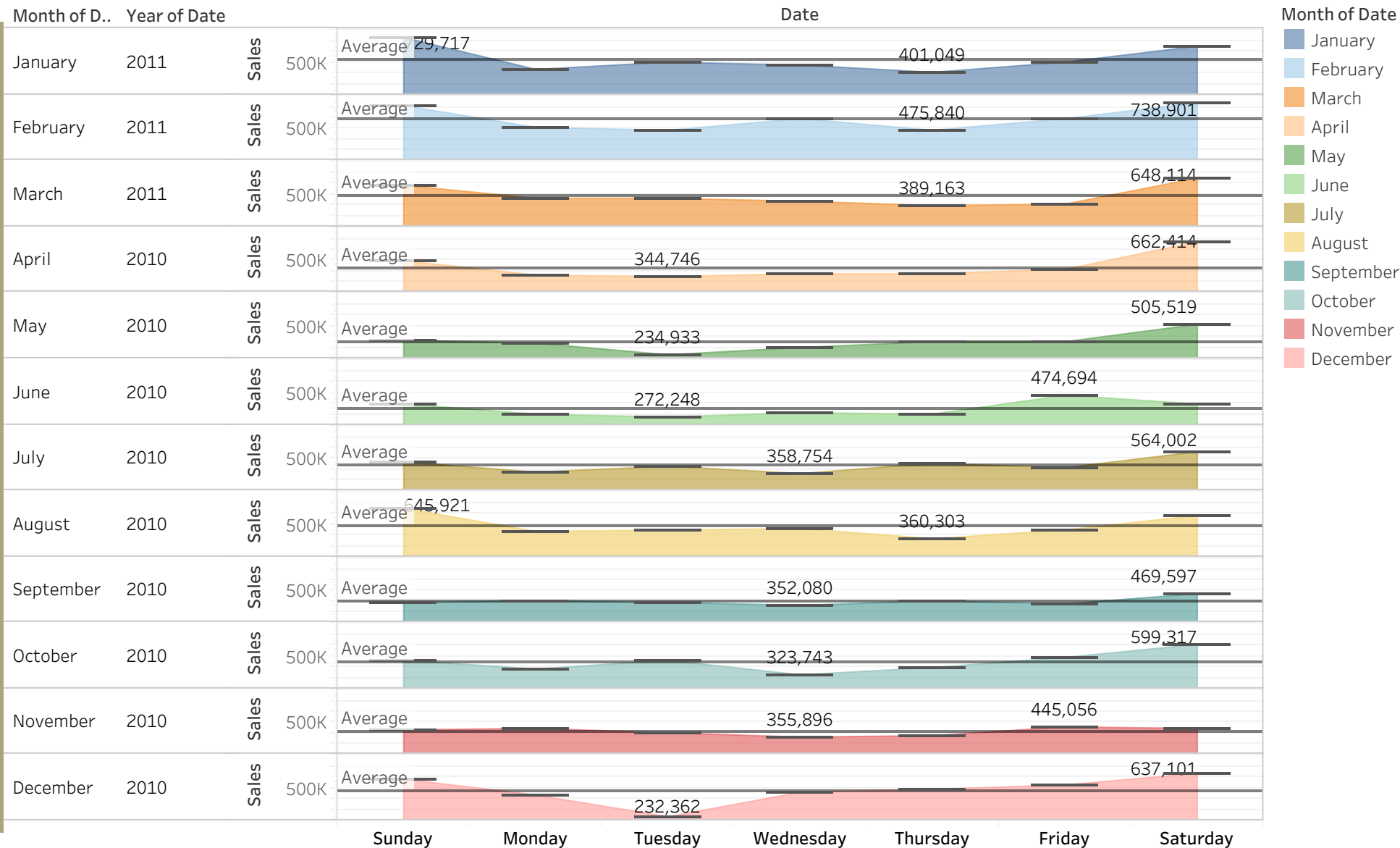
The trend of sum of Sales for Time Hour. Color shows details about Category. The marks are labeled by % of Total Sales.

Hourly  
Sales  
Pattern

# Days of Week Trends

Friday, Saturday and Sunday has good sales which is above weekly average.

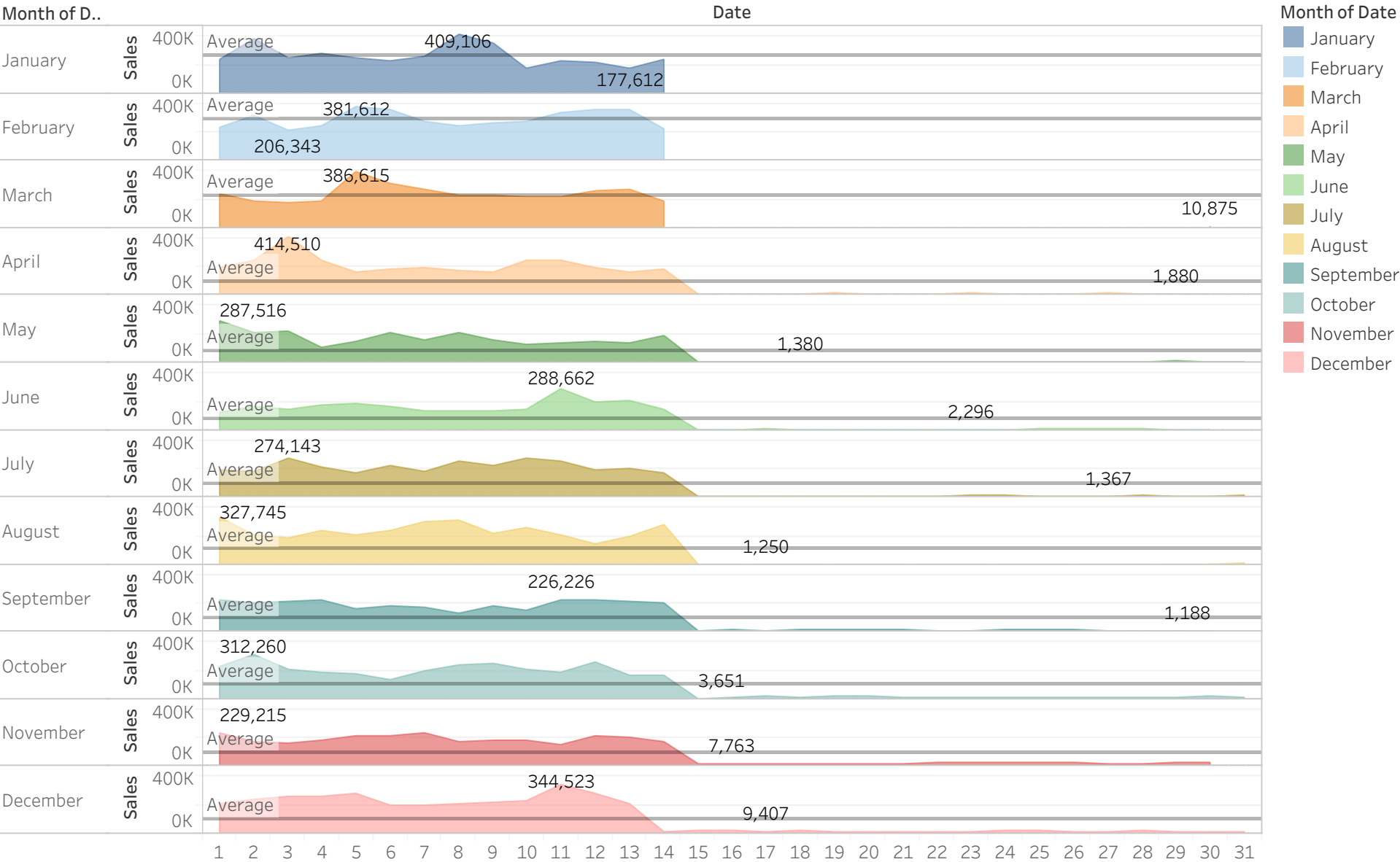
Tuesdays of few months have lowest sale.



Sum of Sales for each Date Weekday broken down by Date Month and Date Year. Color shows details about Date Month. The view is filtered on Date Month, which keeps 12 of 12 members.



# Daywise Month trends

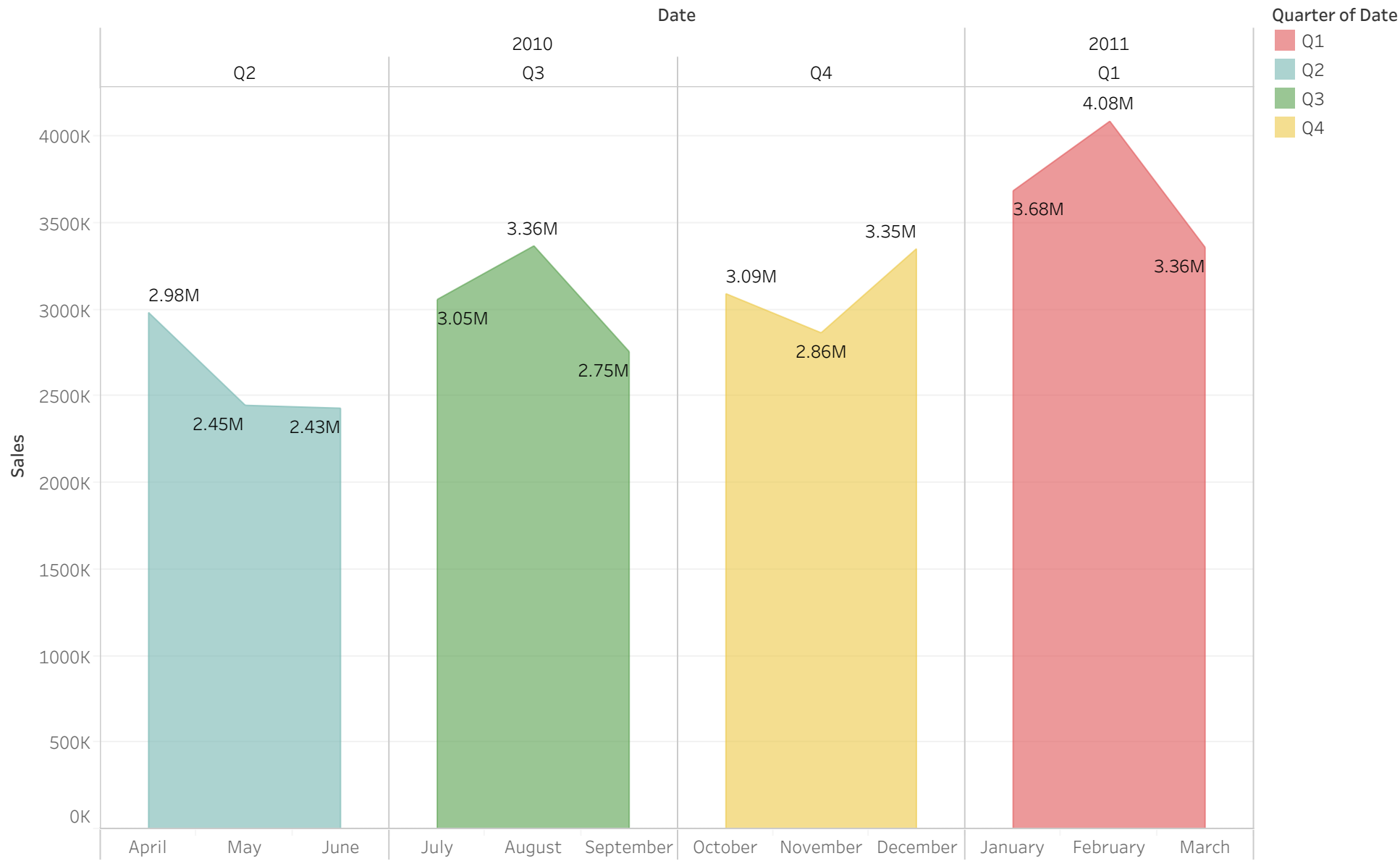


Sum of Sales for each Date Day broken down by Date Month. Color shows details about Date Month. The view is filtered on Date Month, which keeps 12 of 12 members.

There is no prominent sales from 3<sup>rd</sup> week of every month

Quarterly Trends

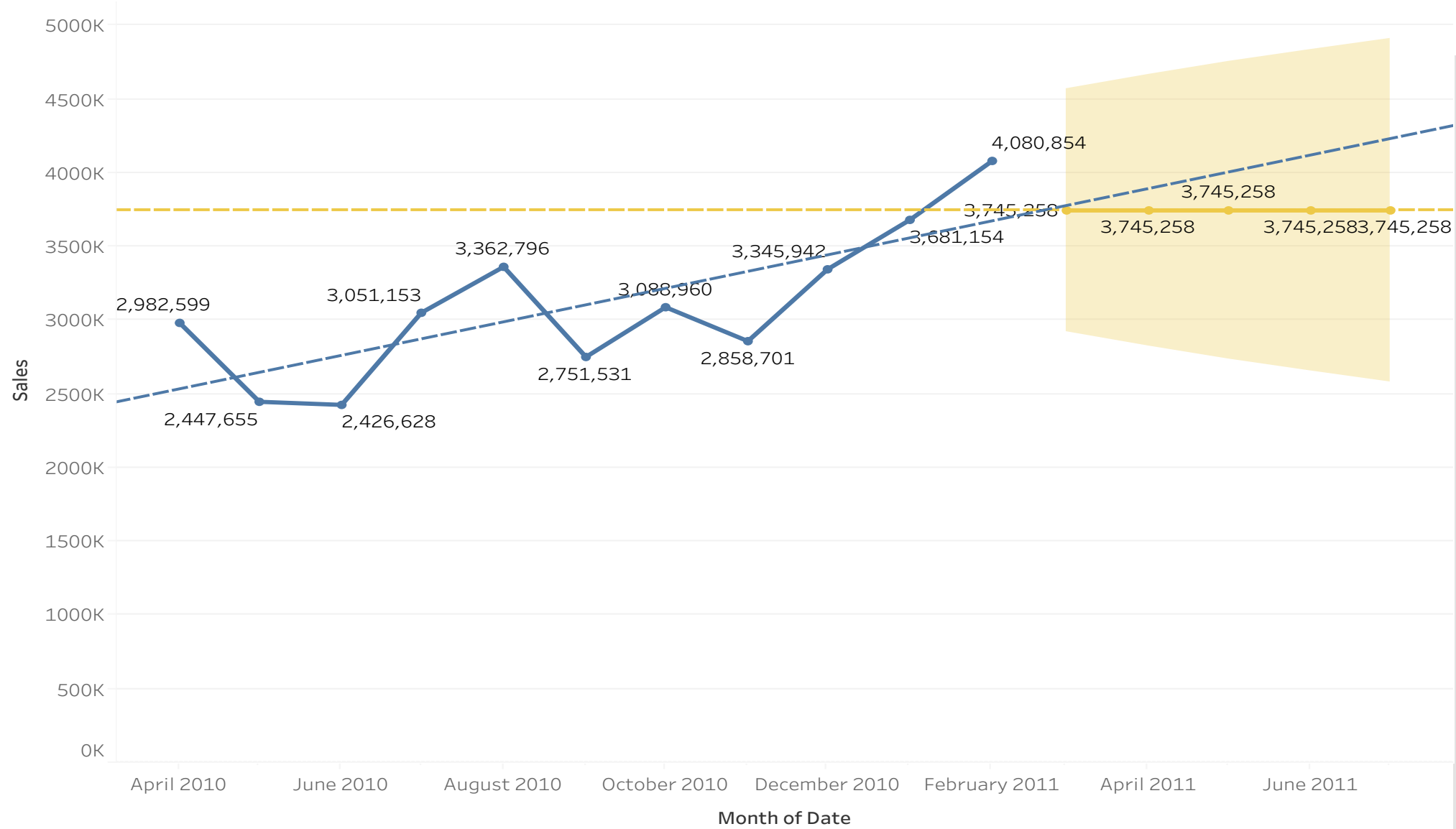
1<sup>st</sup> Quarter is  
the highest  
revenue  
generating  
period



Sum of Sales for each Date Month broken down by Date Year and Date Quarter. Color shows details about Date Quarter. The marks are labeled by sum of Sales.

Sales Forecast

Forecasting the future sales for March to June 2011 based on previous seasonality and trend on sales.

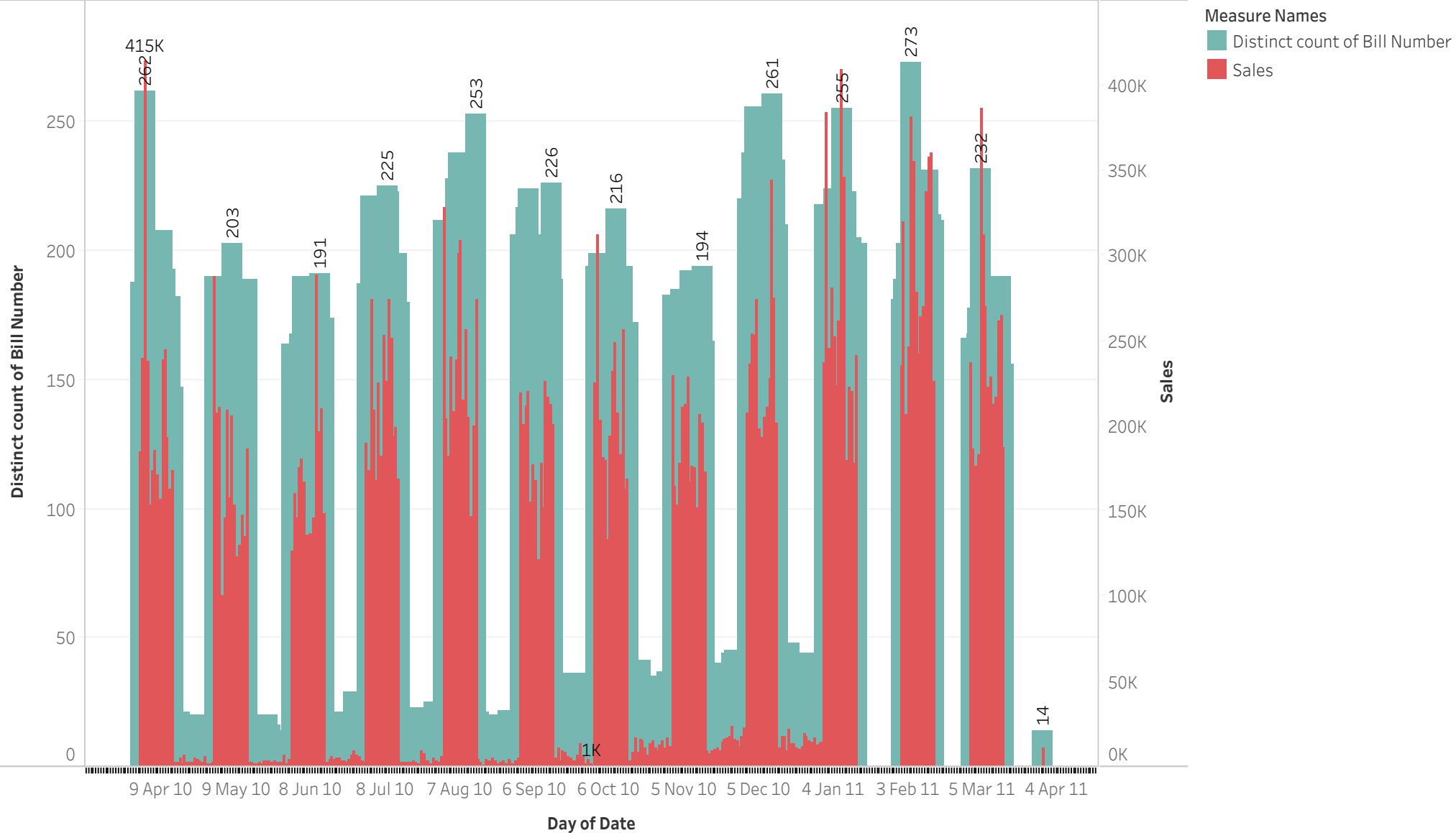


The trend of sum of Sales (actual & forecast) for Date Month. Color shows details about Forecast indicator.

- Forecast indicator**
- Actual
  - Estimate

No of bills and Sales

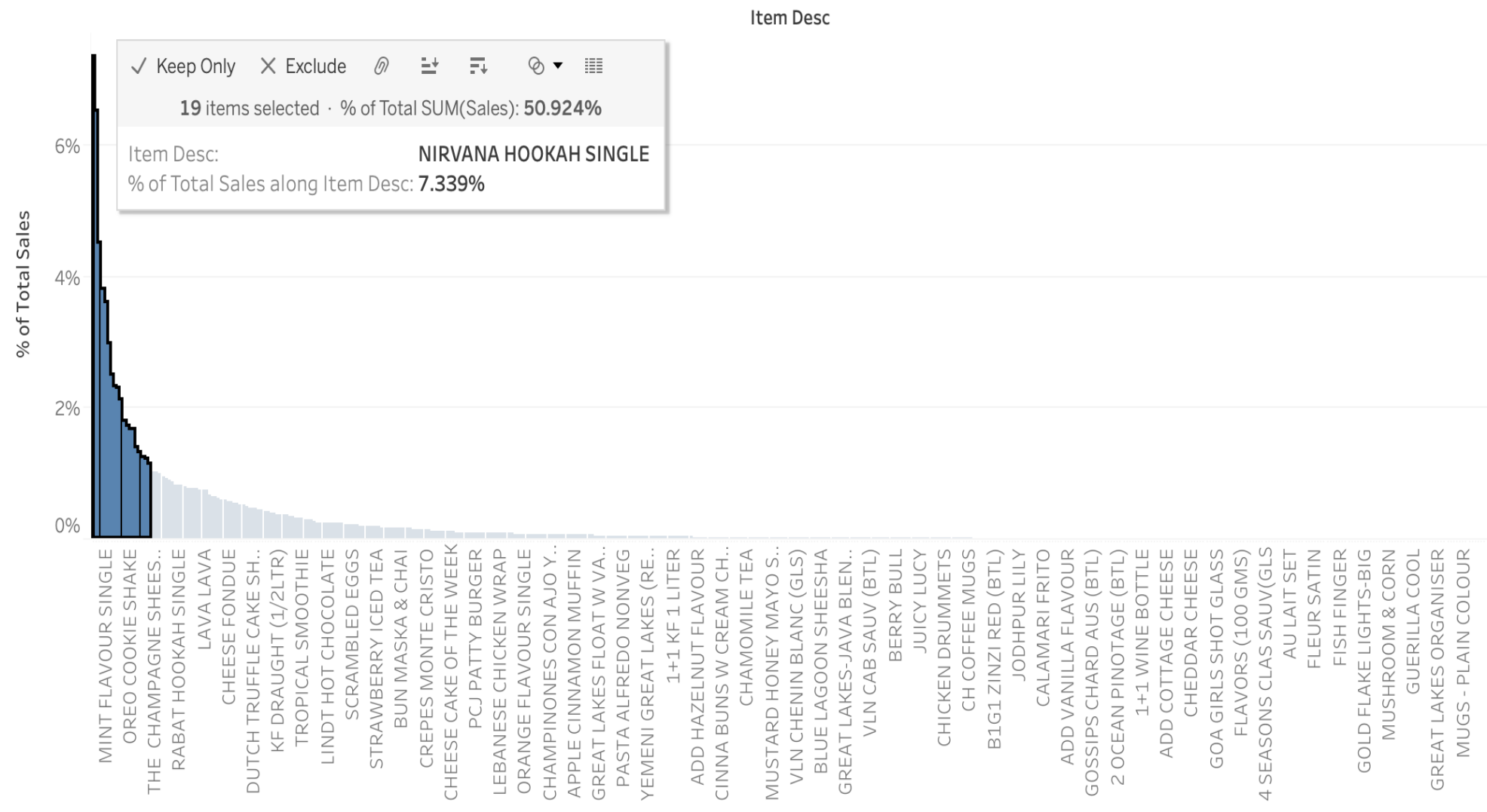
Average sales value per bill is higher during January, February and March 2011



The plots of distinct count of Bill Number and Sales for Date Day. Color shows details about distinct count of Bill Number and Sales. For pane Sum of Sales: The marks are labeled by Sales.

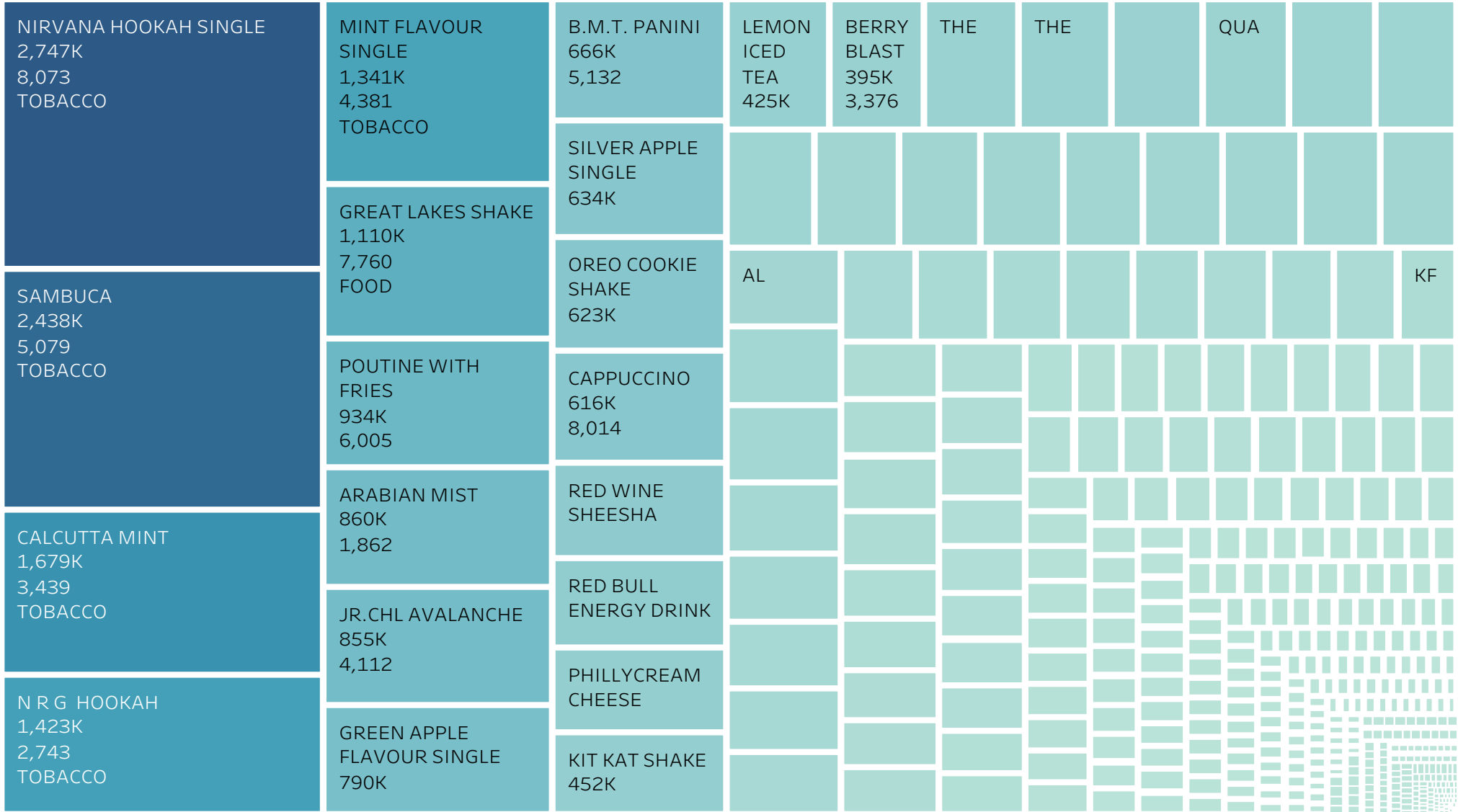
# Contribution % of Each SKUs

19 SKUs out of 453 SKUs contributes 50% of Total Sales



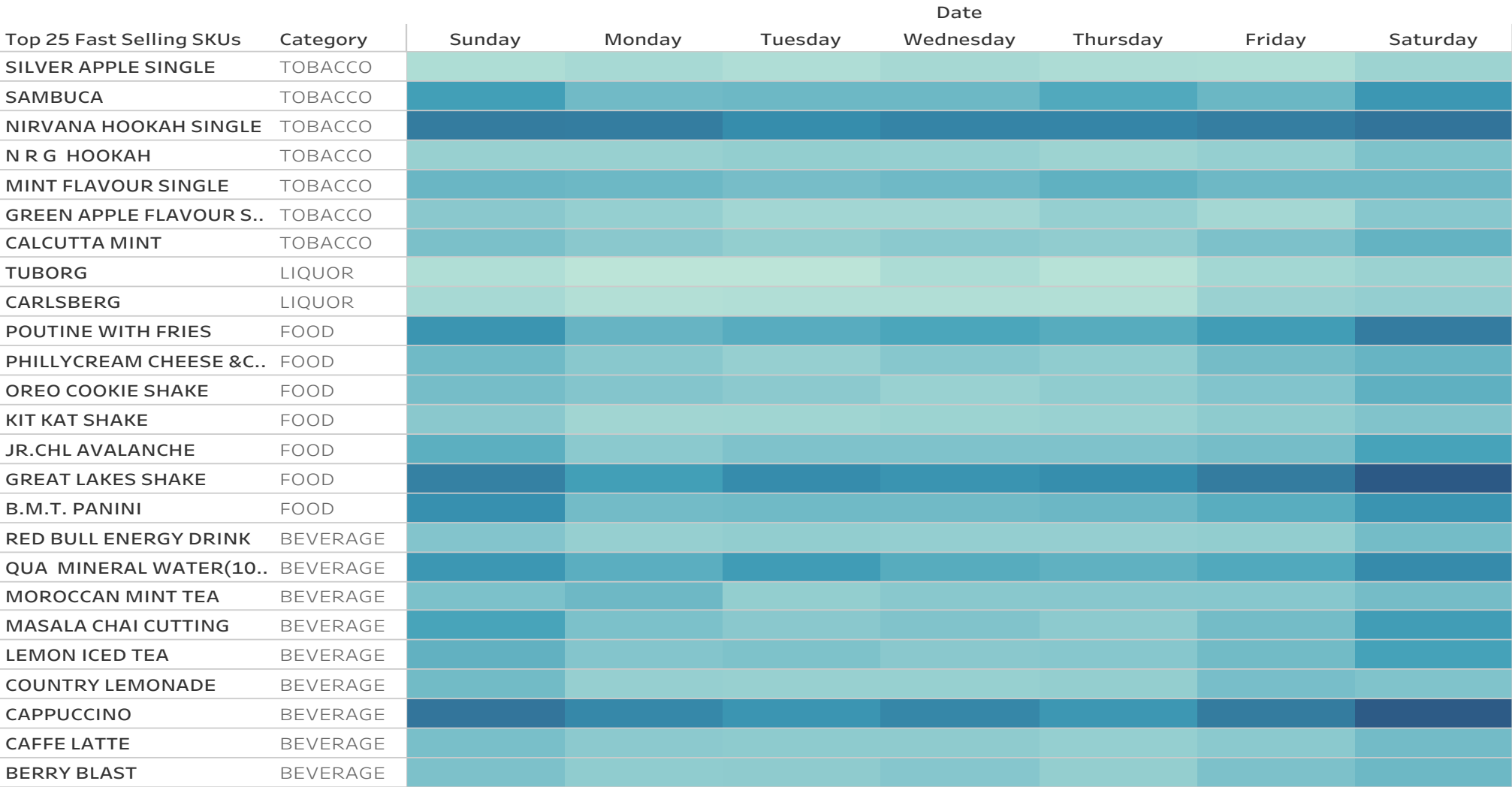
Overview of  
Top Sold  
SKUs

Top Sold SKUs



Item Desc, sum of Sales, sum of Quantity and Category. Color shows sum of Sales. Size shows % of Total Sales. The marks are labeled by Item Desc, sum of Sales, sum of Quantity and Category.

# Top 25 Skus - WeekDays

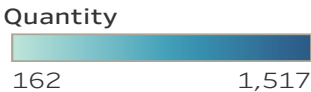


Nirvana Hookah is the fastest selling SKU in Tobacco.

Cappuccino tops in Beverages

Great Lakes Shakes sells best in Food

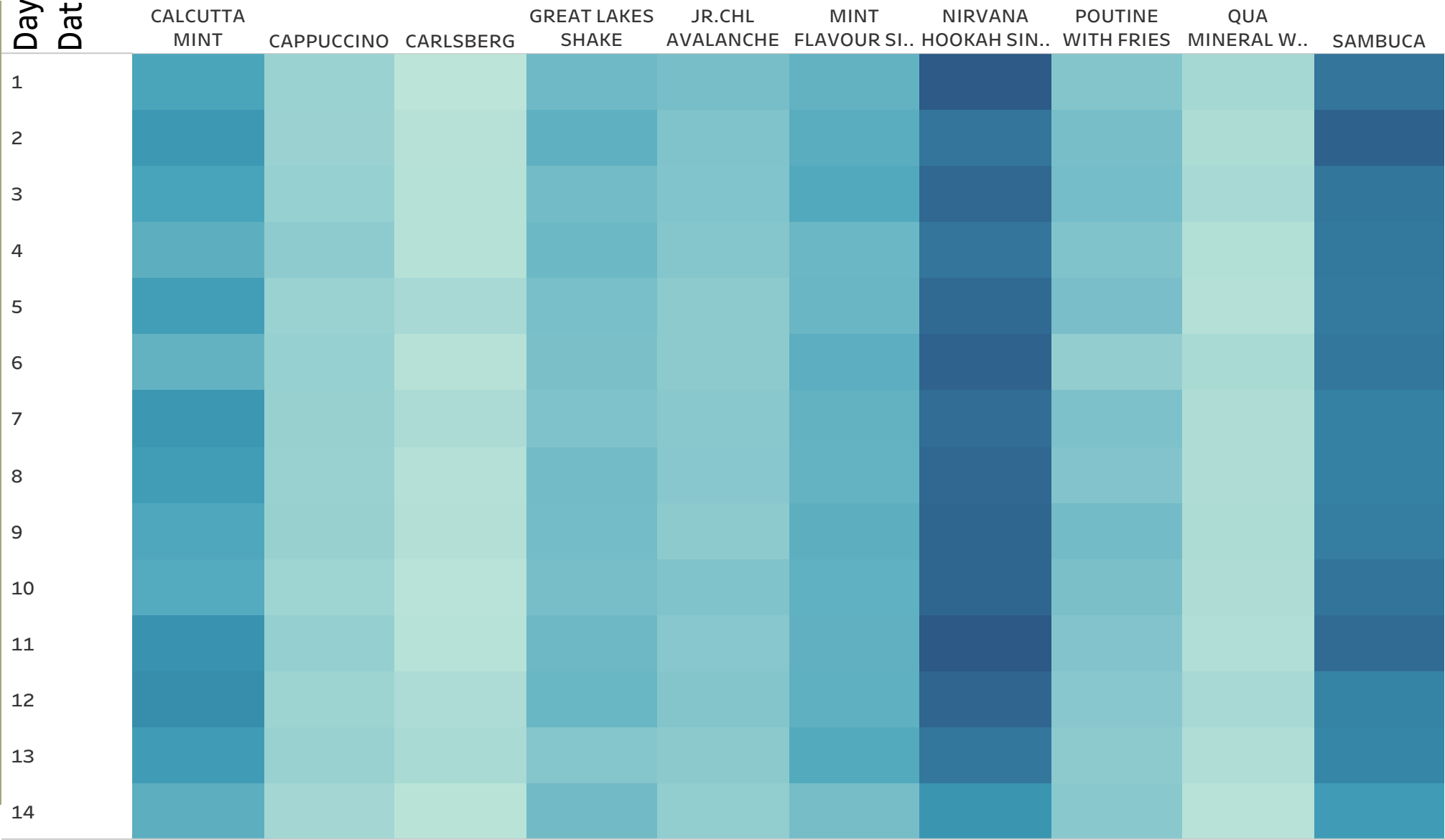
Sum of Quantity (color) broken down by Date Weekday vs. Top 25 Fast Selling SKUs and Category. The view is filtered on Top 25 Fast Selling SKUs, which keeps 25 members.



Hourly sales

Day of  
Date..

Top 10 Fast selling SKUs



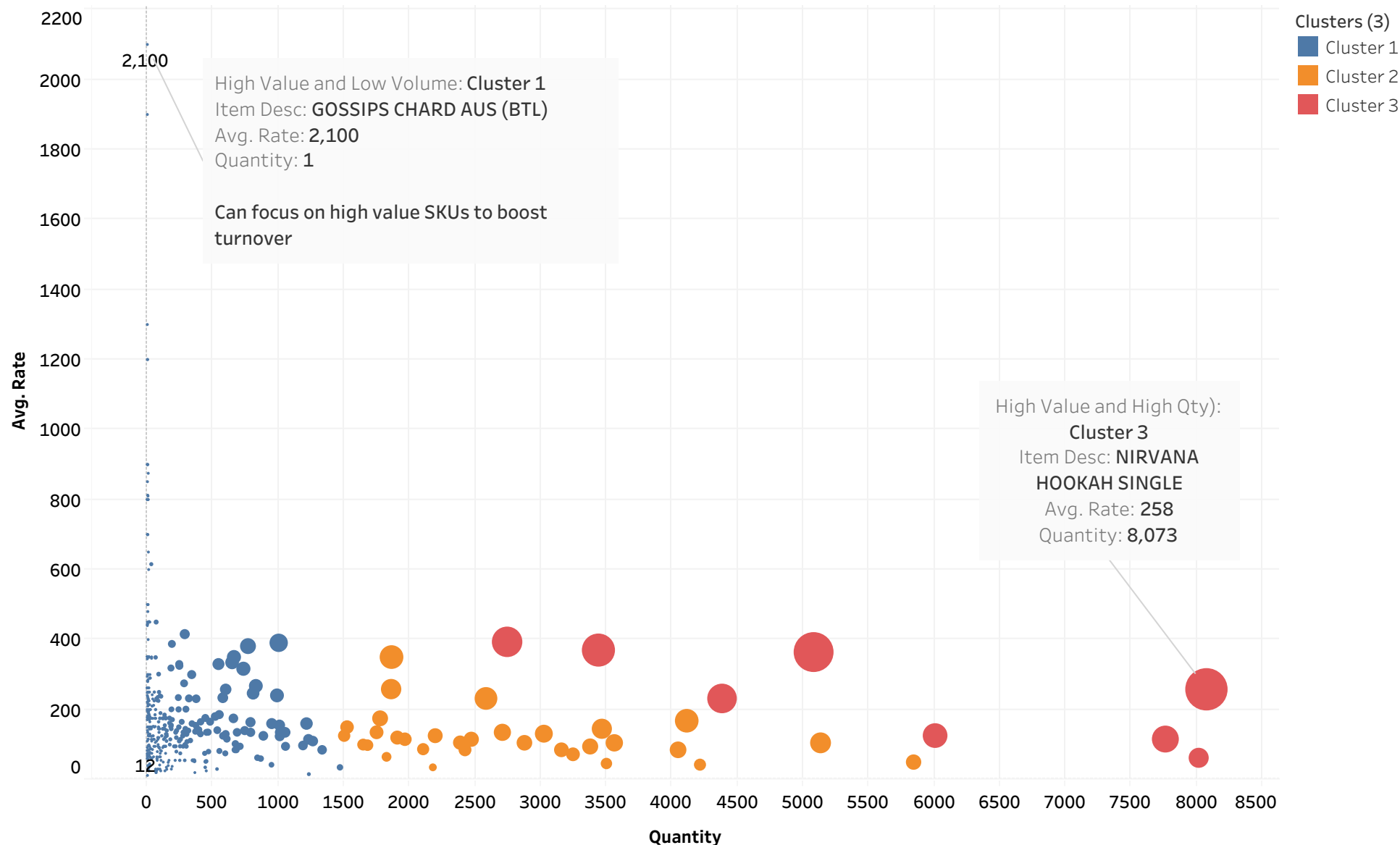
Even Top 10  
Fast selling  
SKUs are not  
sold from third  
week of every  
month

Sum of Sales (color) broken down by Top 10 Fast selling SKUs vs. Date (Sheet21) Day. The view is filtered on Top 10 Fast selling SKUs, which keeps 10 members.



# High value SKUs

Clustering is done based on parameters such as Sales Value, Quantity sold and Average Price of Product



Sum of Quantity vs. average of Rate. Color shows details about Clusters (3). Size shows sum of Sales. Details are shown for Item Desc.

## Top 25 Skus

### Top 25 Fast Selling SKUs



Even Top 25  
SKUs was not  
sold during 3<sup>rd</sup> and  
4<sup>th</sup> Week

Lost multiple days  
sales on Oreo  
cookie shake for  
July 2010.

Date Day for each Top 25 Fast Selling SKUs. Color shows sum of Quantity. The view is filtered on Top 25 Fast Selling SKUs, which keeps 25 members.

## Non Moving SKUs for W3 & W4

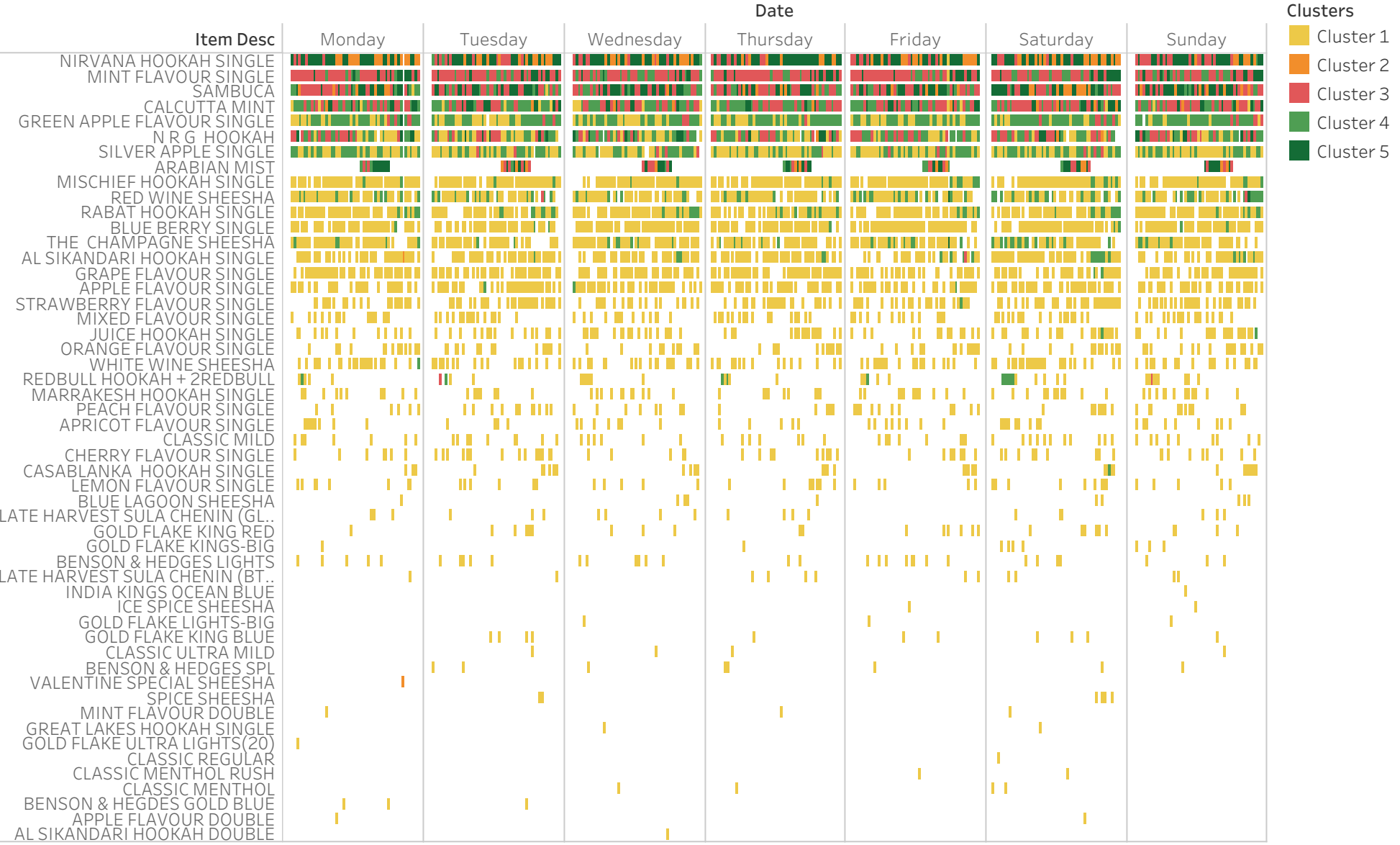
Almost 183 SKUs  
are not sold after  
14<sup>th</sup> day of every  
month which has to  
be examined.

Menu may be  
changed for 3<sup>rd</sup> and  
4<sup>th</sup> week of every  
month.

[illegible]

The view is broken down by Date Day vs. Not sold after 14th day and Item Desc.

Tobacco Sales Pattern



Date Week for each Item Desc broken down by Date Weekday. Color shows details about Clusters. Details are shown for Date Weekday. The data is filtered on Category, which keeps TOBACCO.

Tobacco Sales Overview

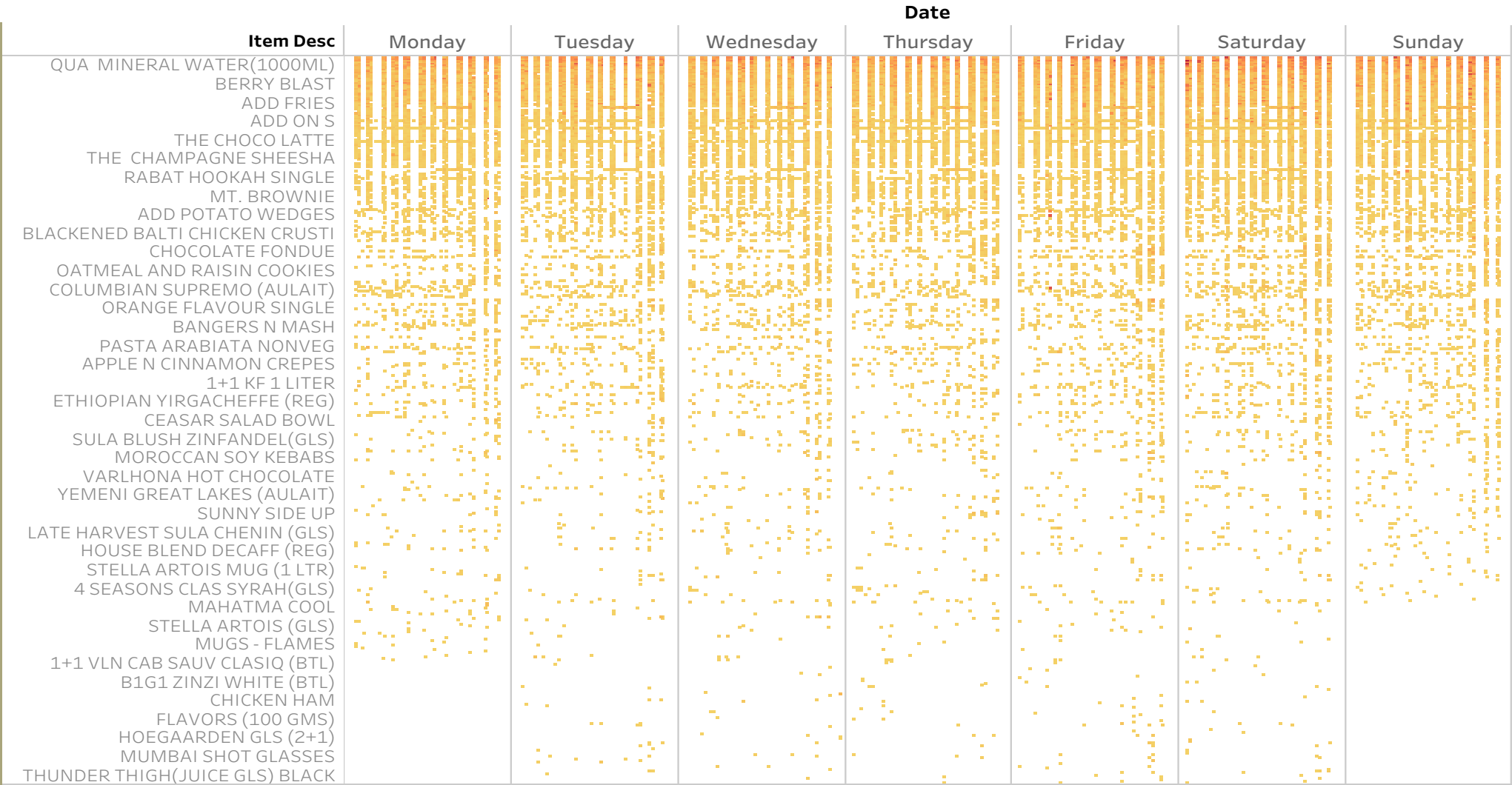
Cluster 5 represents fast moving SKUs and cluster 1 represents slow moving SKUs.

# Overview of Sales Density

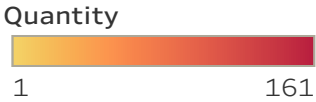
Sales  
Overview of all  
category

Alignment of Item  
Description may  
differ against the  
data points.

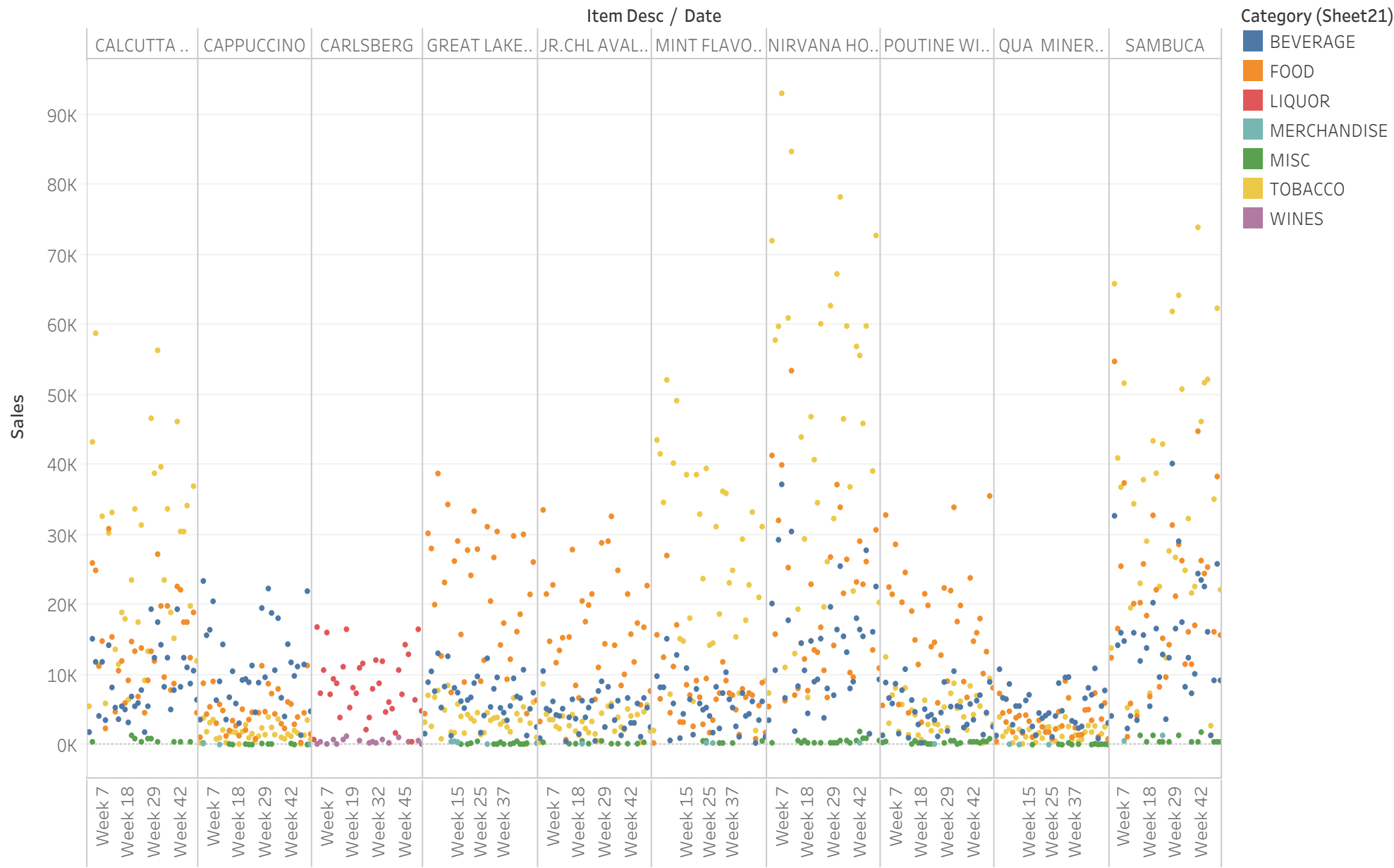
Chart shows Sales  
data points for 12  
months period of  
each day.



Date Week for each Item Desc broken down by Date Weekday. Color shows sum of Quantity (actual & forecast) . Details are shown for Date Weekday. The data is filtered on Category, which keeps 8 of 8 members.



# Category Misclassification



Sum of Sales for each Date Week broken down by Item Desc. Color shows details about Category (Sheet21). The data is filtered on Top 10 Fast selling SKUs, which keeps 10 members.

Few products have been misclassified in different category which can be avoided.

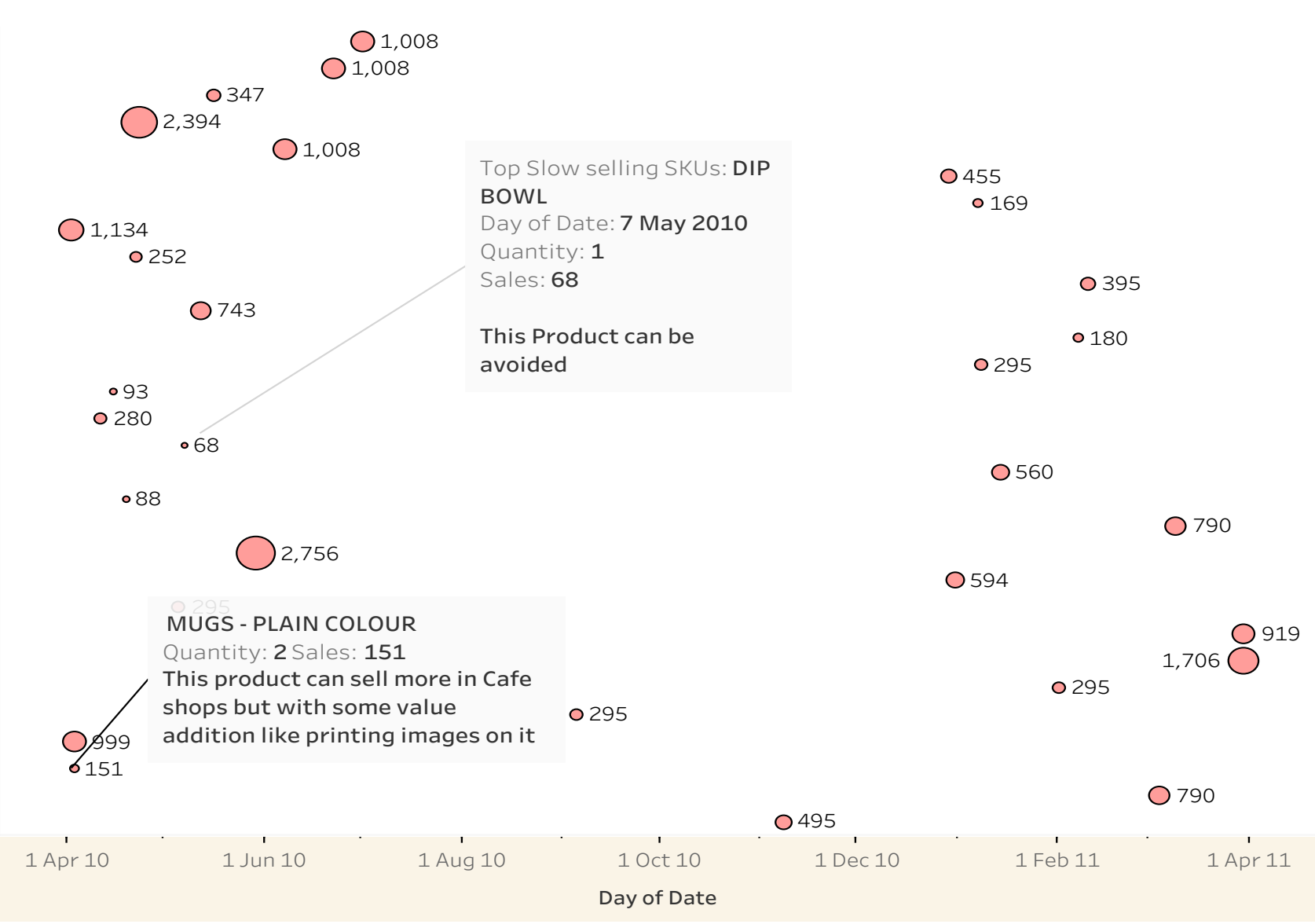
Still we could see the SKUs Sale pattern of Top 10 SKUs

# Top Slow moving Skus

- Top Slow selling SKUs
- 1+1 VLN CAB SAUV (BTL)
- 1+1 VLN SAUV BLANC (BTL)
- 2 AXE TWIST
- 2 OCEAN PINOTAGE (BTL)
- 4 SEASONS CLAS SYRAH(..
- AL SIKANDARI HOOKAH D..
- ASH TRAYS
- B1G1 4SEASON CLAS SAU..
- B1G1 4SEASON CLAS SAU..
- BENARAS BLUE
- CAPONATA
- CH TIN SMALL
- CHAIRMAN COOL
- DECAFFINATE COFFEE FR..
- DHARMATEA LIGHT HOLD..
- DIP BOWL
- ETCHED LEAF TLIGHT HOL..
- GOLD FLAKE ULTRA LIGHT..
- GOLDEN DELIGHT 1.1
- GOSSIPS CHARD AUS (BTL)
- GRAPPO SHEESHA
- GUERRILLA COOL
- HOEGAARDEN GLS (2+1)
- HOEGAARDEN LTR MUGS ..
- KITSCH BLUE
- KONKAN STRIPE
- MODEL-P (IRON ASHTRAY)
- MUGS - PLAIN COLOUR
- UDAIPUR LILY
- ZEN ROCK SQUARE VASE

Few of slow selling SKUs which can be studied and eliminated based on sales pattern.

Almost 35 SKUs has sold less than 2 units but recently sold SKUs has to be considered



Date Day for each Top Slow selling SKUs. Size shows sum of Sales. The marks are labeled by sum of Sales. The view is filtered on Top Slow selling SKUs, which keeps 30 members.

# Suggestion for Combo on Menu

Product  
Recommendation  
using Association  
Rule Algorithm

SET_ID	COMBO_ITEMS	COMBO_PRICE
1	CAPPUCCINO + GREAT LAKES SHAKE	170
2	CAPPUCCINO + MINT FLAVOUR SINGLE	285
3	GREAT LAKES SHAKE + NIRVANA HOOKAH SINGLE	355
4	NIRVANA HOOKAH SINGLE + POUTINE WITH FRIES	370
5	NIRVANA HOOKAH SINGLE + QUA MINERAL WATER	295

Discount shall be given when billed as a set with tax charges extra





Thank You