



*SALES  
FORECASTING  
AND  
INVENTORY  
MANAGEMENT*

THANKS TO STEVE  
MITCHELL FOR MENTORING

## *MOTIVATION*

### Forecasts:

- help reduce costs and improve Customer Service Level.
- drive to better trade-off between cost of excess and cost of shortage in inventory management.
- help build efficient long planning horizon
- facilitate management with data-driven decisions

# *PROJECT SCOPE*

The goal is to use Machine Learning and forecasting techniques to manage inventory of a grocery retailer. It has 54 retail stores and 33 different groups of products. The task is to model the behaviors of 1782 time series for forecasting.

The dataset used come from Corporación Favorita, a large Ecuadorian-based grocery retailer. Thanks to Kaggle for making this dataset available for free.

## *TABLE OF CONTENTS*

- Overview on the demand
- Monthly sales forecast at different stores
- Monthly sale forecast at different stores for different categories of products
- Daily sales forecast at different stores
- Daily sales forecast at different stores for different categories of products
- Improvement on the model
- Conclusion

## *OVERVIEW ON THE DEMAND*

### Weekly seasonality factors/days

Days	Factors
Monday	0.97
Tuesday	0.90
Wednesday	0.94
Thursday	0.79
Friday	0.91
Saturday	1.21
Sunday	1.28

### Global tendency

### Annual seasonality/month

Months	Factors
January	1.08
February	0.98
March	1.08
April	1.11
May	1.15
June	1.15
July	1.21
August	0.96
September	0.76
October	0.79
November	0.79
December	0.94

## *OVERVIEW ON THE DEMAND*

### Sales by store types

Type	Sales share	# store/54
A	32.27	9
B	14.00	8
C	15.00	15
D	32.58	18
E	6.15	4

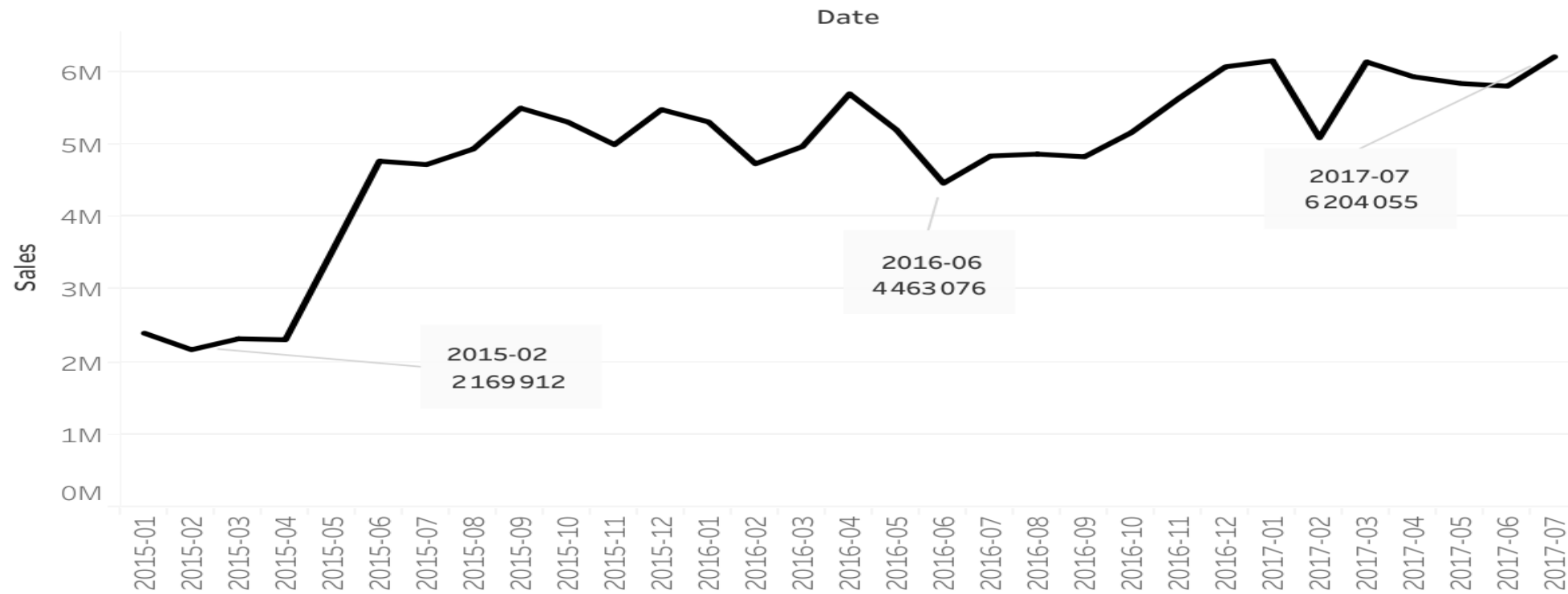
### Sales by product categories

Categories	Sales share
Grocery I	30.32
Beverages	21.30
Produce	13.70
Rest (30 categories)	34.68

- 65.32% of the sales within 3 categories of product
- 2 types of store lead the sales ranking

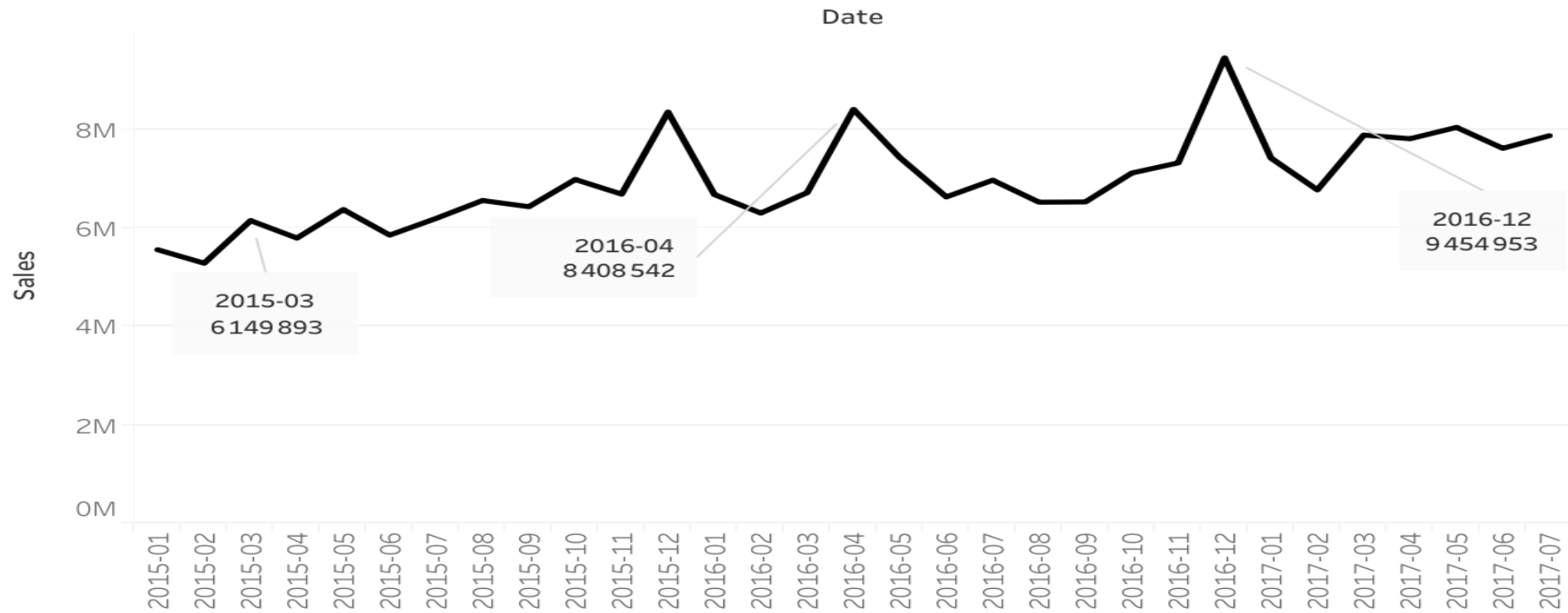
## OVERVIEW ON THE DEMAND

### Beverages growth from 2015-2017



## OVERVIEW ON THE DEMAND

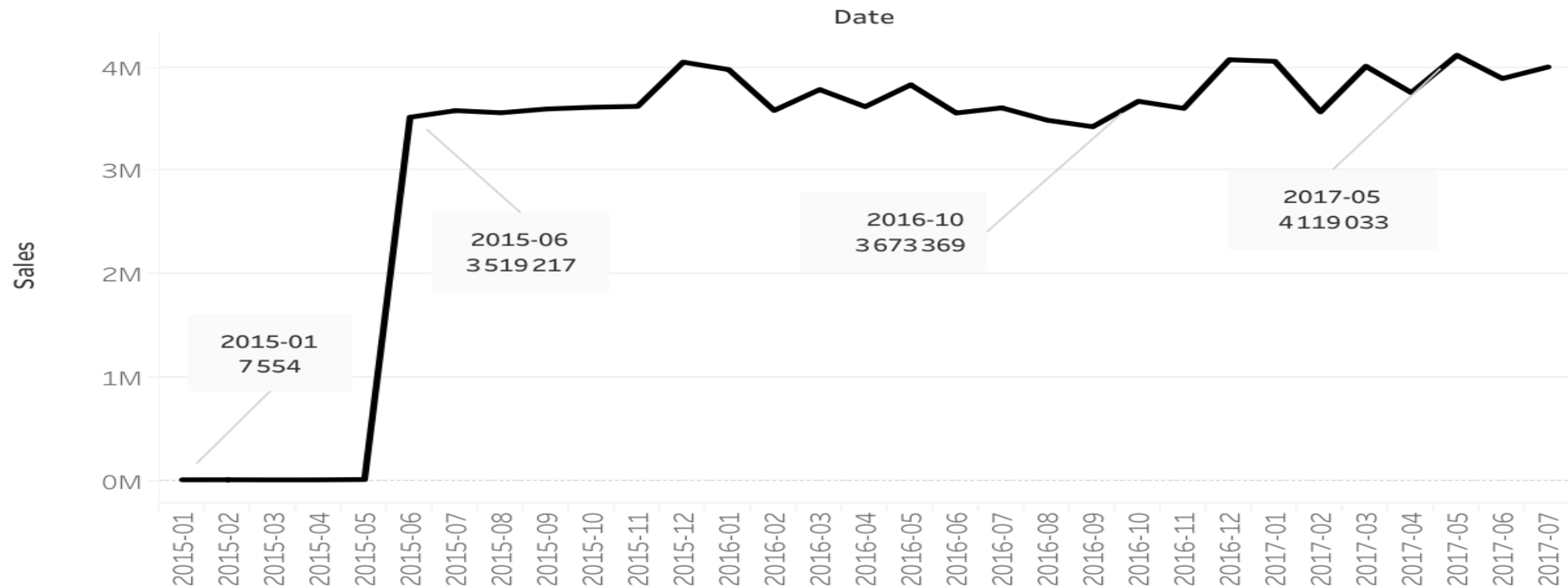
Grocery I growth from 2015-2017





## OVERVIEW ON THE DEMAND

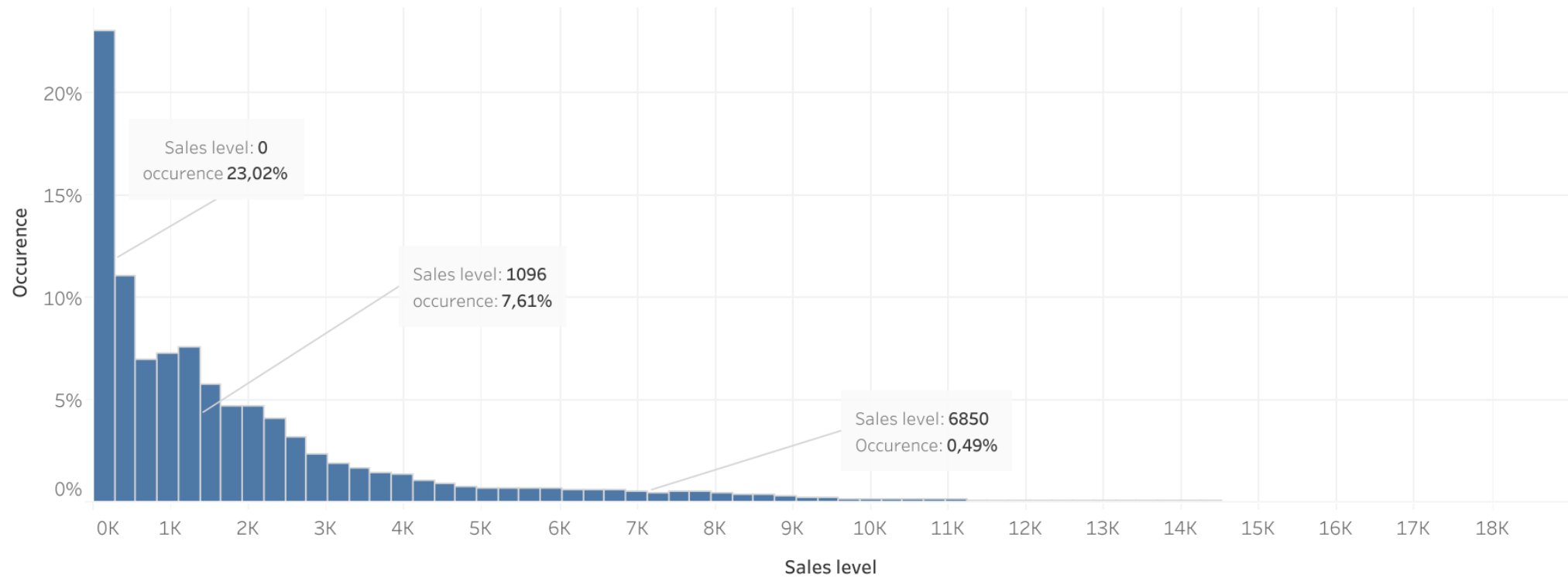
Produce growth from 2015-2017



## OVERVIEW ON THE DEMAND

Produce category

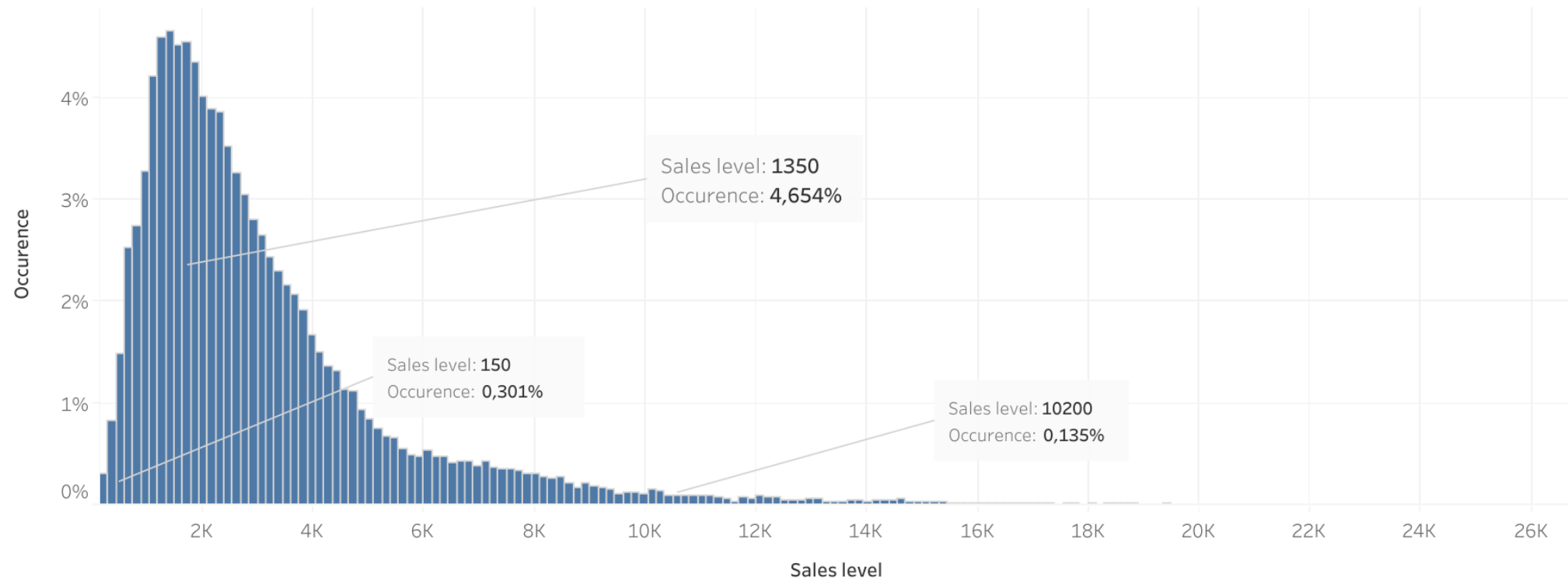
Poisson distribution - slow moving items



## OVERVIEW ON THE DEMAND

Beverages category

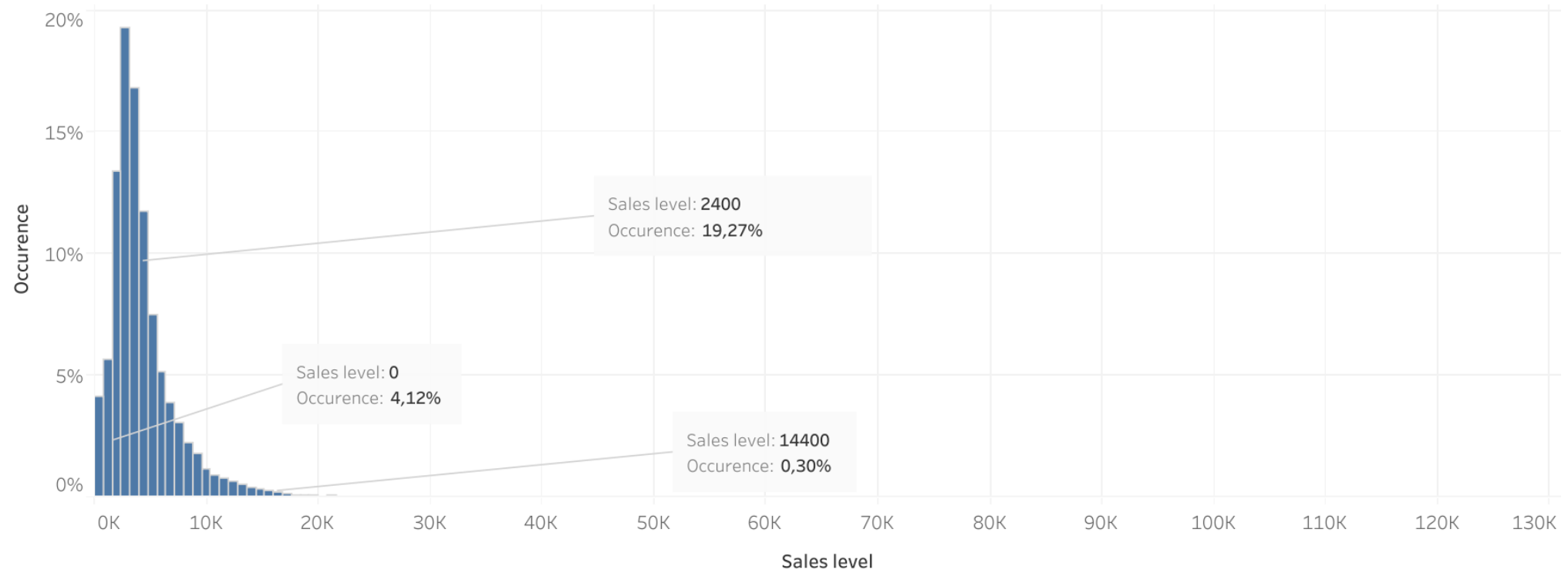
Tends to a normal distribution - medium rate



## OVERVIEW ON THE DEMAND

Grocery I category

Tends to a normal distribution - Medium rate



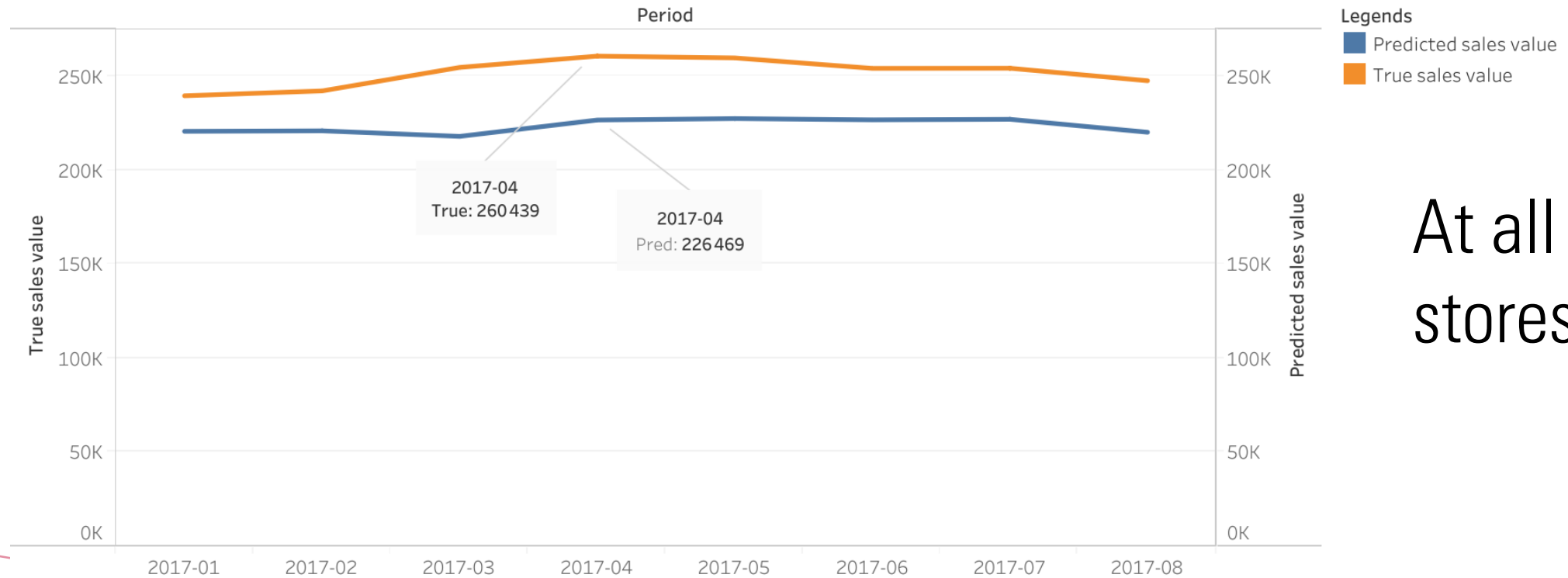


# Monthly sales

# MONTHLY SALES FORECAST AT DIFFERENT STORES

## Random Forest Regressor

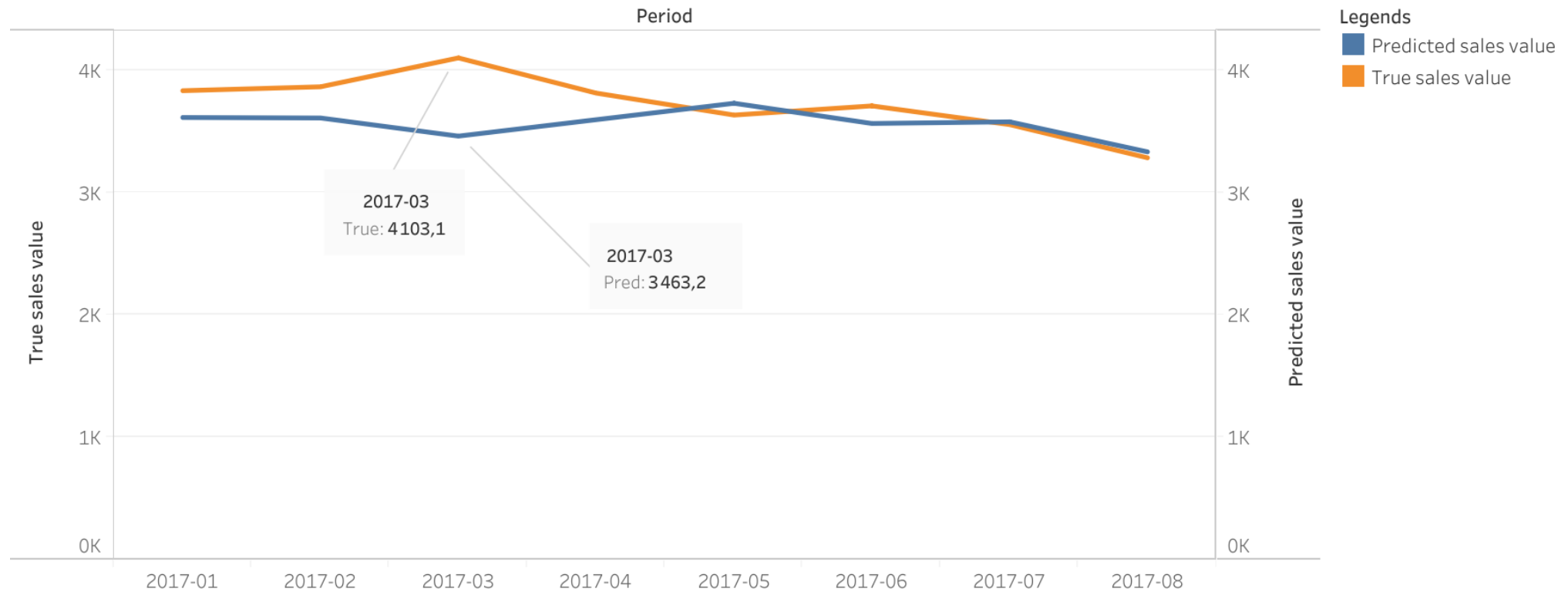
Grocery I - prediction global tendency



At all  
stores

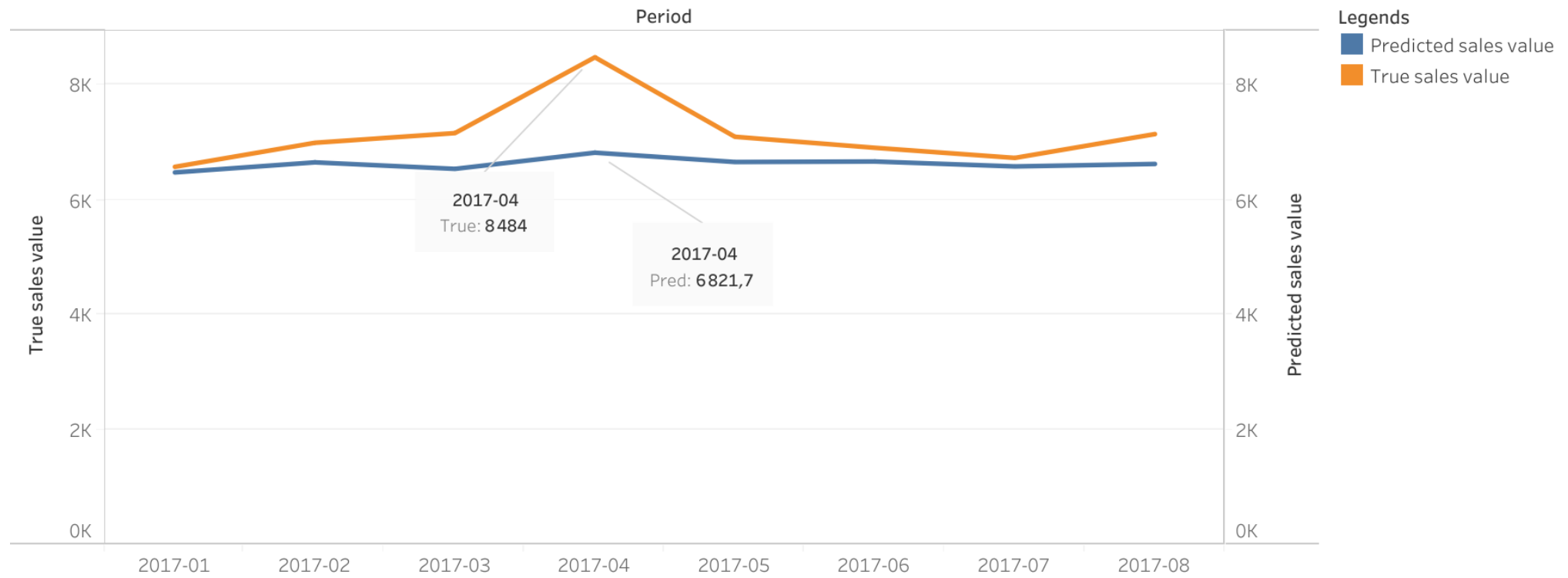
## MONTHLY SALES FORECAST AT DIFFERENT STORES

### Grocery I - store 4



## MONTHLY SALES FORECAST AT DIFFERENT STORES

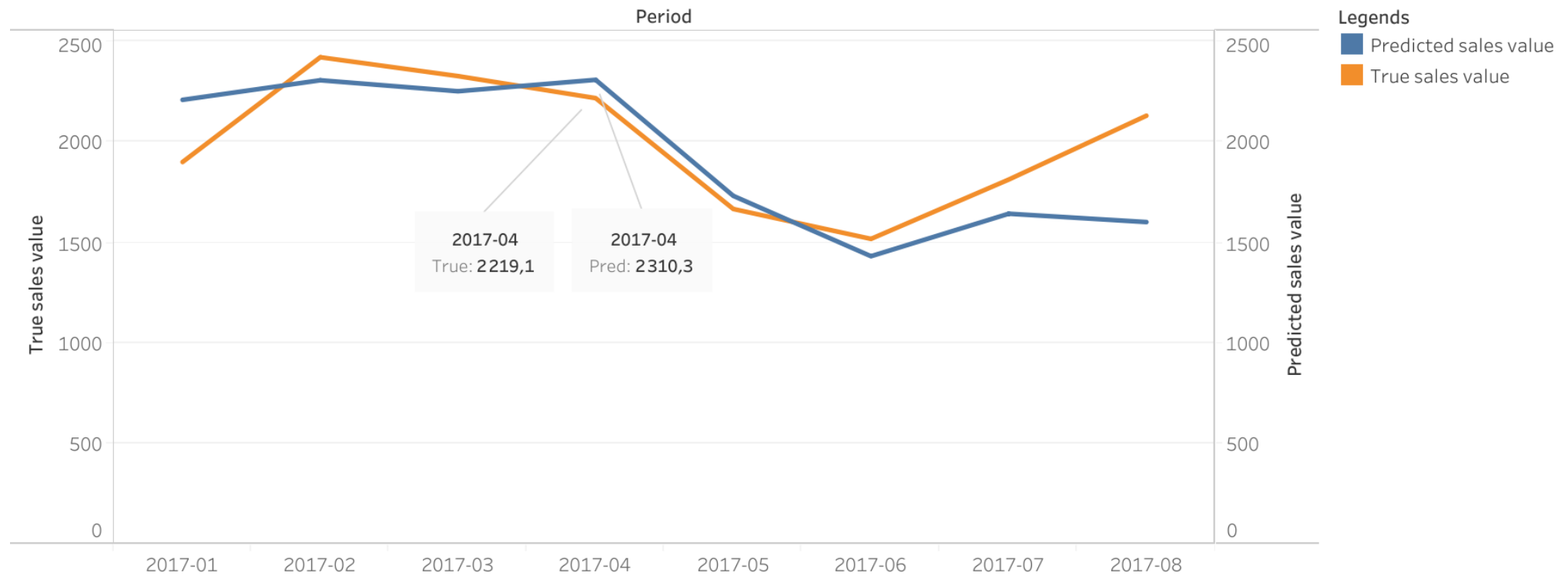
Grocery I - store 9





# MONTHLY SALES FORECAST AT DIFFERENT STORES

## Grocery I - store 25



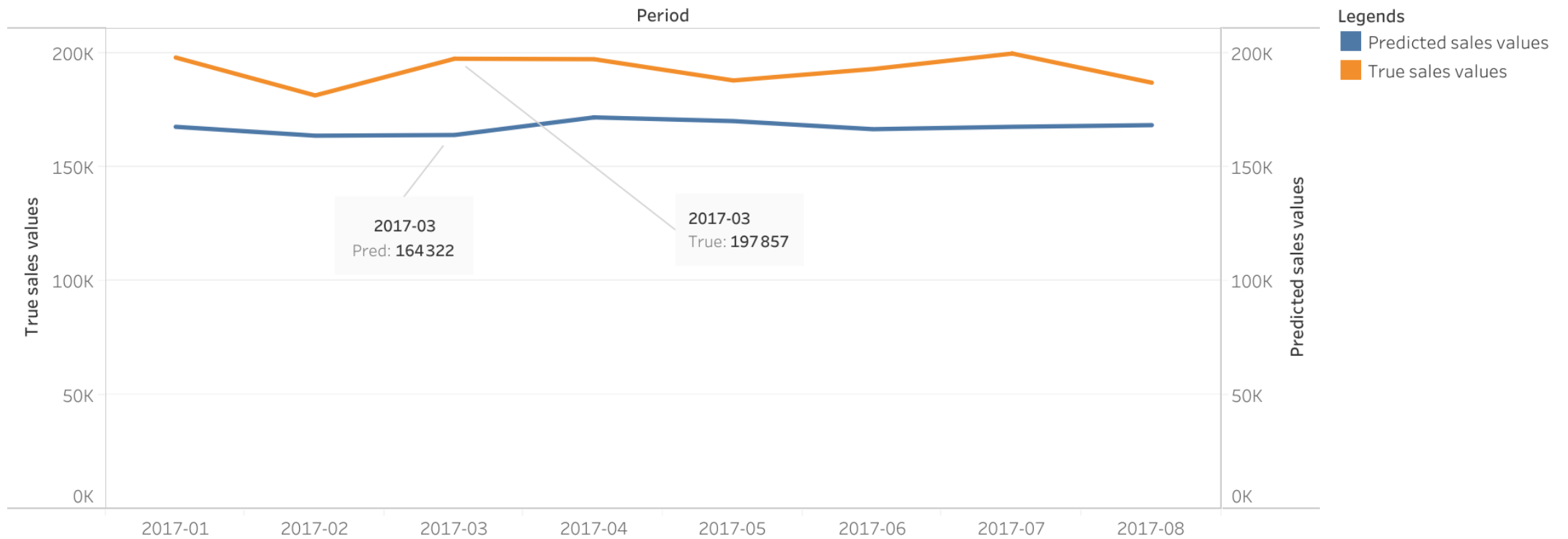
## *MONTHLY SALES FORECAST AT DIFFERENT STORES*

Model performance for “GROCERY I” category by store

Error percentage %	Model accuracy
5	28.90
10	56.87
15	80.09
20	91.20

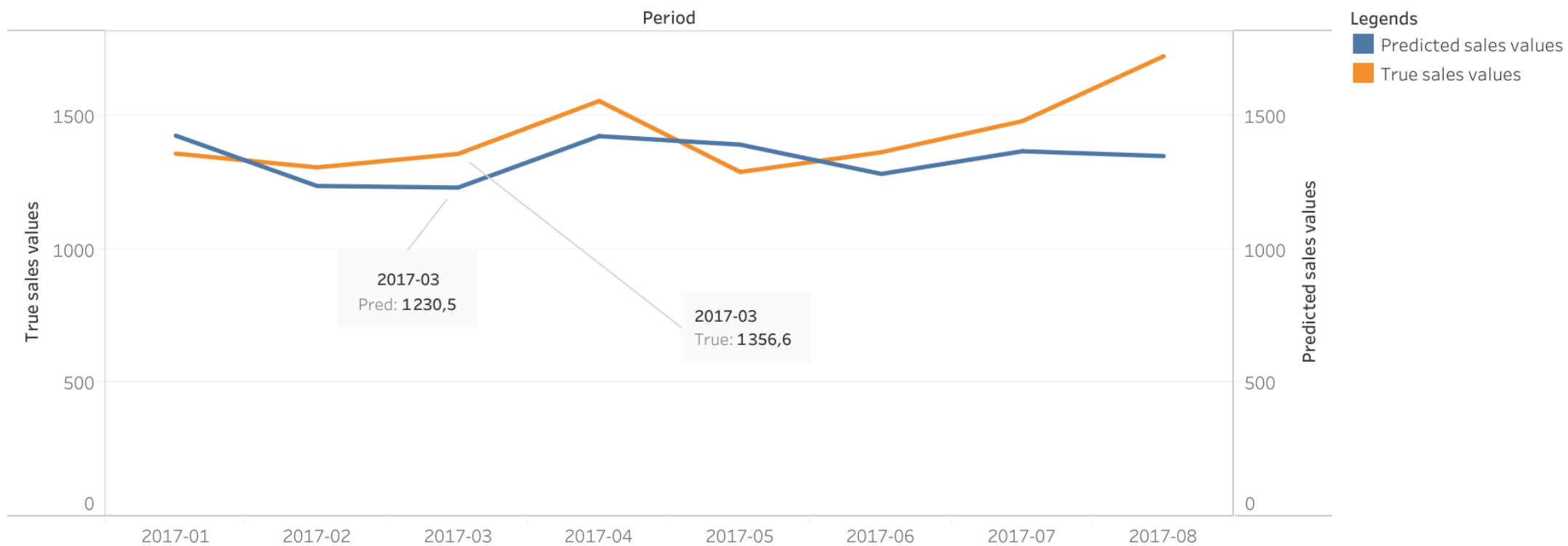
# MONTHLY SALES FORECAST AT DIFFERENT STORES

Beverages - prediction global tendency



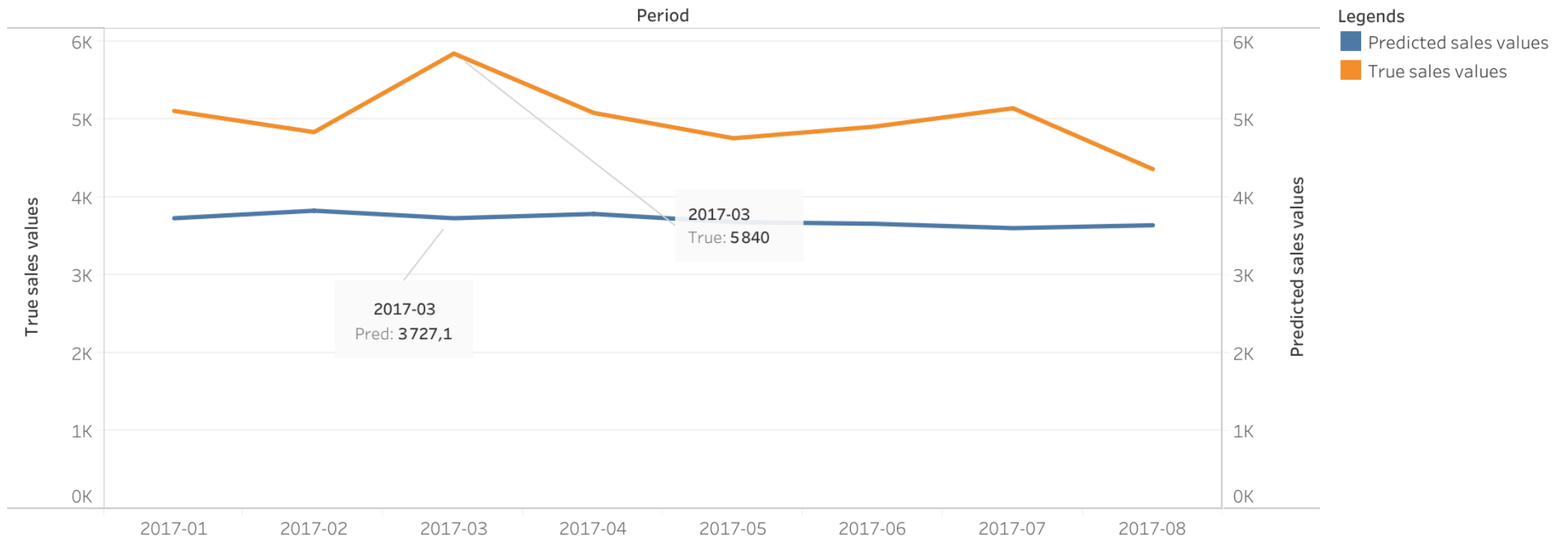
# MONTHLY SALES FORECAST AT DIFFERENT STORES

## Beverages - store 14



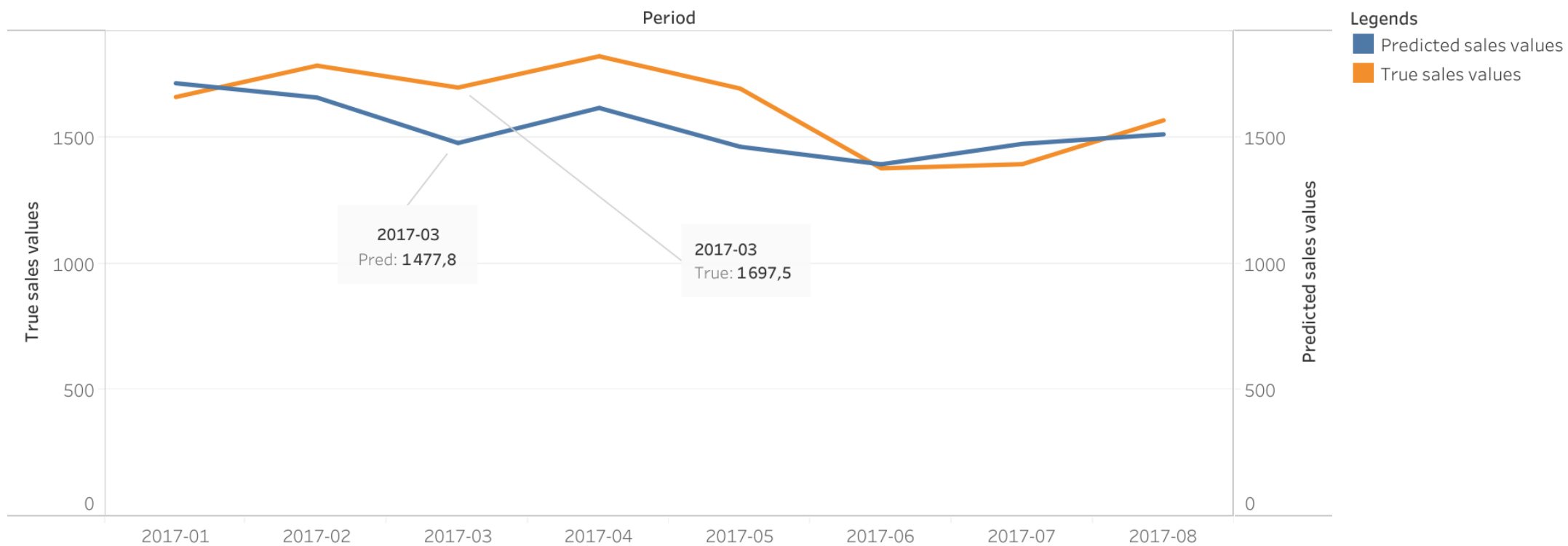
# MONTHLY SALES FORECAST AT DIFFERENT STORES

## Beverages - store 50



# MONTHLY SALES FORECAST AT DIFFERENT STORES

## Beverages - Store 35



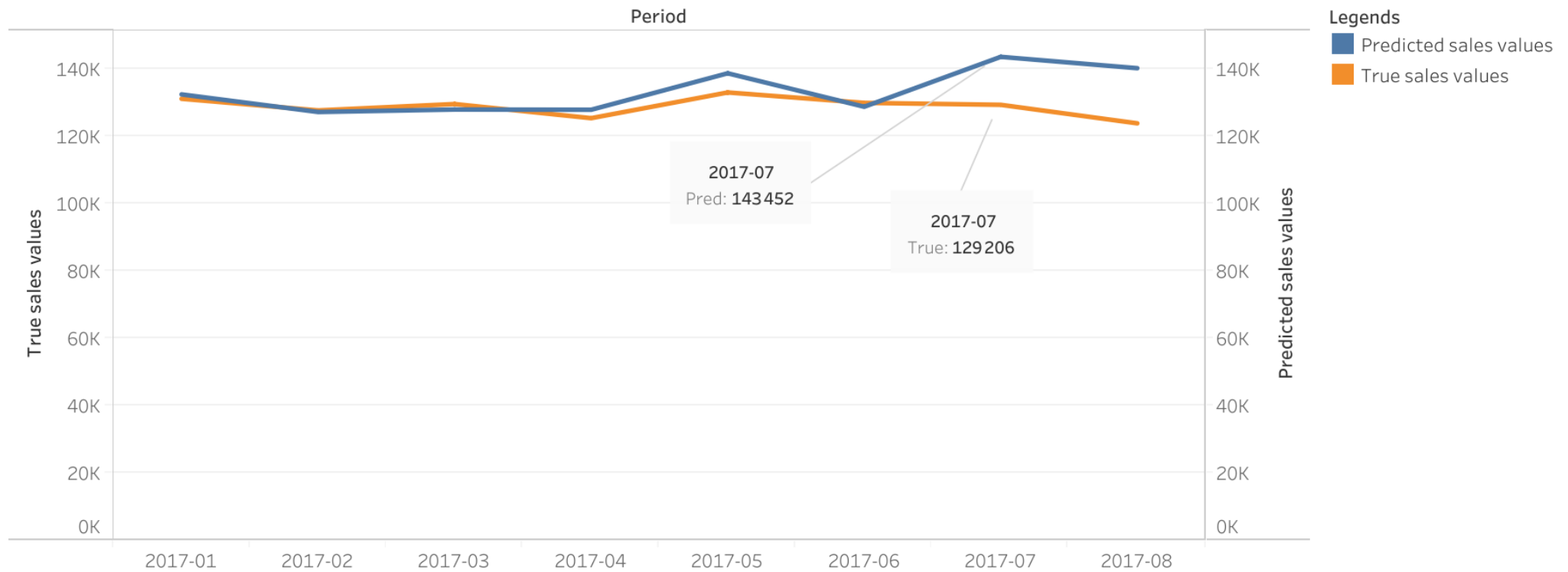
## *MONTHLY SALES FORECAST AT DIFFERENT STORES*

Model performance for “BEVERAGES” category by store

Error percentage %	Model accuracy
5	20.51
10	44.99
15	63.64
20	78.32

# MONTHLY SALES FORECAST AT DIFFERENT STORES

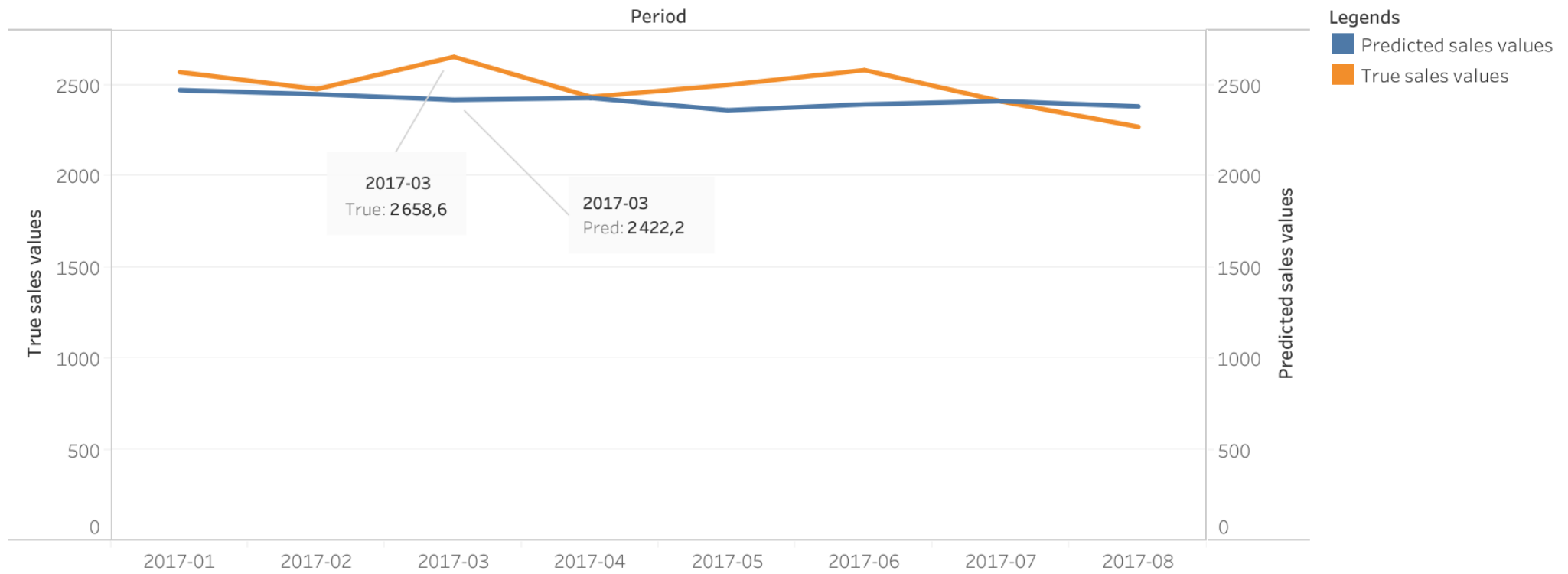
Produce - prediction global tendency





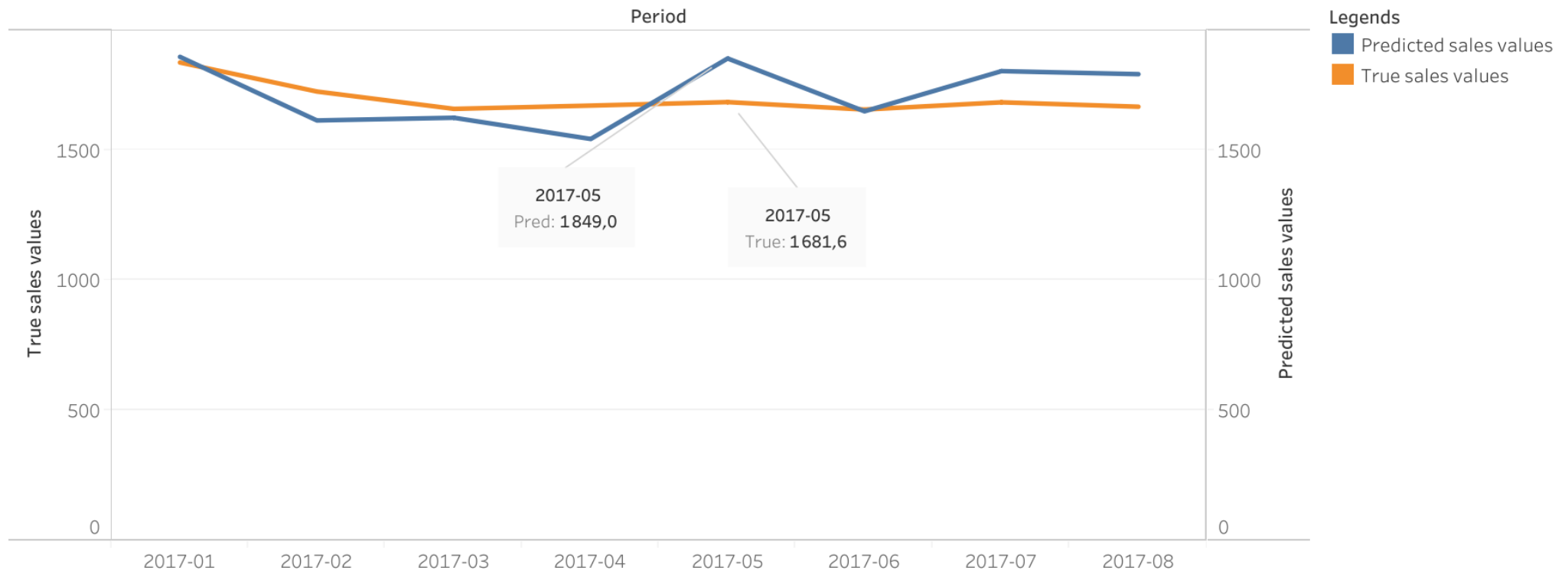
# MONTHLY SALES FORECAST AT DIFFERENT STORES

## Produce - Store 1



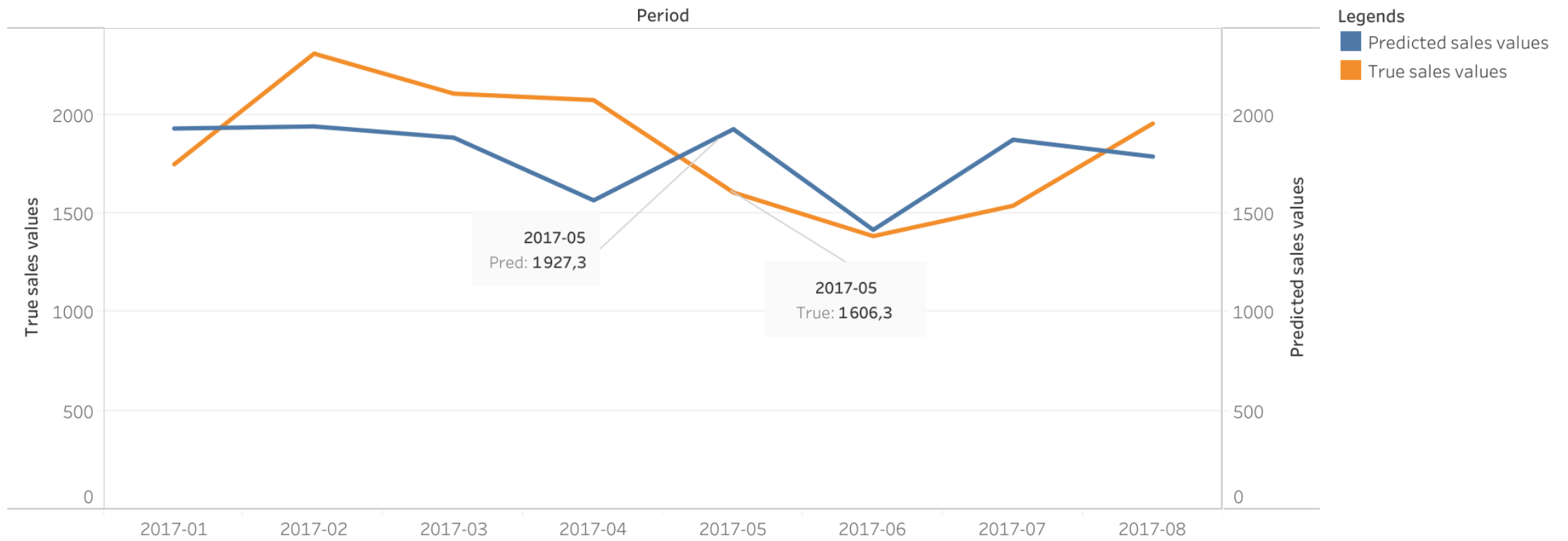
# MONTHLY SALES FORECAST AT DIFFERENT STORES

## Produce - Store 9



# MONTHLY SALES FORECAST AT DIFFERENT STORES

Produce - Store 25



## *MONTHLY SALES FORECAST AT DIFFERENT STORES*

Model performance for “PRODUCE” category demand by store

Error percentage %	Model accuracy
5	40.56
10	59.67
15	71.56
20	76.45

## *MONTHLY SALES FORECAST AT DIFFERENT STORES*

Model performance on predicting monthly sales at each store for financial planning

Error percentage %	Model accuracy
5	28.67
10	57.81
15	69.68
20	78.79
30	90.21


Monthly  
sales  
by store

## *MONTHLY SALES FORECAST AT DIFFERENT STORES*

Model performance on predicting monthly sales for remaining categories at each store for transportation planning

Error percentage %	Model accuracy
5	25.34
10	45.69
15	61.32
20	72.31

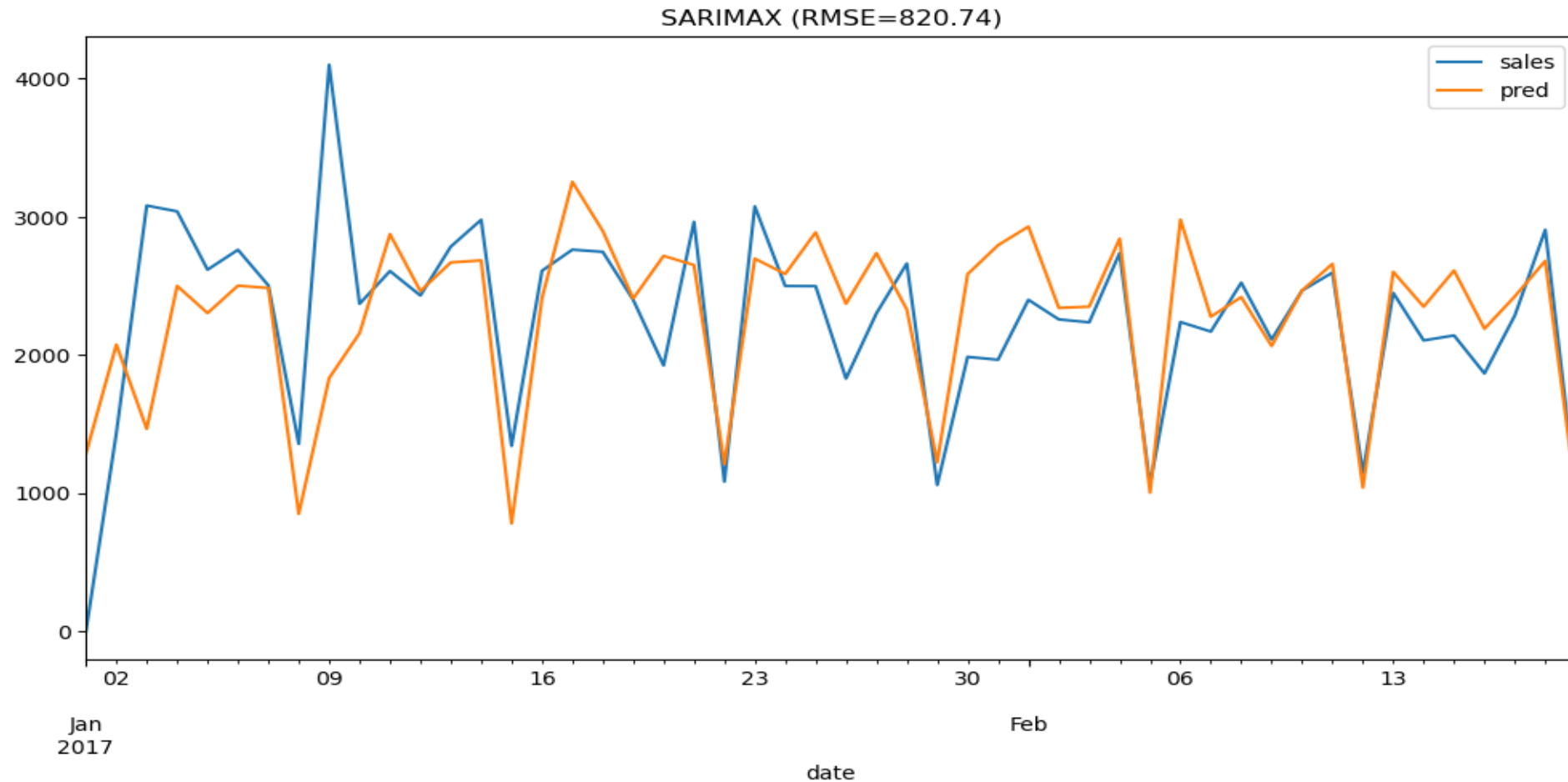
Each  
category  
of  
product  
at each  
store



# Daily sales

## *DAILY SALES FORECAST AT DIFFERENT STORES*

SARIMAX model with no period lead time for "BEVERAGES" category





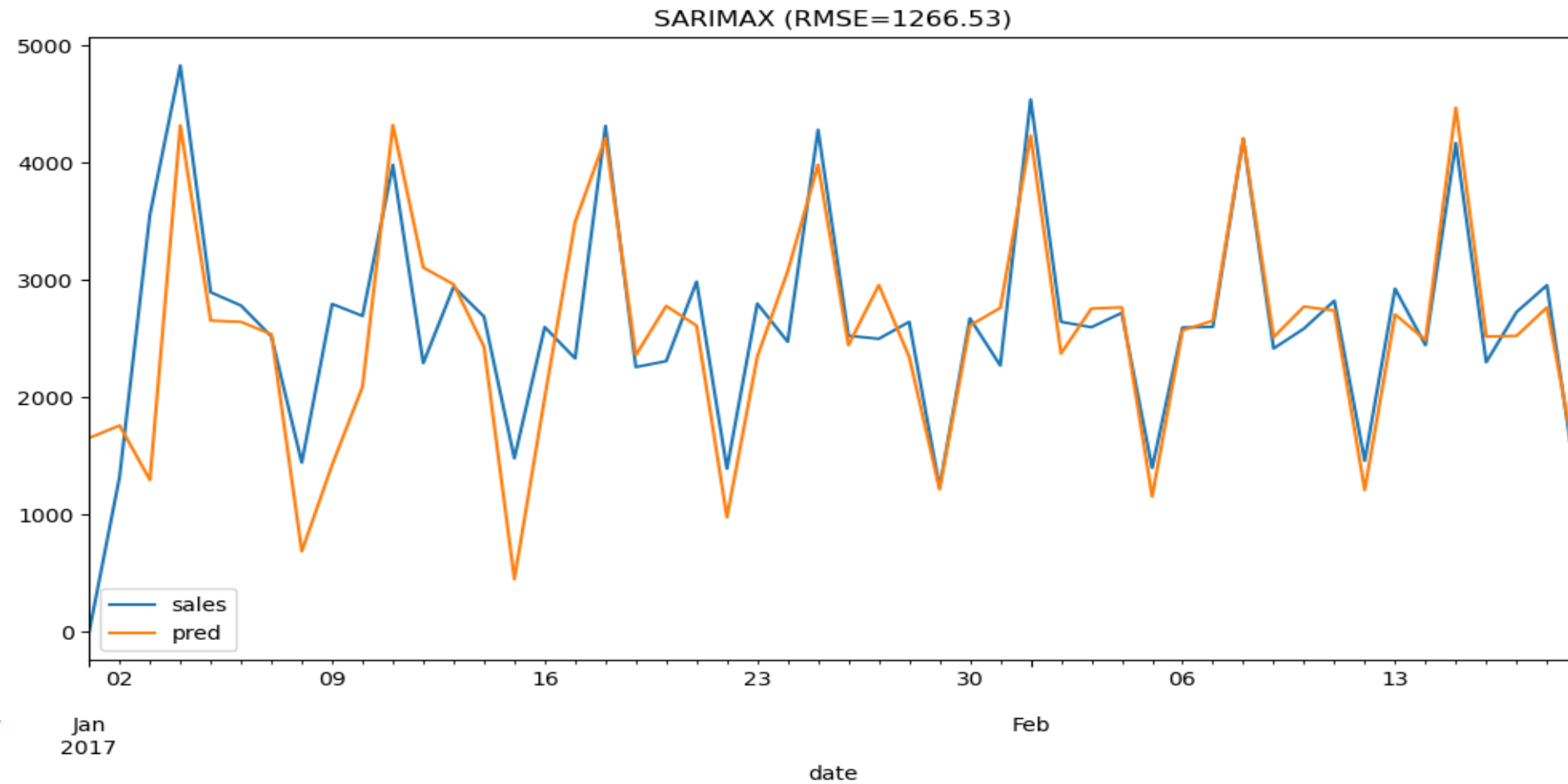
## *DAILY SALES FORECAST AT DIFFERENT STORES*

Model performance on predicting daily sales for “BEVERAGES” category at each stores

Error percentage %	Model accuracy
5	28.19
10	48.90
15	66.96
20	75.77

## DAILY SALES FORECAST AT DIFFERENT STORES

SARIMAX model with no lead time period for "PRODUCE" category



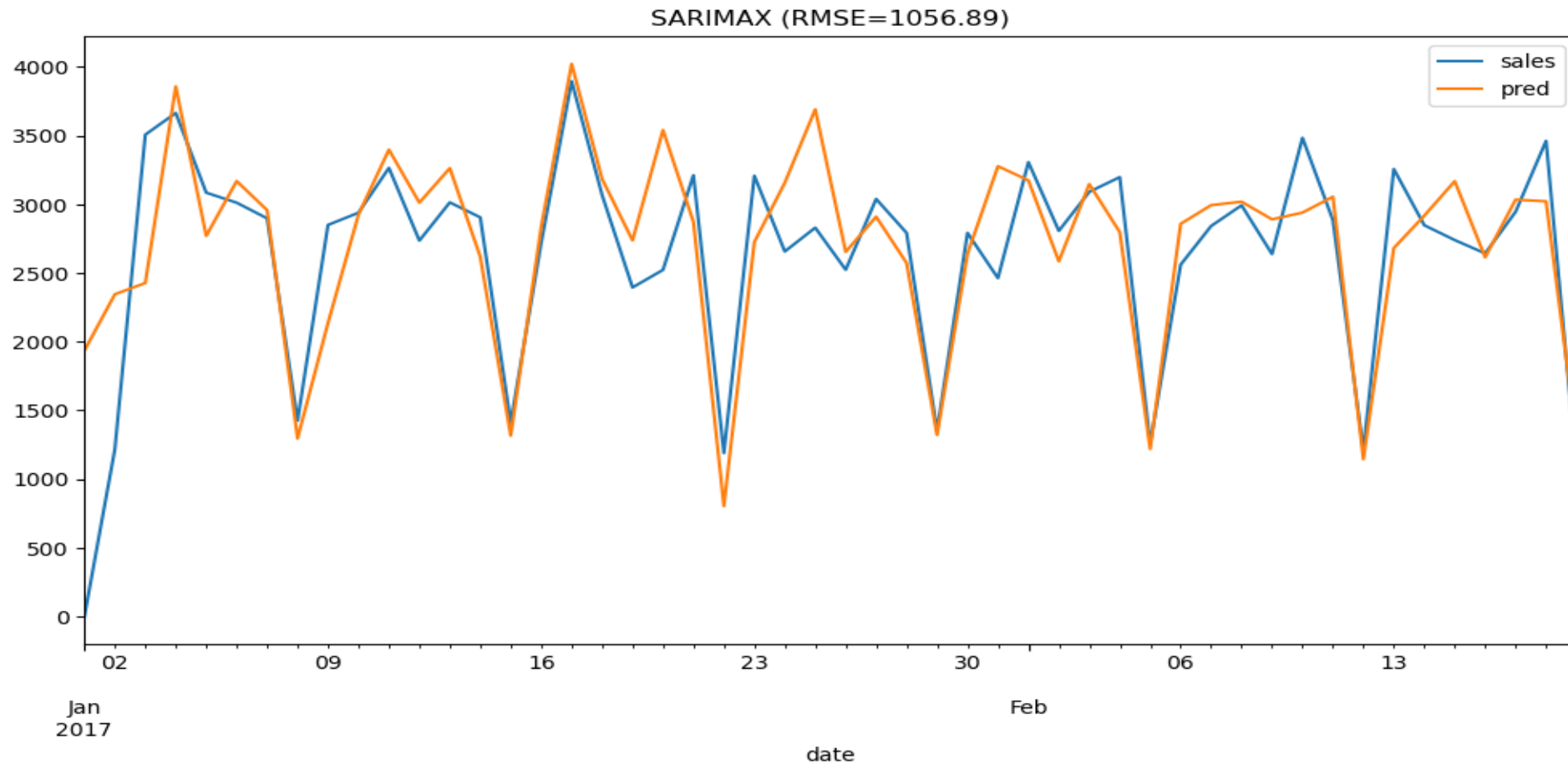
## *DAILY SALES FORECAST AT DIFFERENT STORES*

Model performance on predicting daily sales for “PRODUCE” category at each stores

Error percentage %	Model accuracy
5	28.63
10	54.19
15	70.48
20	79.30

## DAILY SALES FORECAST AT DIFFERENT STORES

SARIMAX model with no lead time period for "GROCERY I" category



## *DAILY SALES FORECAST AT DIFFERENT STORES*

Model performance on predicting daily sales for “PRODUCE” category at each stores

Error percentage %	Model accuracy
5	28.63
10	51.54
15	67.84
20	78.41

## *IMPROVEMENT ON THE MODEL*

- More information on the dataset to handle missing and low values. Those values decrease the model accuracy.
- Case by case analysis. Some series have incompatible behaviors that requires different types of features engineering.



*THANK  
YOU FOR  
YOUR  
TIME*