Retail forecasting and clustering

Vladimir Nikiforov

1. Forecasting

Goal is to calculate sale forecasts for next 28 days:

1. Region (all shops) / ware / day
2. Store / ware / day

Metric: WAPE

Goal metric values:

1. Region / ware / day less than 30%
2. Store / ware/ day less than 50%

It is desirable to apply several different techniques of forecasting and data analysis (ARIMA, random forest, ...).

1. Forecasting. WAPE – what is it?

Weighted Absolute Percent Error (WAPE or WMAPE) is the Sum of Absolute errors divided by the Sum of the Actuals

$$WMAPE = \frac{\sum |Actual - Forecast|}{\sum Actual}$$

WAPE gives you a true picture of forecast quality in an organization and how this will impact the business performance in both Sales and profits.

WAPE can also be construed as the Average Absolute Error divided by the Average Actual quantity

1. Forecasting. Data info

86916 rows * 7 columns:

- Store_id 86916 non-null int64
- SKU_id 86916 non-null int64
- Date 86916 non-null datetime64[ns]
- Promo 15349 non-null float64
- Demand 86916 non-null int64
- Regular_Price 86916 non-null float64
- Promo_Price 15349 non-null float64

There are missing values in Promo/Promo_Price.

Let’s use them to modify Regular_Price.

1. Forecasting. Features

Let’s create new features

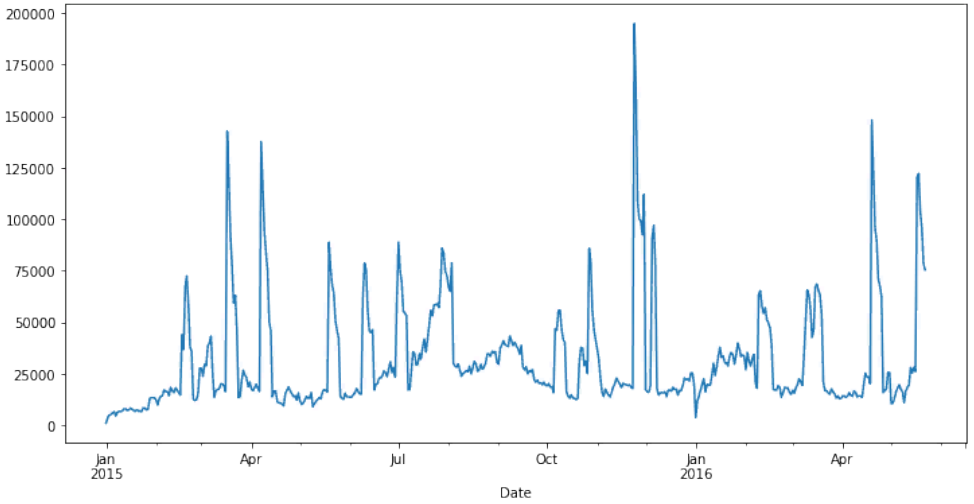
- WEEKDAY – weekday number from Date
- YEARDAY – cos of day number in year from Date
- PRICE – Promo_Price if >0, else Regular_Price
- Demand_lag1_day – Demand value from yesterday
- Demand_lag1_week – Demand value from the same weekday week ago

1. Forecasting. ARIMA

- ARIMA stands for Autoregressive Integrated Moving Average models.
- Univariate (single vector) ARIMA is a forecasting technique that projects the future values of a series based entirely on its own inertia.
- Its main application is in the area of short term forecasting requiring at least 40 historical data points.
- It works best when your data exhibits a stable or consistent pattern over time with a minimum amount of outliers.

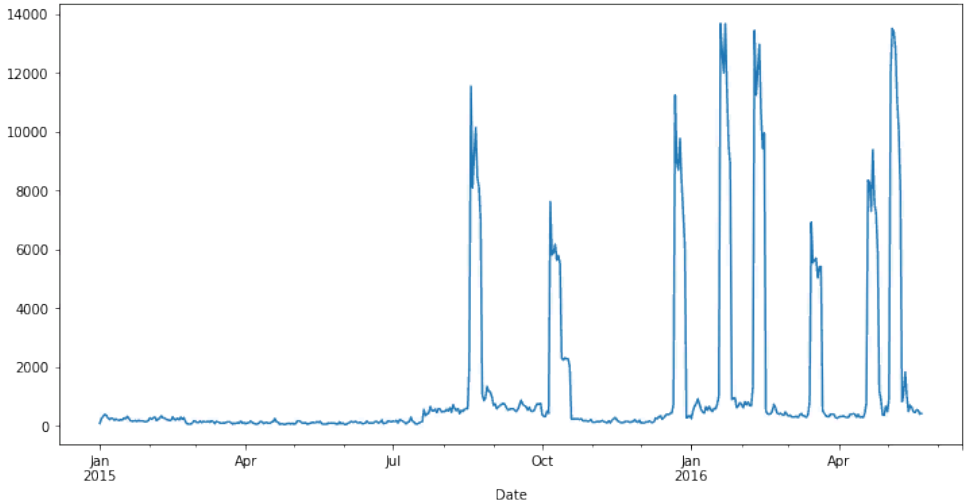
1. Forecasting. ARIMA

SKU_id = 1



WAPE = 21.65%

SKU_id = 2

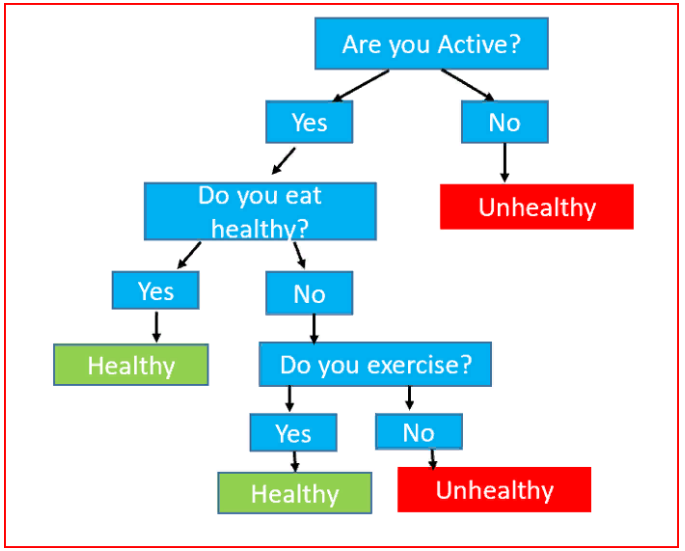


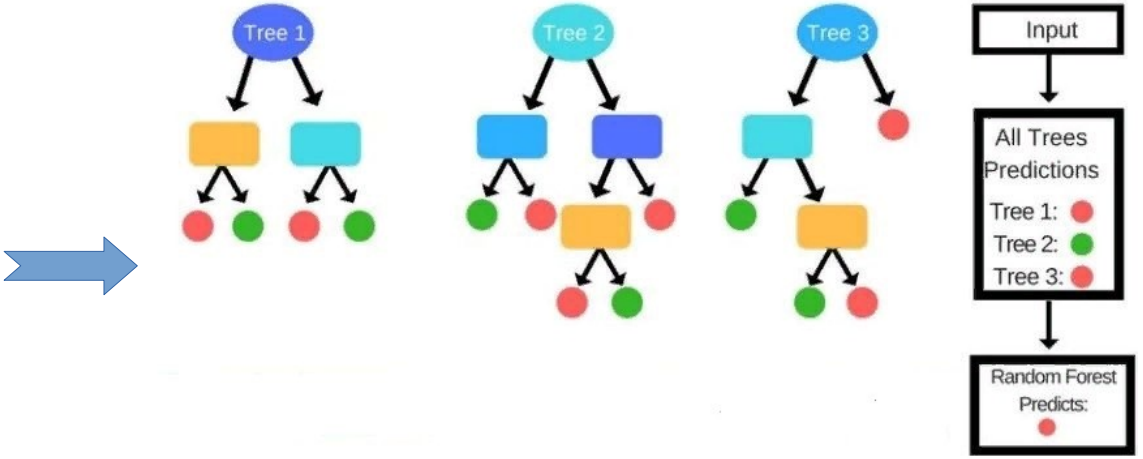
WAPE = 29.85%

** Forecast for each next day made from previous fact-data*

1. Forecasting. Random Forest

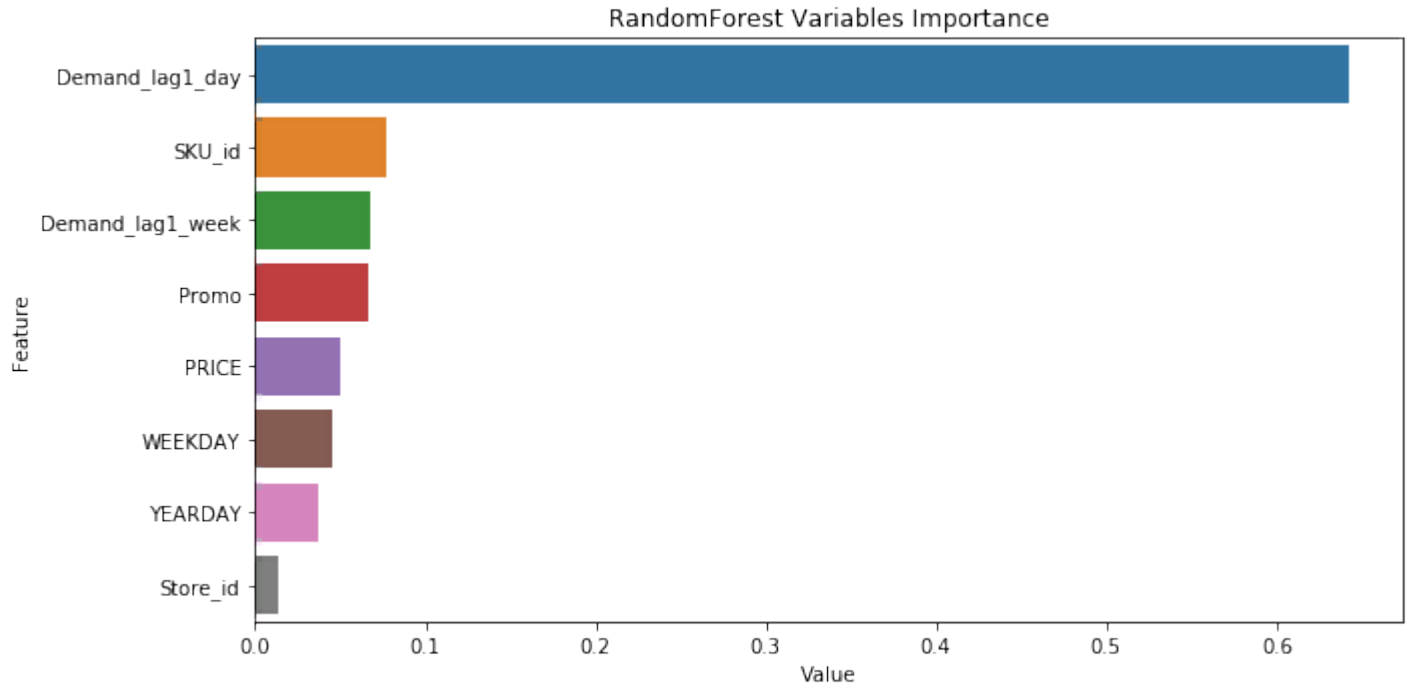
- Random Forest is an ensemble of randomized decision trees.
- Each decision tree gives a vote for the prediction of target variable.
- Random forest chooses the prediction that gets the most vote.
- An ensemble learning model aggregates multiple machine learning models to give a better performance.
- In random forest we use multiple random decision trees for a better accuracy.





1. Forecasting. Random Forest

- Detailed forecast: Store / ware / day
- WAPE = 39.83%



2. Clustering

- Goal is to find clusters in stores data
- Clustering helps divide data into a number of subsets. Each of these subsets contain data that is similar to one another. These subsets are known as clusters.

- We have stores sale data and stores attributes

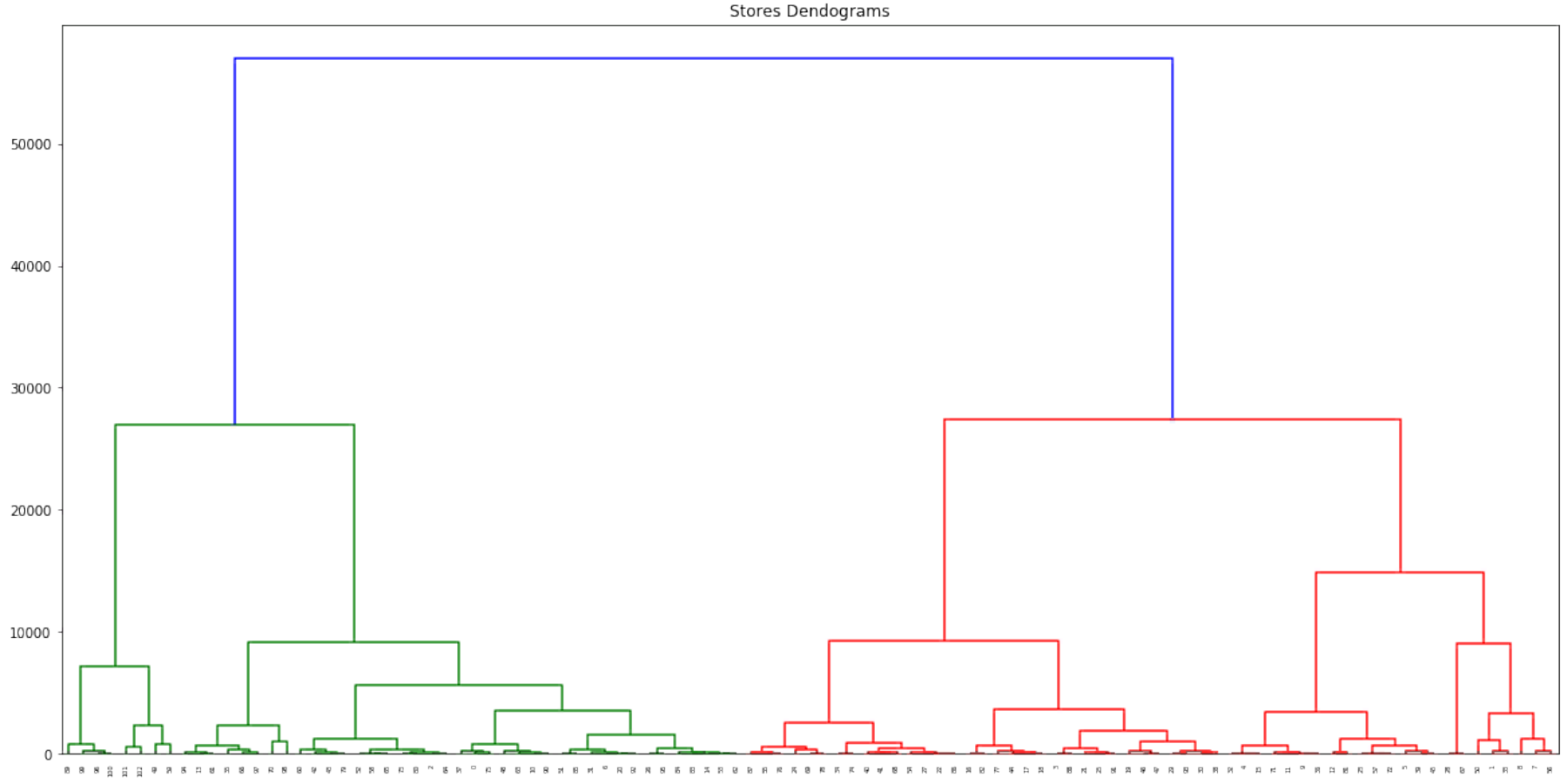
2. Clustering. Features

- We want to find cluster of stores with the same power of sales:
- Let’s create features for each category (level 2 of wares) and calculate **average monthly percent of sold quantity per “number of SKU in category”** from daily sales for each category and “**revenue per 1 square feet**” for each store.

Store	Alc	Bac	KidFood	Diabet	Cond	PetFood	Water	Add	Tab	Kid	Br	Revenue_per_square
1	0.300275	2.149165	1.230385	1.727069	1.949320	9.762331	1.385337	0.822255	2.220865	0.576590	8.037128	9768.996297
2	0.545191	7.027134	2.620226	3.511610	4.151190	29.575847	2.453708	1.677054	2.521745	1.273234	13.103581	23029.120716
3	0.371701	2.758914	1.234908	1.911515	2.180084	11.014311	1.825030	1.030597	2.720372	0.891680	6.346894	12003.246645
4	0.609614	3.694375	1.432687	2.460331	2.599241	13.640608	2.328957	1.254898	4.949867	0.851365	11.329470	16512.405695
5	0.763618	4.533633	2.211038	3.147070	3.318566	21.678563	2.489573	1.586891	4.475934	1.376133	11.670952	19422.691949

2. Clustering. Results

Hierarchical clustering method provide us clusters’ power

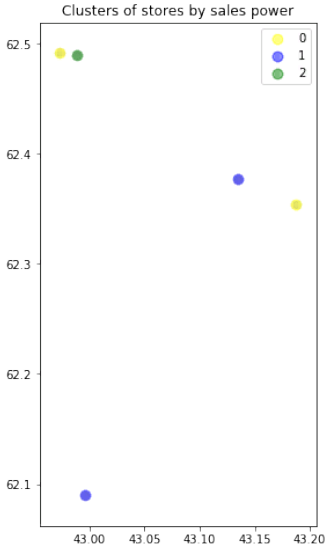


2. Clustering. Results

- Based on dendrogram we can split our stores to 4 big clusters or 9 small clusters.
- We can provide cluster label to each store.

4 clusters

Store	Cluster
1	1
2	0
3	1
4	2
5	0



9 clusters

Store	Cluster
1	0
2	3
3	0
4	1
5	4