KAMIL KANDZIA
KANDZIA.KAMIL@GMAIL.COM

Business case

What should be taken into account when granting credit?

Available data

1. ***store_sales_per_category.csv*** contains weekly sales volume (in Euro's) for stores in scope for 7 different products.
2. ***store_distances_anonymized.csv*** contains distances in kilometres between each pair of stores that are within 5 kilometres of each other.
3. ***gdata_anonymized.csv*** contains the number of high education institutions that are within 5 km radius from each store in scope.

Prophet – forecasting procedure

Why is this a good choice for forecast analysis?

Flexibility

The measurements do not need to be regularly spaced

Facebook uses it for producing reliable forecasts for planning and goal setting

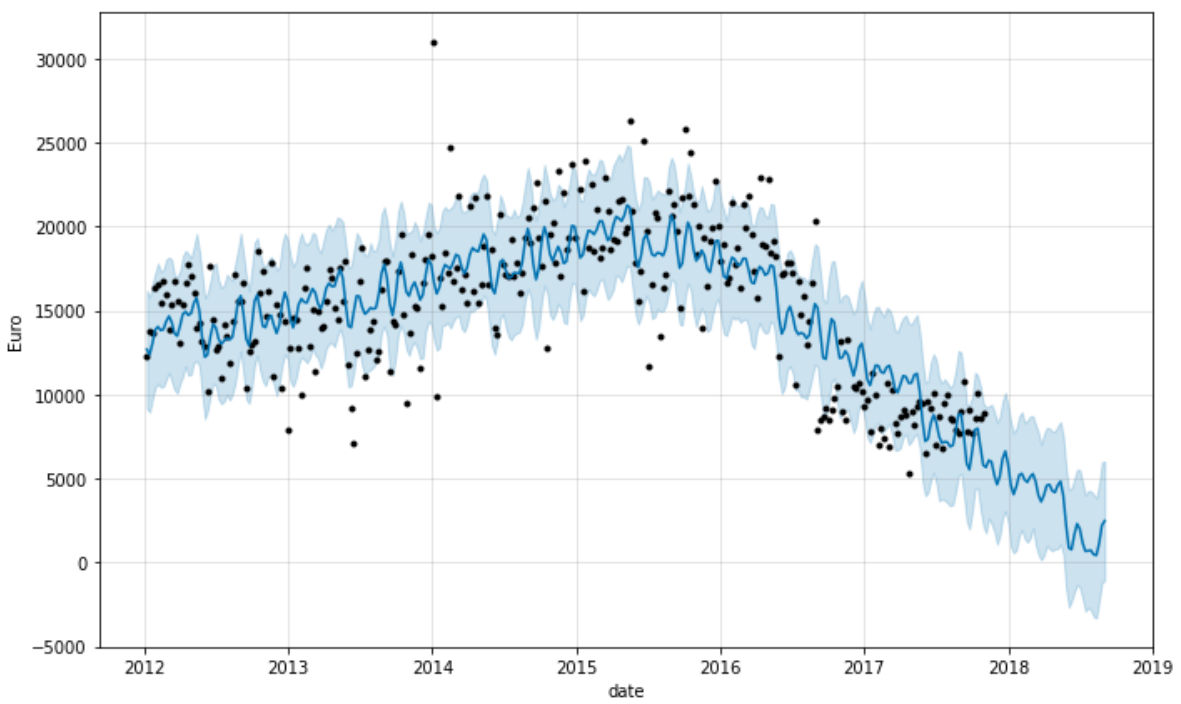
Prophet

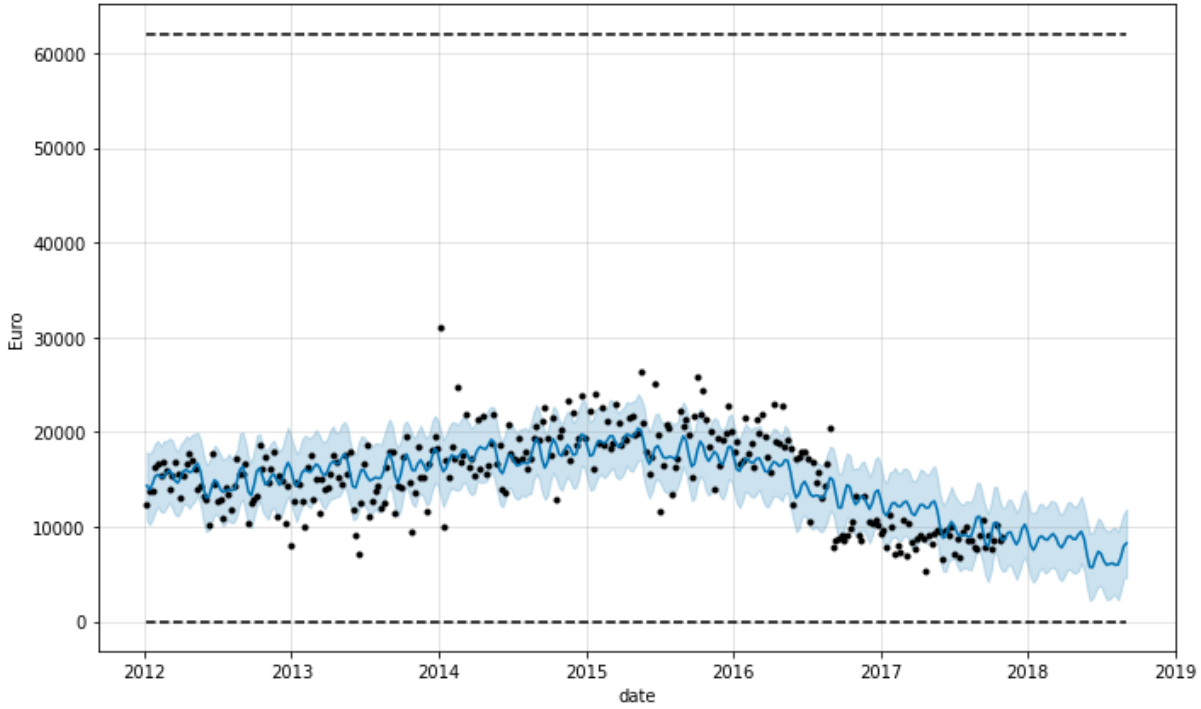
Predicted value: total sales of alcohol per week.

Additional regressors: sales of each type of alcohol per week.

Due to the lack of a given country, it is not possible to add additional regressors such as days off.

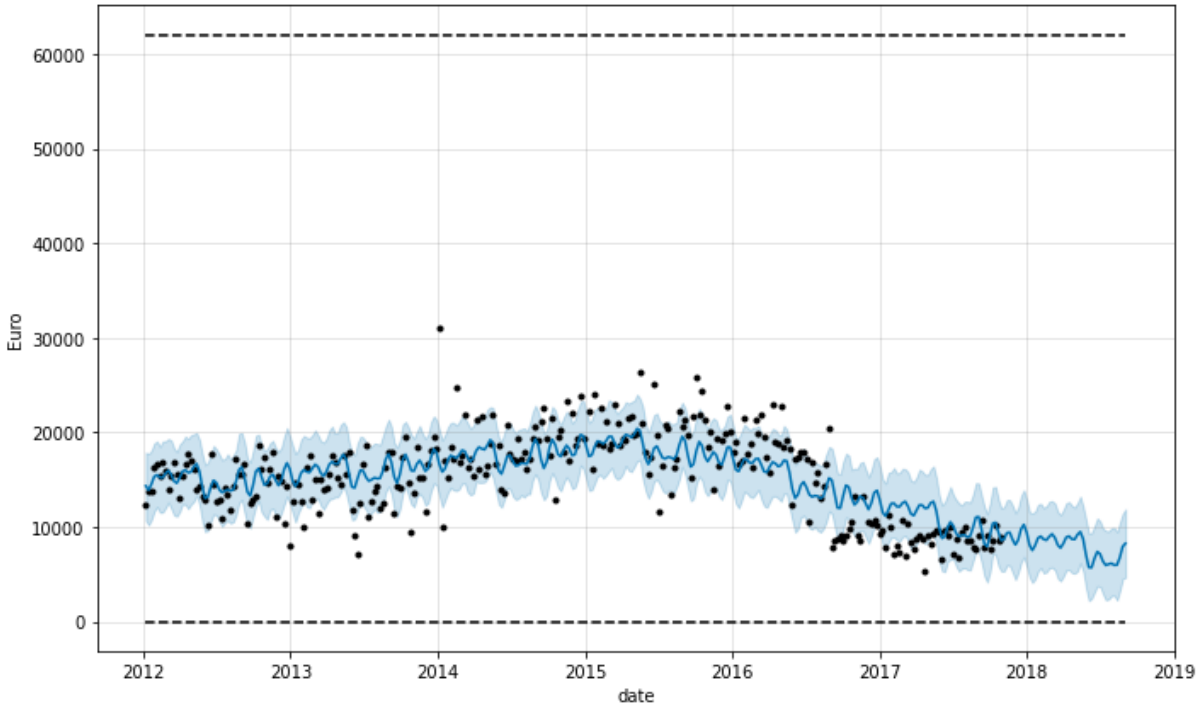
Linear and logistic model



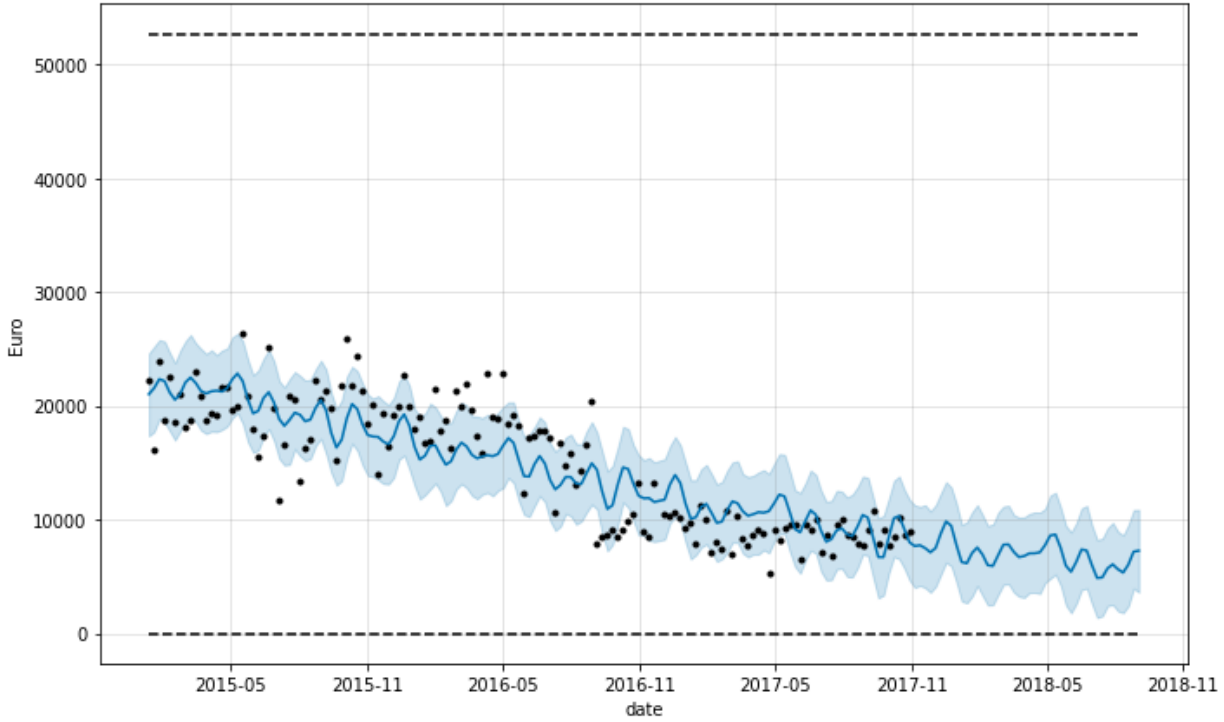


Weekly sales volume (in Euro's) for chosen store in scope for all products.

Logistic model

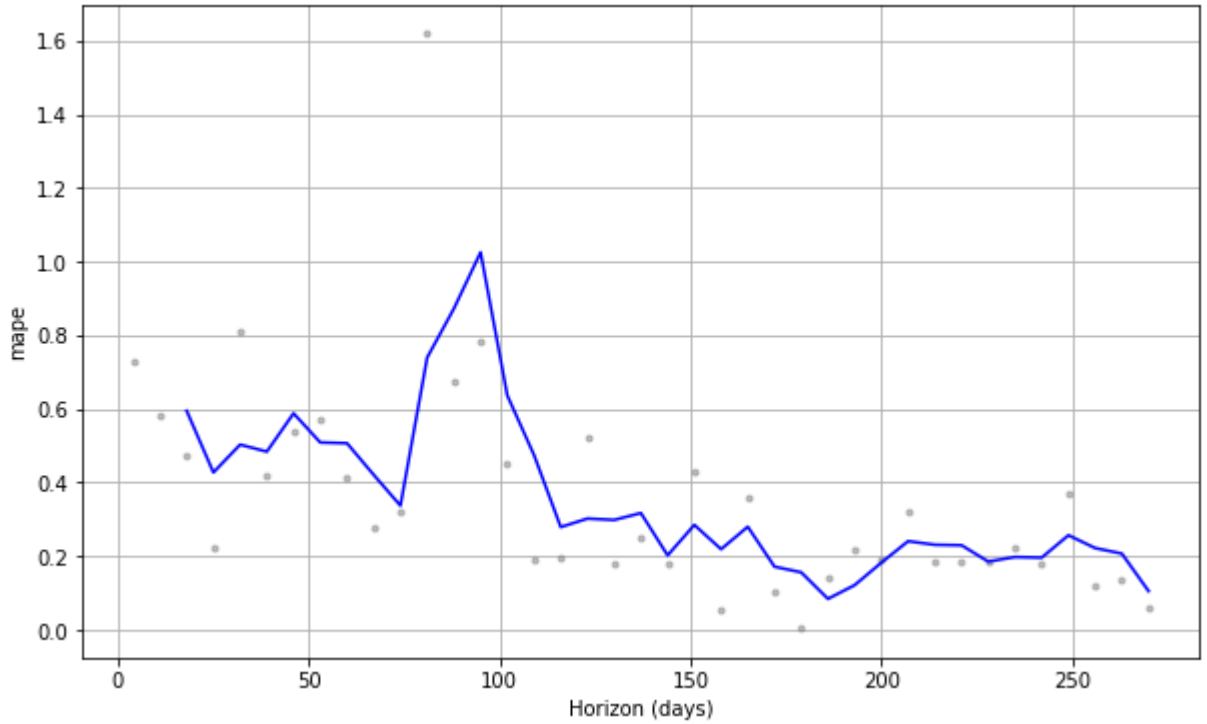


Long term



Short term

Logistic model



	ds	yhat	yhat_lower	yhat_upper	y	cutoff
0	2017-02-05	12162.989849	8394.608666	15872.575036	7045.34	2017-02-01
1	2017-02-12	12685.330166	8814.235970	16666.447738	8012.02	2017-02-01
2	2017-02-19	10859.415055	6997.236242	14567.385923	7362.05	2017-02-01
3	2017-02-26	13113.931197	9195.357805	16756.076362	10714.25	2017-02-01
4	2017-03-05	12465.775646	8600.902722	16058.511970	6886.44	2017-02-01

MAPE (Mean Absolute Percentage Error) in the time horizon

Cutoff 2017-02-01

How to analyse the competitiveness of a given store?

Find stores with similar features

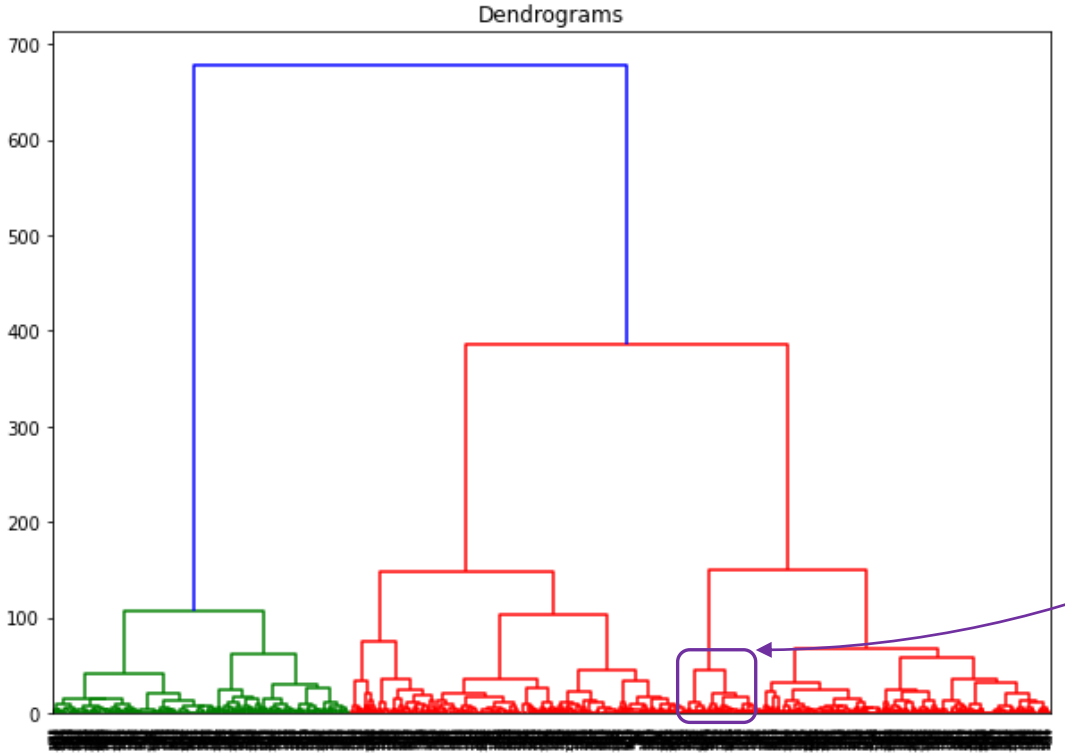
Check how many competitors are in the same group for each store

Compare the competition in the same group with the median/medium for the market

How to find similar stores?

	university or college	foodstores or supermarkets or gorceries	restaurant	churches	gym	stadium	store_id
0	0	0	4	4	0	0	1856
1	0	0	20	15	3	0	1857
2	0	0	4	6	0	1	1858
3	0	0	0	4	0	0	1859
4	0	0	1	0	0	0	1860

How to find similar stores?



Stores with similar parameters

How to find similar stores?

	university or college	foodstores or supermarkets or gorceries	restaurant	churches	gym	stadium	store_id	cluster
26	5	7	20	20	20	3	0	135
65	5	7	20	20	20	3	39	135
66	5	7	20	20	20	3	40	135
210	4	6	20	20	20	3	184	135
301	4	6	20	20	20	3	275	135
326	4	6	20	20	20	3	300	135
374	5	7	20	20	20	3	348	135
441	4	7	20	20	20	3	415	135
549	5	7	20	20	20	3	523	135
550	5	7	20	20	20	3	524	135
627	4	7	20	20	20	3	601	135
808	4	7	20	20	20	2	782	135

Position of the store against the competitors

Number of stores in the same cluster	Median of the competitor stores (within 5km radius from each store) in the same cluster	Number of competitor stores (within 5 km radius from each store) of chosen store
33	41	28

Interpretation

If the level of competition is lower than for the median/average, this means that the store potentially has less risk of competition.

Recommendations

Long (+4 years) and short (1-2 year) term analysis

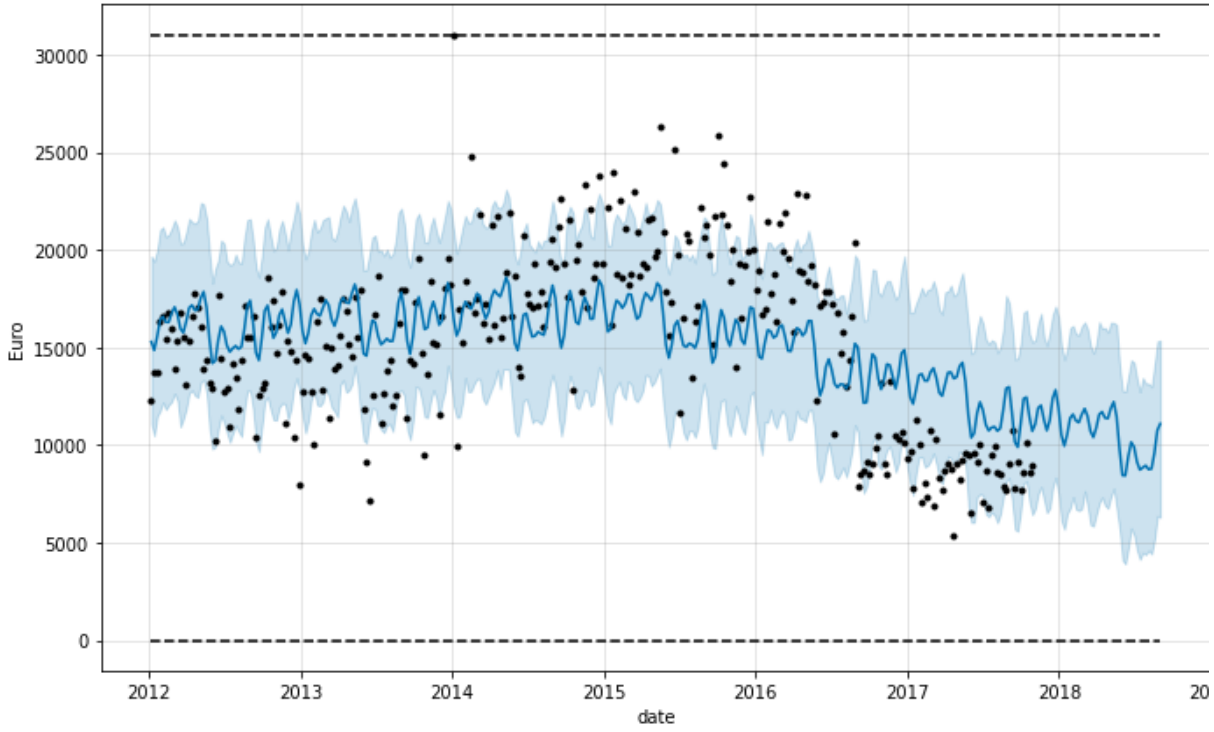
The recommended period to forecast (for store sales) is at least 2 years

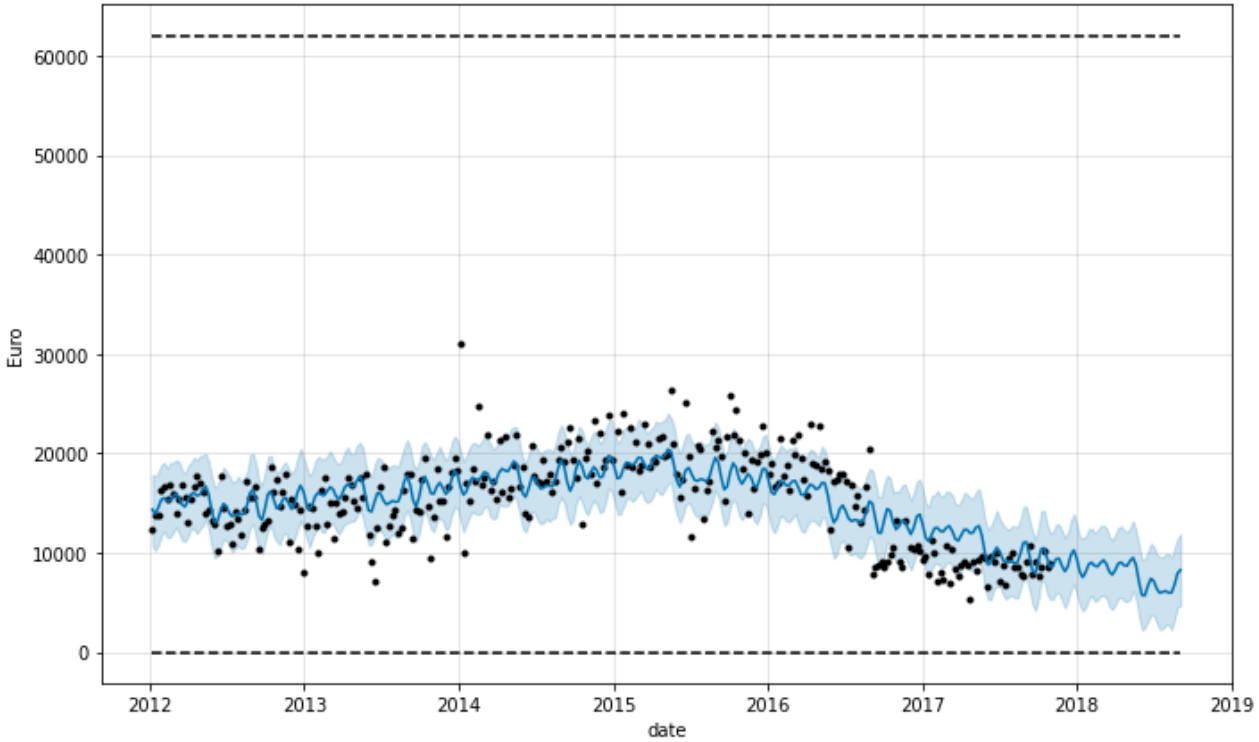
Additional regressors could improve forecasted data

Clustering is a good approach to finding similar stores

End of the presentation

Bonus slide: different cap values





Bonus slide: cross validation