Aalto-yliopisto
Perustieteiden korkeakoulu

Recognizing campaigns that cause cannibalization (presenting the results)

Petra Huttunen

11.06.2018

Instructor: *Tuomas Viitanen, TkT, Relex Solutions*

Supervisor: *Fabricio Oliveira*

Systeemianalyysin laboratorio

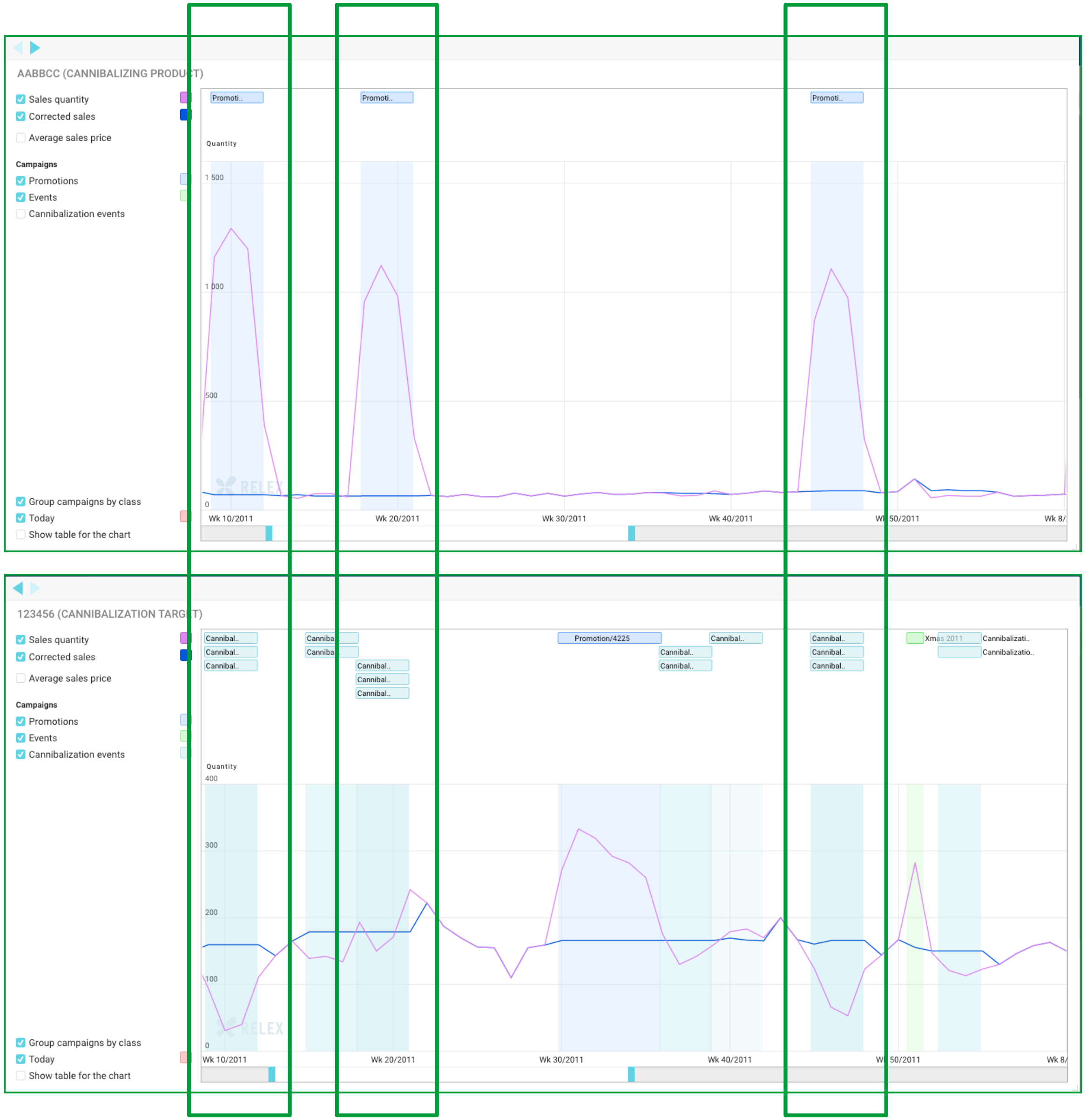
Cannibalization

A phenomenon where one product diverts sales from a substitute product, causing the sales of the substitute product to drop.

(Copulsky, 1976)

Research questions

- Which promotions cause cannibalization?
- Can we predict before the product is promoted, whether that promotion will cause cannibalization or not?


AABBCC (CANNIBALIZING PRODUCT)
Sales quantity
Corrected sales
Average sales price
Campaigns
Promotions
Events
Cannibalization events
Group campaigns by class
Today
Show table for the chart
Quantity
Wk 10/2011
Wk 20/2011
Wk 30/2011
Wk 40/2011
Wk 50/2011
123456 (CANNIBALIZATION TARGET)
Promotion/4225
Xmas 2011

Findings from literature

- Products with similar attributes attract similar customers and can therefore divert sales from each other. (Mason and Milne, 1994)
- If one product is made more appealable to the customer because of a promotion, the customer can choose that product over another and cannibalization will occur. (Dawes, 2012)
- Promotions have different effects on product sales based on marketing, how good the offer is and how much the price is discounted. (González-Benito et al, 2010)

Method

- Dataset contained information about the campaign and cannibalizing product, cannibalization event and cannibalization relationship
- The goal was to fit a regression model to the data
  - Dependent variable: additional sales of cannibalization event
- Data was real customer data from 2013 to 2016
  - 2013-2015 was used as analysis period
  - Data from 2016 was predicted by using the model fitted to analysis data
- Product groups used in analysis were beef, chicken, coca cola and frozen potatoes

Regression models in general

- A regression model consists of three parts
  - Dependent variable y
  - Systematic part of the model $f(x;\beta)$
    - Function of the independent variable(s) x
  - Residuals ε of the model
- The goal is to choose parameter β so that the residuals are as small as possible
  - One option is to use ordinary least squares method

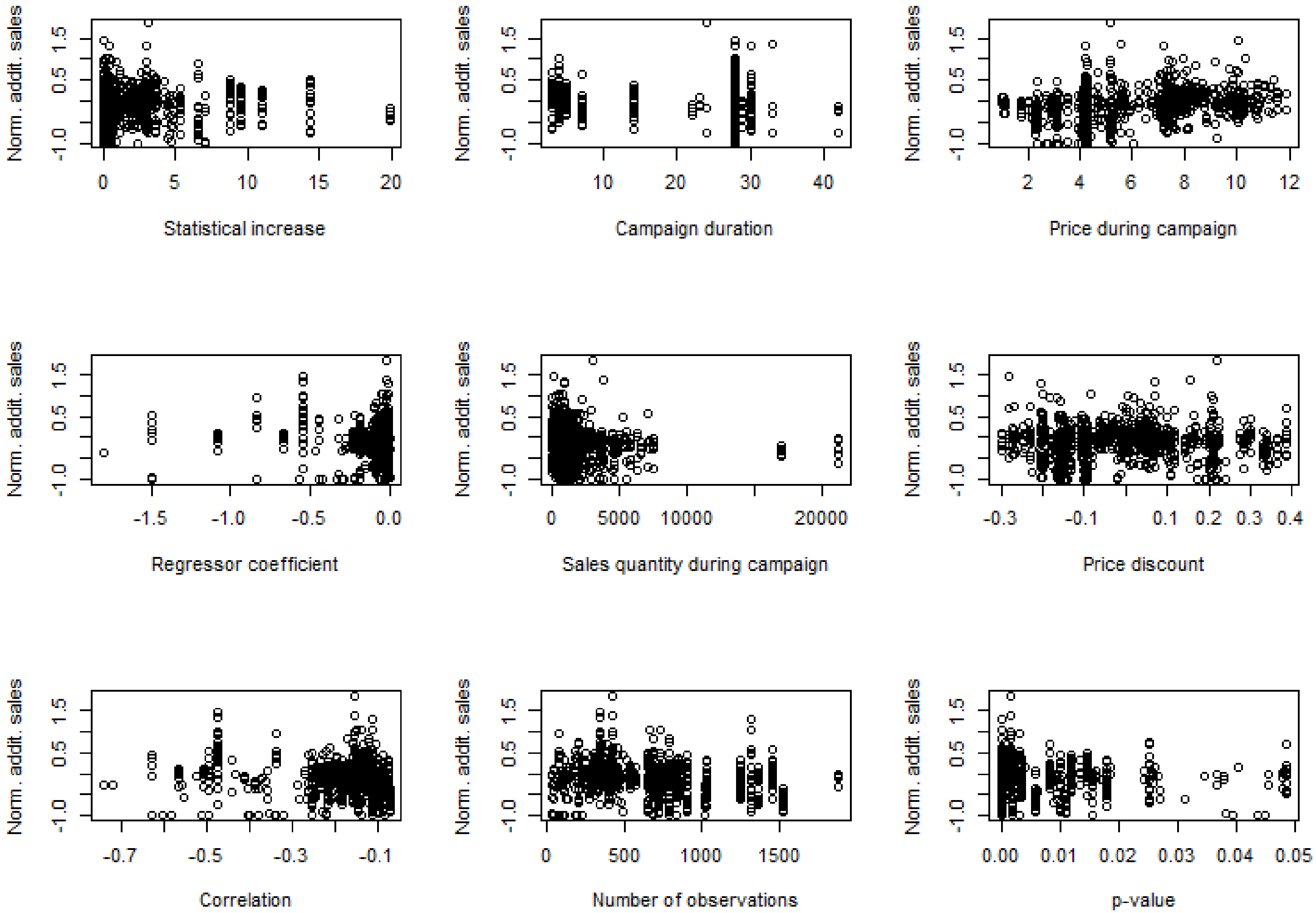
$$y = f(x;\beta) + \varepsilon$$

$$\min_{\beta_0 \ldots \beta_k} \sum_{i=1}^{n} \varepsilon_i = \sum_{i=1}^{n} (y - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_k x_{ik})^2$$

Independent variables

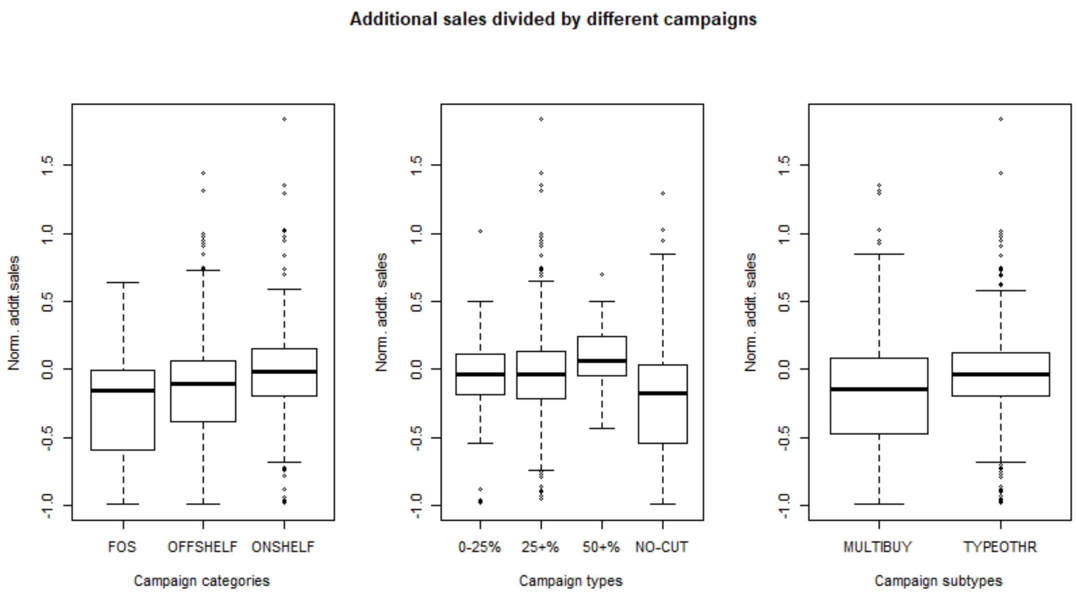
- Product group
- Campaign types
  - Category, type, subtype
- Sales metrics
  - Sales quantity, baseline sales, additional sales
- Statistical increase of campaign (=relative sales increase)
- Campaign duration
- Price metrics
  - Price discount, price during campaign, price before campaign
- Information about the relationship
  - Regressor coefficient, correlation, p-value…

Correlations between the dependent and independent variables



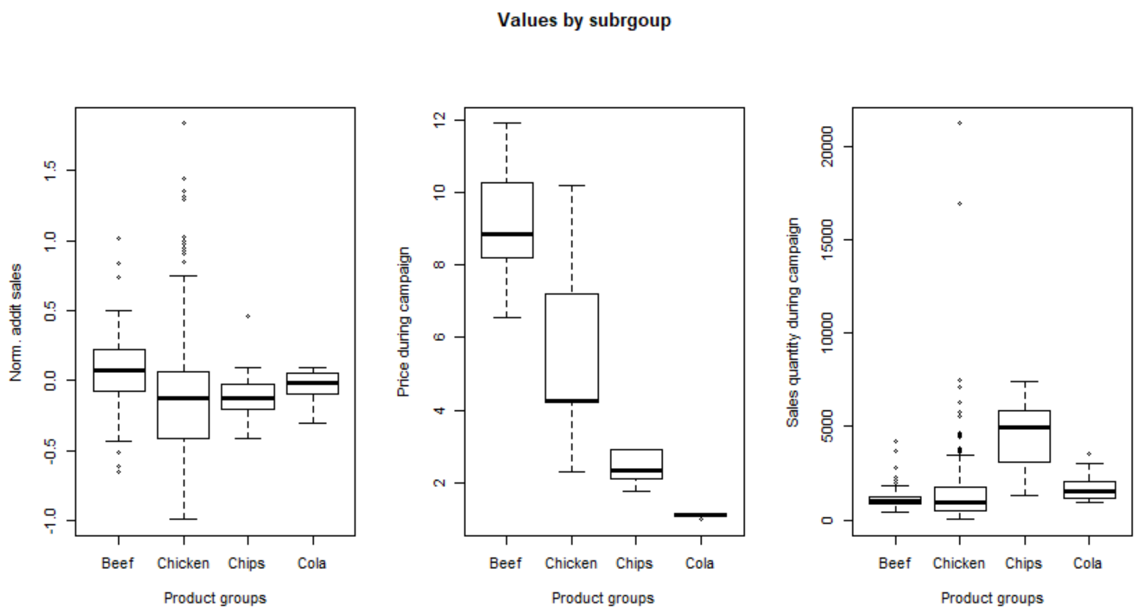
Categorical variables

- Can we distinguish differences between campaigns in terms of cannibalization?
  - Campaigns seem to act in a very similar way



Categorical variables

- How do the product groups differ in terms of sales and prices?
  - Chicken seems to have biggest variation



Fitting the regression model

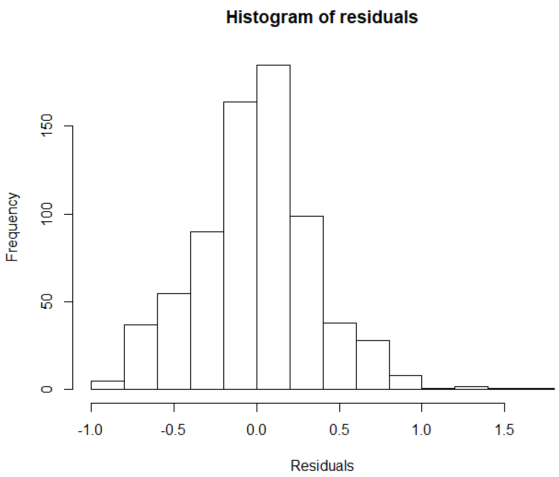
- Model was first complied with all of the variables
  - Variables were eliminated based on significance and multicollinearity
- Linear variables and non-linear variables
- Log-model
- Normalised vs. real data

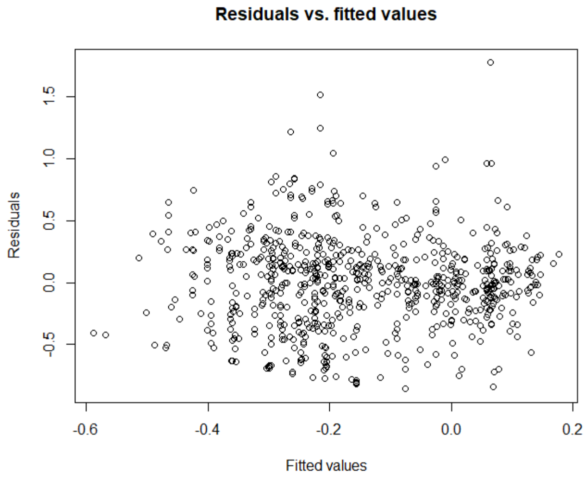
Fitted regression model

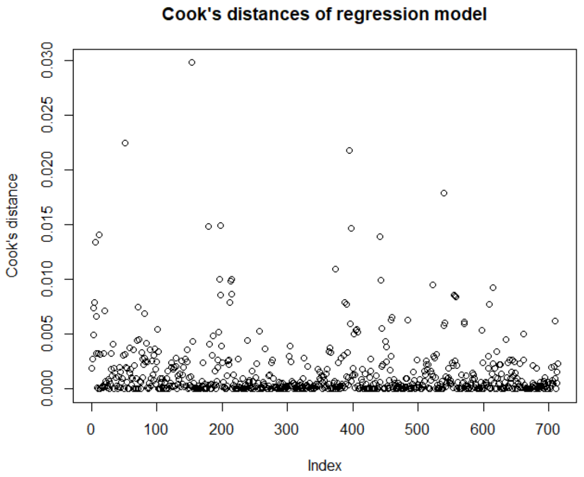
Coefficient	Estimate	P-value
Intercept	0.25	0.0034
Category: offshelf	0.065	0.13
Category: onshelf	0.20	3.13e-5
Statistical increase	0.02	1.80e-6
Campaign duration	-0.0058	0.0010
Correlation	1.1	9.72e-7
Number of observations	-0.00029	2.44e-9
P-value	-3.4	0.036

Goodness of the model

- Adjusted R-squared 15.63%
- All VIF-values were under 2
- No significant outliers (based on Cook’s distance)

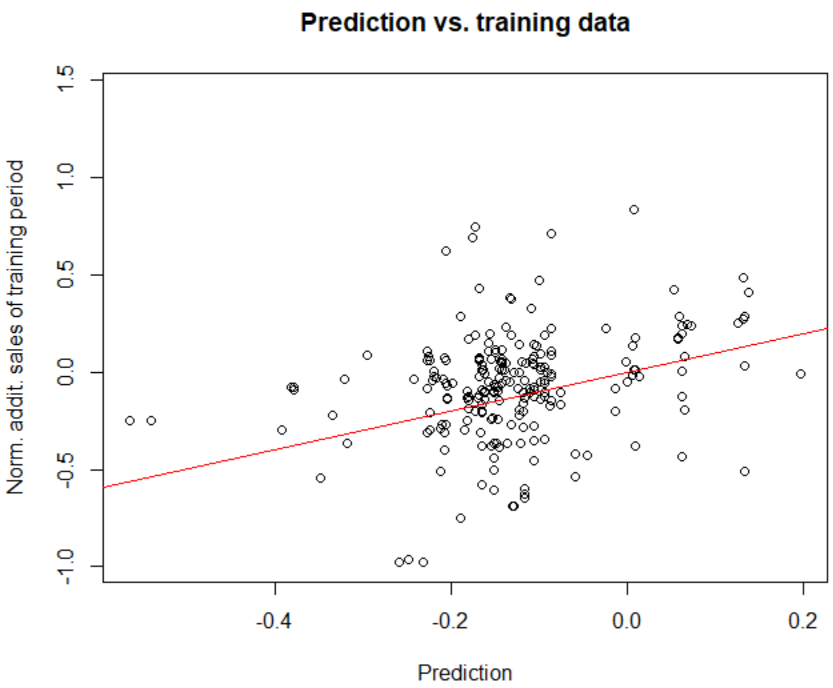






Prediction of the model

- The model was used to predict the normalised additional sales of training data
- The results were compared with training data
- Dataset contained a lot of noise



Issues with dataset

- Data contained noise and exceptions
- Exceptions were mostly caused by
  - Holiday events
  - Overlapping campaigns
  - Accuracy of baseline sales
    - Back-to-back campaigns
- Not all exceptions could be cleaned from the data

Conclusions

- The regression model had a low coefficient of determination
  - The model build based on analysis data didn’t explain the training data
- Confirmation that certain attributes do have an affect on cannibalization
  - Campaign categories
  - Strength of the cannibalization relationship
  - Statistical increase and campaign duration were interesting
- Other significant variables in various models
  - Price during campaign
  - Sales quantity
  - Subgroups

Future prospects

- Missing some possibly important information
  - Marketing plans for campaigns
  - Brand and quality of products
- Cannibalization was recognized only within product groups

References

- Copulsky, W., 1976. Cannibalism in the Marketplace. *The Journal of Marketing*, pp.103-105.
- Mason, C.H. and Milne, G.R., 1994. An approach for identifying cannibalization within product line extensions and multi-brand strategies. *Journal of Business Research*, *31*(2-3), pp.163-170.
- Dawes, J.G., 2012. Brand-pack size cannibalization arising from temporary price promotions. *Journal of Retailing*, *88*(3), pp.343-355.
- González-Benito, Ó., Loyola-Galván, Z.I. and Muñoz-Gallego, P.A., 2010. Inter-size and inter-brand competition analysis within a product category: Scope of cannibalization effects. *Journal of Brand Management*, *17*(4), pp.254-265.