



Aalto-yliopisto  
Perustieteiden  
korkeakoulu

# Determining cannibalization rates from sales data (topic presentation)

*Olli Herrala*

*11.04.2018*

*Advisor: Mikko Ervasti, Sellforte Solutions Ltd.*

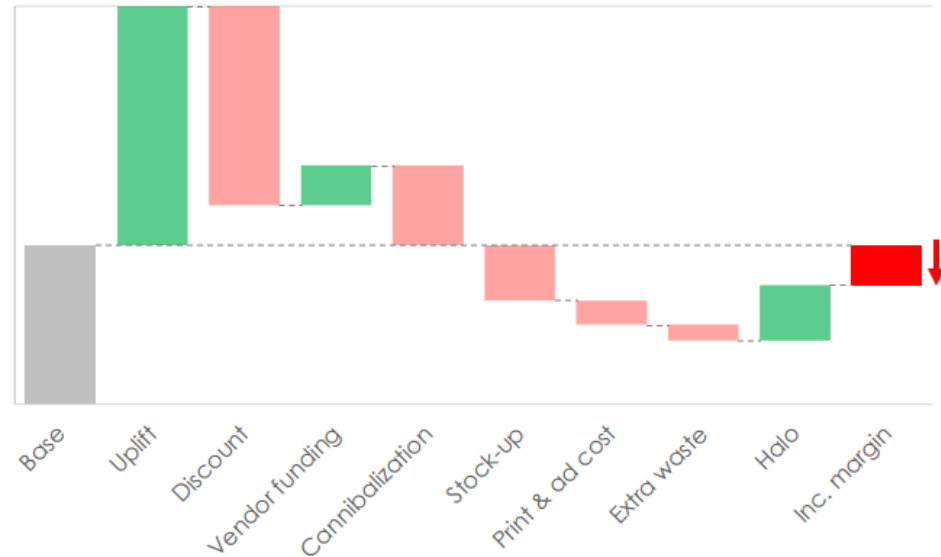
*Supervisor: Fabricio Oliveira*

Työn saa tallentaa ja julkistaa Aalto-yliopiston avoimilla verkkosivuilla. Muilta osin kaikki oikeudet pidätetään.

# Background

- Extremely price driven retail landscape in Finland
- Promotion efficiency
  - Looking at uplift and discount is not enough

Incremental margin (€M)



[https://www.sellforte.com/pdf/Sellforte\\_Stockmann\\_Delicatessen\\_Promotion\\_Effectiveness\\_Case\\_Study\\_2017.pdf](https://www.sellforte.com/pdf/Sellforte_Stockmann_Delicatessen_Promotion_Effectiveness_Case_Study_2017.pdf)

- Some of the promotion uplift comes from decreased sales in other similar products. This is known as cannibalization

# Goals

- An overview of previous methods
- An insight on how strongly different products cannibalize each other on item level (as opposed to e.g. category level)
- A robust method that can handle simulated data with added noise

# Definitions

- Only cannibalization and complementarity (“negative cannibalization”) considered, stock-up and seasonality not considered
- The method is not intended to work on data with multiple consecutive and simultaneous promotions, as such method would be significantly more difficult to implement

# Sources

- An Implemented System for Improving Promotion Productivity Using Store Scanner Data (Abraham & Lodish, 1993)
- Actual customer data for developing the methods
- Simulated data replicating the customer data for the thesis results
- Coworkers at Sellforte

# Methods and tools

- Pandas library in Python is used to process the data
- Form a matrix  $C$ , where elements  $C_{ij}$  specify cannibalization rates for product  $j$  cannibalizing product  $i$ 
  - Determine weekly promotion uplift  $U$  and cannibalization  $D$  from data by comparing to last week
  - Assume form  $UC = D$
  - Least squares approximation for  $C$

# Schedule

- 2-3/2018: familiarization with tools and data
- 11.4.2018: topic presentation
- 4-5/2018: building and validating the model
- 11.6.2018: final presentation
- 6-7/2018: writing the thesis