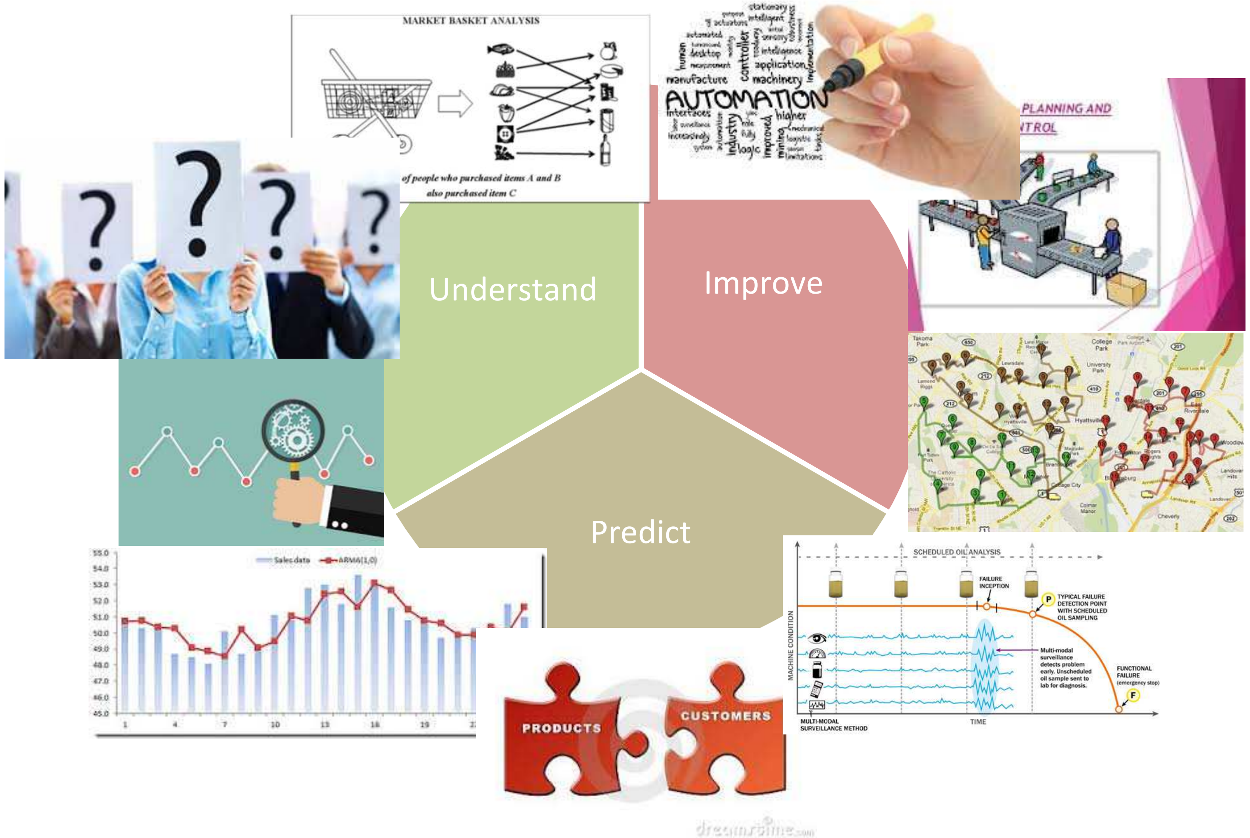
Customer Intelligence in der Praxis

IT-Logix AG
Sotiris Dimopoulos, PhD

Zürich, 11. September 2018

Data Science & AI : Possibilities?



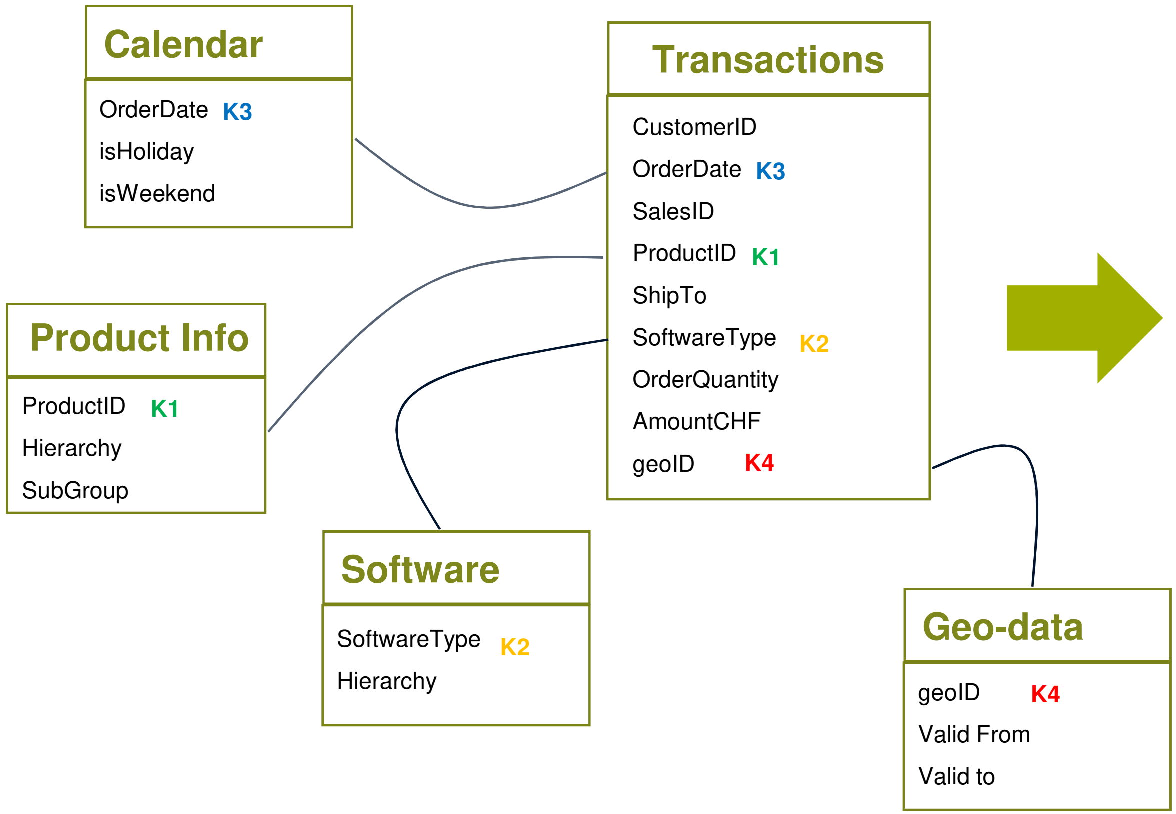
Outlook

- Customer Dimensions
  - ❑ Use Machine Learning to extract new customer dimensions
- Customer Geo-Analytics
  - ❑ Leverage geo-data to understand your customers
- Predictive Customer Analytics
  - ❑ Predict quantities for your customers

Customer Dimensions

Customer Dimensions

- Applicable to: Online shops / Non-profit organizations / B2C
- Goal: Produce new customer dimensions based on transactional data



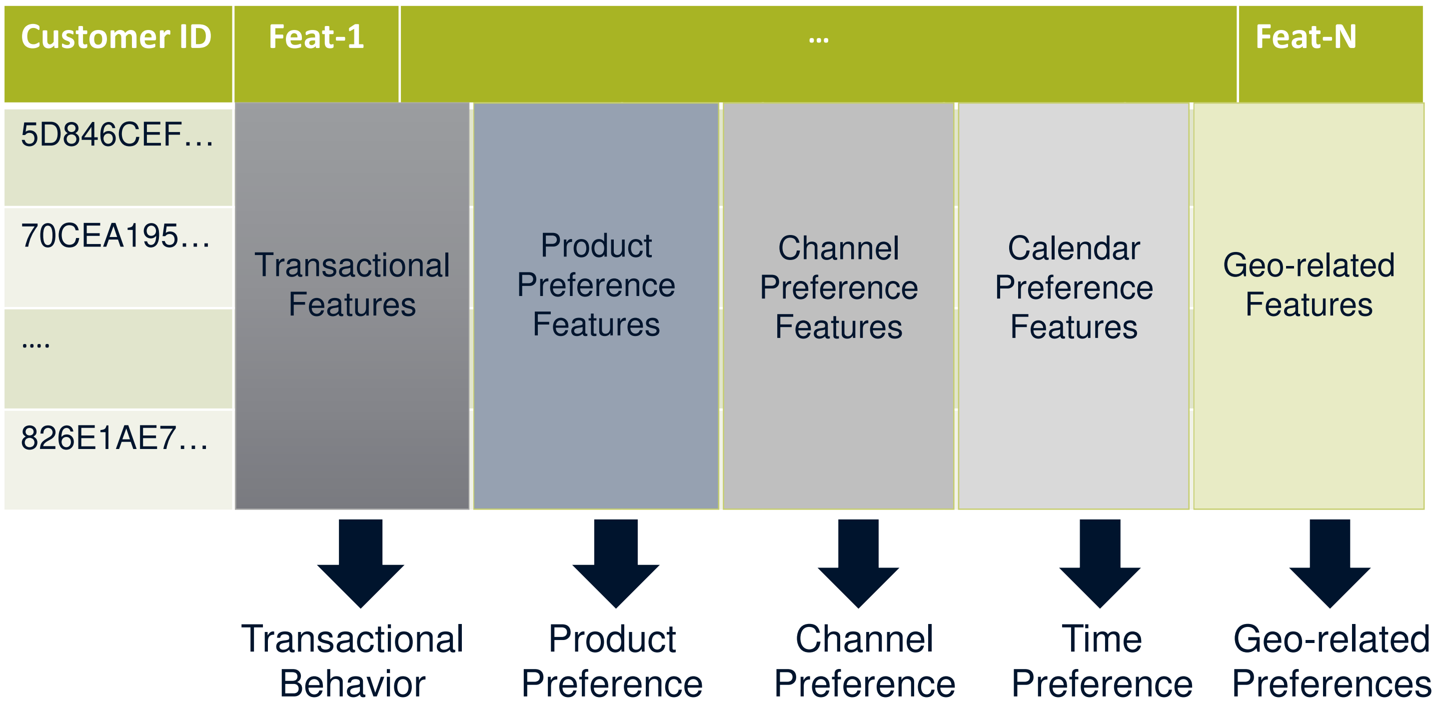
Customer ID	feat1	…	featN
5D846CEF…			
70CEA195…			
….			
826E1AE7…			

- Table / One row for each customer
- More dimensions for each customer

BOAK

Customer Dimensions

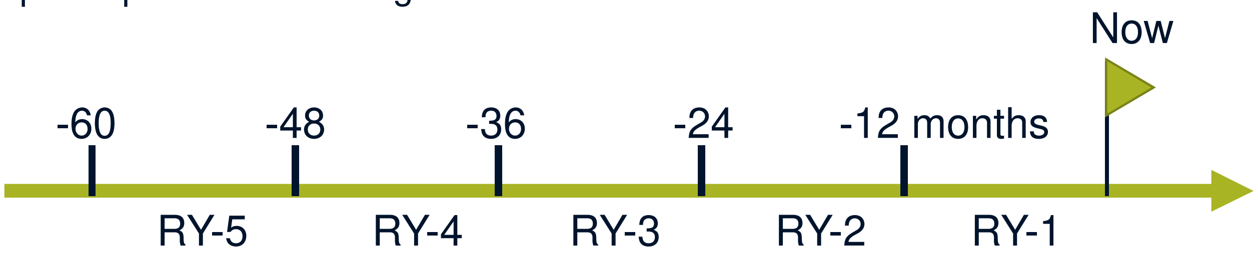
- Extraction of more than 100 customer dimensions
- Great variety of customer dimensions



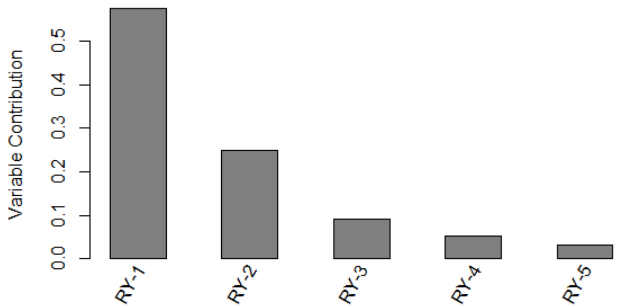
BOAK

Customer Dimensions: Examples

- Example 1: predictive scoring



CUSTOMER	32 CHF	0 CHF	45 CHF	35 CHF	98 CHF

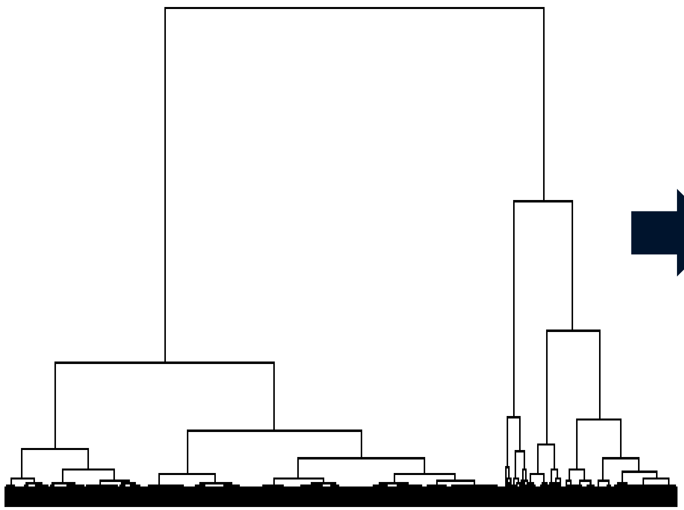


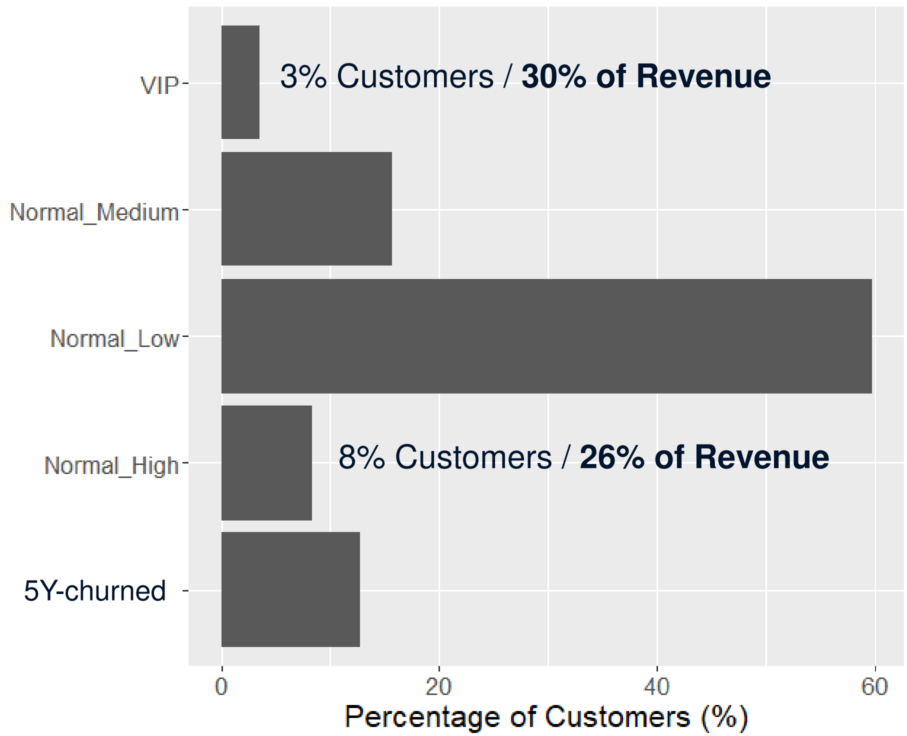
Rolling Year	Weight
Rolling Year 1	0.577
Rolling Year 2	0.249
Rolling Year 3	0.091
Rolling Year 4	0.053
Rolling Year 5	0.030

CUSTOMER SCORE	70.316

Customer Dimensions: Examples

- Example 2: Classes

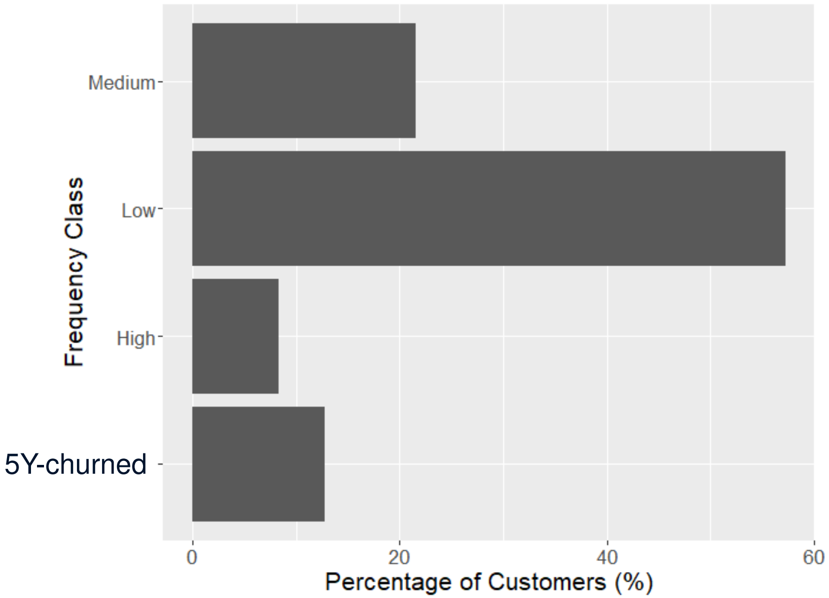


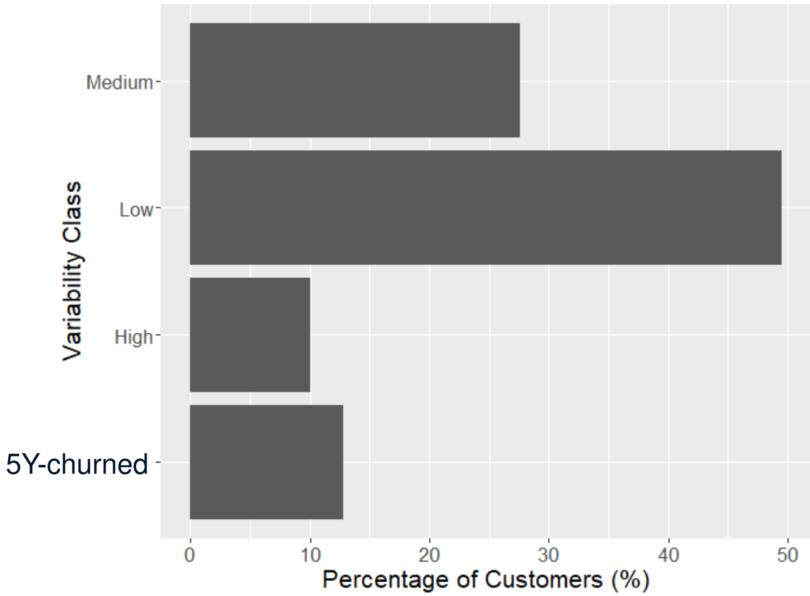


BOAK

Customer Dimensions: Examples

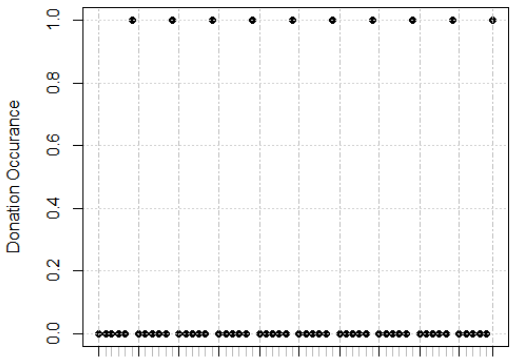
- Example 2: Classes



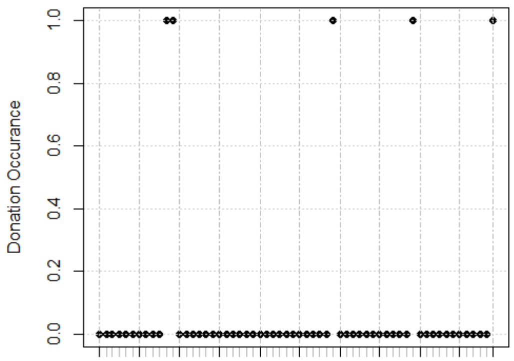


Customer Dimensions: Examples

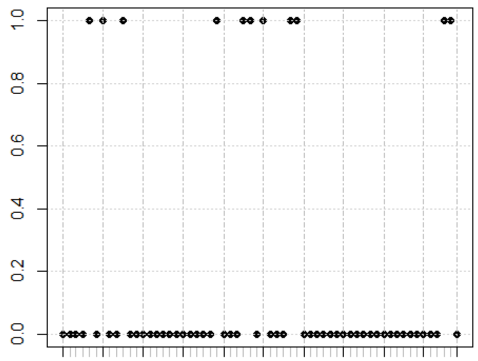
- Example 3: Periodicity



Periodicity Class	Customer Base (%)
Strong	0.5%



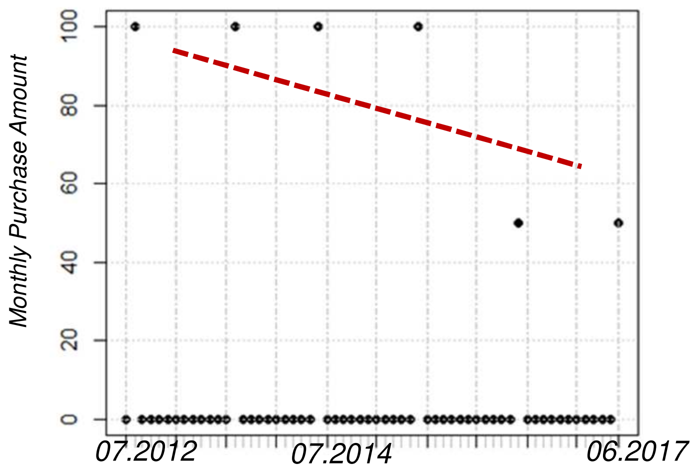
Periodicity Class	Customer Base (%)
Weak	5%

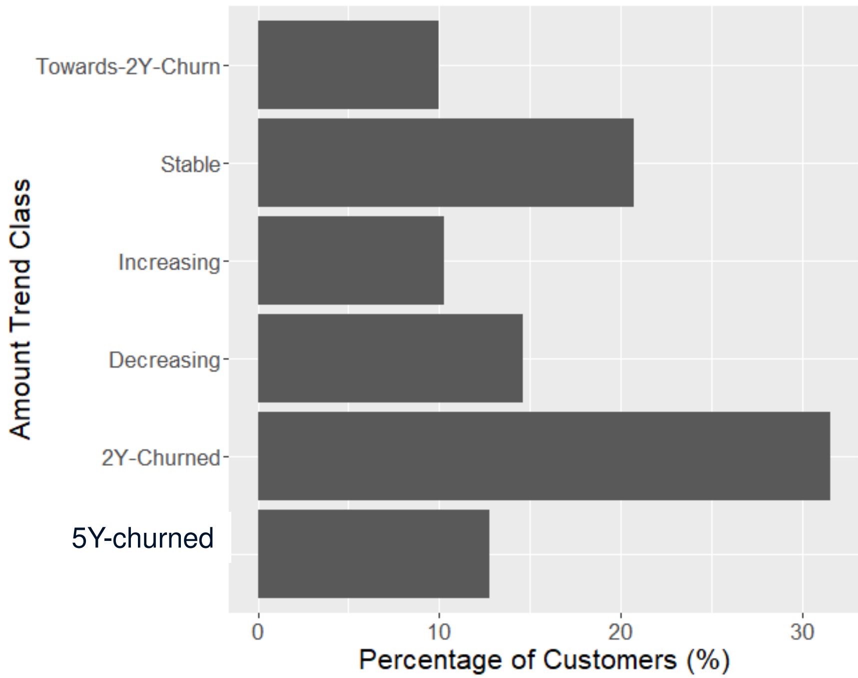


Periodicity Class	Customer Base (%)
Aperiodic	82%

Customer Dimensions: Examples

- Example 4: Trends





BOAK

Customer Dimensions: Use

- Customer Behavior Table

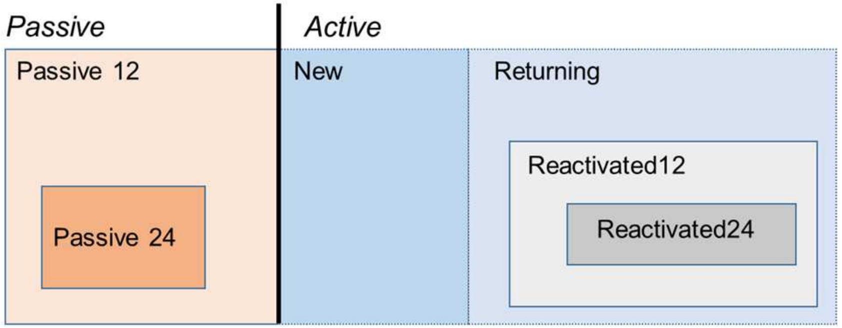
<table>
<tr><th>Customer ID</th><th>Feat-1</th><th colspan="3">…</th><th>Feat-87</th></tr>
<tr><td>5D846CEF…</td><td rowspan="4">Transactional Features</td><td rowspan="4">Product Preference Features</td><td rowspan="4">Channel Preference Features</td><td rowspan="4">Calendar Preference Features</td><td rowspan="4">Geo-related Features</td></tr>
<tr><td>70CEA195…</td></tr>
<tr><td>….</td></tr>
<tr><td>826E1AE7…</td></tr>
</table>

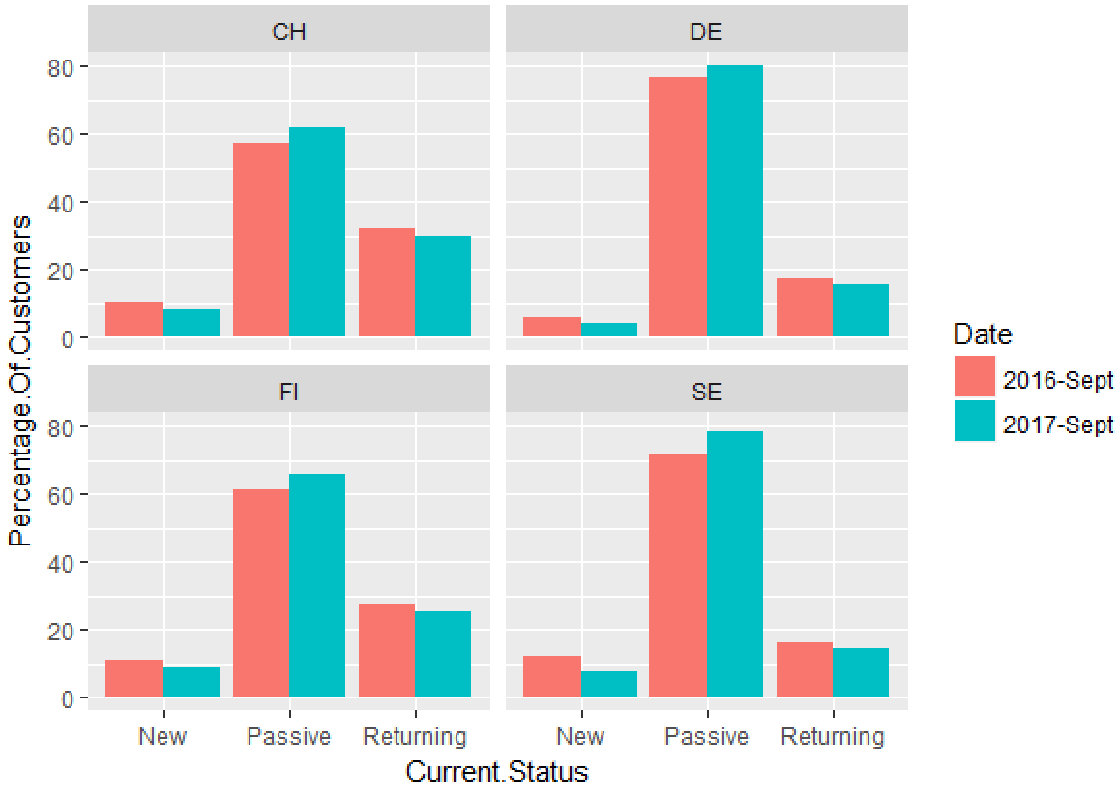
- Descriptive statistics
  - E.g. for reports/dashboards/pivot tables, etc
- Customer segments
  - Create customer groups on-the-fly, monitoring
- Input for predictive models
  - e.g. churn, retention, offer acceptance, promotion of trips, etc

BOAK

Customer Dimensions: Descriptive Statistics

- Descriptive statistics





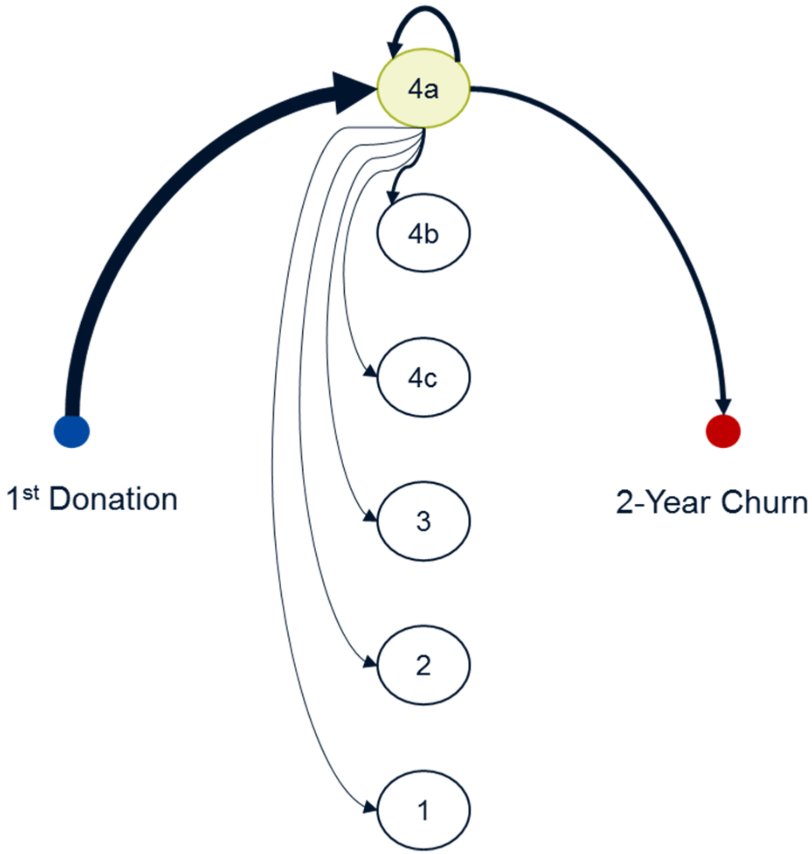
Customer Dimensions: Customer segments

- Customer Segments

ID	Features	Population (%)	Revenue (%)
4a	Amount_Group= Normal_low Nr. Of Reactions 5Y = 1	30%	5.5%
4b	Amount_Group= Normal_low Nr. Of Reactions 5Y > 1 Frequency = Low Periodicity = Aperiodic	33%	14%
4c	Amount_Group= Normal_Low Nr. Of Reactions 5Y > 1 Frequency = Medium, High | Periodicity = weak, strong	22%	14%
3	Amount_Group = Normal_Medium	13%	28%
2	Amount_Group = Normal_High	2%	10%
1	Amount.Class = Special	1%	30%

ID	Features	Population	Revenue (%)
3p	Periodicity = Strong,Weak Frequency = Low	2.6%	5.4%

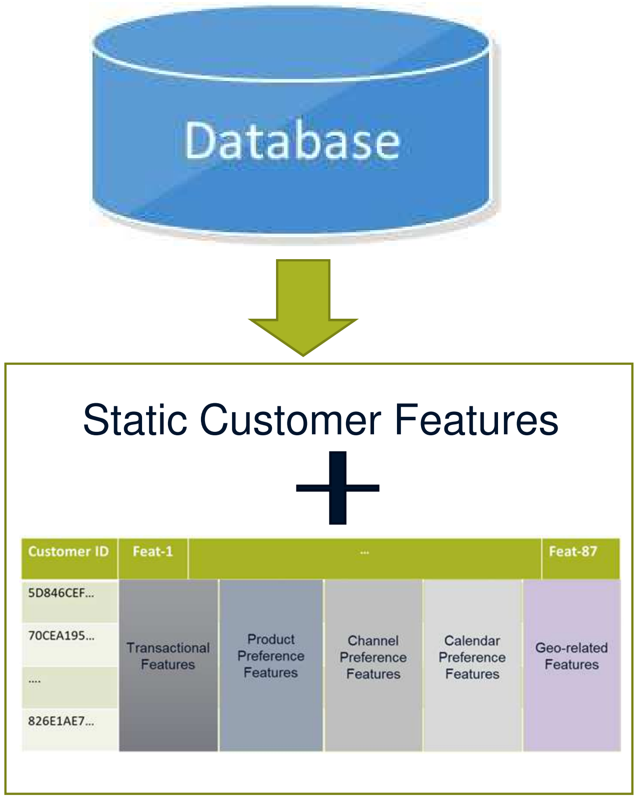
ID	Features	Population	Revenue (%)
3f	Frequency = Medium, High	7.2%	15 %



BOAK

Customer Dimensions: Predictive models

- Predictive models



Predictive Model for Active Customers

Performance:

Accuracy = 0.864

Specificity = 0.824

Sensitivity = 0.903

AUC = 0.913

Probability of retaining Active- / Passive-Status after 12 months

	A	B	C
1	CustomerIdentifier	Prob. Active	Prob. Passive
2	!avaratnam.karthiga@gmail.com	0.971922517	0.028077483
3	000_000@thebc.ch	0.970012963	0.029987037
4	007christophmueller@gmail.cor	0.962658048	0.037341952
5	007danny@web.de	0.721094251	0.278905749
6	00826chenxi@163.com	0.999678135	0.000321865
7	00r1994@gmail.com	0.945259809	0.054740191
8	01020017350@naver.com	0.998696387	0.001303613
9	01029102123@hanmail.net	0.429482877	0.570517123
10	01042young@naver.com	0.996441901	0.003558099
11	01047179203@hanmail.net	0.417009234	0.582990766
12	0125alex@gmail.com	0.361113429	0.638886571
13	01z0121184@gmail.com	0.998056054	0.001943946
14	03451224041-0001@t-online.de	0.417009234	0.582990766
15	0505angela@gmail.com	0.998741686	0.001258314
16	0510dorud@naver.com	0.998755455	0.001244545
17	0620607@gmail.com	0.761701345	0.238298655
18	0763693639@romandie.com	0.903373599	0.096626401

Customer Geo-Analytics

Geo-data in Switzerland

- BFS as a source of information

Regional Level

State (Kanton)

District (Bezirk)

Community (Gemeinde)

Hectare (100m x 100m)

Features / Attributes

- Age groups
- Housing Information
- Country of origin
- Marital status
- Language
- Work sectors
- Family and kids
- Religion
- Political opinions
- Commuting
- Urbanization
- Educational level
- Income
- etc

Geo-data in Switzerland

› BFS as a source of information

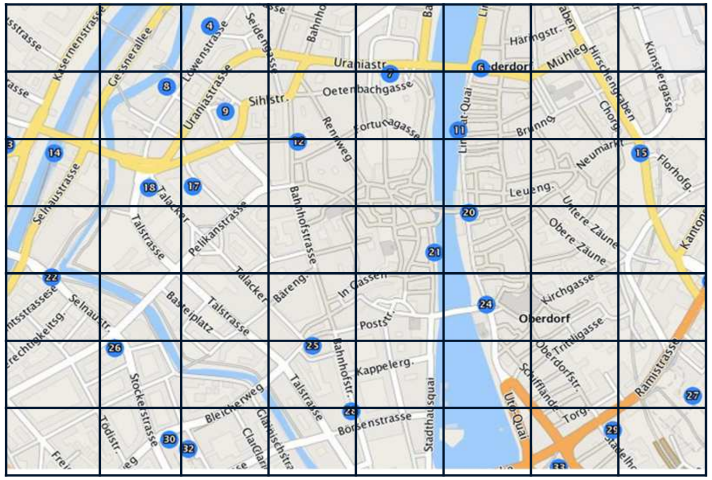
<u>Regional Level</u>

State
(Kanton)

District
(Bezirk)

Community
(Gemeinde)

Hectare
(100m x 100m)



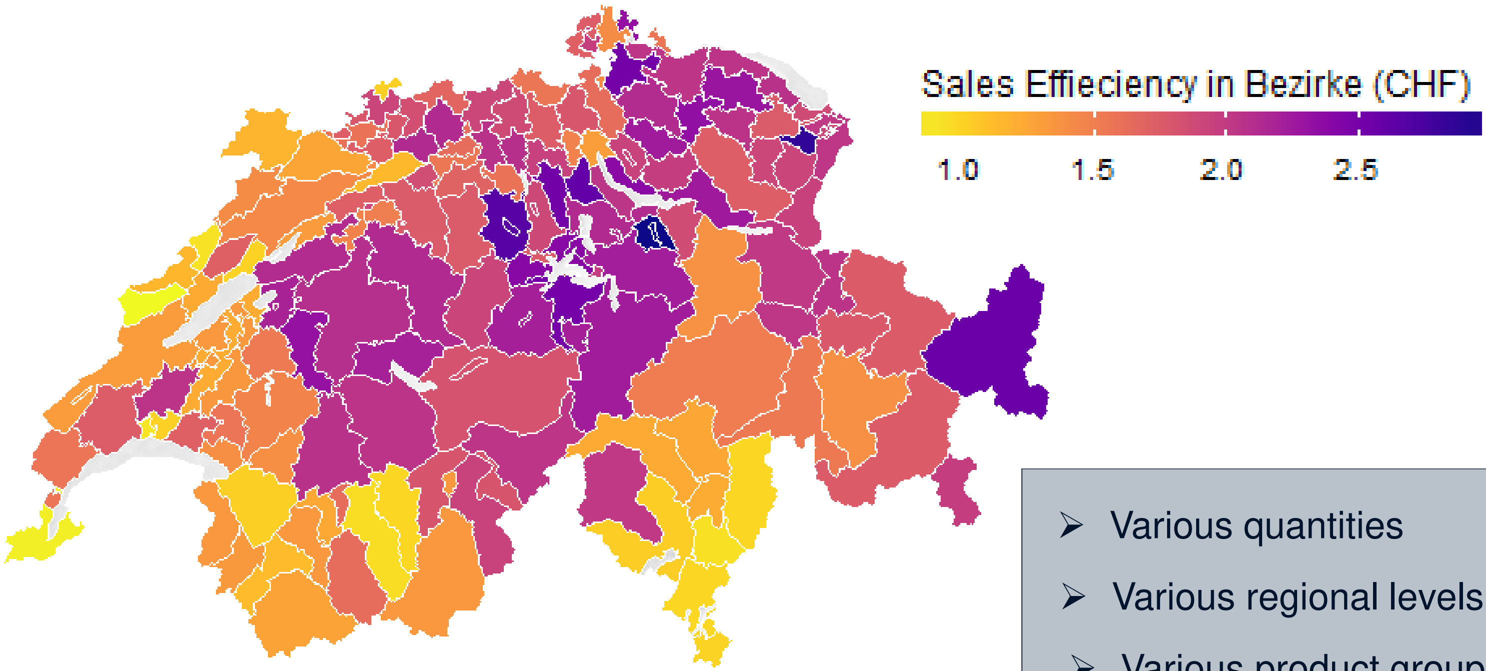
<u>Features / Attributes</u>

- Age groups
- Housing Information
- Country of origin
- Marital status
- Language
- Work sectors
- Family and kids
- Religion
- Political opinions
- Commuting
- Urbanization
- Educational level
- Income
- etc

BOAK

Customer Geo-Analytics: Results

- Quantification sales efficiency at the regional level

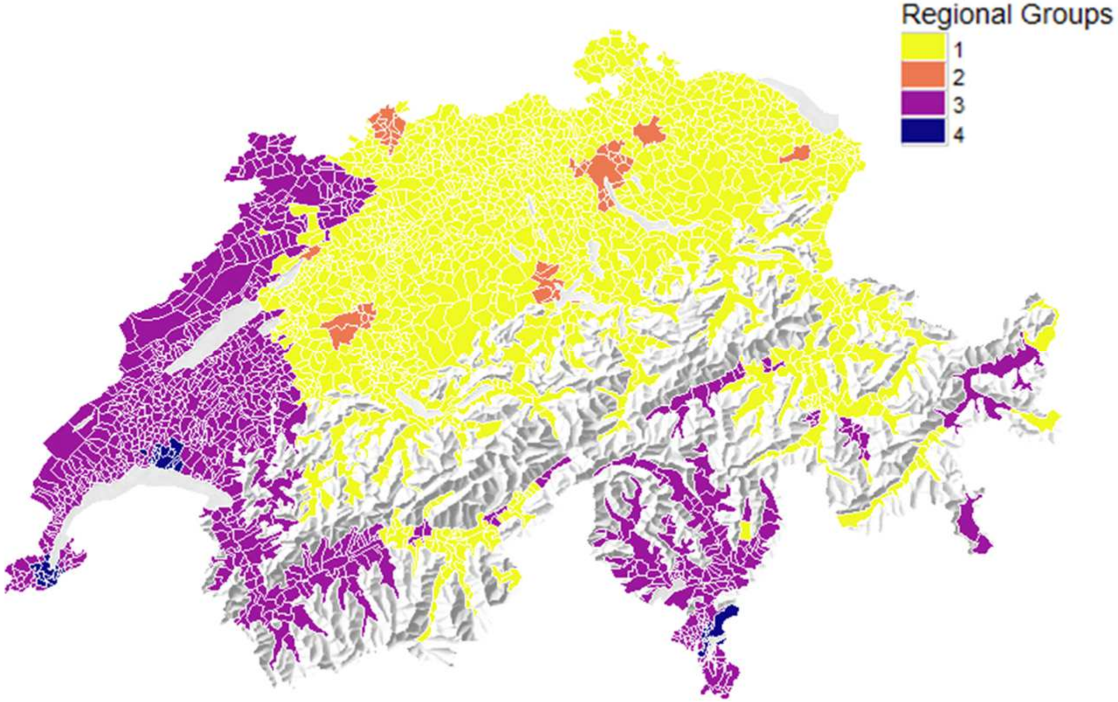


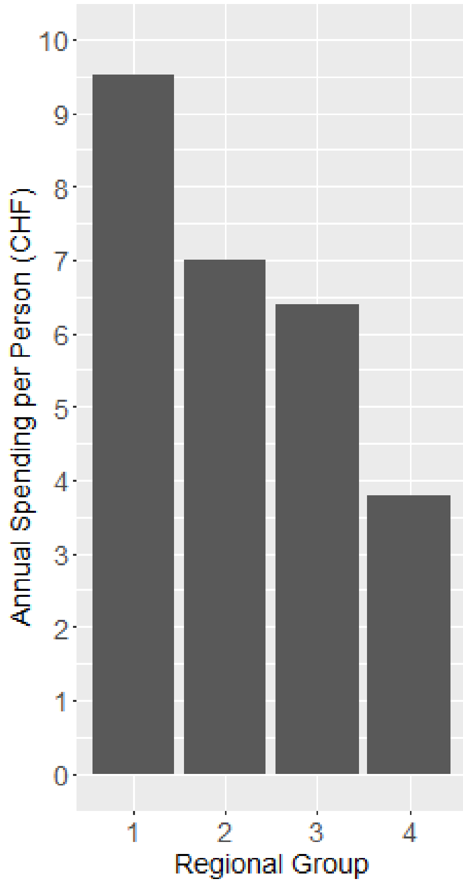
Annual Revenue Efficiency (= Annual Revenue / Population) of a product-group for all Bezirke in 2016

- Various quantities
- Various regional levels
- Various product groups

Customer Geo-Analytics: Results

- Regional grouping

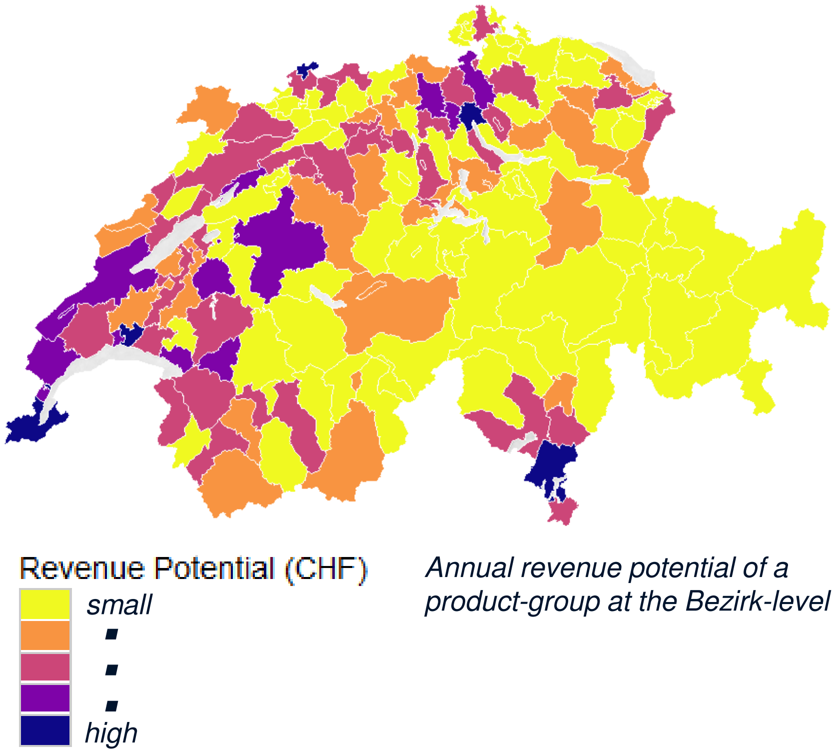




BOAK

Customer Geo-Analytics: Results

- Annual revenue potential



Annual revenue potential of a product-group at the Bezirk-level

Which regions should be targeted first?

Kanton	Revenue Potential (in '000s CHF)	Cumulative Percentage (%)
Vaud	569	14.2
Genève	558	28.1
Zürich	539	41.6
Ticino	347	50.3
…	…	…

Annual revenue potential of a product-group at the Kanton level

BOAK

Customer Geo-Analytics: Diversity of Geo-data in CH

- BFS as a source of information

Regional Level

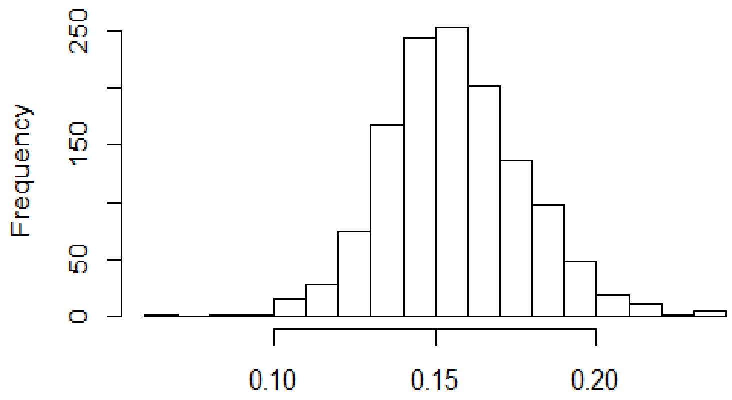
State (Kanton)

District (Bezirk)

Community (Gemeinde)

Hectare (100m x 100m)

Quota of persons of age 0-14 across political communities in CH



Diversity is the key to informativeness!

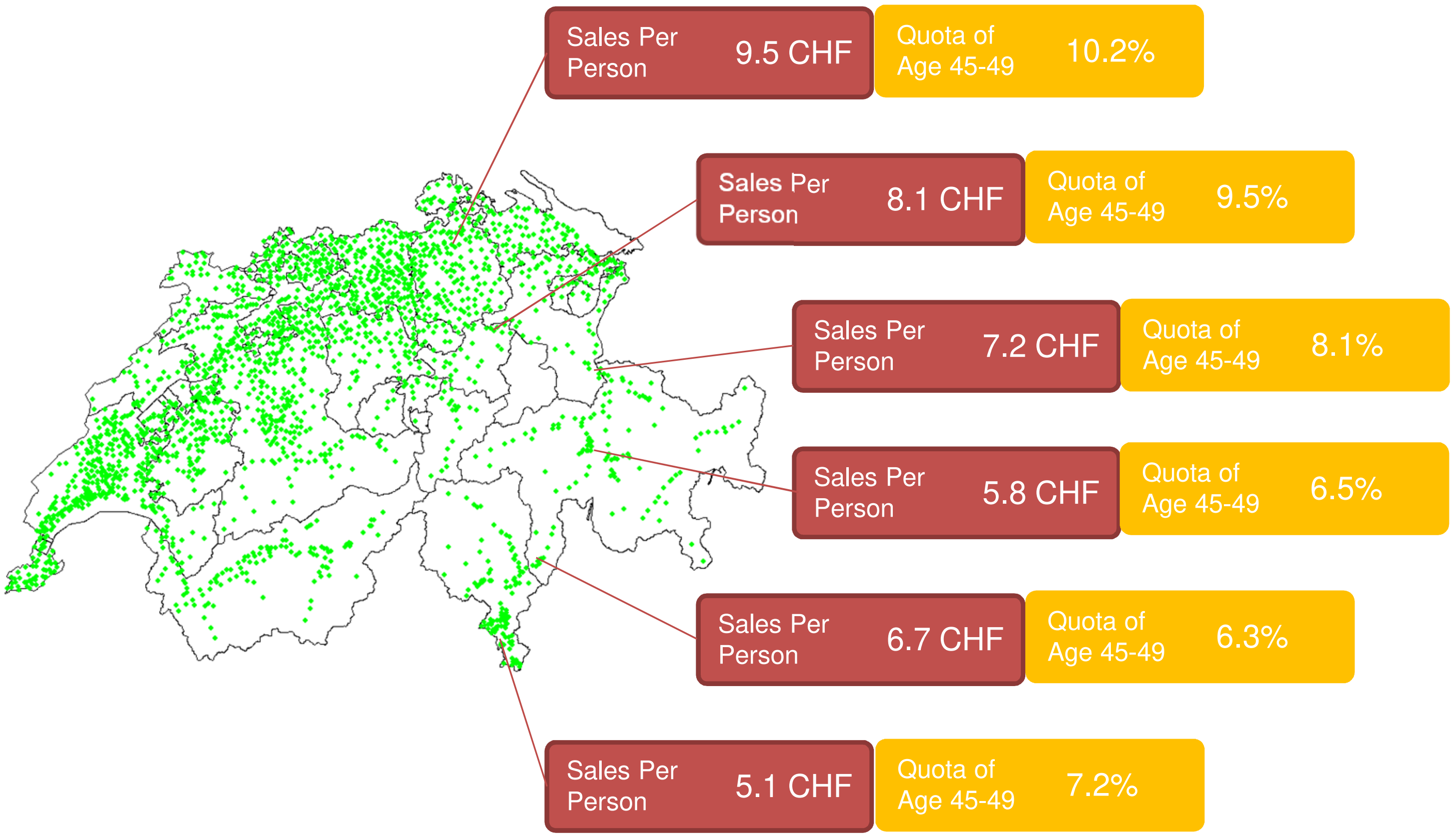
Features / Attributes

- **Age groups**
  - **0-4**
  - **5-9**
  - **10-14**
  - 15-19
  - 20-24
  - ...
  - 89 and more
- Housing Information
- Country of origin
- Marital status
- Language
- Work sectors
- etc

BOAK

Customer Geo-Analytics: Results

- Analysis for age-groups: exploit regional diversity



Customer Geo-Analytics: Results

- Analysis for age-groups: exploit regional diversity

Independent Variables

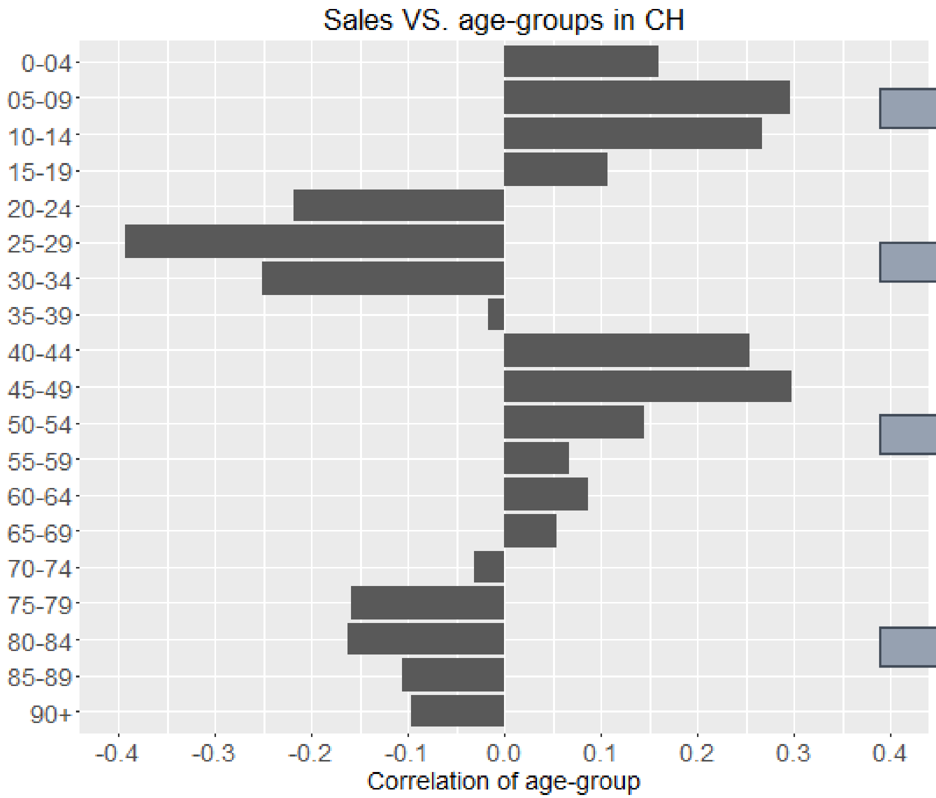
Response Variable

141

Region ID	Name	Age Demographics						Sales
1006	Bezirk Sense	0.143895	0.181273	0.978240	0.372817		0.544616	7.28
1002	District de la Glâne	0.135870	0.191506	0.942876	0.396533		0.399811	6.28
109	Bezirk Uster	0.185181	0.225751	0.994931	0.310582	•••	1.055710	5.57
110	Bezirk Winterthur	0.151051	0.475483	0.976424	0.284481		0.676363	6.35
1402	Bezirk Reiat	0.179851	0.406456	0.856140	0.251028		0.810297	2.81
1107	Bezirk Lebern	0.204975	0.353950	0.974240	0.178009		1.159469	11.37
	⋮			⋮				⋮
1403	Bezirk Schaffhausen	0.126046	0.202425	0.298306	0.377231	1.007837	0.675612	7.61

Customer Geo-Analytics: Results

- Sales Efficiency for age-groups



Families with kids

Young people

middle-aged people

Old people

Predictive Customer Analytics

Predictive Customer Analytics: Business Perspective

- Client: Retailer of drinking products
- Logistics Department for **Ho**tels-**Re**staurants-**Ca**tering (HoReCa)

Orders/Deliveries

Calendar

Weather

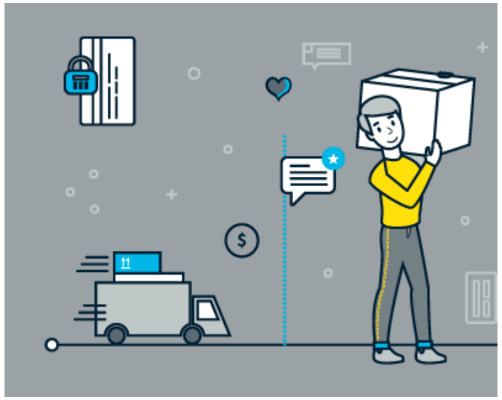
Regions

Marketing

Predictive models

Reco-Engine

Business Processes



Understand customers

Avoid out-of-stock situations

Provide good recommendations

BOAK

Predictive Customer Analytics: Predictive models

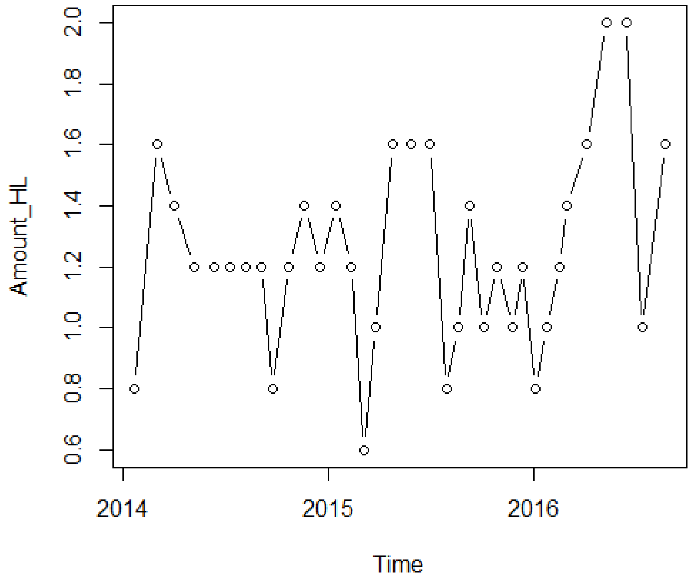
- Analysis per customer per product

We analyzed 45'360 time-series

Example: Customer ID = 77 294 025

- 43 deliveries in total
- 29 different products

Product ID	Nr. Of Deliveries
10041	35
10099	32
10152	29
10379	25
10476	24
10514	23
10601	19
10409	14
10153	12
10975	12
10606	10
10448	9
11542	8
11543	7
…	…



- Deliveries start from 2014-01-22
- Amount_HL in [0.6, 2] HL
- Time interval between consecutive orders in between 14 to 42 days

BOAK

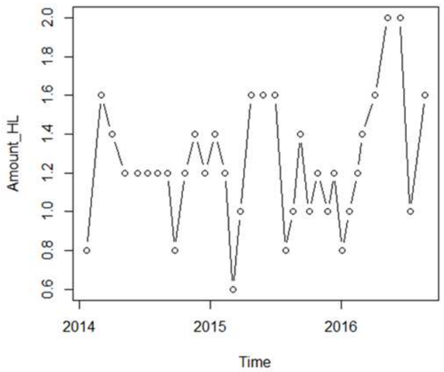
Predictive Customer Analytics: Predictive models

- Engaged and non-engaged customers

  - **Gr.1: 972 out of 1579 Customers (62%)**

  - Patterns exist for at least 1 product of these customers

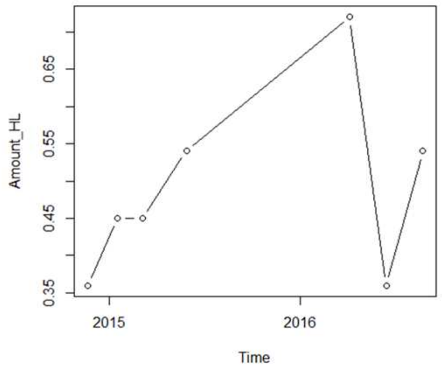
  ✓ ***88%*** of the total deliveries
  ✓ ***83%*** of the total express-deliveries



  - **Gr.2: 607 out of 1579 Customers (38%)**

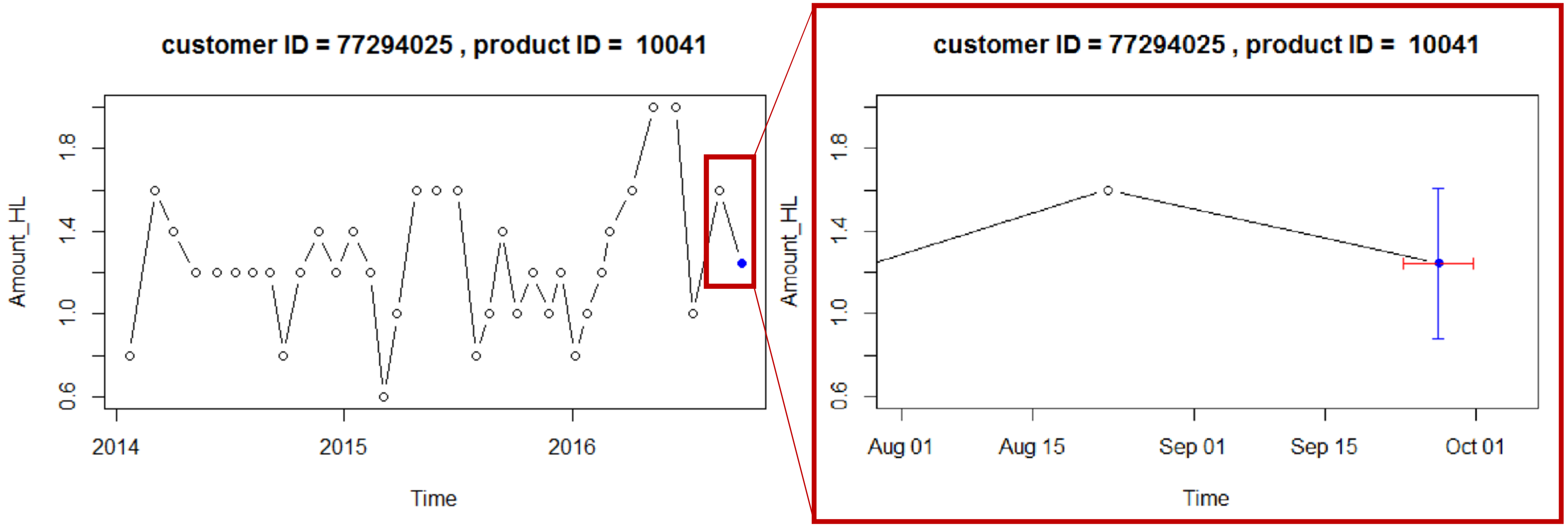
  - Patterns do not exist (at least based on the history of deliveries)

  - Example: Between all consecutive orders
    Min Day Difference = 49
    Max Day Difference = 313



Predictive Customer Analytics: Predictive models

- Engaged customers (Gr.1): Main Products - Predictions



Predict order-date of customer based on the main product

Predictive Customer Analytics: Business Perspective

- Client: Retailer of drinking products
- Logistics Department for **Ho**tels-**Re**staurants-**Ca**tering (HoReCa)

Orders/Deliveries

Calendar

Weather

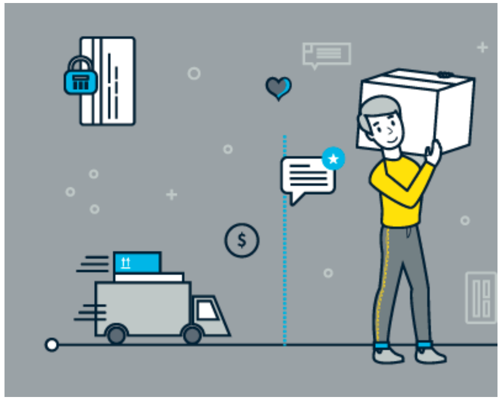
Regions

Marketing

Predictive models

Reco-Engine

Business Processes

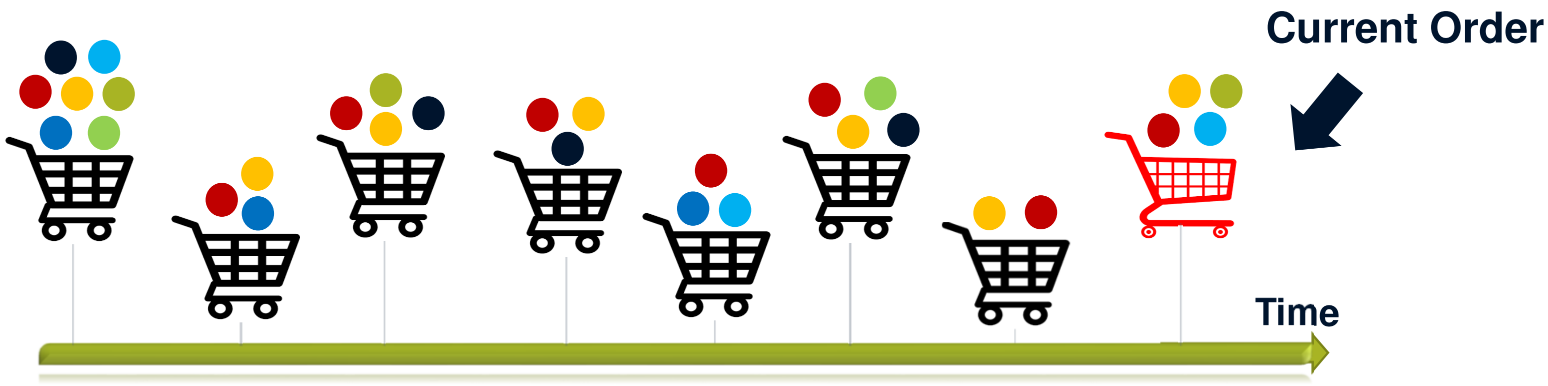


Understand customers
Avoid out-of-stock situations

Provide good recommendations

Predictive Customer Analytics: Recommendations

- "Past orders" and "current order" of a customer



- *Given the current order*, what products shall we recommend???

1. Intra recommendations → Based on *current customer's* history of orders & product-consumption

* Recommended products have been ordered in the past

2. Inter recommendations → Based on orders from *similar customers* and *popular combinations* of products

* Recommended products have NOT been ordered in the past

BOAK

DS Products @IT-Logix

DS Products @IT-Logix

- Data Science Workshop
  - Explore possibilities with data science
- Auditing of AI/ML models
  - Assess quality of existing AI/ML models
- ML/AI Tutorials and Hands-on Sessions
  - Learn basic concepts of ML and AI
- Requirements Engineering with Data Science
  - Outlier detection via content-based screening
  - Identify important missing quantities

Data Science for Business

Zürich, Switzerland · 568 members · Public group

Organized by
Sotiris Dimopoulos

17
OCT

Wednesday, October 17, 2018

Machine Learning in Enterprise Applications: A Perspective from SAP

Hosted by Sotiris Dimopoulos
From Data Science for Business
Public group

You're going 37 people going

Share:

Details

Agenda:

18:00 Introduction & welcome, Dr. Sotiris Dimopoulos, IT Logix

18:10 Machine learning in enterprise applications: example use cases and approaches, Dr. Sean Kask, SAP

19:00 Discussion and apéro

Short Description of the Talk:

Organizer tools

Wednesday, October 17, 2018
6:00 PM to 8:00 PM
Add to calendar

Au Premier Bar (Hauptbahnhof)
Bahnhofplatz 15 · Zürich

How to find us
Meeting room "Les Trouvailles"

Wir freuen uns auf angeregte Gespräche mit Ihnen

- Sotiris Dimopoulos, PhD
  Senior Data Science Consultant

Twitter LinkedIn Xing YouTube Blog

IT-LOGIX

BOAK