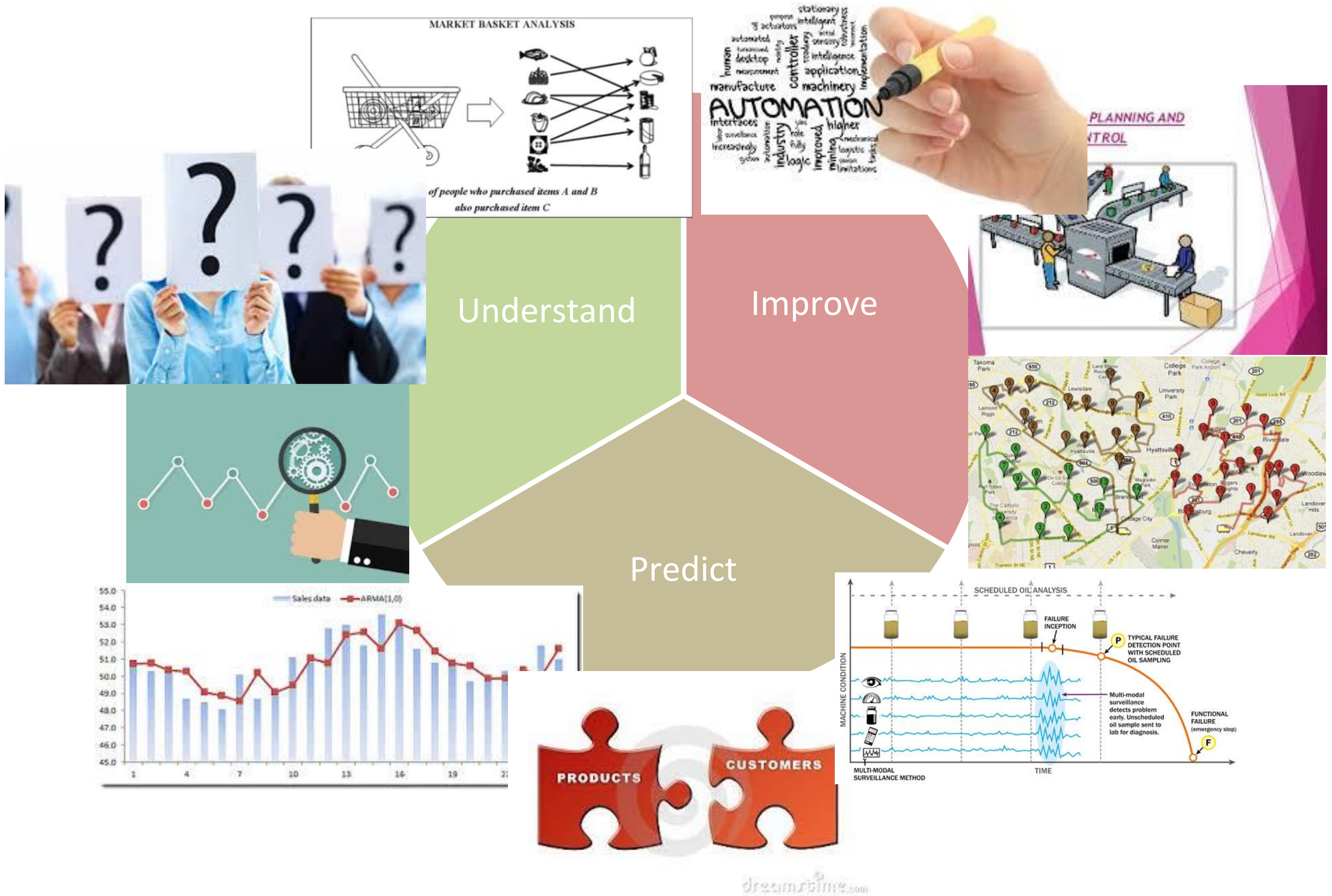
Customer Intelligence in der Praxis

IT-Logix AG
Sotiris Dimopoulos, PhD

Zürich, 02.07.2018

Data Science & AI : Possibilities?



Outlook

- Customer Dimensions
  - ❑ Use Machine Learning to extract new customer dimensions

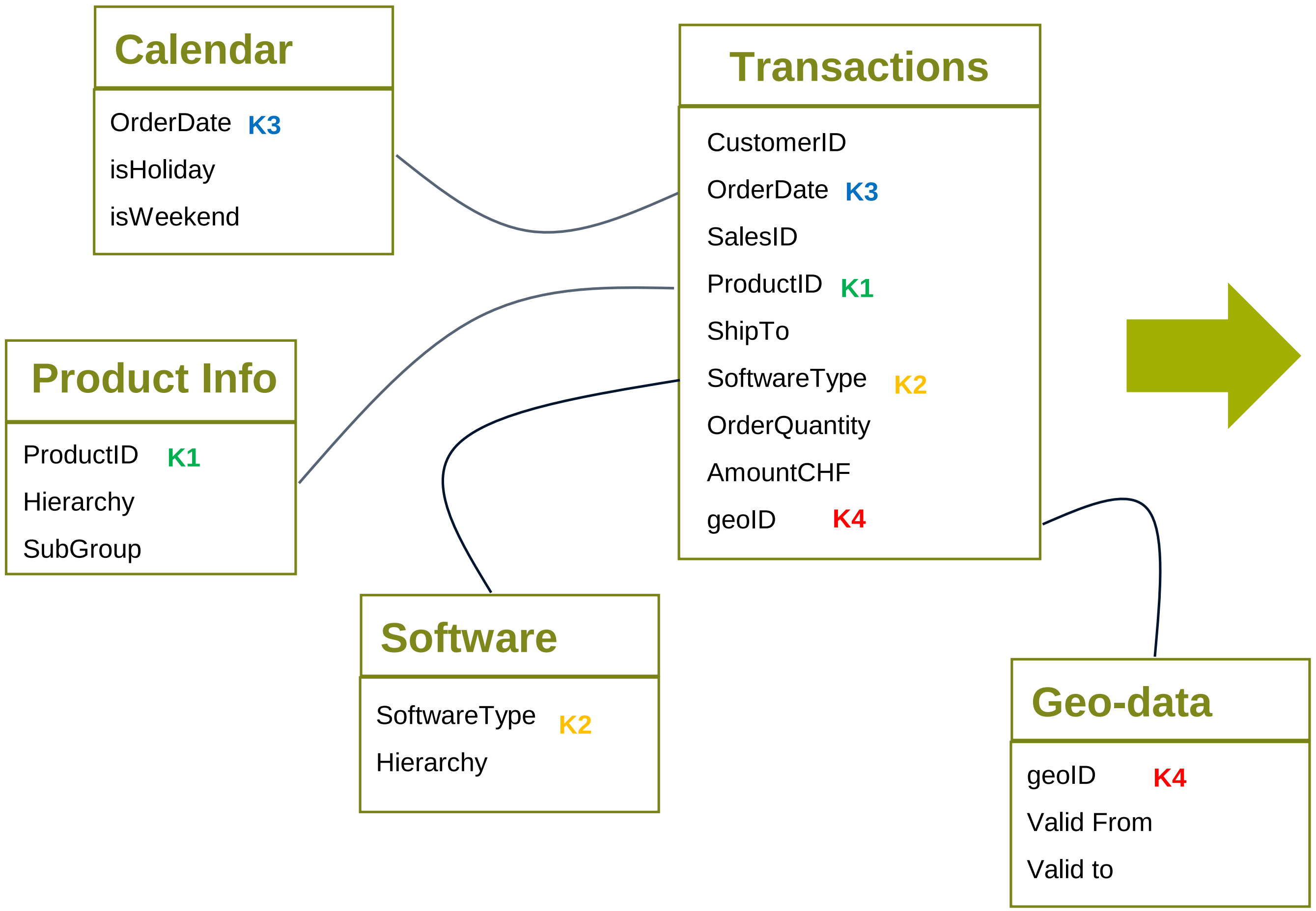
- Customer Geo-Analytics
  - ❑ Leverage geo-data to understand your customers

- 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

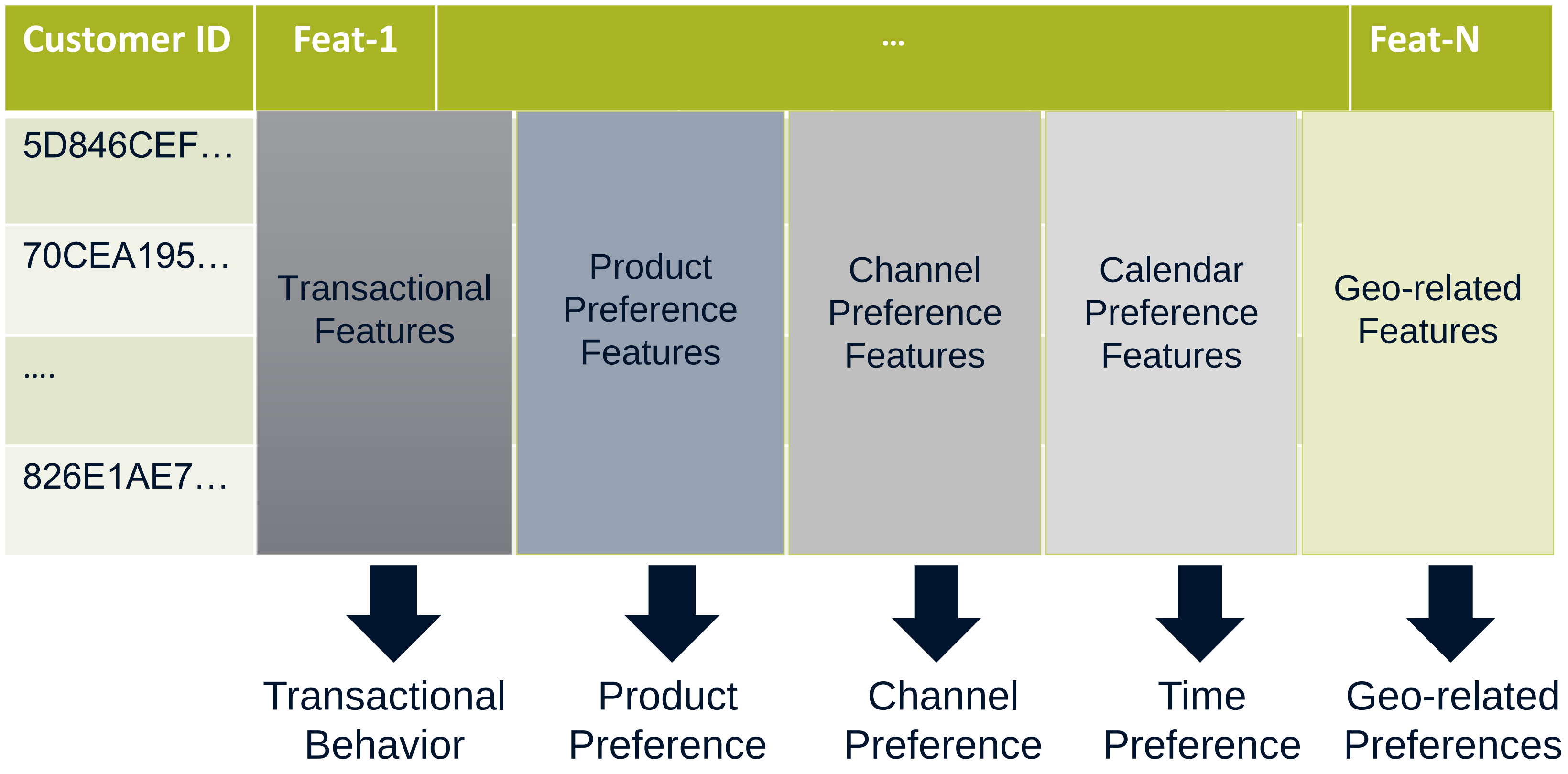


Customer ID	feat1	…	featN
5D846CEF…			
70CEA195…			
….			
826E1AE7…			

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

Customer Dimensions

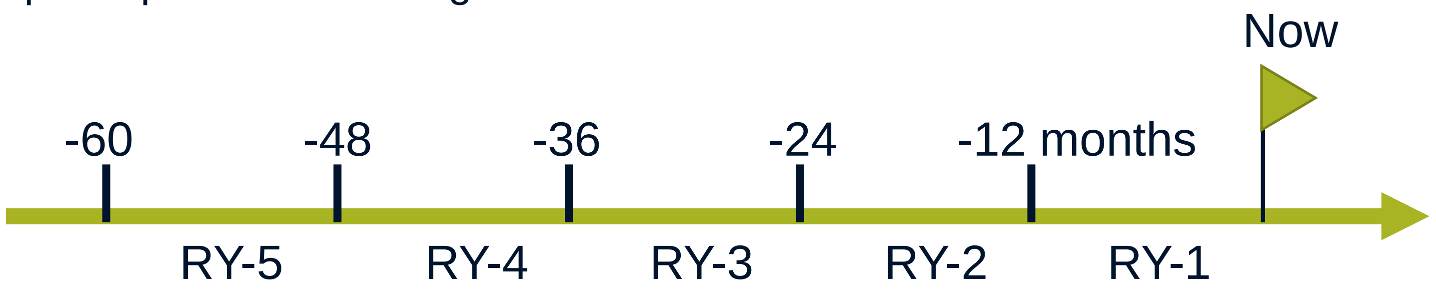
- Extraction of more than 100 customer dimensions
- Great variety of customer dimensions



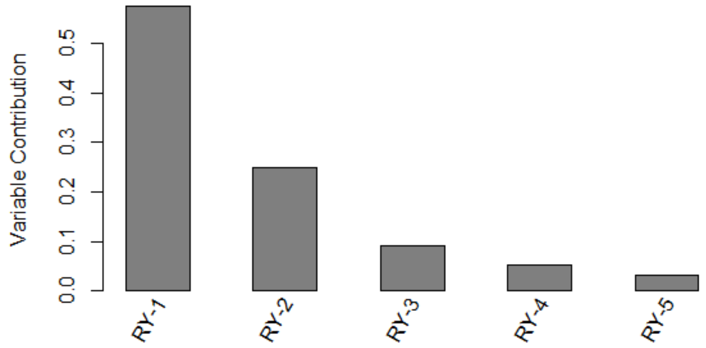
MAKE BI

Customer Dimensions: Examples

- Example 1: predictive scoring



CUSTOMER	32 CHF	0 CHF	45 CHF	35 CHF	98 CHF

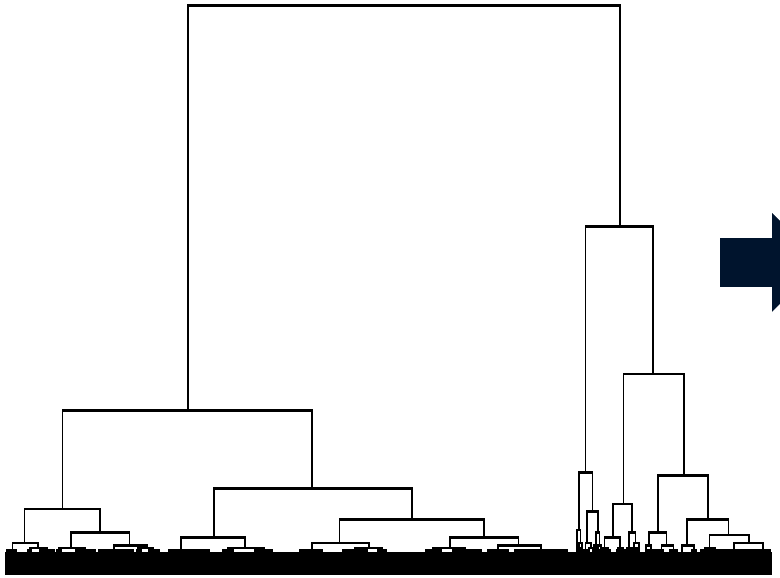


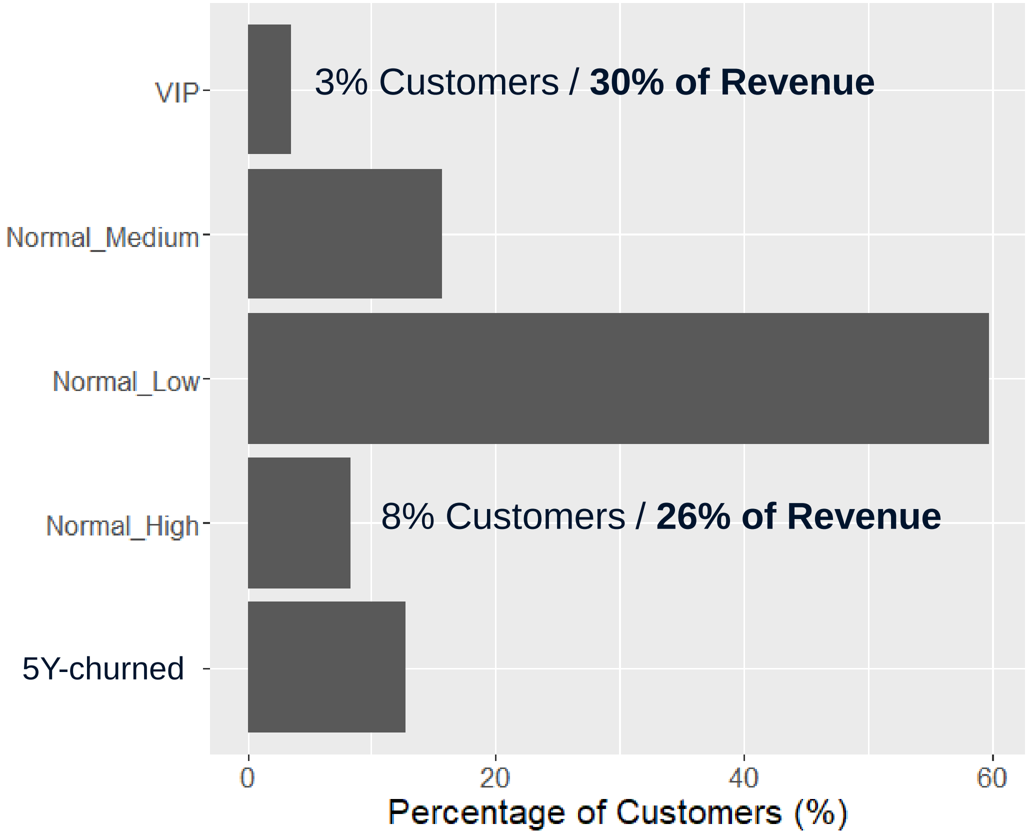
Rolling Year	Weight
Rolling Year 1	0.577
Rolling Year 2	0.249
Rolling Year 3	0.091
Rolling Year 4	0.053
Rolling Year 5	0.030

CUSTOMER SCORE	70.316

Customer Dimensions: Examples

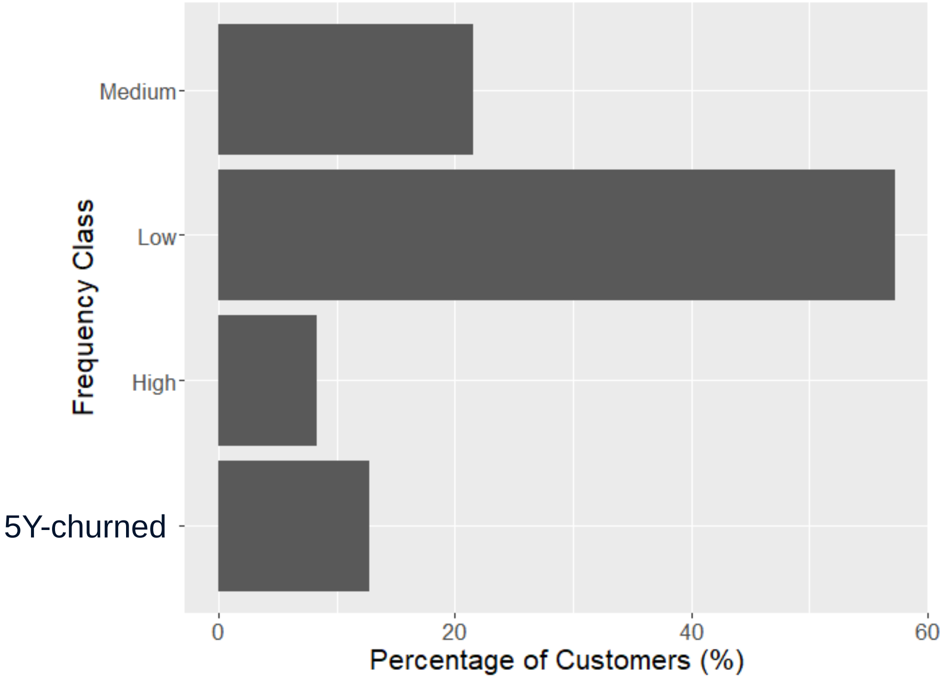
- Example 2: Classes

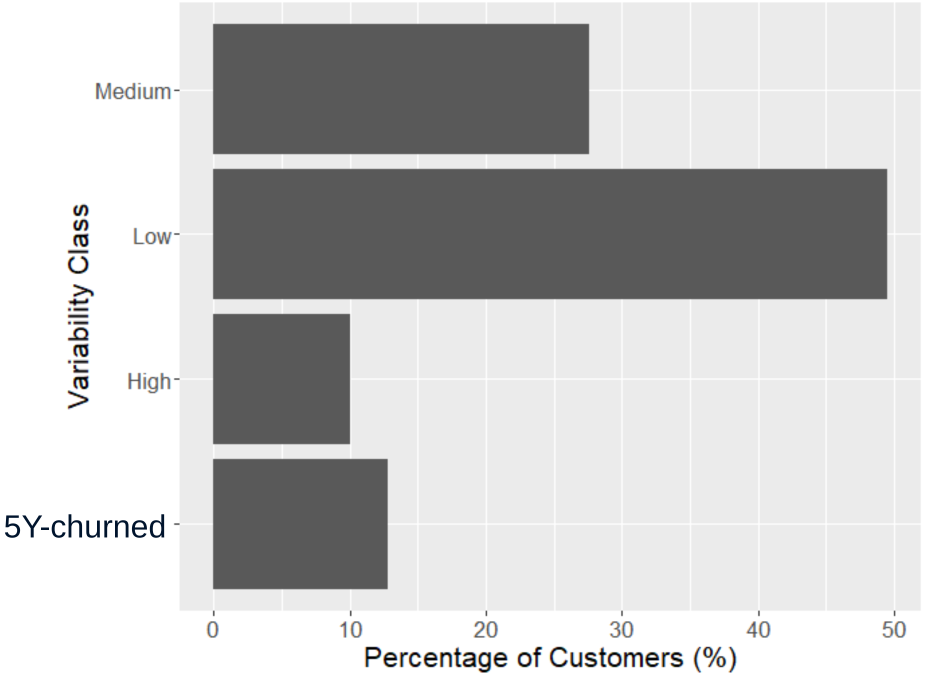




Customer Dimensions: Examples

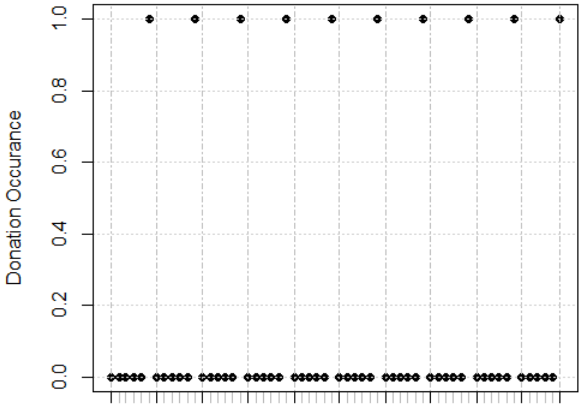
- Example 2: Classes



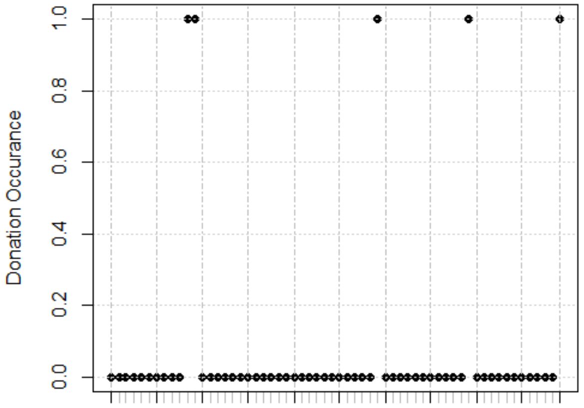


Customer Dimensions: Examples

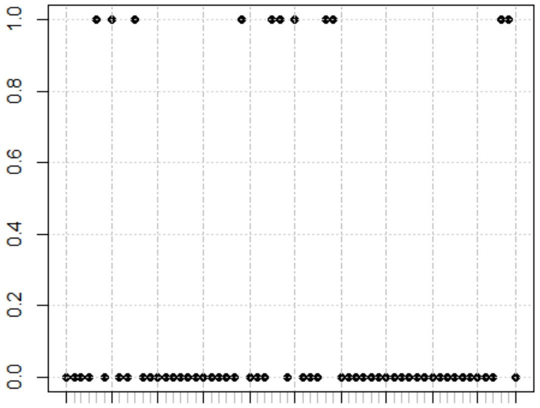
- Example 3: Periodicity



Periodicity Class	Customer Base (%)
Strong	0.5%



Periodicity Class	Customer Base (%)
Weak	5%

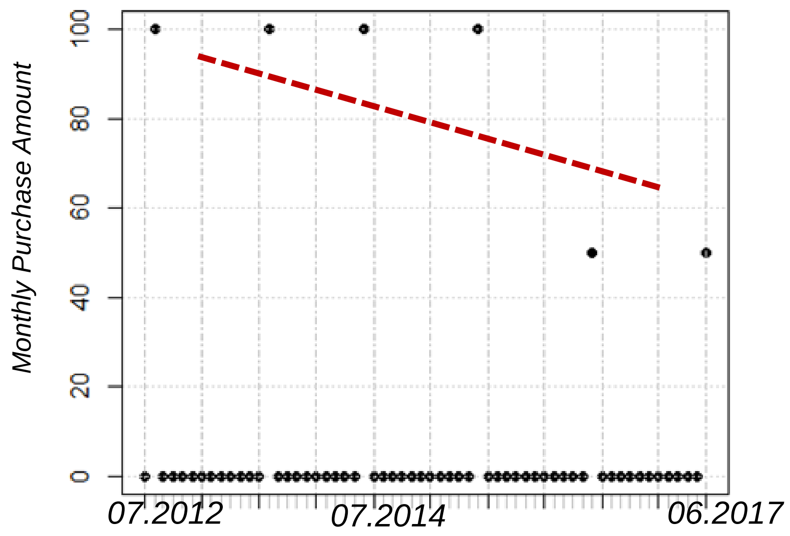


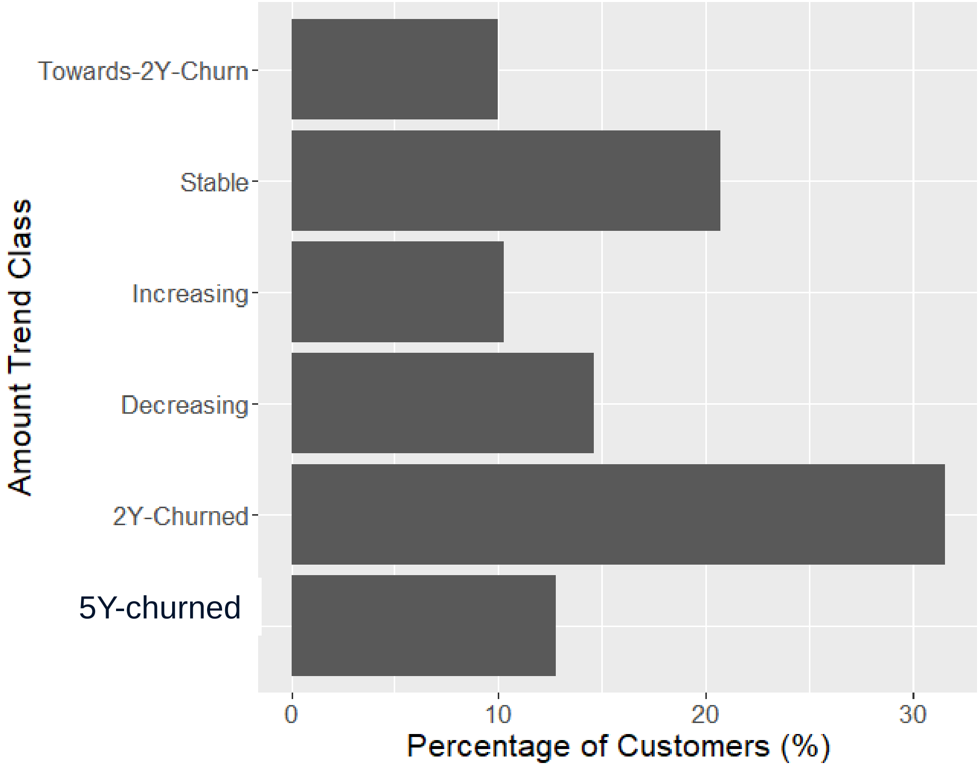
Periodicity Class	Customer Base (%)
Aperiodic	82%

MAKE BI

Customer Dimensions: Examples

- Example 4: Trends





MAKE BI

Customer Dimensions: Use

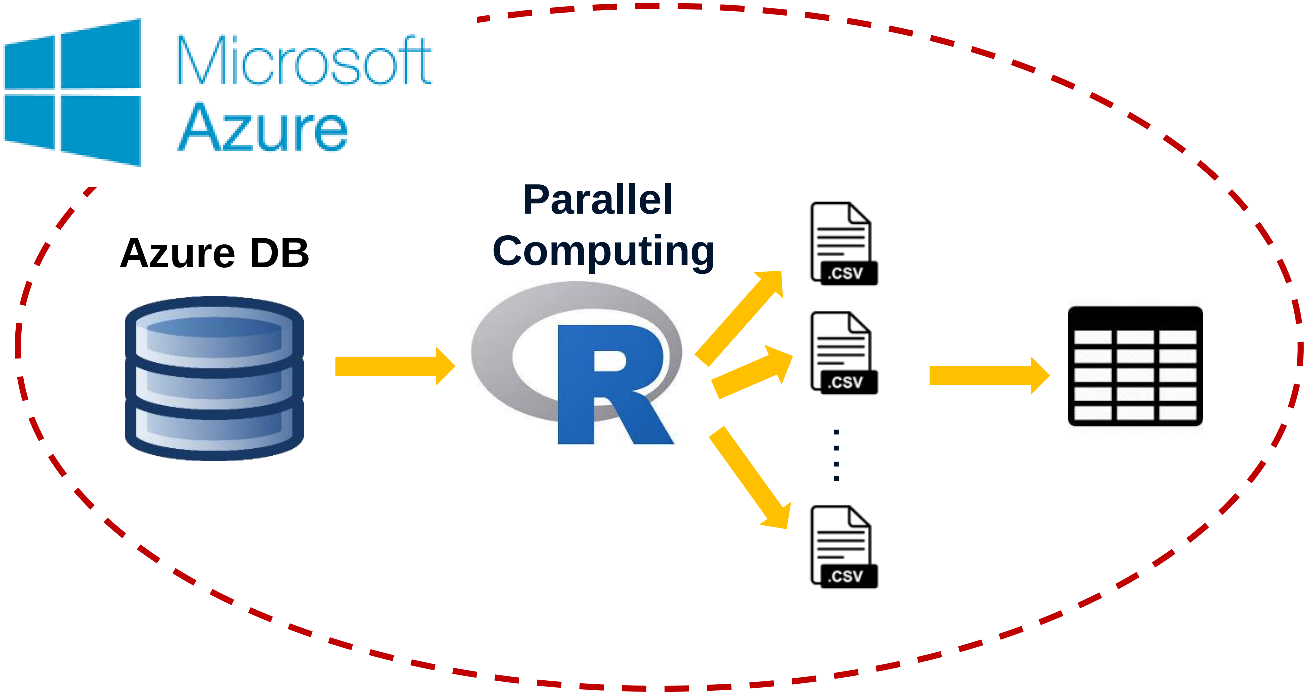
- Customer Behavior Table

Customer ID	Feat-1	...			Feat-87
5D846CEF...	Transactional Features	Product Preference Features	Channel Preference Features	Calendar Preference Features	Geo-related Features
70CEA195...					
....					
826E1AE7...					

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

Customer Dimensions: Operationalization

› Technology and Pricing



- ❑ Azure Batch for scalable parallel computing
- ❑ Compute cost = 4CHF / month, for analyzing 2.5M customers

Customer Geo-Analytics

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>

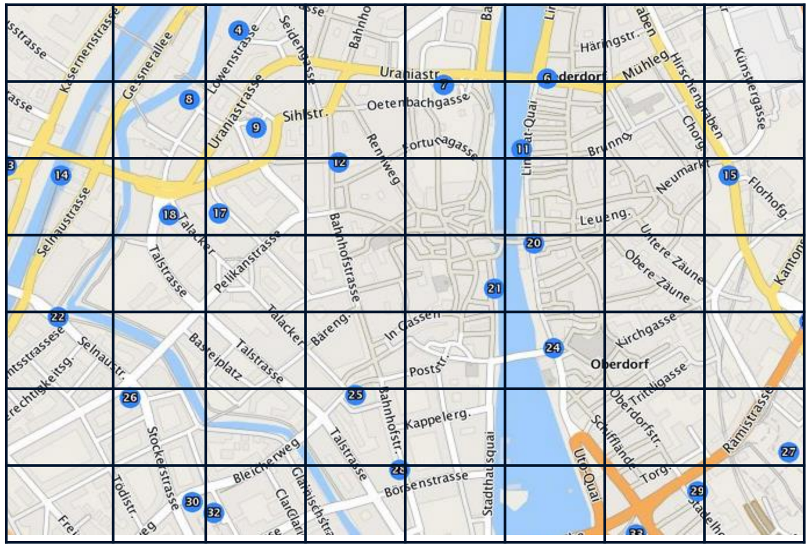
- Age groups
- Housing Information
- Country of origin
- Marital status
- Language
- Work sectors
- Family and kids
- Religion
- Political opinions
- Commuting
- Urbanization
- Educational level
- Income
- etc

Geo-data in Switzerland

› BFS as a source of information

Regional Level

- State (Kanton)
- District (Bezirk)
- Community (Gemeinde)
- **Hectare (100m x 100m)**

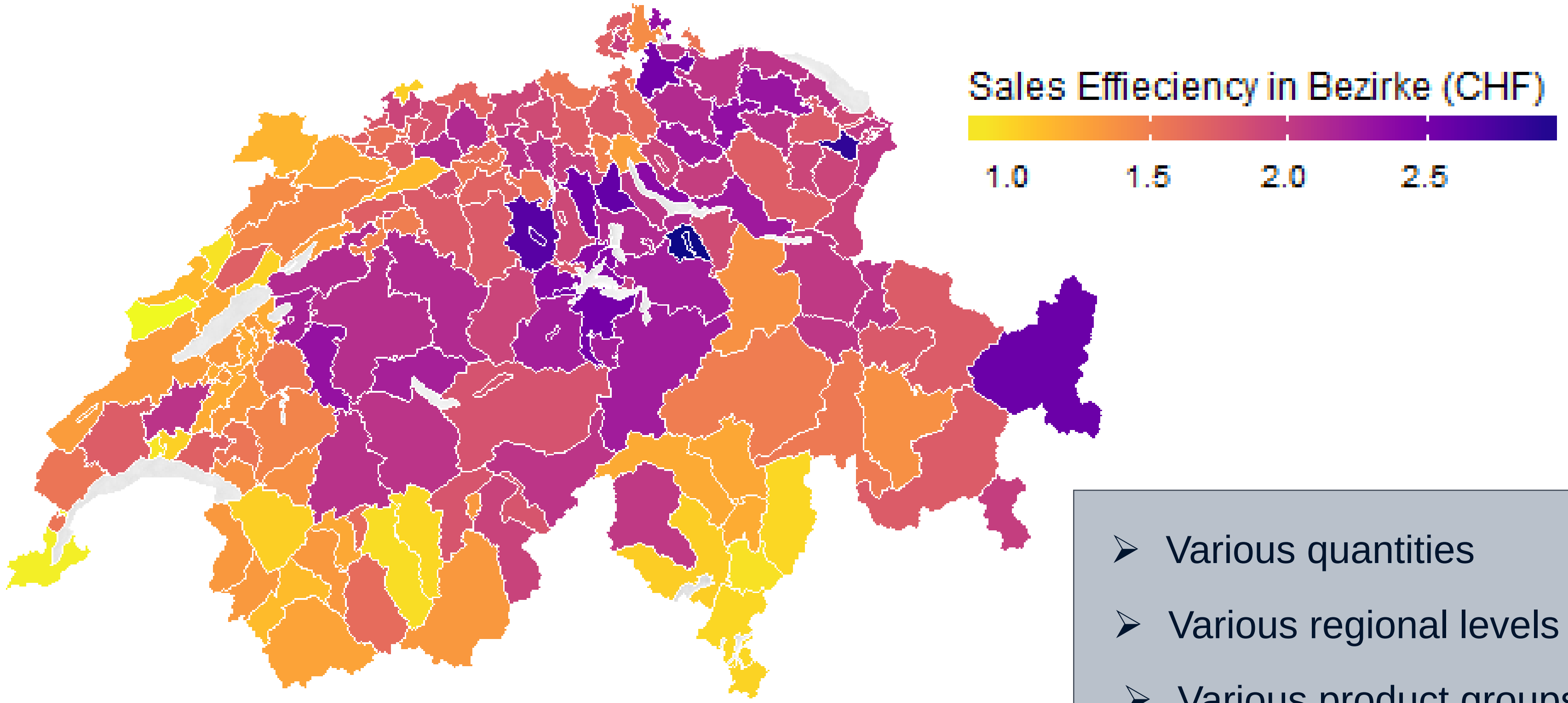


Features / Attributes

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

Customer Geo-Analytics: Results

- Quantification sales efficiency at the regional level



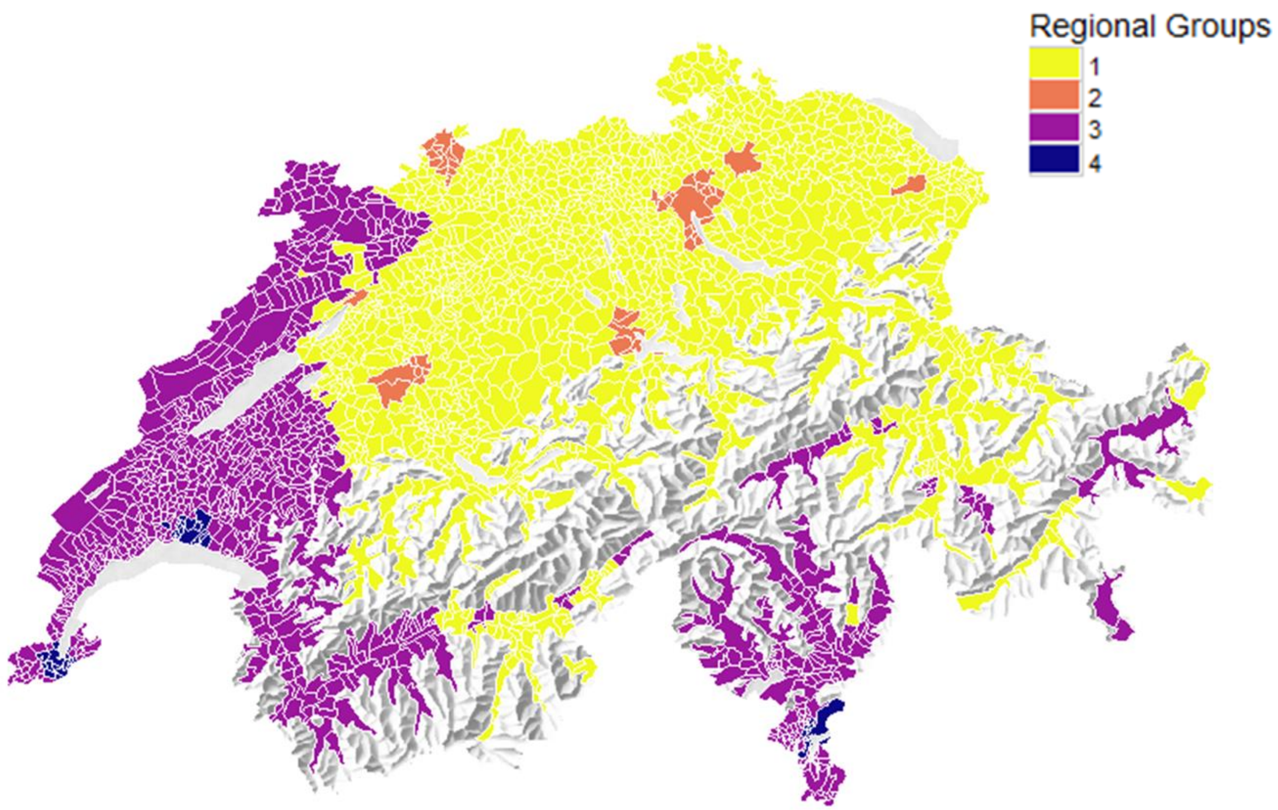
Annual Revenue Efficiency (= Annual Revenue / Population) of a product-group for all Bezirke in 2016

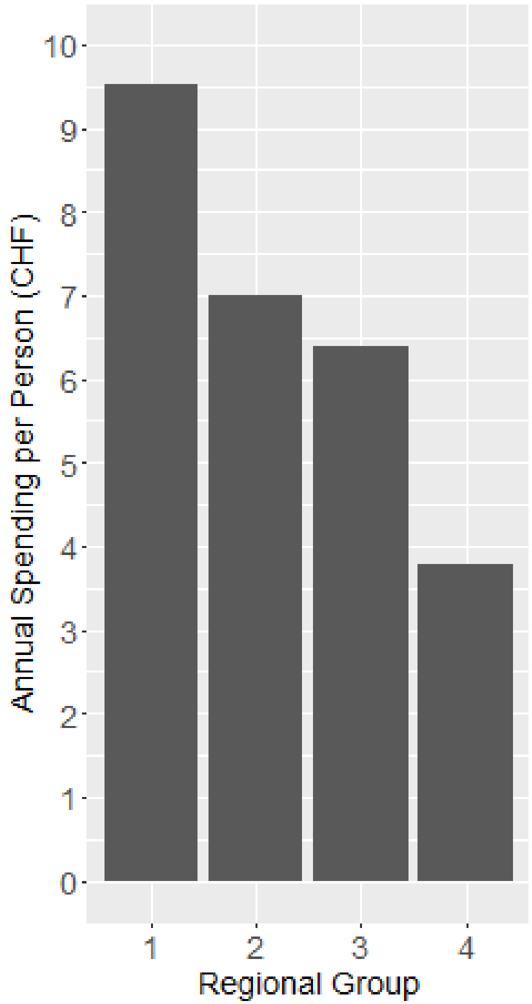
- Various quantities
- Various regional levels
- Various product groups

MAKE BI

Customer Geo-Analytics: Results

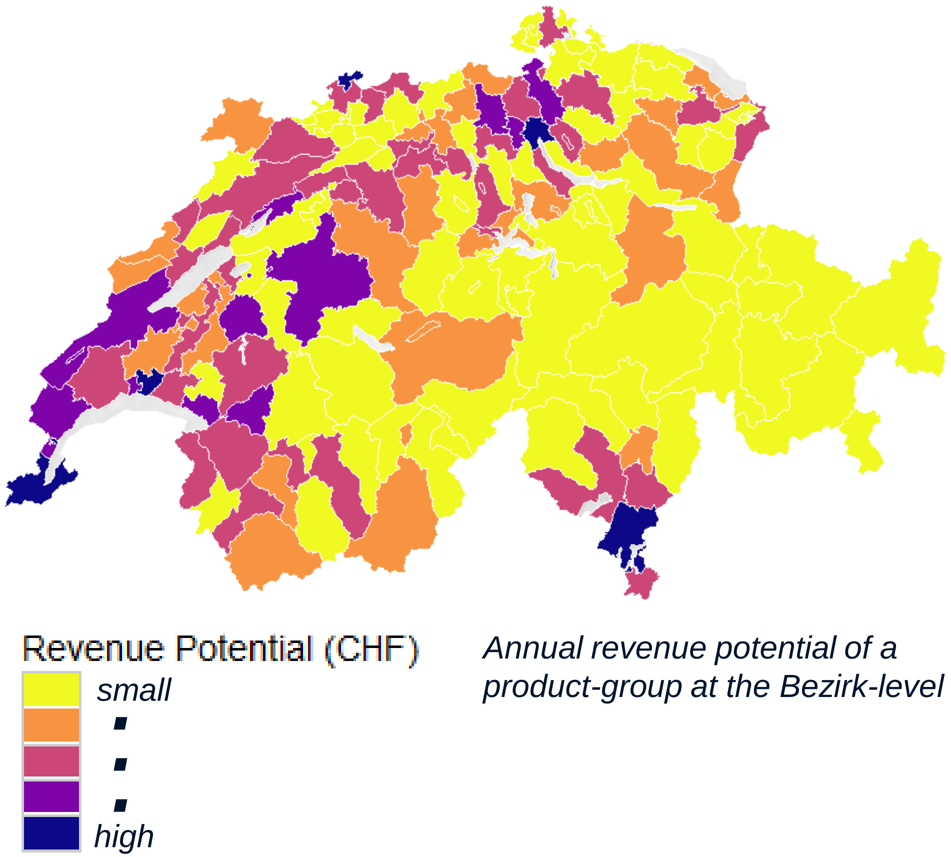
- Regional grouping





Customer Geo-Analytics: Results

- Annual revenue potential



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

<u>Which regions should be targeted first?</u>

KANTON_NAME	Revenue Potential (CHF)	Cumulative Percentage (%)
Vaud	5.69*X	14.2
Genève	5.58*X	28.1
Zürich	5.39*X	41.6
Ticino	3.47*X	50.3
…	…	…

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

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

- BFS as a source of information

<u>Regional Level</u>

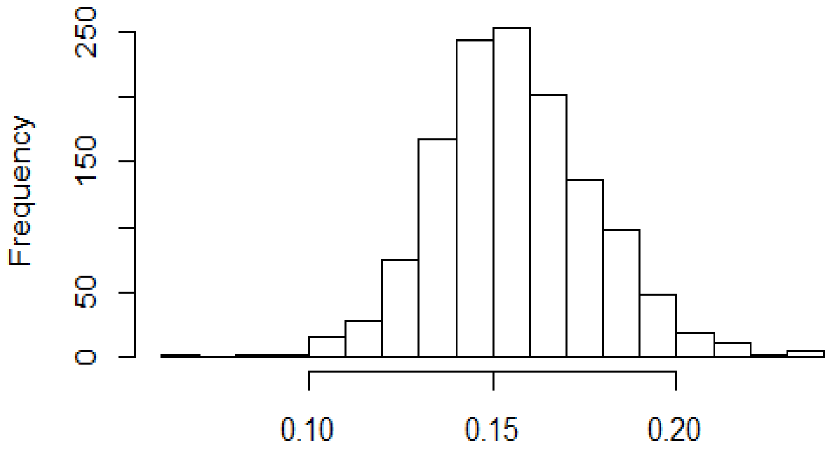
State
(Kanton)

District
(Bezirk)

Community
(Gemeinde)

Hectare
(100m x 100m)

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



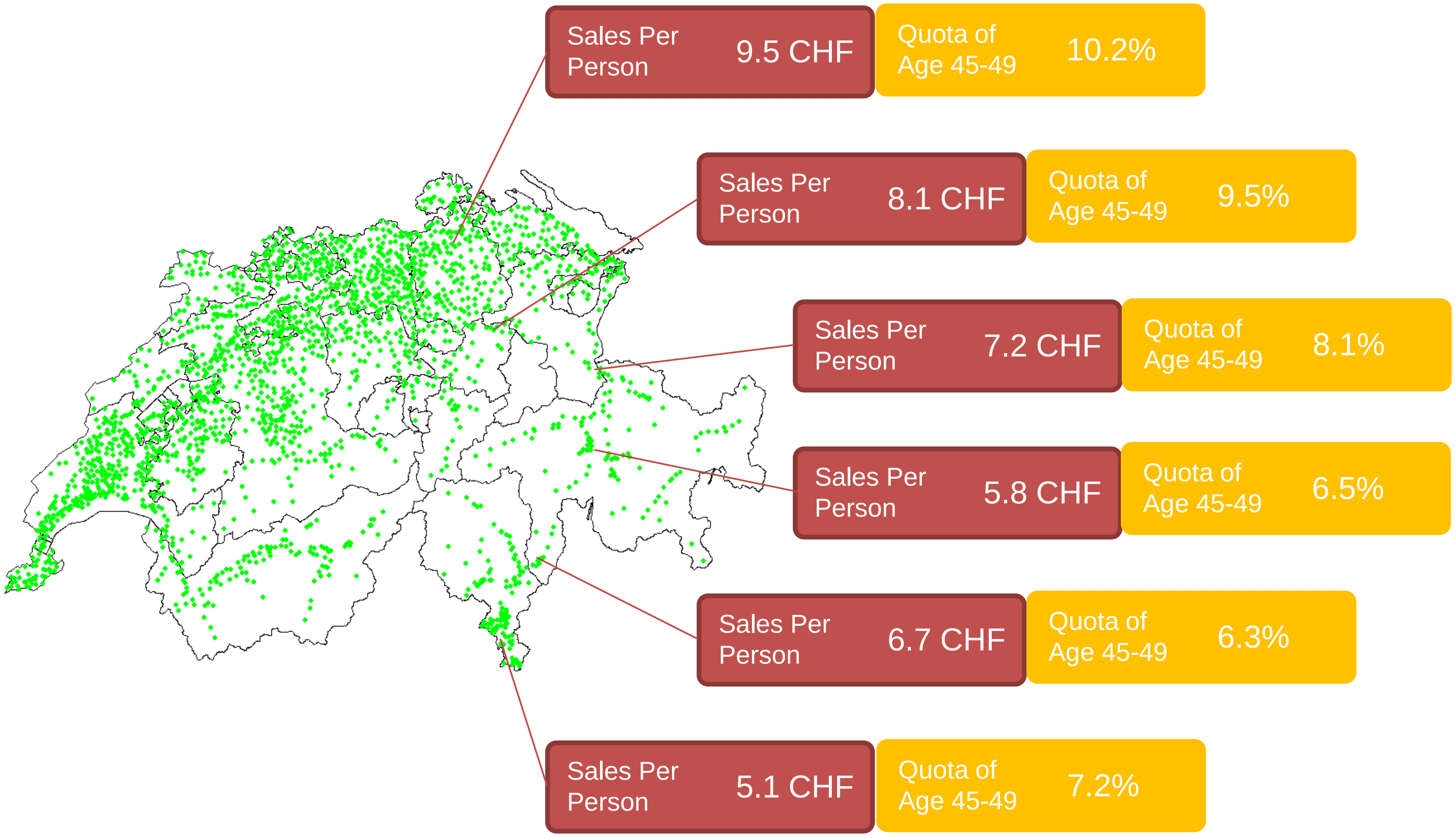
Diversity is the key to informativeness!

<u>Features / Attributes</u>

- **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

Customer Geo-Analytics: Results

- Analysis for age-groups: exploit regional diversity



Customer Geo-Analytics: Results

- Analysis for age-groups: exploit regional diversity

Independent Variables

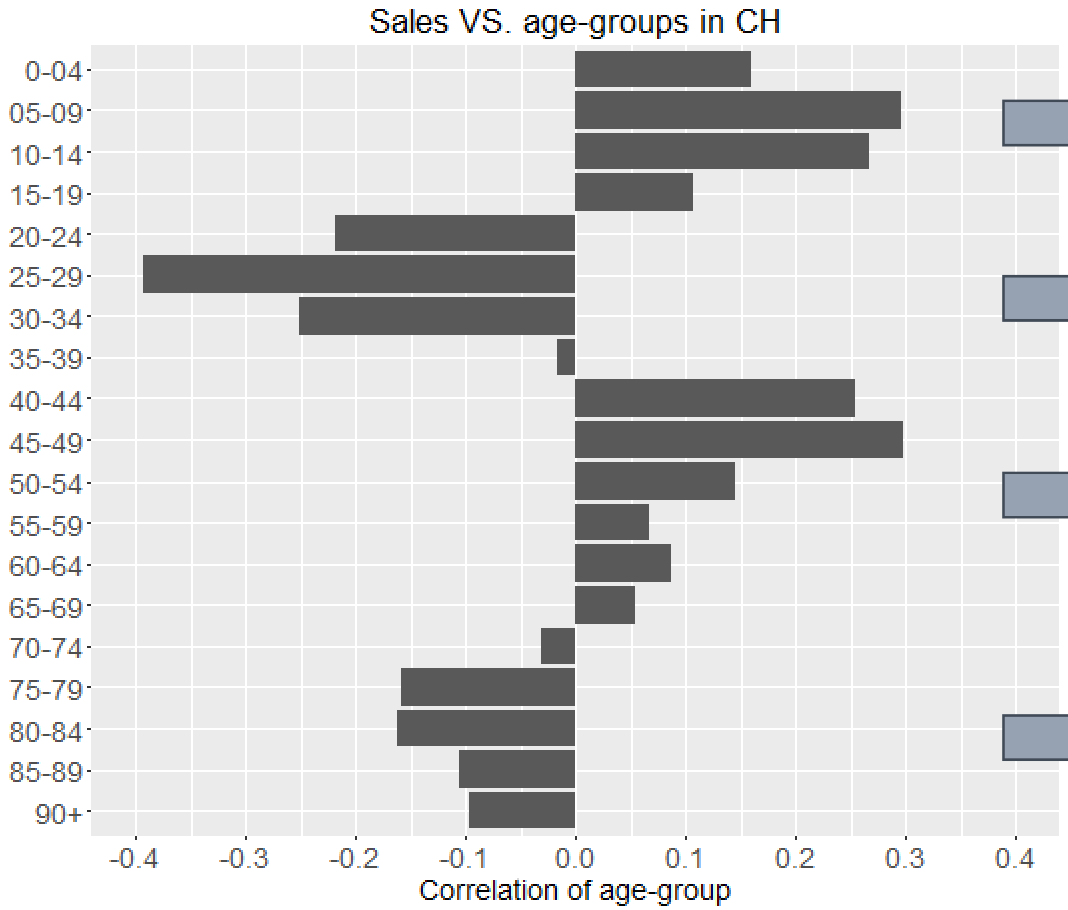
Response Variable

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

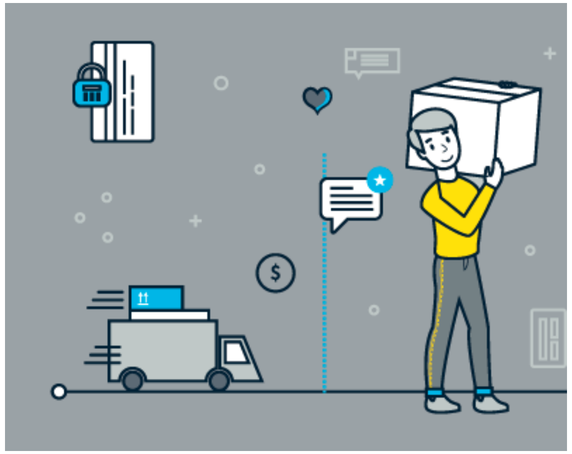
Regions

Marketing

Predictive models

Reco-Engine

Business Processes



Understand customers
Avoid out-of-stock situations

Provide good recommendations

Predictive Customer Analytics: Predictive models

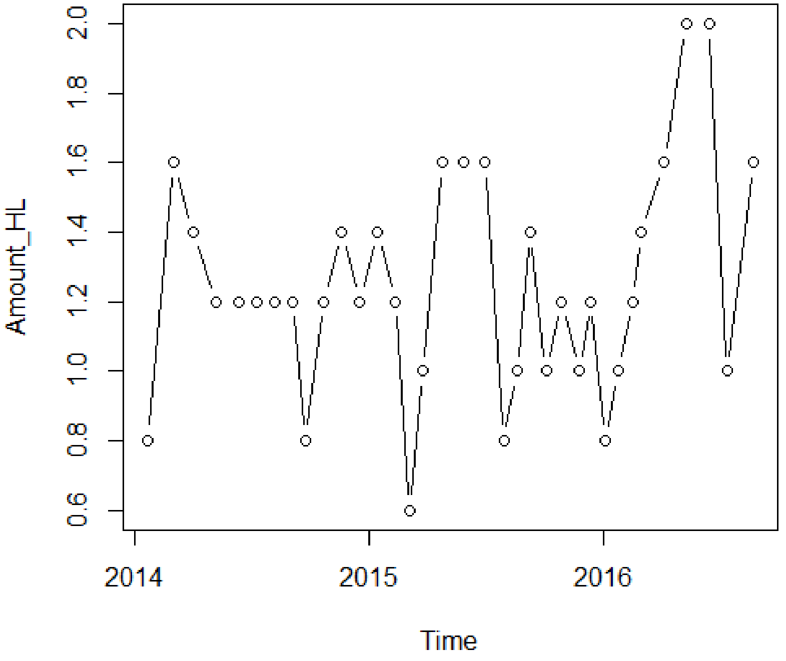
- Analysis per customer per product

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

MAKE BI

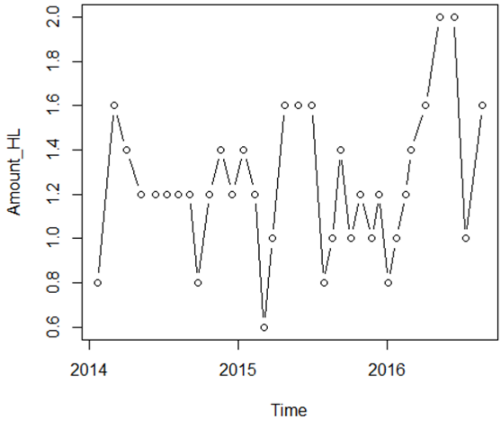
Predictive Customer Analytics: Predictive models

- Engaged and non-engaged customers

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

- Patterns exist for at least 1 product of these customers

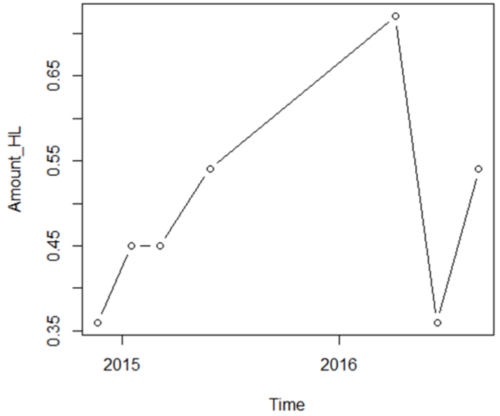
✓ ***88%*** of the total deliveries
✓ ***83%*** of the total express-deliveries



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

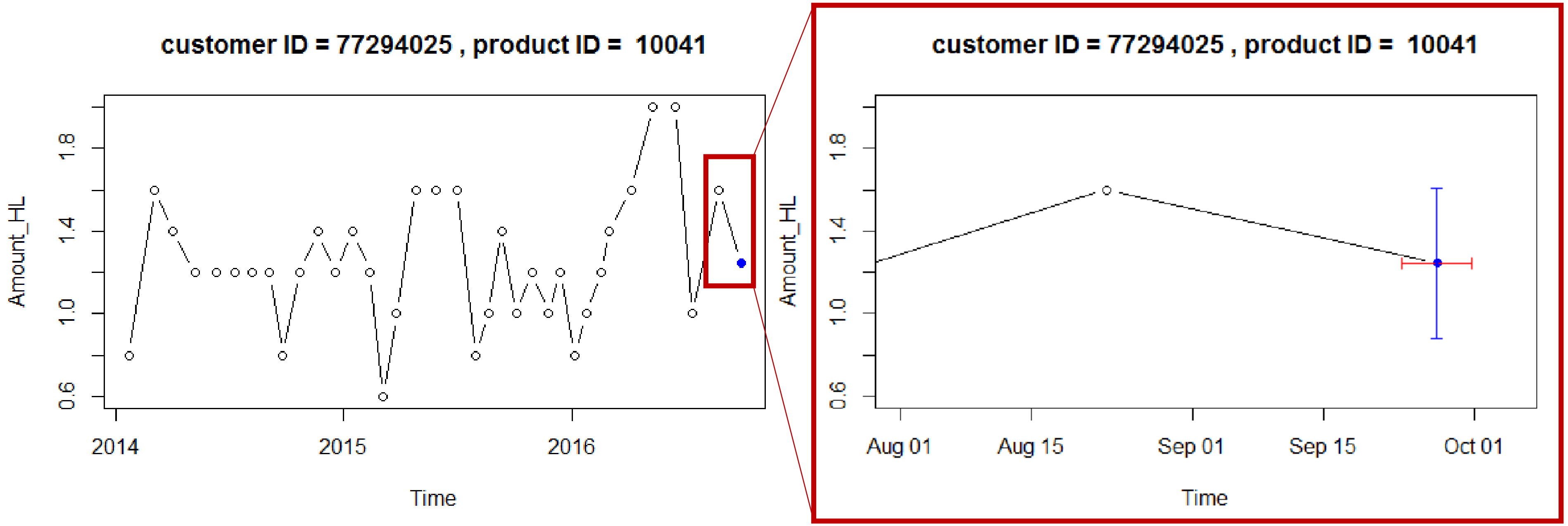
- Patterns do not exist (at least based on the history of deliveries)

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



Predictive Customer Analytics: Predictive models

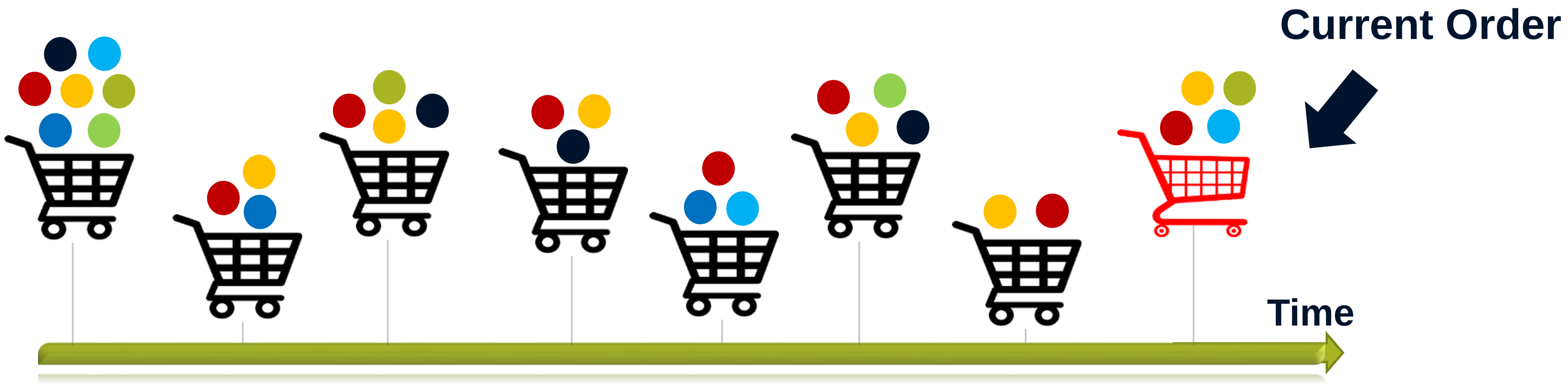
- Engaged customers (Gr.1): Main Products - Predictions



Predict order-date of customer based on the main product

Predictive Customer Analytics: 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

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

Wir freuen uns auf angeregte Gespräche mit Ihnen

- Dr. Sotiris Dimopoulos
  Senior Data Science Consultant

Twitter LinkedIn Xing YouTube Blog

IT-LOGIX

MAKE BI