Machine Learning with Calculation Views

A Use Case for Customer Data Segmentation

Machine Learning with SAP HANA

About me

Machine Learning with SAP HANA

- Benedict Baur
- Freelance SAP consultant and developer
- Focus area:
  - SAP Business Warehouse (BW/4 HANA),
  - SAP HANA Modelling with Calculation Views
  - Machine Learning (SAP HANA, Python)
- Trainer for Calculation Views (for Brandeis Consulting)
- Author of the book „Machine Learning mit SAP HANA" (Espresso Tutorials, 2022)
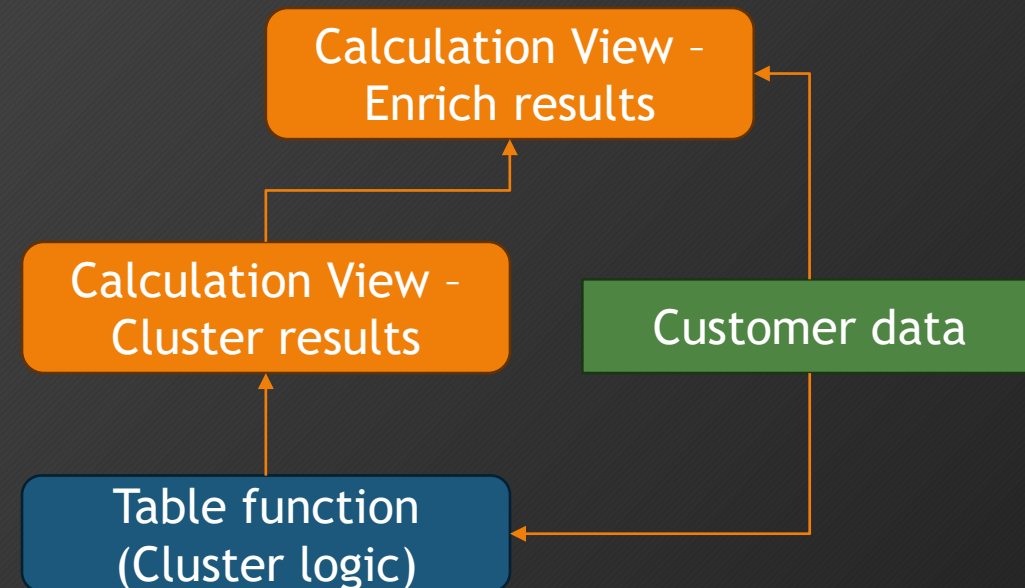- LinkedIn: kurzelinks.de/drbb

Motivation: Customer Data segmenation

- Starting point
  - Data set of customers of a company (Bank, Insurance, Telecommunication)
  - Data set contains several attributes (age, location, sales amount) and information about behavior of customers
- Goal: Segment data set in groups (*clusters*) of similar customers
  - Similarity calculation using attributes of the input set
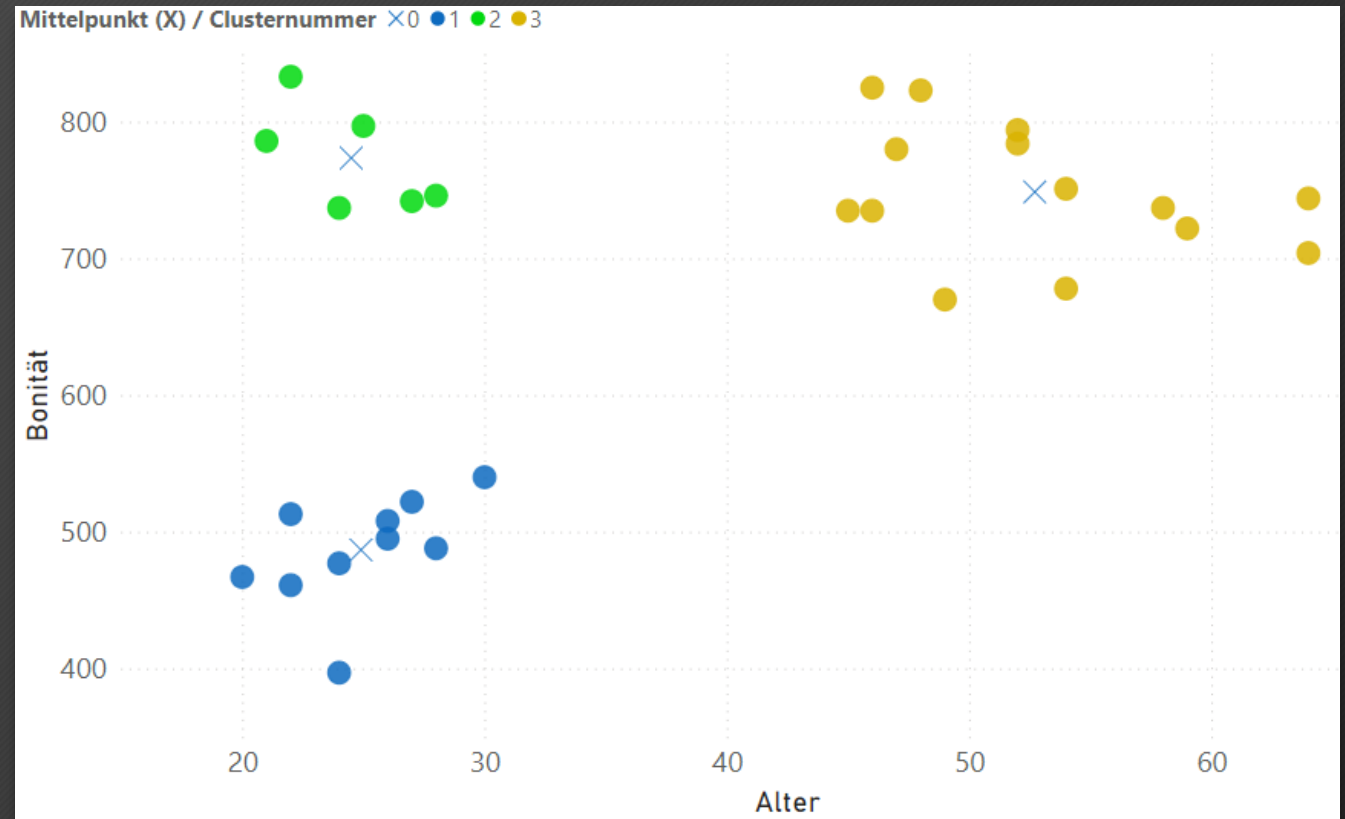
HANA native implementation strategy

- SAP HANA built-in cluster algorithms: K-Means Clustering, Hierarchical Clustering, DBScan, etc.

- Implement Cluster algorithm in table function using SQLScript

- Integrate table function in Calculation View
  - Trigger Ad-Hoc execution of Cluster algorithm
  - Read results from snapshot table (new feature in Business Application Studio)

- Enrich results from Cluster algorithm with original customer attributes in Calculation View

Machine Learning with SAP HANA

Calculation View – Enrich results

Calculation View – Cluster results

Customer data

Table function (Cluster logic)

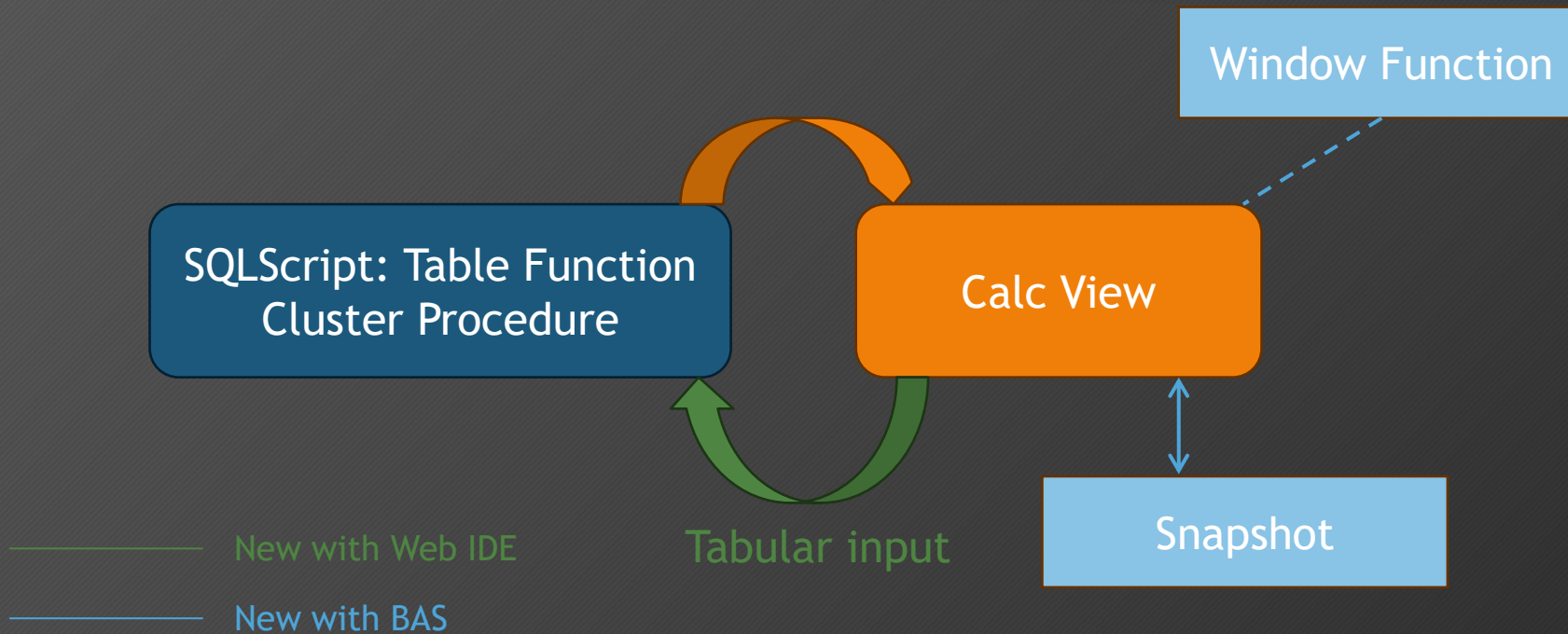# New features in web based modeling of Calculation Views

New with Web IDE (XSA)
- Improved integration of table functions:
  - Support for tabular input parameters
  - Supply input data for table function from Calculation View
  - Filtering of input data, Adapt mapping of input columns

New with Business Application Studio (HANA Cloud/XSA)
- Snapshot function
  - Save result of Calculation View in a snapshot table
  - Access both Snapshot data and base view (real-time data) via Interface View
- New node type: Window function
  - Lead, Lag, Cumulating Sum
  - Random sampling, Binning (Ntile), statistical measures (e.g. percentiles, median)

Machine Learning with SAP HANA

New features in web based modeling of Calculation Views

Window Function

SQLScript: Table Function
Cluster Procedure

Calc View

Snapshot

Tabular input

New with Web IDE

New with BAS

Machine Learning with SAP HANA

Live demo

- Store results of Cluster algorithm as snapshot
- Analyze results of Cluster algorithm in Calculation View
  - Random sampling using Window functions
  - Binning of column

Machine Learning with SAP HANA