Rotman

Master of Management Analytics

INTRO TO SAS VIYA VISUAL TOOLSET

SAS Bootcamp 1 (2 hours)

August 9, 2019 Prepared by Jay Cao / TDMDL

Rotman School of Management
UNIVERSITY OF TORONTO

Plan and Goal

- Plan
  - SAS Viya big picture overview
  - Hands-on with SAS Viya visual toolset

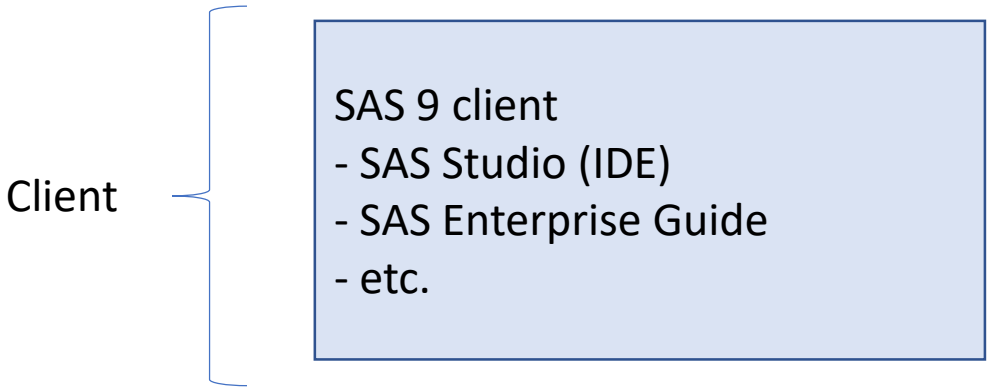
- Goal: learn the basics of
  - SAS Viya visual toolset
  - data processing and modeling work flow

What's SAS

- An analytics software suite
  - Retrieve data from different sources
  - Manage and manipulate data
  - Perform analytics
  - Report and visualize results

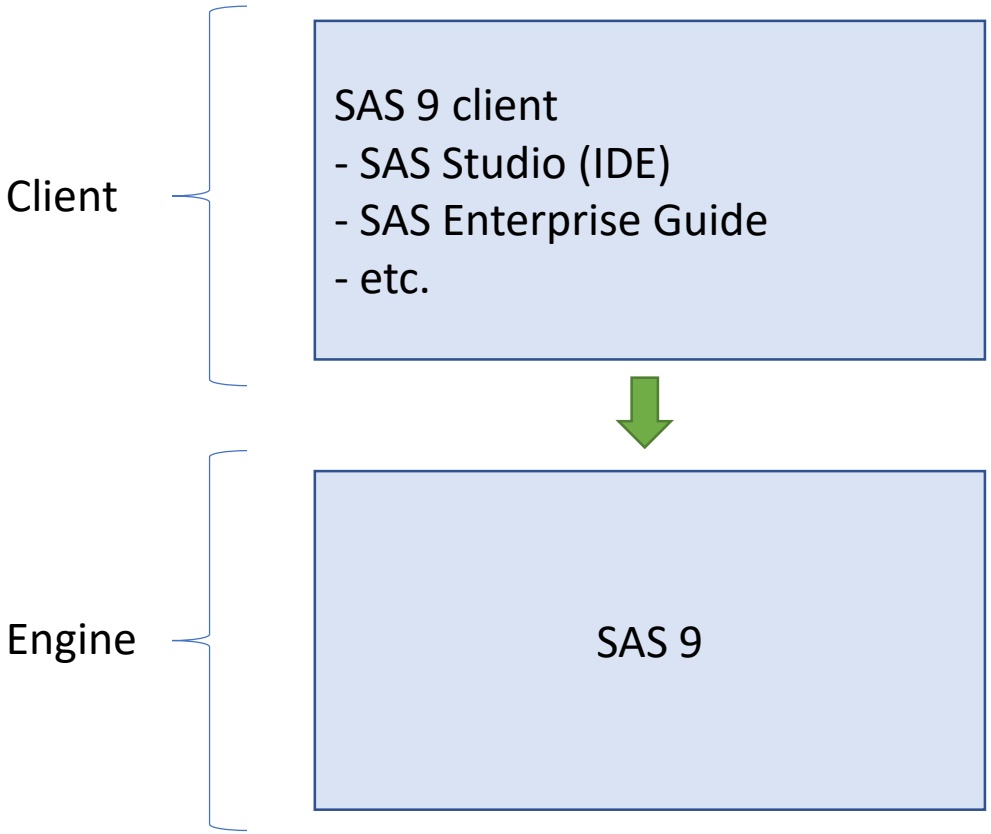
- Widely used by big companies in major industries
  - Health, banking, finance, government, etc.

SAS 9 vs SAS Viya: Two SAS Engines



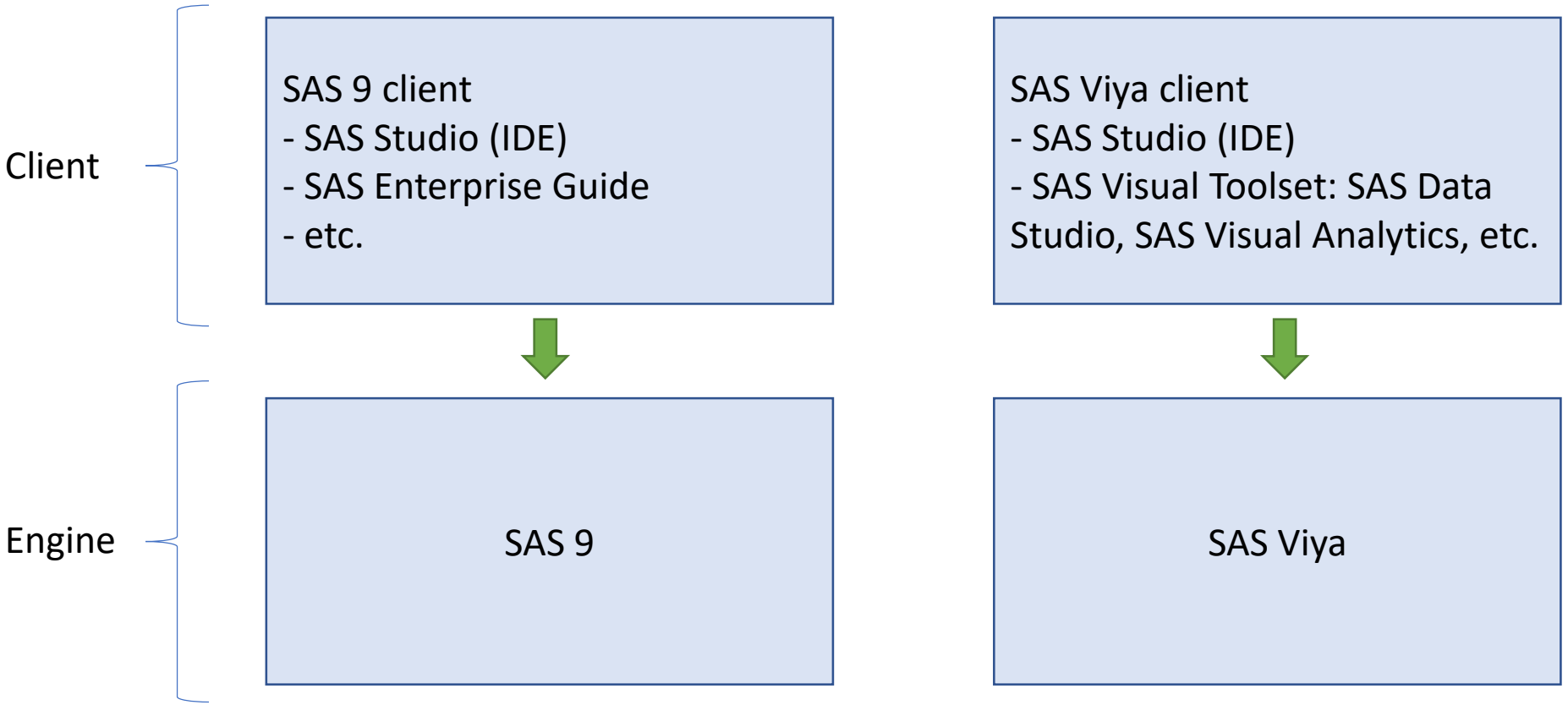
Ref. 1) SAS 9 and SAS Viya; 2) SAS 9 and SAS Viya Functional Comparison

SAS 9 vs SAS Viya: Two SAS Engines



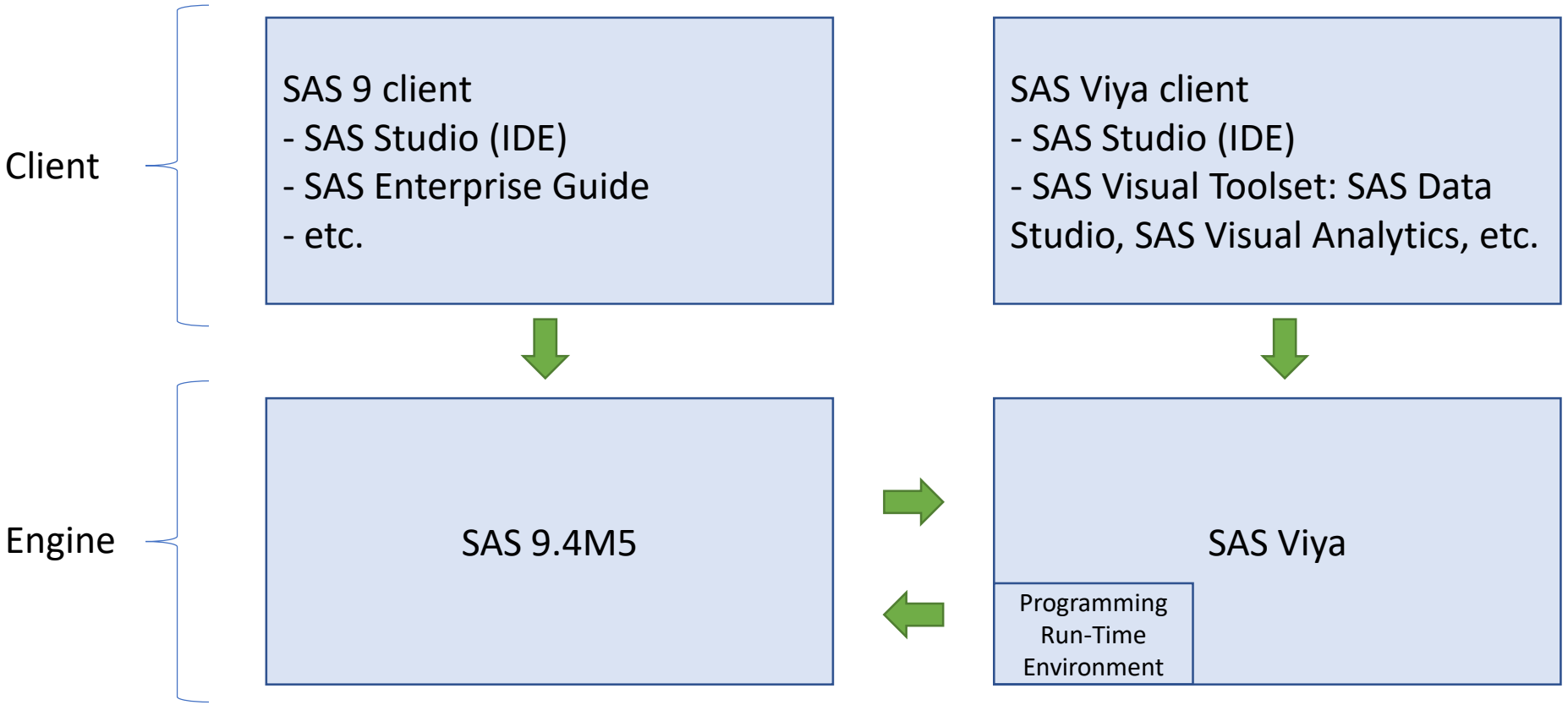
Ref. 1) SAS 9 and SAS Viya; 2) SAS 9 and SAS Viya Functional Comparison

SAS 9 vs SAS Viya: Two SAS Engines



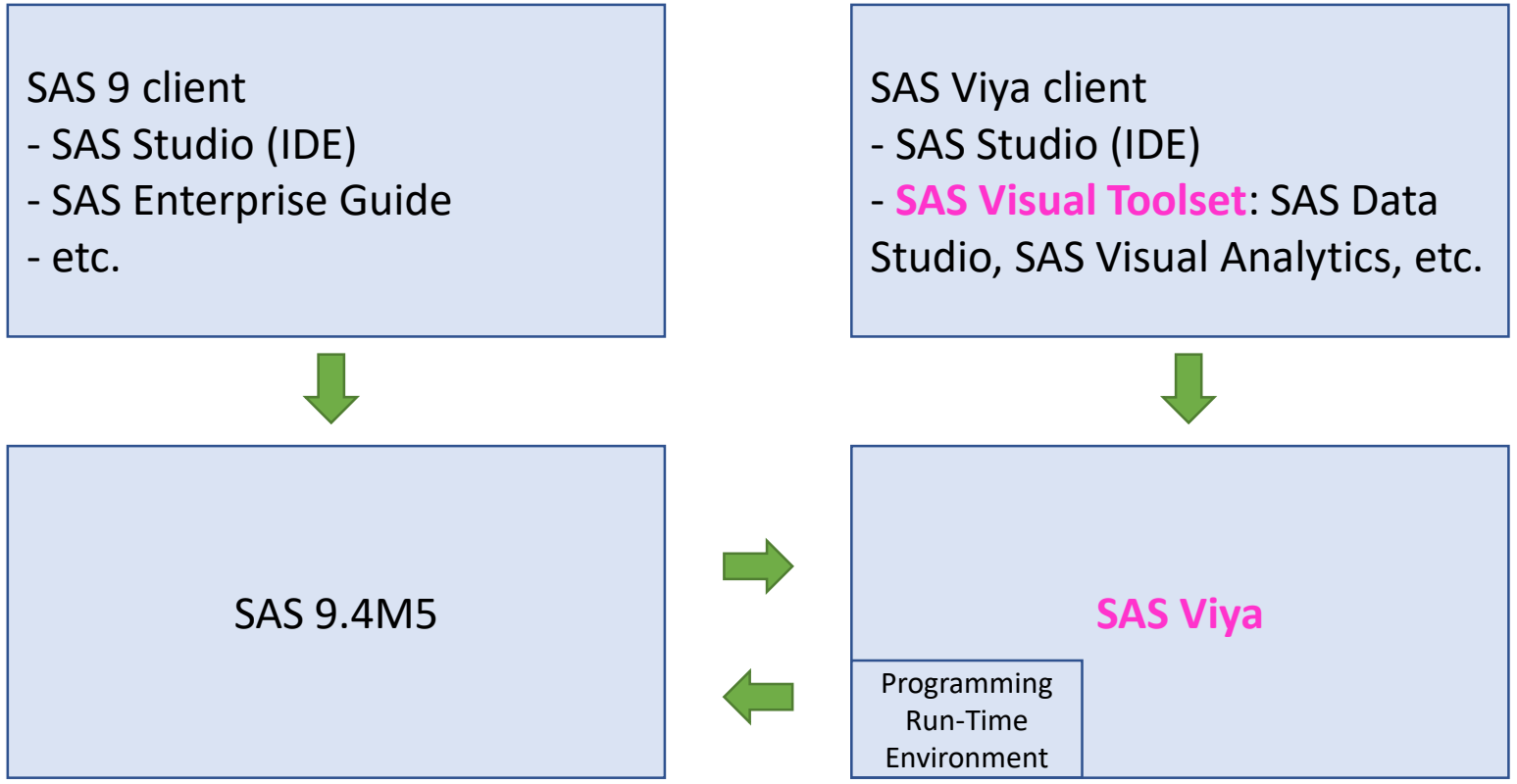
Ref. 1) SAS 9 and SAS Viya; 2) SAS 9 and SAS Viya Functional Comparison

SAS 9 vs SAS Viya: Two SAS Engines



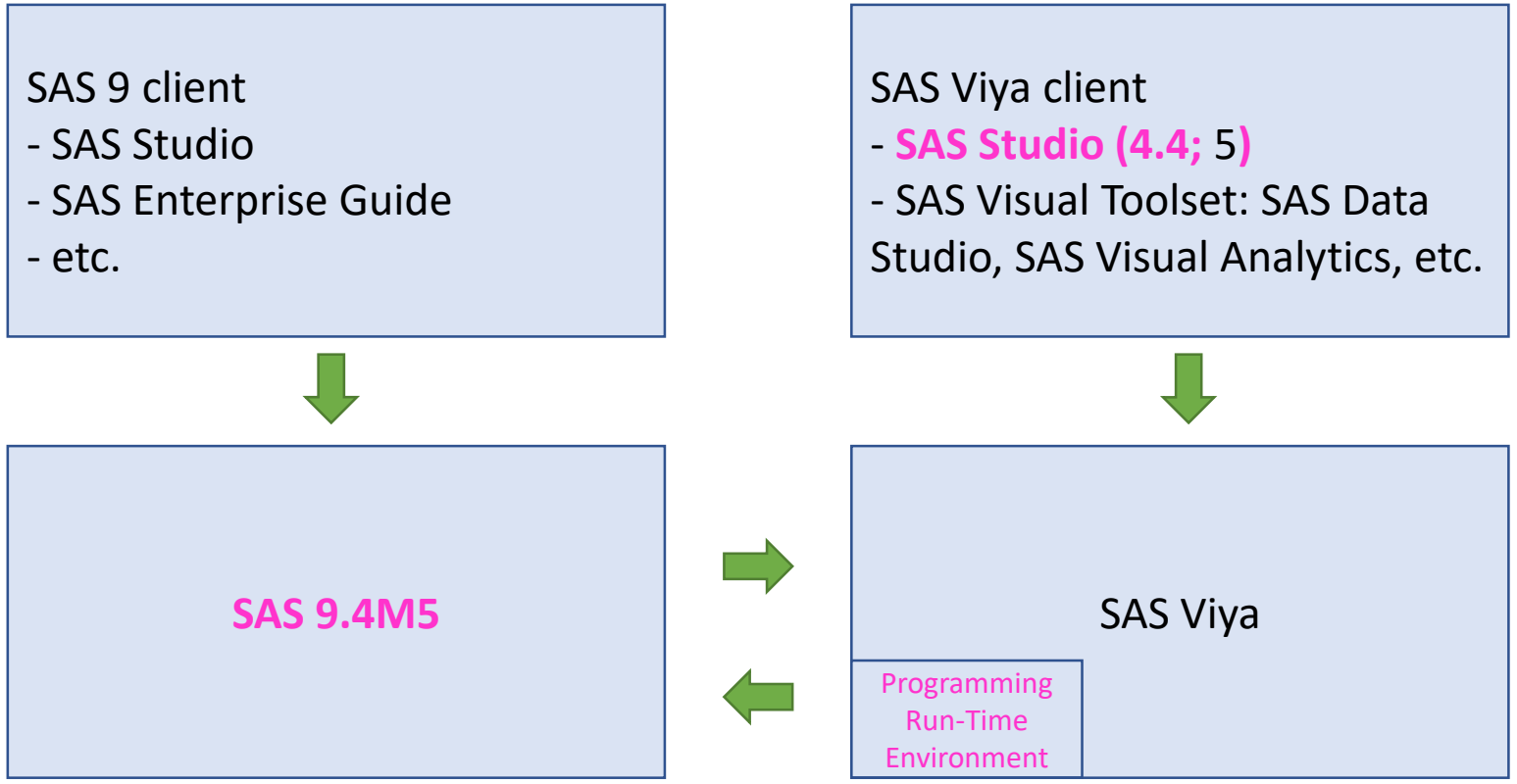
Ref. 1) SAS 9 and SAS Viya; 2) SAS 9 and SAS Viya Functional Comparison

SAS 9 vs SAS Viya (Workshop 1)



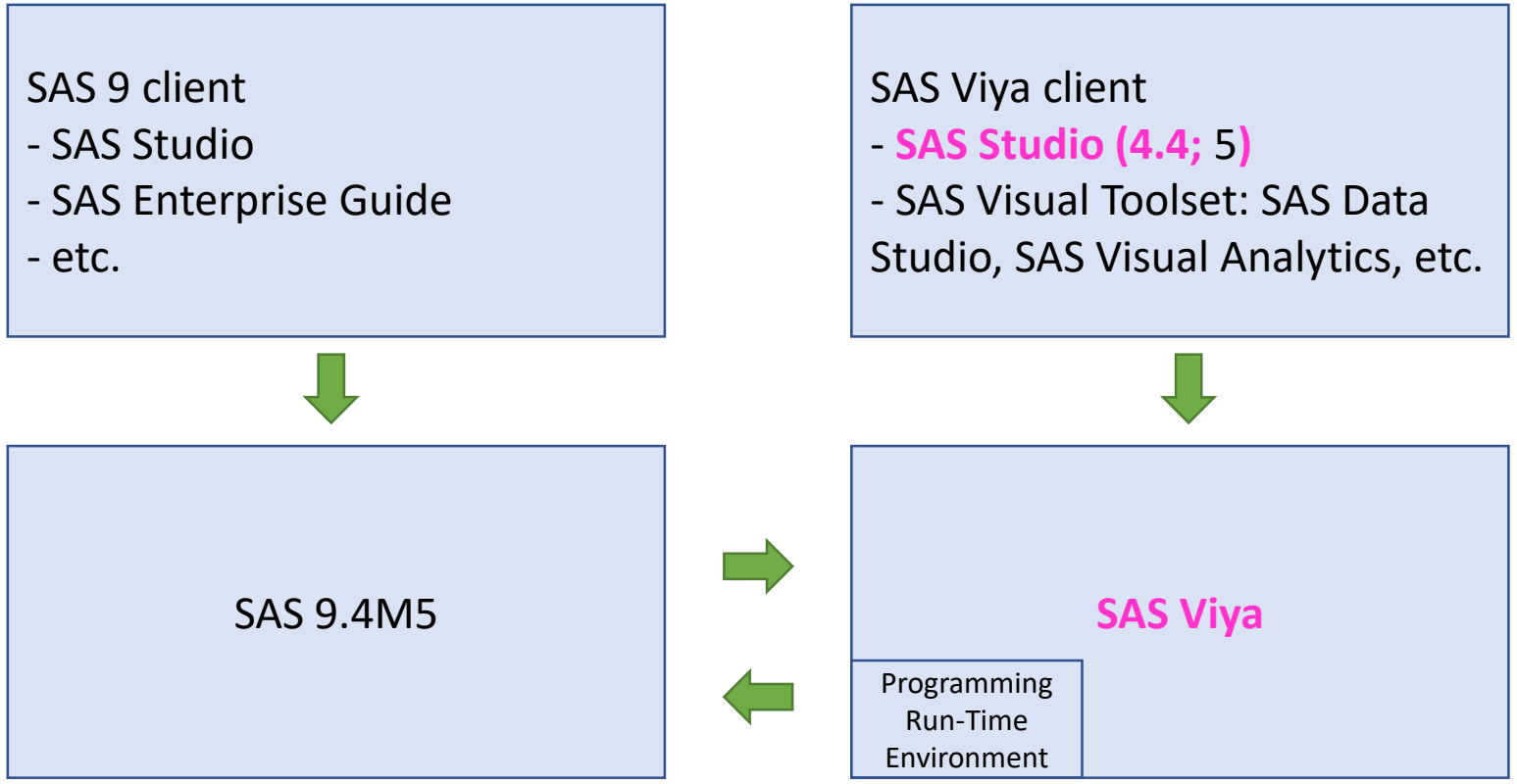
Ref. 1) SAS 9 and SAS Viya; 2) SAS 9 and SAS Viya Functional Comparison

SAS 9 vs SAS Viya (Workshop 2)



Ref. 1) SAS 9 and SAS Viya; 2) SAS 9 and SAS Viya Functional Comparison

SAS 9 vs SAS Viya (if time we have time)



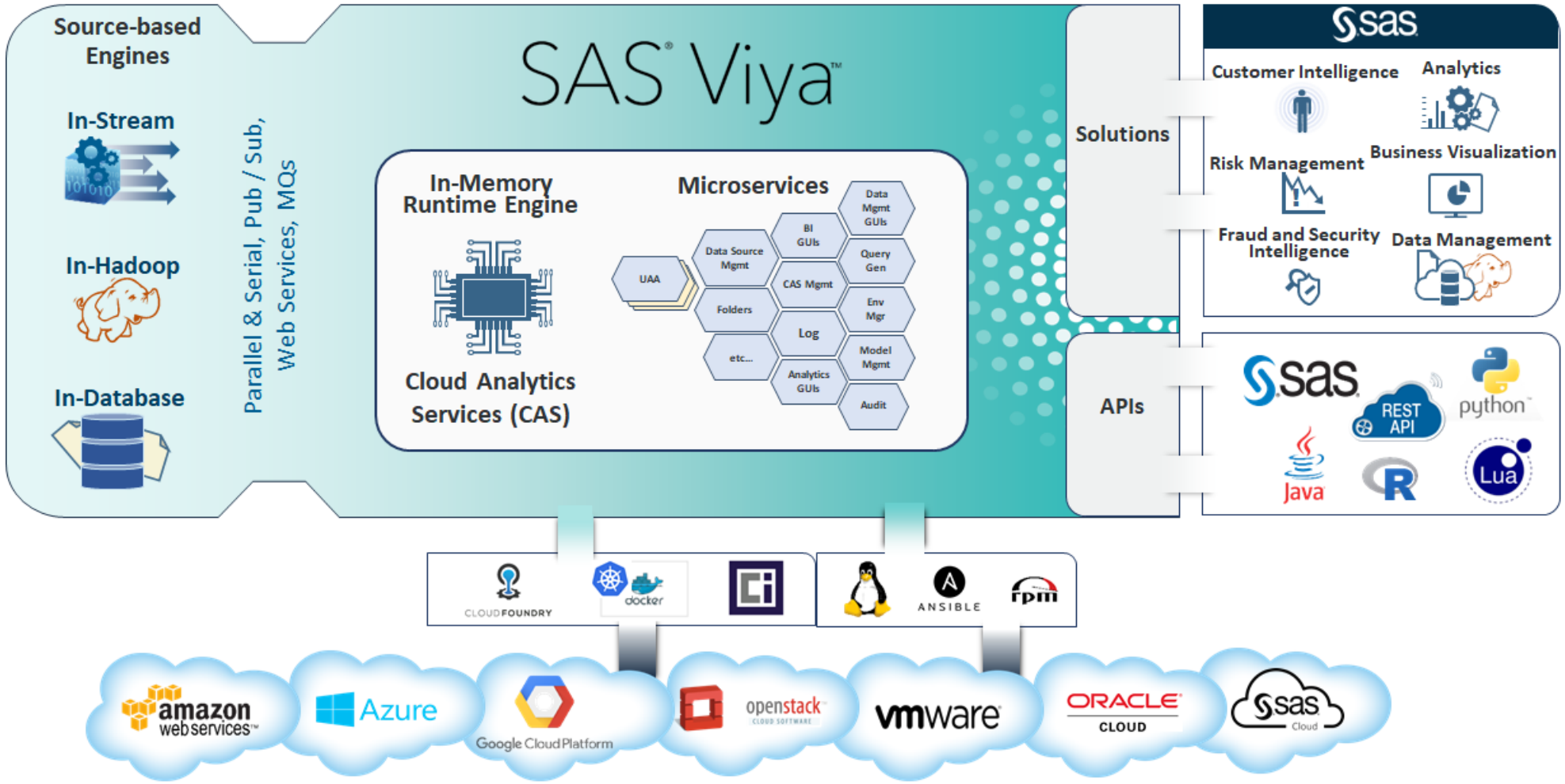
Ref. 1) SAS 9 and SAS Viya; 2) SAS 9 and SAS Viya Functional Comparison

SAS Viya

- In-memory analytics engine
  - fast
- Cloud-enabled
  - Elastic, scalable, and fault-tolerant
- Interface to open-source programming languages (APIs)
  - Python (Python-SWAT), R (R-SWAT), etc.

Ref. https://www.sas.com/content/dam/SAS/en_us/doc/overviewbrochure/sas-viya-108233.pdf

SAS Viya Cloud Architecture



Ref. https://blogs.sas.com/content/hiddeninsights/2019/01/11/how-to-get-your-head-in-the-cloud/#prettyPhoto

SAS Viya Visual Toolset Overview - SAS Drive

- A hub for SAS visual tools
  - SAS Data Prepare
  - SAS Visual Analytics, SAS Visual Statistics etc.
  - SAS Model Studio

- A place to organize and share content

SAS Viya Visual Toolset Overview - Tools

- Data Manipulation
  - SAS Data Prepare (Data Explorer, Data Studio, and Lineage Viewer)
- Analytics & Modeling
  - SAS Visual Analytics
  - SAS Visual Statistics
  - …
- Model Building, Pipeline and Comparison
  - SAS Model Studio (SAS Visual Data Mining and Machine Learning, etc.)

SAS Data Prepare

- SAS Data Explorer – Manage Data
  - Load and profile data
- SAS Data Studio – Prepare Data
  - Manipulate/transform data
- SAS Lineage Viewer – Explore Lineage
  - Visualize objects in Viya (data, transformation process, report, visualization, model, etc.)

Ref. https://support.sas.com/en/software/sas-data-preparation-support.html

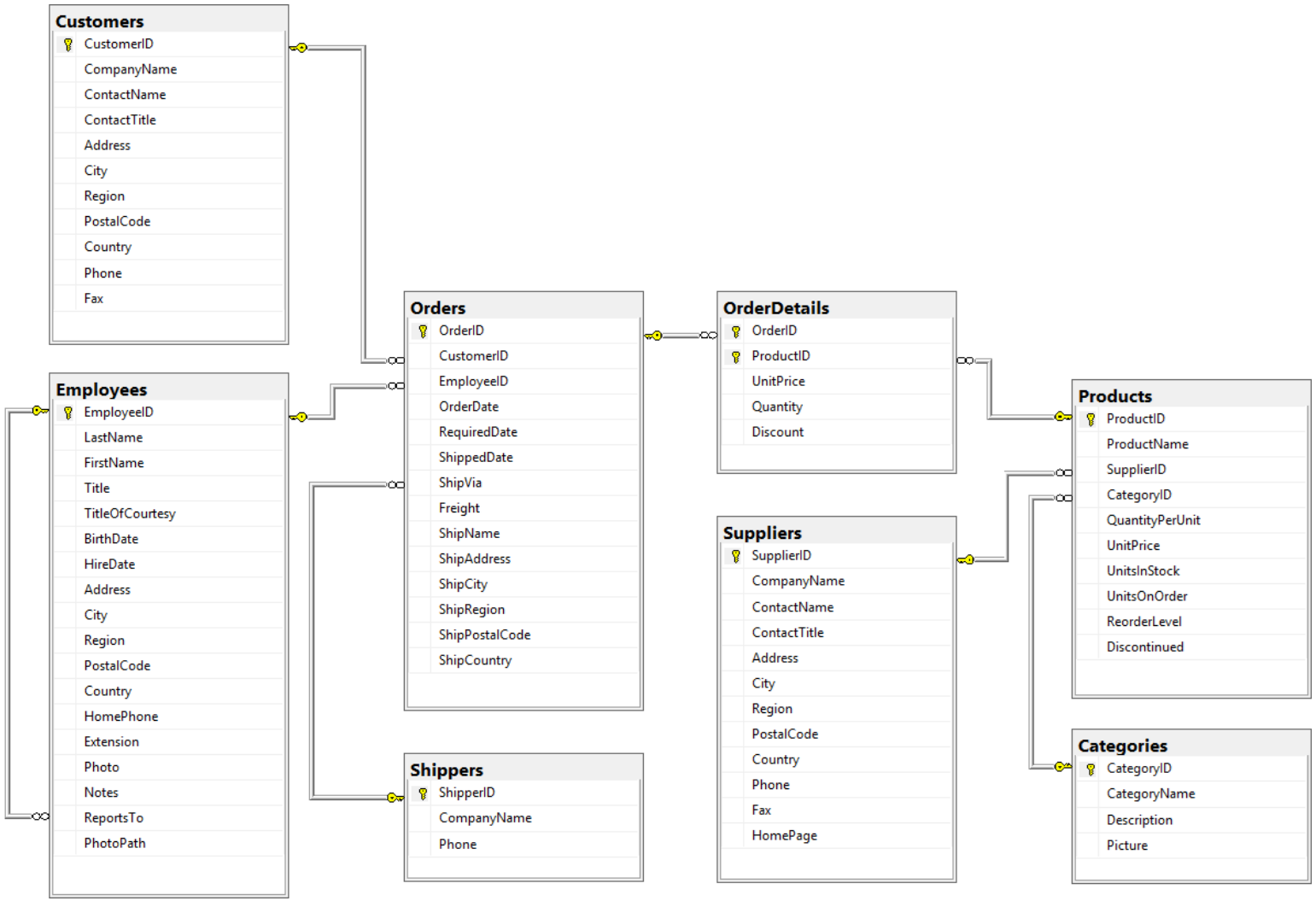
SAS Data Explorer & Data Studio – Q1

- Q1. Who are hired in 2011?
  - Use the Northwind Employees dataset
    - all datasets used today are on course website: https://tdmdal.github.io/mma-sas/

- Steps
  - Load the employee dataset
  - Convert HireDate from VARCHAR type to NUMERIC/DATETIME type (or NUMERIC/DATE)
  - Filter to obtain employees hired in 2011

Ref. informat: ANYDTDTMw.; ANYDTDTEw.

The Northwind Dataset



SAS Data Explorer & Data Studio – Q2

- Q2. How old are the employees when they were hired?
  - Use the Employees dataset

- Steps
  - Load the employee dataset
  - Convert…
  - Calculate… (Hint: use Custom Transforms -> Calculated column; and see Ref.)

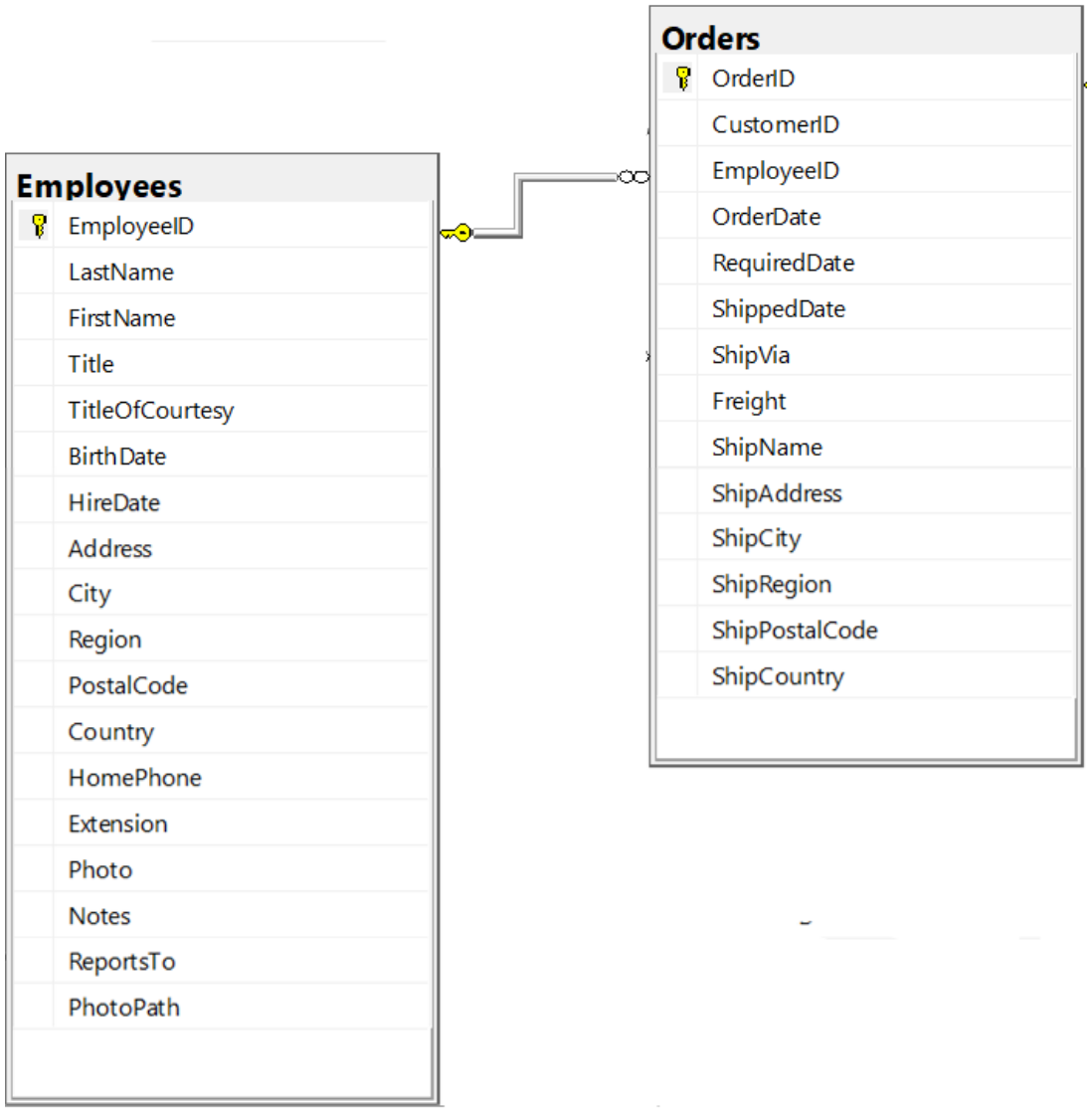
Ref. YRDIF(); INT()

SAS Data Explorer & Data Studio – Q3

- Q3. Who fulfilled the order 10248 (OrderID = 10248)
  - Use the Employees and Orders datasets

- Steps
  - Load Orders datasets
  - Join Orders table with Employees table using EmployeeID
  - Filter the resulting table to show the order with OrderID = 10248

Join Employees and Orders datasets



Join – Inner Join

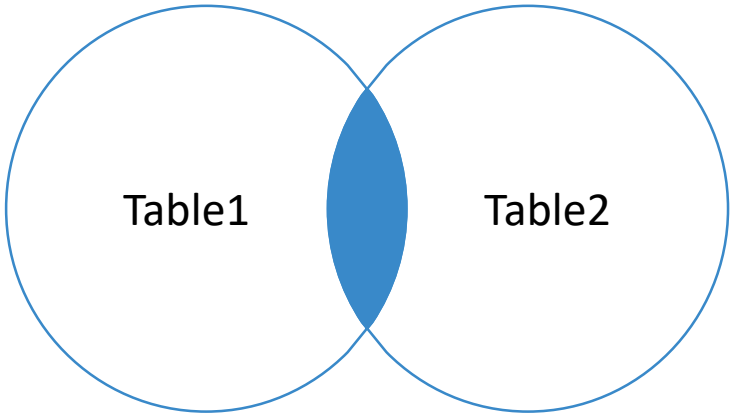


Table1

pk	t1c1
1	a
2	b

Table2

fk	t2c1
1	c
1	d
3	e

pk	t1c1	fk	t2c1
1	a	1	c
1	a	1	d

Join – Left (Outer) Join

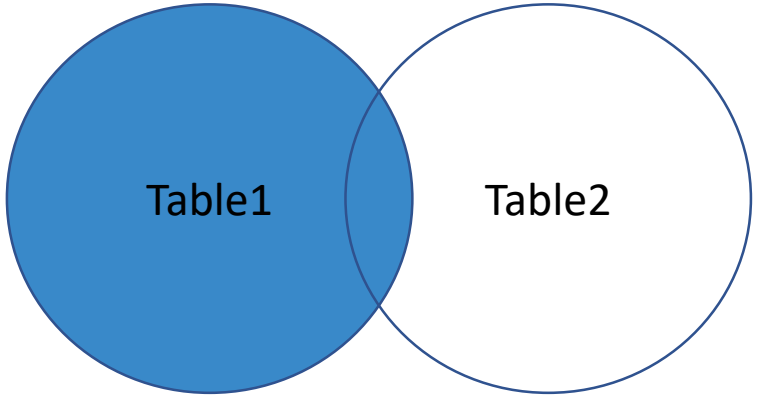


Table1

pk	t1c1
1	a
2	b

Table2

fk	t2c1
1	c
1	d
3	e

pk	t1c1	fk	t2c1
1	a	1	c
1	a	1	d
2	b	null	null

Join – Right Outer Join*

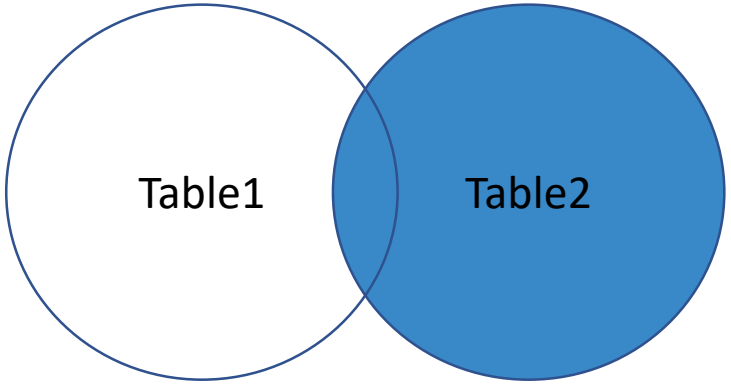


Table1

pk	t1c1
1	a
2	b

Table2

fk	t2c1
1	c
1	d
3	e

pk	t1c1	fk	t2c1
1	a	1	c
1	a	1	d
null	null	3	e

Join – Full (Outer) Join

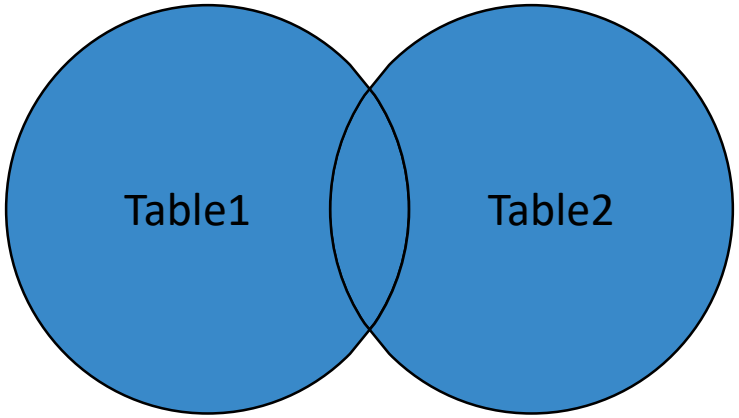


Table1

pk	t1c1
1	a
2	b

Table2

fk	t2c1
1	c
1	d
3	e

pk	t1c1	fk	t2c1
1	a	1	c
1	a	1	d
2	b	null	null
null	null	3	e

SAS Data Explorer & Data Studio – Q4

- Q4 What's the total revenue by year
  - Use Orders and OrderDetails datasets
  - When calculating revenue, ignore discount

- Steps?

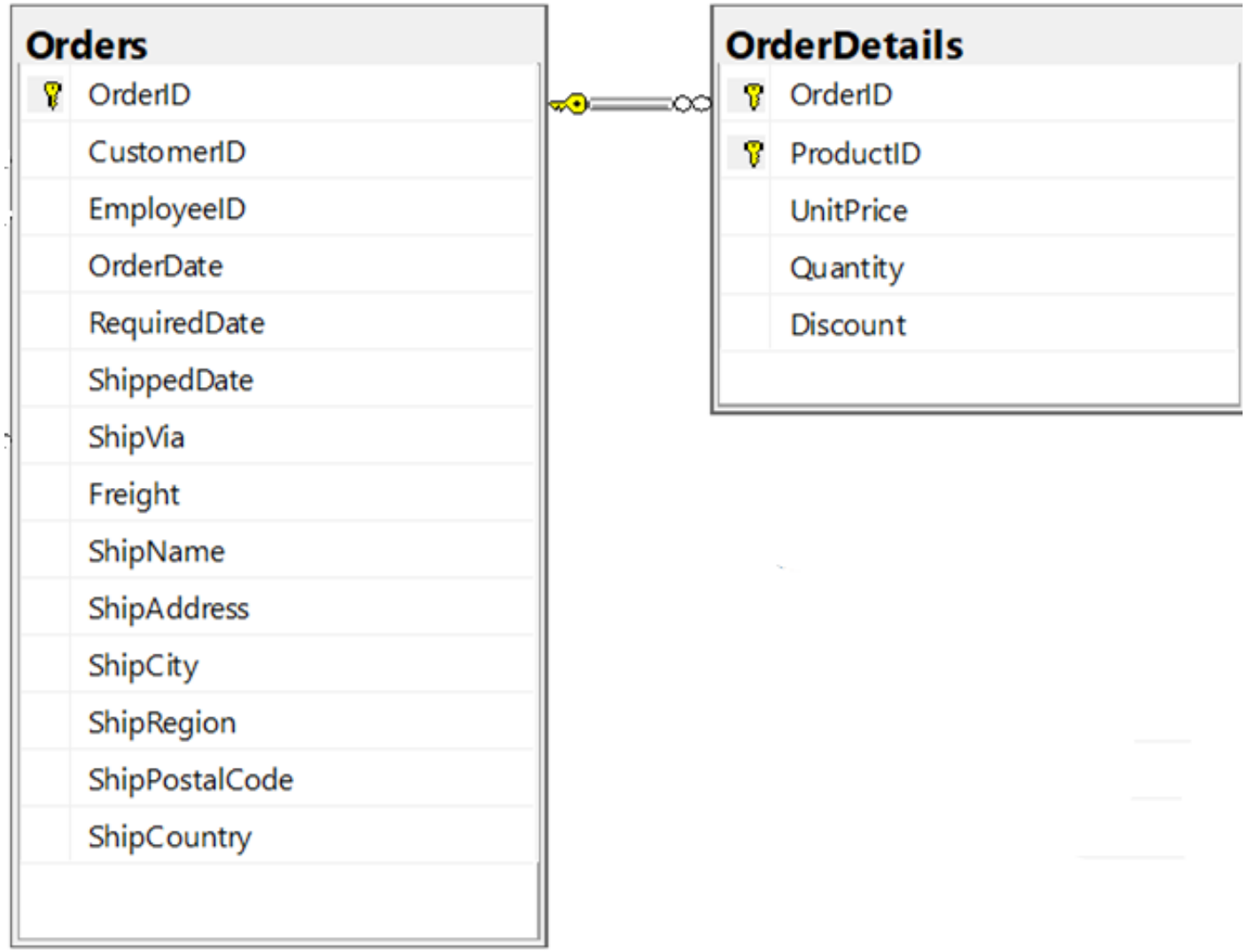
SAS Data Explorer & Data Studio – Q4

- Q4 What's the total revenue by year
  - Use Orders and OrderDetails datasets
  - When calculating revenue, ignore discount

- Steps
  - Load data
  - Join Orders and OrderDetails by OrderID
  - Calculate sales for each item in an order
  - Convert OrderDate to NUMERIC/DATE and extract year
  - Aggregate by year (how? CASL and FEDSQL;)

Ref. https://blogs.sas.com/content/sgf/2018/03/16/sas-data-studio-code-transform-part-1/

Join Orders and OrderDetails



SAS Visual Analytics – Q5

- Produce a bar plot
  - Revenue by country and year (together)
  - Revenue by country and employee (home work)

- Steps
  - in Data Studio
    - Join datasets…
    - Convert…
  - in Visual Analytics

Ref. https://support.sas.com/en/software/visual-analytics-support.html

SAS Visual Statistics

- Get started on your own using these video tutorials
  - Getting Started with SAS Visual Statistics (7m)
  - Building a Logistic Regression Model (5m)
  - Building a Clustering Model (4m)

- Build a churning model
  - Use this dataset: Telco Customer Churn

Ref. https://support.sas.com/en/software/visual-statistics-support.html

SAS Model Studio

- Get started on your own using these video tutorials
  - Getting Started with Data Mining and Machine Learning Pipelines
  - Managing Variables
  - Feature Engineering
  - The Variable Selection
  - Imputing Missing Values

- Another churning dataset for practice
  - churn_telecom.sas7bdat (in SAS data format)

Ref. https://support.sas.com/en/software/visual-data-mining-and-machine-learning-support.html