
SAS MELBOURNE USER GROUP

3 JUNE 2015



Melbourne Users Exploring SAS working with Partners & Banking

MEET OUR SMUG COMMITTEE 2015



- **Gary Anderson – Chairperson**



- **Elisa Young**



- **Flavio Giurco**



- **Leslie Shroot**



- **Neville Ford**

**“WE WANT TO HEAR
FROM YOU...
WHAT IS YOUR SAS
STORY?”**

Please introduce yourself to the committee during our networking session this evening

VOTE FOR TODAY'S BEST PRESENTATION...

DON'T FORGET TO COMPLETE THE SMUG SURVEY

By completing the survey you could win...

SAS goodies – a USB speaker



Using SAS Within a robust Data Integration Framework

Why I Don't Just Write Code !



Spoiler Alert !

- Who am I
Where do I come from
Why I consider myself a hardcore programmer
- What's a Robust DIF ?
- Using DI - SAS Data Integration Studio
- My Call To Action...

What's a Robust DIF ?

- Have you ever wished your colleagues **kept track** of their **code** ?
- Have you ever wished you had a **data lineage** or a **catalogue** of your **reports** ?
- Have you ever wished that when your sources change stuff, the **consequential impacts** are part of a report and NOT part of three weeks of trawling through custom code ?

What's a Robust DIF ?



Copy Paste - %include - macros - eGuide - Libraries - SAS DI

The Evolution of Code

What's a Robust DIF ?

- Well documented
- Maintainable
- Portable
- Standard framework

Case Study

- Import SAS package and using sashelp.class:
 - Filter out all people with a 'p' in their name
 - Apply 10% extra to the height of all females
 - Sort by descending weight
 - Output to a DB
 - Perform Impact Analysis



So what have you just seen ?

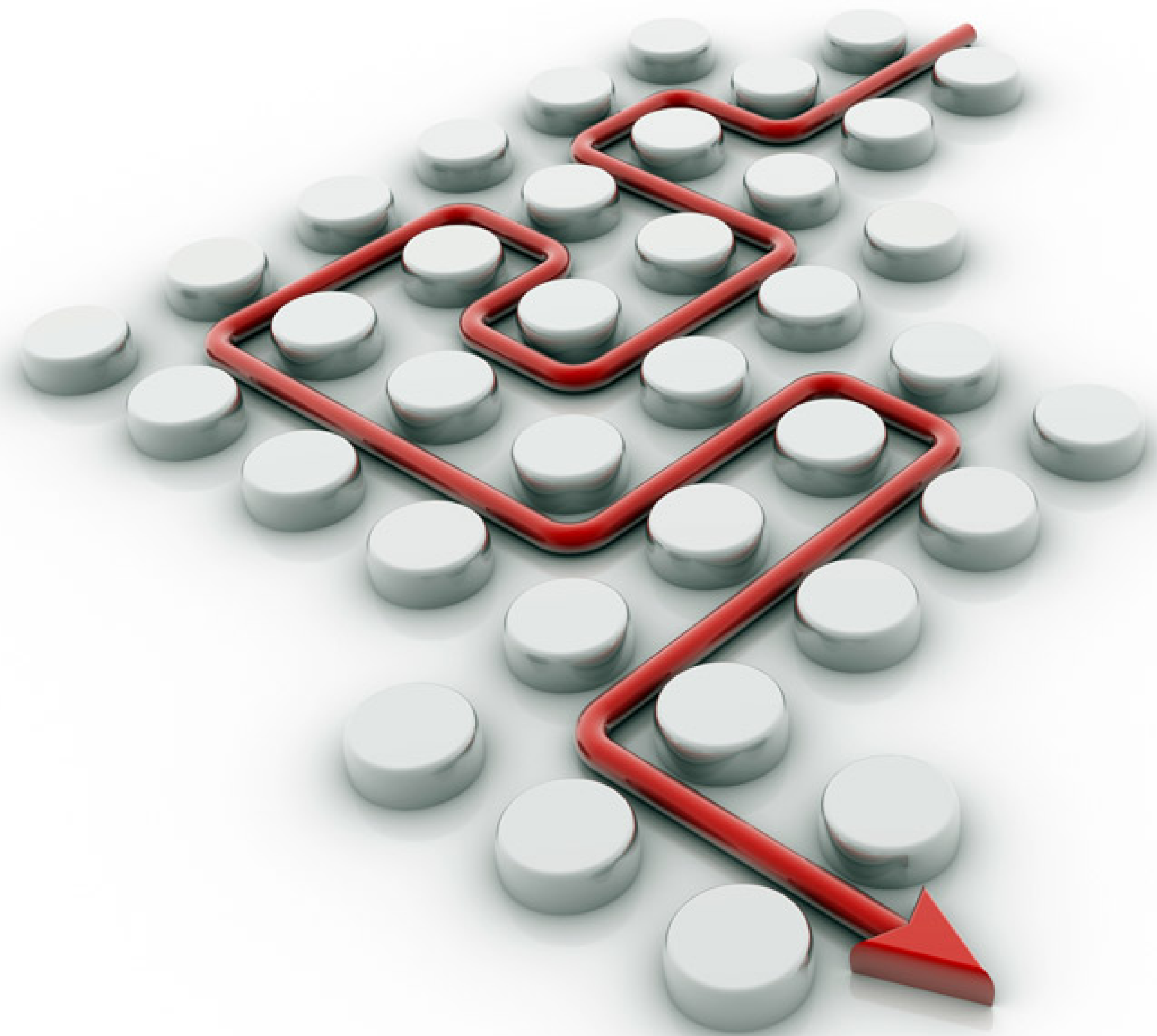
- Importing a SAS package
- Writing code by point and clicks
- Data Lineage
- An example of a Data Integration Framework

! CALL TO ACTION !

Adoption of DI Studio to support the hardening of SAS solutions

AND

Read generated code because you 'wrote' it



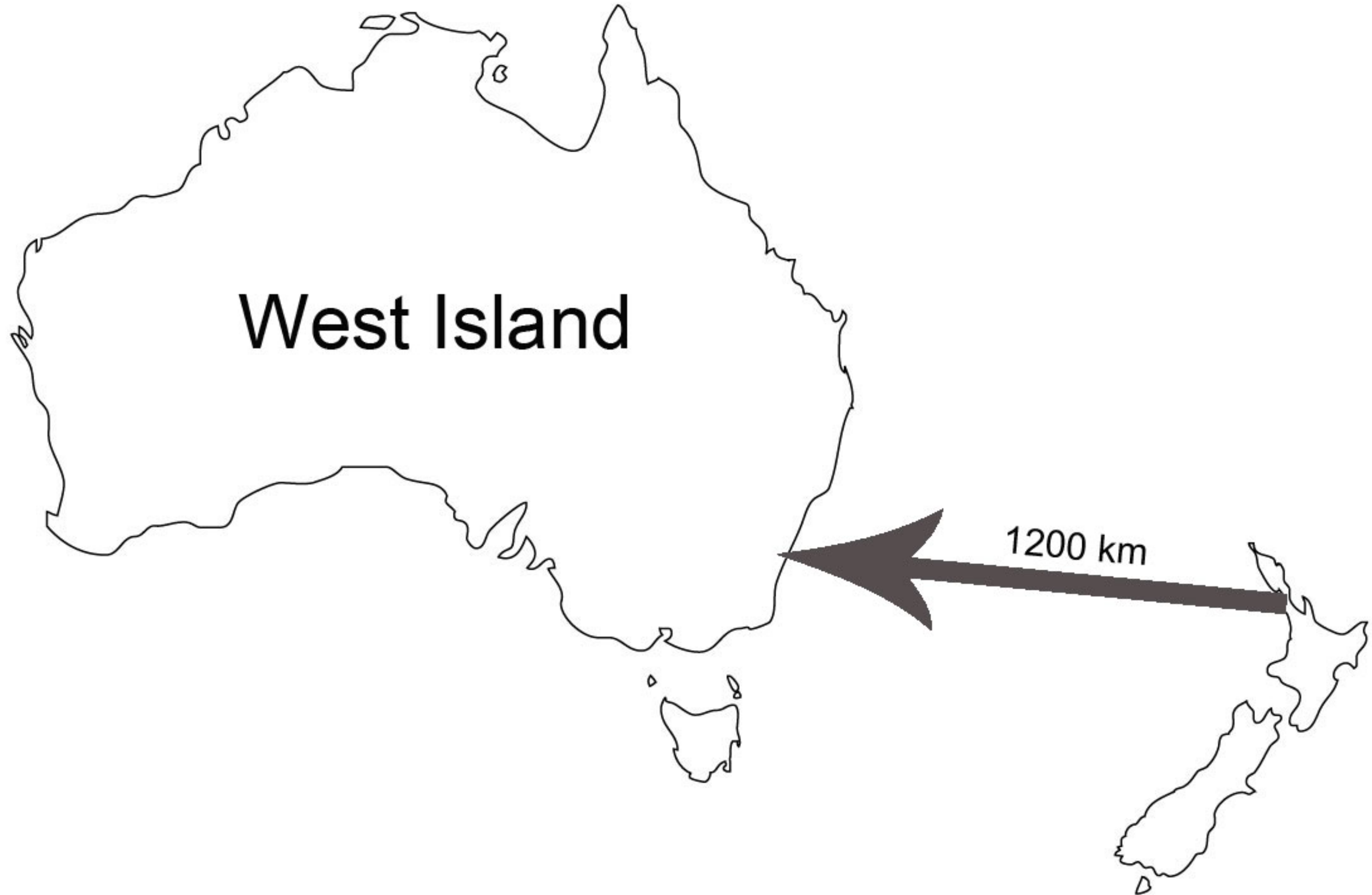
Richard Carson
Senior SAS Designer



Bridging the Gap

Between SAS Credit Risk Management for Banking
and APRA regulatory reporting





West Island

1200 km





Skill
Strength
Strategy



Style
Support
Passion

Passion + Skill + Strategy

Offensive
Strategy



Defensive
Strategy



Banks are Taking Risk

The officers and directors of the bank are men of proven integrity and ability, and it is also under the direct supervision of the United States Treasury Department, giving double assurance of a sound, satisfactory administration of its affairs.

**New
Accounts
Are
Invited**



Officers:

T. W. STUBBLEFIELD,
President.
FRANCIS A. BLUNDON,
First Vice President.
M. HOLZHEIMLEIN,
Second Vice President.
WM. H. DE LASHMUTT,
Cashier.

Board of Directors:

W. H. Bross.
Isaac H. Ford.
M. Holzheimlein.
E. F. Anderson.
Fred H. Geiser.
W. H. De Lashmutt.
Jesse H. Hamburgh.
Julian P. Dodge.
Robert Allen.
Jesse Cook.
Howard Boyd.
T. W. Stubblefield.
C. H. Kies.
A. R. Bross.
W. G. Kistner.
H. Ellsworth Banta.
Francis A. Blundon.
D. H. Buckingham, Jr.
J. R. Clemens.
Ed. J. Hammett.
Jacob Lieber.
J. V. Stuckert.

Counsel, Howard Boyd.



Australian Prudential Regulation Authority

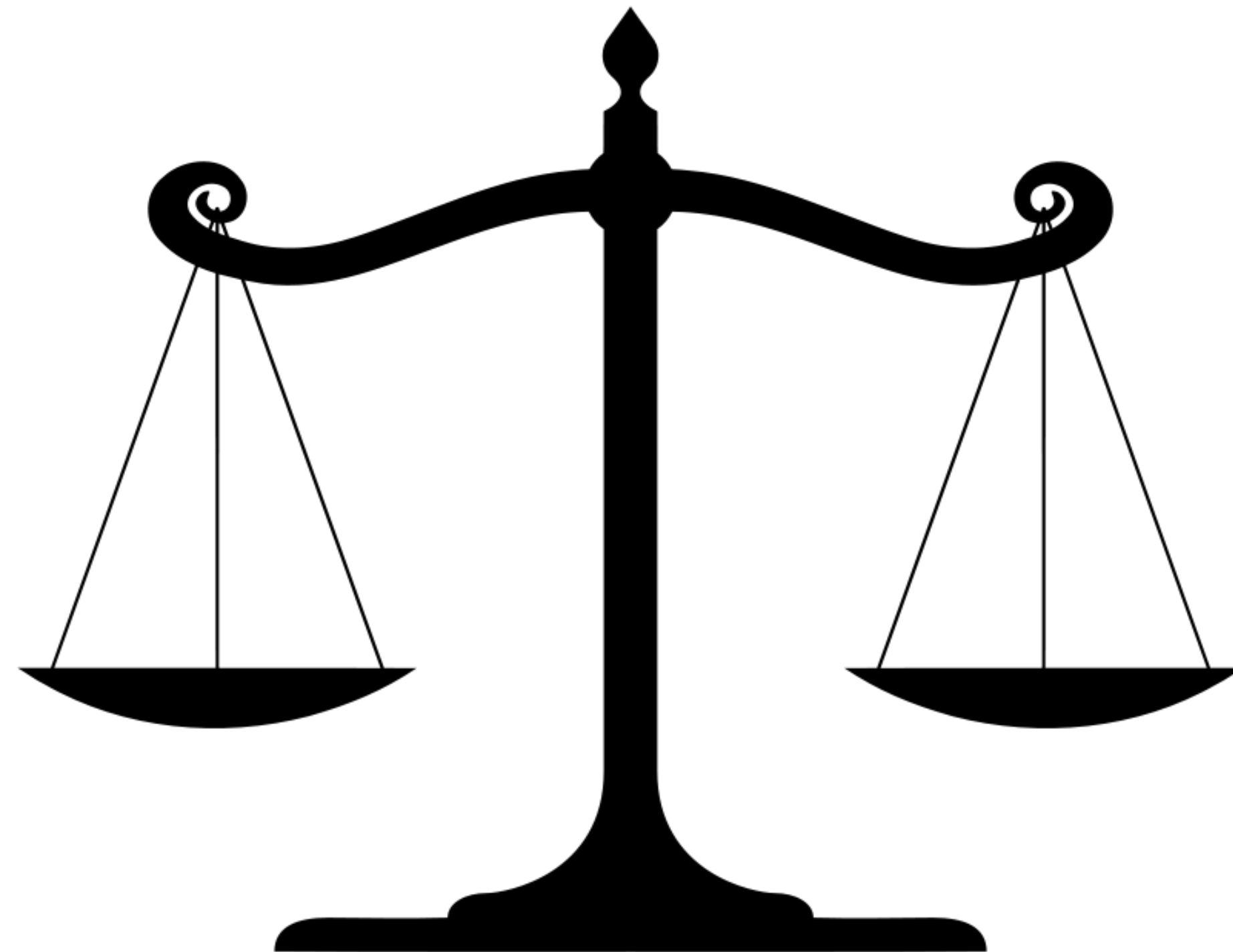


Accreditation Authority
STD & IRB Reporting Specifications
Public Disclosure Requirements

The Great Credit Risk Quandary

Standard Approach







- Less Computationally Intensive
- Potentially Higher Capital Requirement
- Less Informative

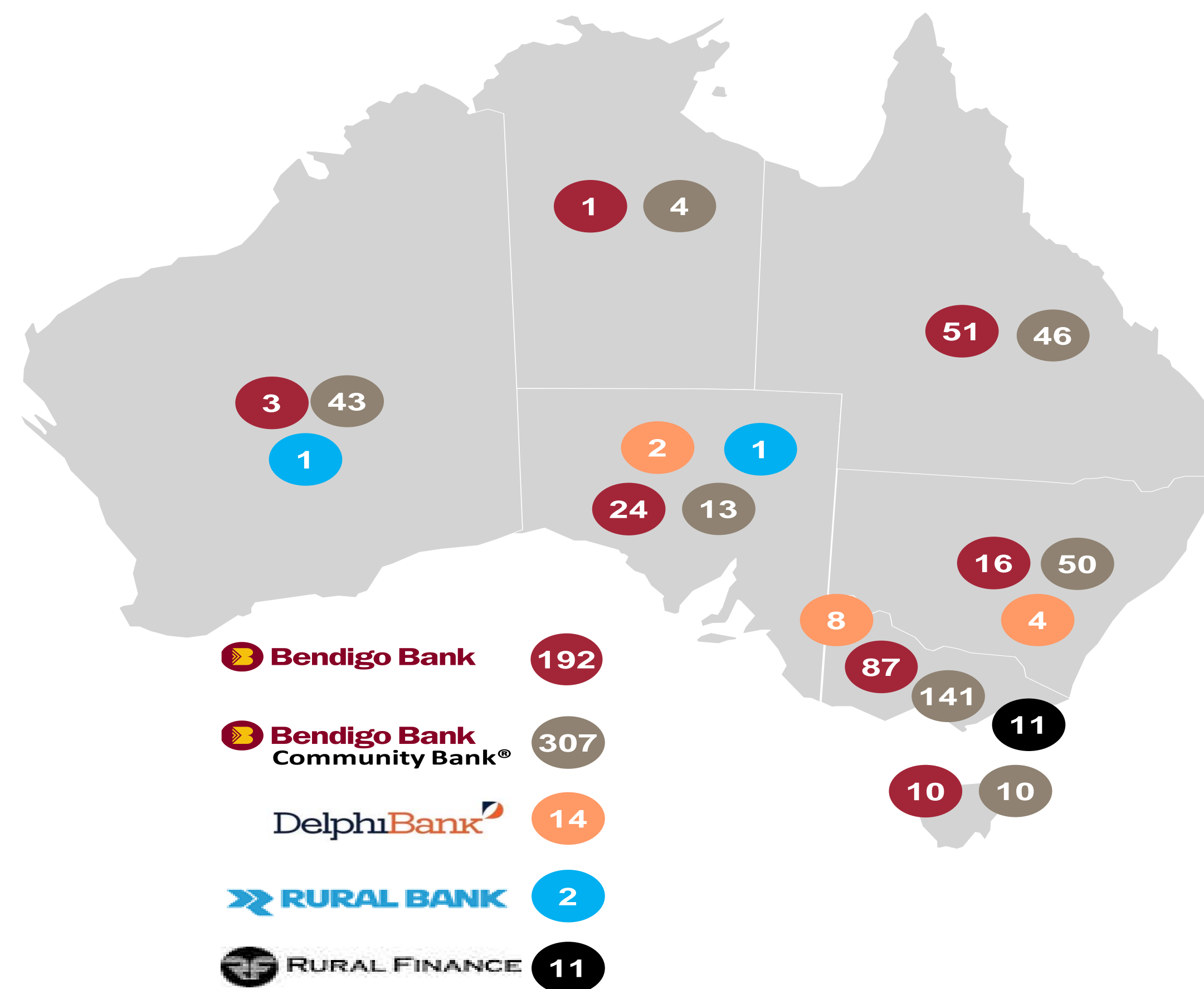


IRB Approach

- More Computationally Intensive (\$\$\$)
- Potentially Better Capital Utilisation
- Detailed Risk Analysis

DIVERSE RETAIL BANK

Business	Segment
 Bendigo and Adelaide Bank	Domestic bank operating five distinct brands
 Bendigo Bank	Consumer and SME banking
 Adelaide Bank	Third party channel
 Bendigo Wealth	Margin lending, wealth management, superannuation
 RURAL BANK RURAL FINANCE	Agri-business banking
 Delphi Bank	Consumer and SME banking for Hellenic communities



We're aiming to be Australia's most customer connected bank

GOALS

Business

- Contribute towards our customer connected vision
- Generate analytics that provide meaningful insights
- Contribute towards value creation
- Improve responsiveness and efficiencies
- Build a strategic, adaptive and enduring capability

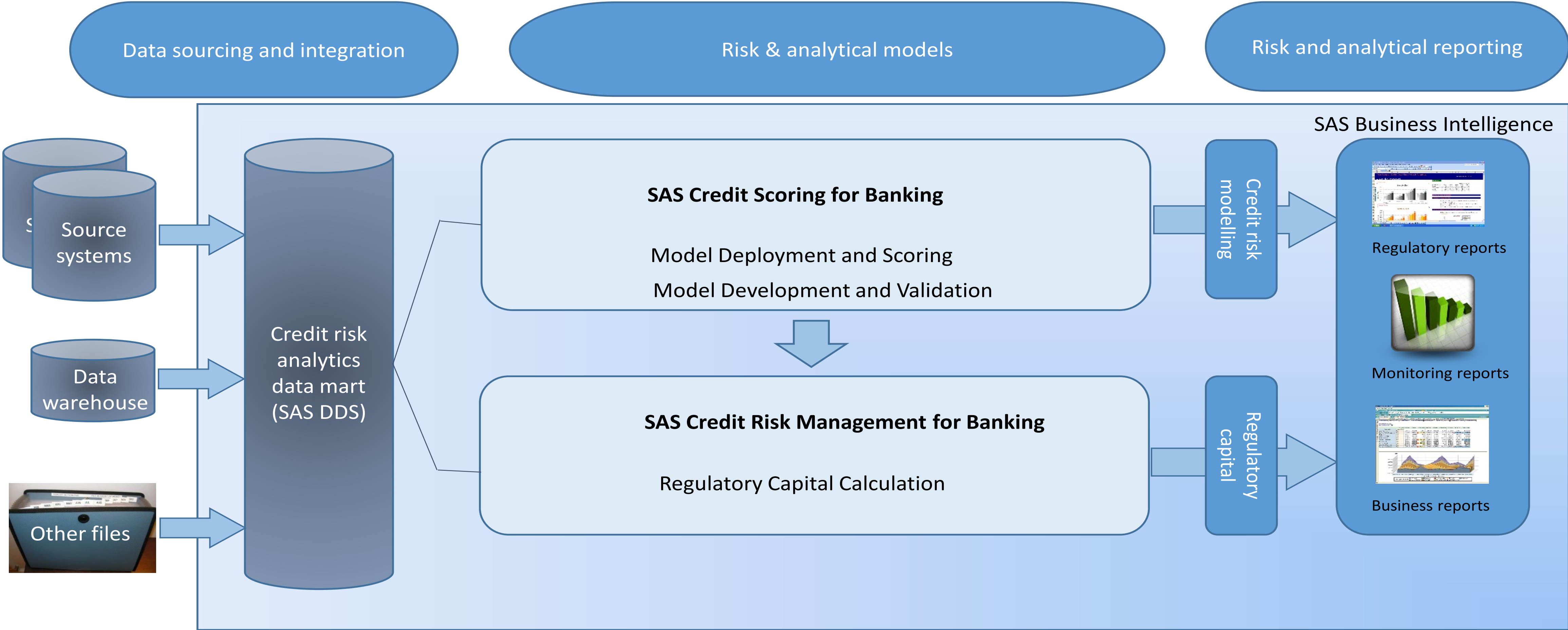
Regulatory

- Establish a stable and efficient platform for credit risk modelling and reporting
- Support business growth and regulatory changes
- Promote confidence in, and integrity of, data and information
- Establish auditability and traceability of end to end process
- Support accreditation goals

supporting and enabling a customer connected strategy

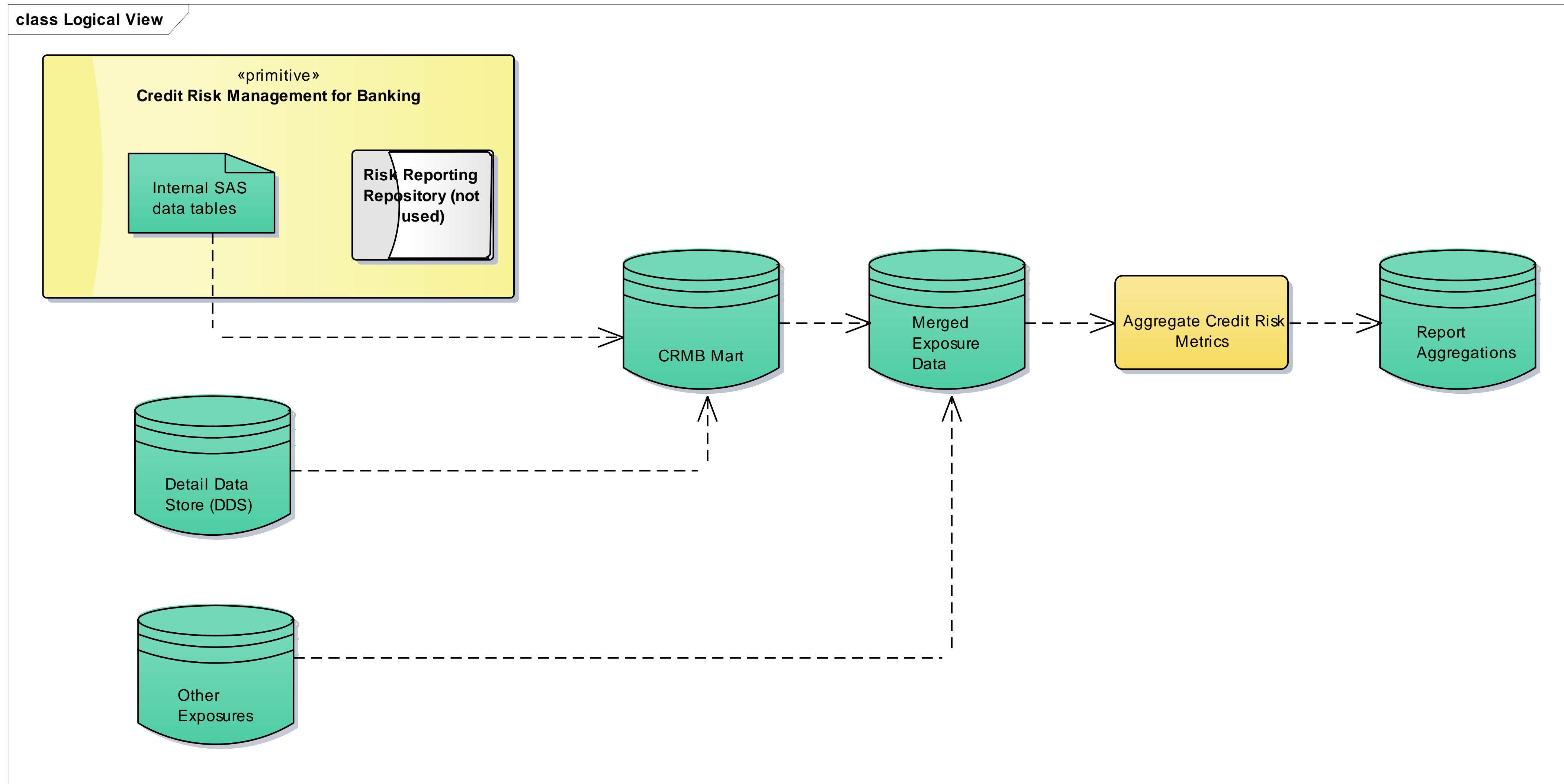
CREDIT RISK BUSINESS PLATFORM

Contemporary and scalable





Credit Risk Reporting Store



DI Studio implementation of Aggregation process

1. Consistent with other processes in Credit Risk Reporting Store
2. Job will be run as a step within the Reporting process flow
3. Preserves Data Lineage

However...

Default DI Studio options for Summary Statistics do not produce an output table

Base SAS view of aggregation process

```
LIBNAME ossalesa BASE "S:\workshop\Ori onStar" ;
```

```
Proc SQL ;  
  Create table Sales_Analysis_Dec2005 As  
  Select *  
  From   ossalesa.sales_analysis  
  Where  Year   =   2005  
        And Month = 12  
        ;  
Quit ;
```

```
Proc Means NoPrint  
  Data   =   Sales_Analysis_Dec2005  
  ;  
  Class  Vendor  
        ProductCategory  
        ProductSubcategory  
        ProductName  
  ;  
  Var    TotalSales  
  ;  
  Id     Year  
        Month  
  ;  
  Output Out = Sales_Summary_Dec2005  
        Sum =  
        / ways  
  ;  
  Types  Vendor  
        Vendor * ProductCategory  
        ProductCategory * ProductSubcategory * ProductName  
  ;  
Run ;
```

SAS Data Integration Studio 4.6 - ST

File Edit View Check Outs Actions Debug Tools Window Help

New [Icons] Workspace Server

Transformations
Inventory
Folders Checkouts

Imzce4

- Create REPSTG.FINANCIAL_CLASSES
- New Job 33031
- SMUG**
- Test CRMB

Basic Properties

Name	Value
Name	SMUG
Description	Demonstration of R...
Folder Location	/Products/Bendigo B...
Checked Out By	
Tables Loaded	
Last Modified By	Richard Carson
Metadata Modified	18/05/2015 1:01:04...
Metadata Created	18/05/2015 1:01:04...
Metadata Archived	
Metadata Archived As	
Logical Type	Job
Metadata ID	A5WR1DEU.A700002F
Usage Version	1.0

SMUG

Up Run Stop [Icons]

Diagram Code Log Output

Details

Status Warnings and Errors Statistics Control Flow

[Icons] Clear All

Order	Name	Status	Details

SASApp admzce4@BBLDTL as Richard Carson sas-metadata.bbld...

SAS Data Integration Studio 4.6 - ST

File Edit View Check Outs Actions Debug Tools Window Help

New [Icons] Workspace Server

Inventory

Folders Checkouts

Transformations

SQL

- Create Table
- Delete
- Execute
- Extract
- Insert Rows
- Join
- Merge

Basic Properties

Name	Value
Name	Extract
Description	
Node Type	Extract
User Written	No
Diagnostic Mode	No
Checkpoint	No
Run always when re...	No
Control Order	
Metadata ID	A5WR1DEU.\$0000000

SMUG *

Up Run Stop [Icons]

Diagram Code Log Output

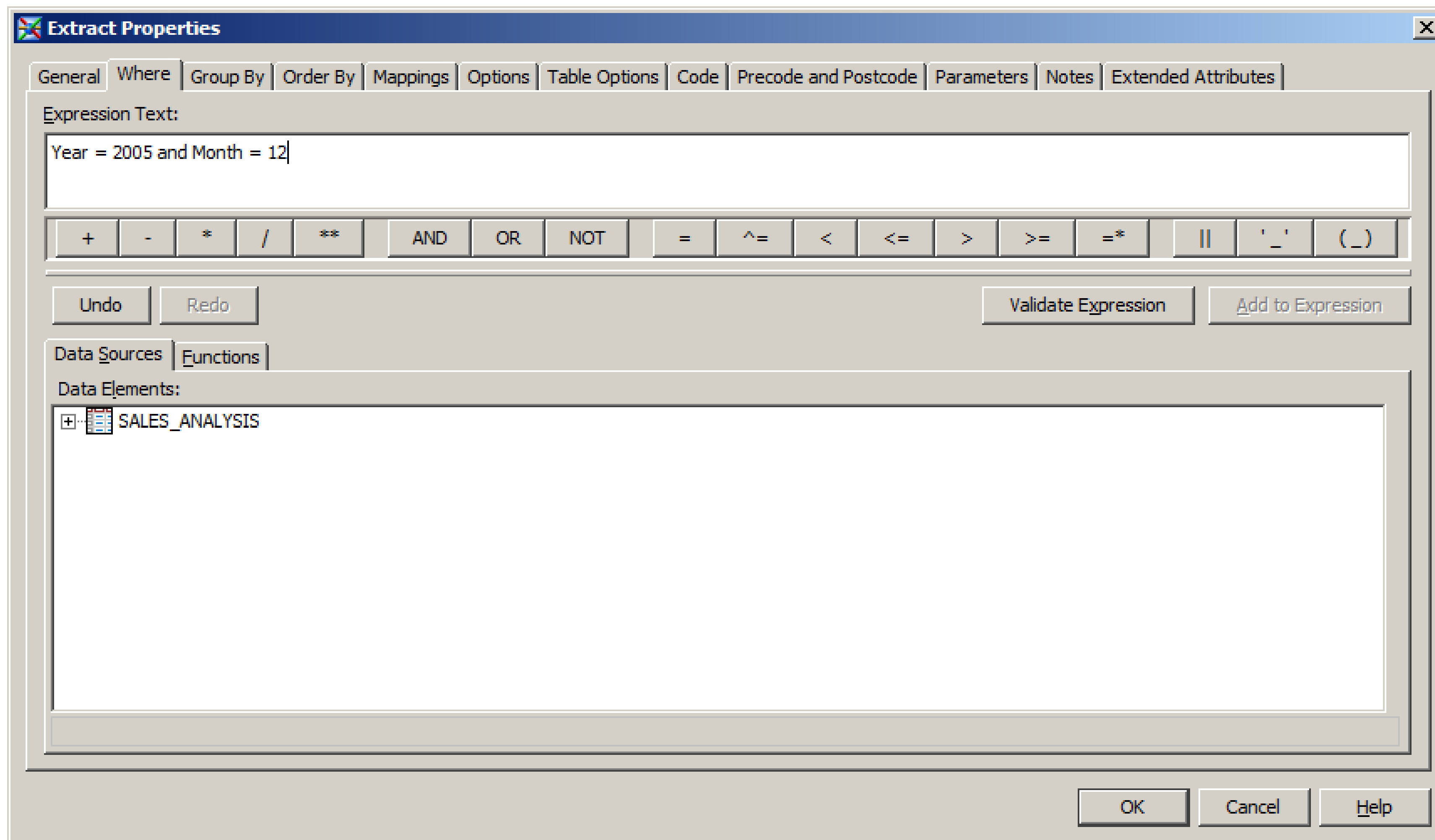
Details

Mappings Status Warnings and Errors Statistics Control Flow

[Icons] Clear All

Order	Name	Status	Details

SASApp admzce4@BBLDTL as Richard Carson sas-metadata.bbld...



SAS Data Integration Studio 4.6 - ST

File Edit View Check Outs Actions Debug Tools Window Help

New [Icons] Workspace Server

Inventory

Folders Checkouts

Transformations

SQL

- Create Table
- Delete
- Execute
- Extract
- Insert Rows
- Join
- Merge

Basic Properties

Name	Value
Name	Extract
Description	
Node Type	Extract
User Written	No
Diagnostic Mode	No
Checkpoint	No
Run always when re...	No
Control Order	1
Metadata ID	A5WR1DEU. \$0000000

SMUG *

Up Run Stop [Icons]

```

    graph LR
      SALES_ANALYSIS[SALES_ANALYSIS] --> Extract[Extract]
  
```

Diagram Code Log Output

Details

Mappings Status Warnings and Errors Statistics Control Flow

Last Run: 21/05/2015 11:00:43 PM Clear All

Order	Name	Status	Details
1	Precode	Completed successfully	
2	Extract	Completed successfully	
3	Postcode	Completed successfully	
	SMUG	Completed successfully	

Completed successfully

SASApp admzce4@BBLDTL as Richard Carson sas-metadata.bbld...

SAS Data Integration Studio 4.6 - ST

File Edit View Check Outs Actions Debug Tools Window Help

New [Icons]

Inventory [Dropdown]

Folders Checkouts [Dropdown]

Transformations [Dropdown]

SQL [Dropdown]

- Create Table
- Delete
- Execute
- Extract
- Insert Rows
- Join
- Merge

Basic Properties

Name	Value
Name	Extract
Description	
Folder Location	
Checked Out By	Not applicable
Table Name	W16X1JX
Library	WORK
DBMS	SAS View
Number of Rows	Row count is disab...
Number of Columns	11
Last Modified By	
Metadata Modified	
Metadata Created	
Logical Type	
Metadata ID	A5WR1DEU.\$000...

View Data: Extract

Go to row [Input] Go [Icons]

#	Vendor	CustomerID	ProductID	ProductName	ProductCategory	ProductSubcategory
1	Outdoor Rec...	100	101001	Basketball Jerseys ...	Team Sports	Basketball
2	Outdoor Rec...	100	101001	Basketball Jerseys ...	Team Sports	Basketball
3	Outdoor Rec...	100	101001	Basketball Jerseys ...	Team Sports	Basketball
4	Outdoor Rec...	100	101001	Basketball Jerseys ...	Team Sports	Basketball
5	Outdoor Rec...	100	101001	Basketball Jerseys ...	Team Sports	Basketball
6	Outdoor Rec...	100	101001	Basketball Jerseys ...	Team Sports	Basketball
7	Outdoor Rec...	100	101001	Basketball Jerseys ...	Team Sports	Basketball
8	Outdoor Rec...	100	101001	Basketball Jerseys ...	Team Sports	Basketball
9	Outdoor Rec...	100	101001	Basketball Jerseys ...	Team Sports	Basketball
10	Outdoor Rec...	101	101001	Basketball Jerseys ...	Team Sports	Basketball
11	Outdoor Rec...	101	101001	Basketball Jerseys ...	Team Sports	Basketball
12	Outdoor Rec...	101	101001	Basketball Jerseys ...	Team Sports	Basketball
13	Outdoor Rec...	101	101001	Basketball Jerseys ...	Team Sports	Basketball
14	Outdoor Rec...	101	101001	Basketball Jerseys ...	Team Sports	Basketball
15	Outdoor Rec...	101	101001	Basketball Jerseys ...	Team Sports	Basketball
16	Outdoor Rec...	101	101001	Basketball Jerseys ...	Team Sports	Basketball
17	Outdoor Rec...	101	101001	Basketball Jerseys ...	Team Sports	Basketball
18	Outdoor Rec...	102	101001	Basketball Jerseys ...	Team Sports	Basketball
19	Outdoor Rec...	102	101001	Basketball Jerseys ...	Team Sports	Basketball
20	Outdoor Rec...	102	101001	Basketball Jerseys ...	Team Sports	Basketball
21	Outdoor Rec...	103	101001	Basketball Jerseys ...	Team Sports	Basketball
22	Outdoor Rec...	103	101001	Basketball Jerseys ...	Team Sports	Basketball

SASApp | admzce4@BBLDTL as Richard Carson | sas-metadata.bbld...

SAS Data Integration Studio 4.6 - ST

File Edit View Check Outs Actions Debug Tools Window Help

New [Icons] Workspace Server

Inventory

Folders Checkouts

Transformations

- Correlations
- Distribution Analysis
- Forecasting
- Frequency
- One-Way Frequency
- Summary Statistics
- Summary Tables

Archived

Basic Properties

Name	Value
Name	Summary Statistics
Description	Provides data summ...
Node Type	Summary Statistics
Transform Version	22
User Written	No
Diagnostic Mode	No
Checkpoint	No
Run always when re...	No
Control Order	
Metadata ID	A5WR1DEU. \$0000067

SMUG *

Up Run Stop [Icons]

```

    graph LR
      SALES_ANALYSIS[SALES_ANALYSIS] --> Extract[1 Extract]
      Extract --> Summary_Statistics[Summary Statistics]
  
```

Diagram Code Log Output

Details

Mappings Status Warnings and Errors Statistics Control Flow

Last Run: 21/05/2015 11:00:43 PM Clear All

Order	Name	Status	Details
1	Precode	Completed successfully	
2	Extract	Completed successfully	
3	Postcode	Completed successfully	
	SMUG	Completed successfully	

Completed successfully

Summary Statistics (Generated transformation) SASApp admzce4@BBLDTL as Richard Carson sas-metadata.bbld...

Inventory

Folders Checkouts

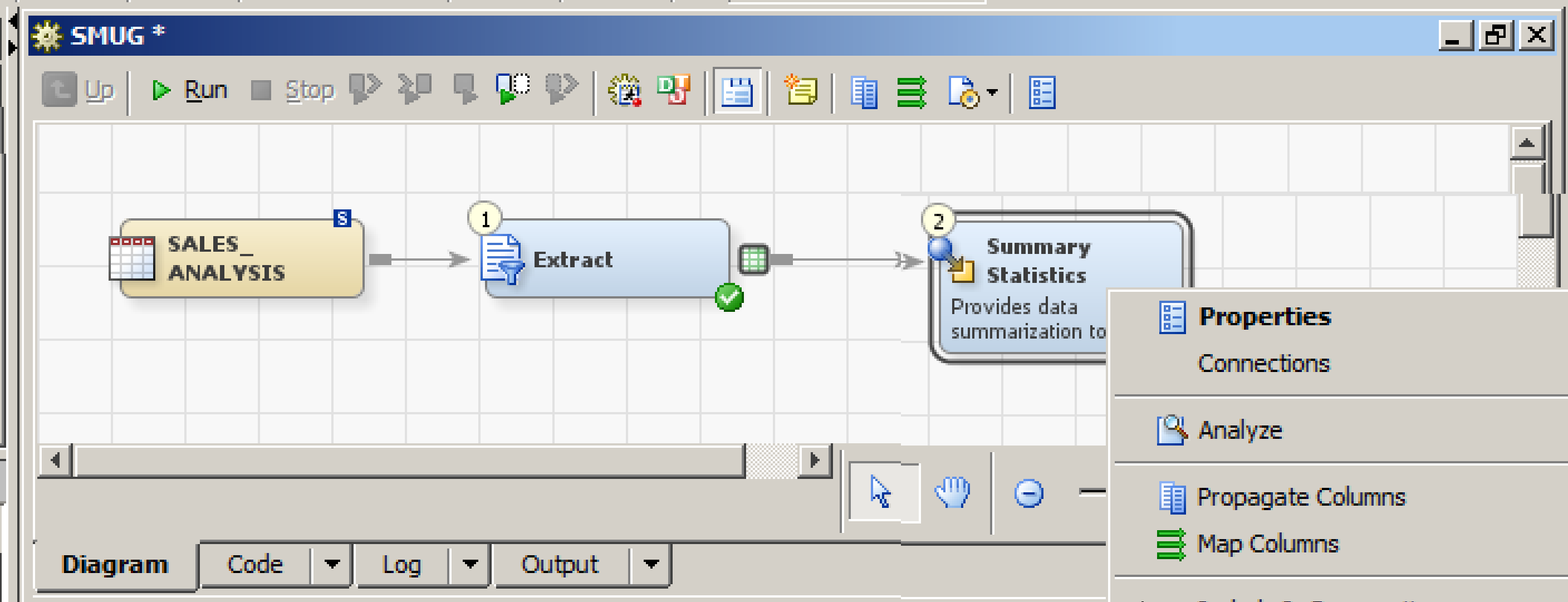
Transformations

- Correlations
- Distribution Analysis
- Forecasting
- Frequency
- One-Way Frequency
- Summary Statistics
- Summary Tables

Archived

Basic Properties

Name	Value
Name	Extract
Description	
Folder Location	
Checked Out By	Not applicable
Table Name	W16X1JX
Library	WORK
DBMS	SAS View
Number of Rows	Row count is disab...
Number of Columns	11
Last Modified By	
Metadata Modified	
Metadata Created	
Logical Type	
Metadata ID	A5WR1DEU.\$000...



Details

Columns Status Warnings and Errors Statistics Control Flow

Order	Name	Status	Last Run:
1	Precode	Completed successfully	
2	Extract	Completed successfully	
3	Postcode	Completed successfully	
	SMUG	Completed successfully	

Completed successfully

Properties

- Connections
- Analyze
- Propagate Columns
- Map Columns
- Include In Propagation
- Include In Mappings
- History
- Collect Diagnostics
- Run From Selected Transformation
- Run To Selected Transformation
- Run Selected Transformations
- Assign as Restart-Point
- Exclude From Run
- Ports
- Add Work Table
- Expand Ctrl+E
- Delete
- Find In

SAS Data Integration Studio 4.6 - ST

File Edit View Check Outs Actions Debug Tools Window Help

New Summary Statistics Properties

General Mappings Options Table Options Code Precode and Postcode Parameters Notes Extended Attributes

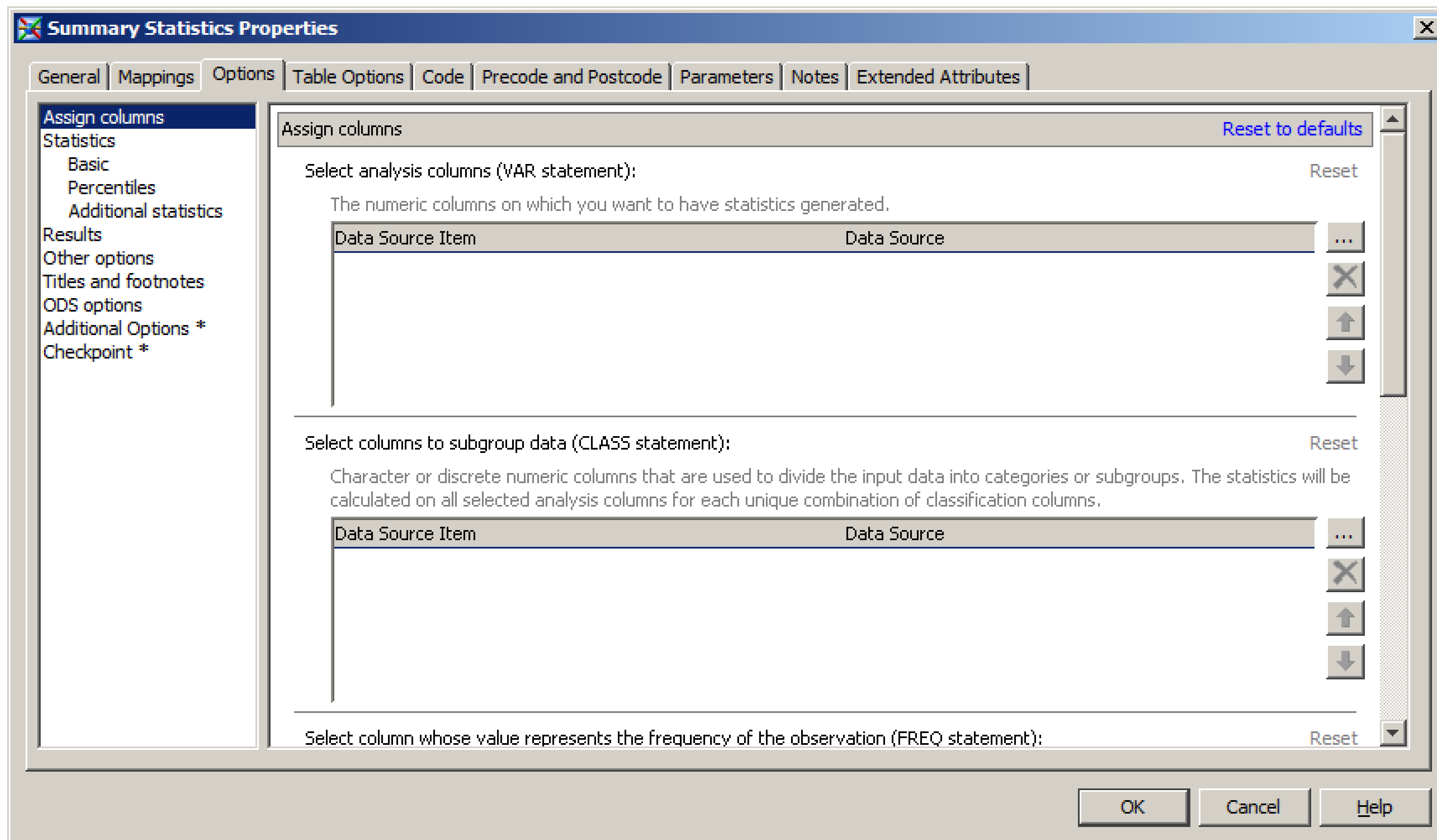
Source ta... Target table: Summary Statistics_OUTPUT0 (W1EV1NF)

#	Color	#	Expression	Table	Table Description	Type	Length	Informat
1	Ver	1		Summary Statistic...		Character	30	(None)
2	Cus	2		Summary Statistic...		Character	26	(None)
3	Pro	3		Summary Statistic...		Character	11	(None)
4	Pro	4		Summary Statistic...		Character	16	(None)
5	Pro	5		Summary Statistic...		Numeric	8	(None)
6	Pro	6		Summary Statistic...		Numeric	8	(None)
7	Mor	7		Summary Statistic...		Numeric	8	DOLLAR12.2
8	Yea	8		Summary Statistic...		Numeric	8	(None)
9	Tot	9		Summary Statistic...		Numeric	8	(None)
10	Tot	10		Summary Statistic...		Numeric	8	(None)
11	Pro							

OK Cancel Help

Completed successfully

Summary Statistics (Generated transformation) SASApp admzce4@BBLDTL as Richard Carson sas-metadata.bbld...



SAS Data Integration Studio 4.6 - ST

File Edit View Check Outs Actions Debug Tools Window Help

New Summary Statistics Properties

General Mappings Options Table Options Code Precode and Postcode Parameters Notes Extended Attributes

Assign columns

Statistics

- Basic
- Percentiles
- Additional statistics

Results

Other options

Titles and footnotes

ODS options

Additional Options *

Checkpoint *

Basic P

Name

Descrip

Node T

Transf

User V

Diagn

Checkp

Run al

Contro

Metad

Browse...

Select additional columns to include in the output data set (ID statement): [Reset](#)

The columns that you assign are included as additional columns in the output data set.

Data Source Item	Data Source
Month [Month]	/Extract_A5WR1DEU.\$00000005
Year [Year]	/Extract_A5WR1DEU.\$00000005

Select columns to obtain separate analysis on each discrete value (BY Statement): [Reset](#)

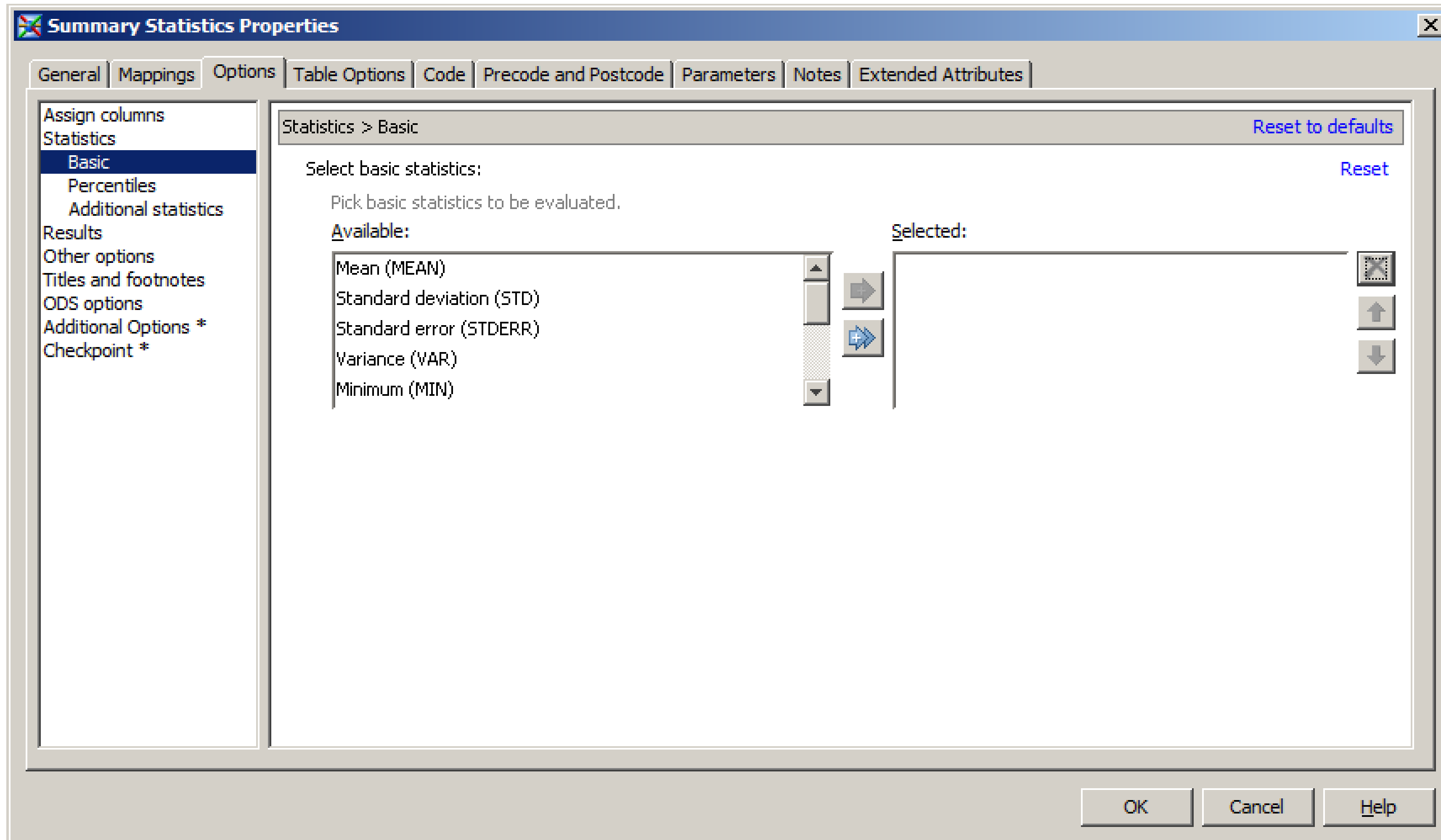
Specifies one or more columns to sort the table by. Analyses are performed on each group.

Data Source Item	Data Source

OK Cancel Help

Completed successfully

SASApp admzce4@BBLDTL as Richard Carson sas-metadata.bbld...



Summary Statistics Properties

General | Mappings | Options | Table Options | Code | Precode and Postcode | Parameters | Notes | Extended Attributes

Assign columns
Statistics
Basic
Percentiles
Additional statistics
Results
Other options
Titles and footnotes
ODS options
Additional Options *
Checkpoint *

Results [Reset to defaults](#)

Specify sort order for class columns in the output (ORDER): [Reset](#)
Specifies the sort order to create the unique combinations for the values of the class columns in the output.

Suppress all displayed output (NOPRINT): [Reset](#)
Suppresses all displayed output. Select this option if you want to create an output data set only.

Specify additional output statistics (OUTPUT statement): [Reset](#)
Specifies which statistics to store in the output data set. ID group, maximum id and minimum id specifications may be specified here.

Combinations of classification columns (WAYS statement): [Reset](#)
Specify one or more integers that define the number of class columns to combine to form all unique combinations of class columns.

OK Cancel Help

Summary Statistics Properties

General | Mappings

Assign columns
Statistics
Basic
Percentiles
Additional statistics
Results
Other options
Titles and footnotes
ODS options
Additional Options *
Checkpoint *

Quick Note:

Types Vendor
Product

Notes
Assigned:

Documents
Assigned:

General | Mappings | **Options** | Table Options | Code | Precode and Postcode | Parameters | Notes | Extended Attributes

Other PROC MEANS options:
Specify any PROC MEANS options that could not be specified elsewhere.

Specify BY statement options: Reset
Alternative way to specify by statement.

Specify CLASS statement options: Reset
Specify options for CLASS statement.

Specify OUTPUT statement options: Reset
Specify additional options for OUTPUT statement.
autolabel ways inherit

Specify TYPES statement options: Reset
Identifies which of the possible combinations of class columns to generate.
Vendor * ProductCategory * ProductSubcategory * ProductName

Specify options for columns (VAR statement): Reset

OK Cancel Help

SAS Data Integration Studio 4.6 - ST

File Edit View Check Outs Actions Debug Tools Window Help

New [Icons]

Transformations
Inventory
Folders Checkouts

admzce4

- Create REPSTG.FINANCIAL_CLASSES
- New Job 33031
- SMUG
- Test CRMB

SMUG

View Data: Summary Statistics_OUTPUT0 (74 rows)

Go to row Go [Icons]

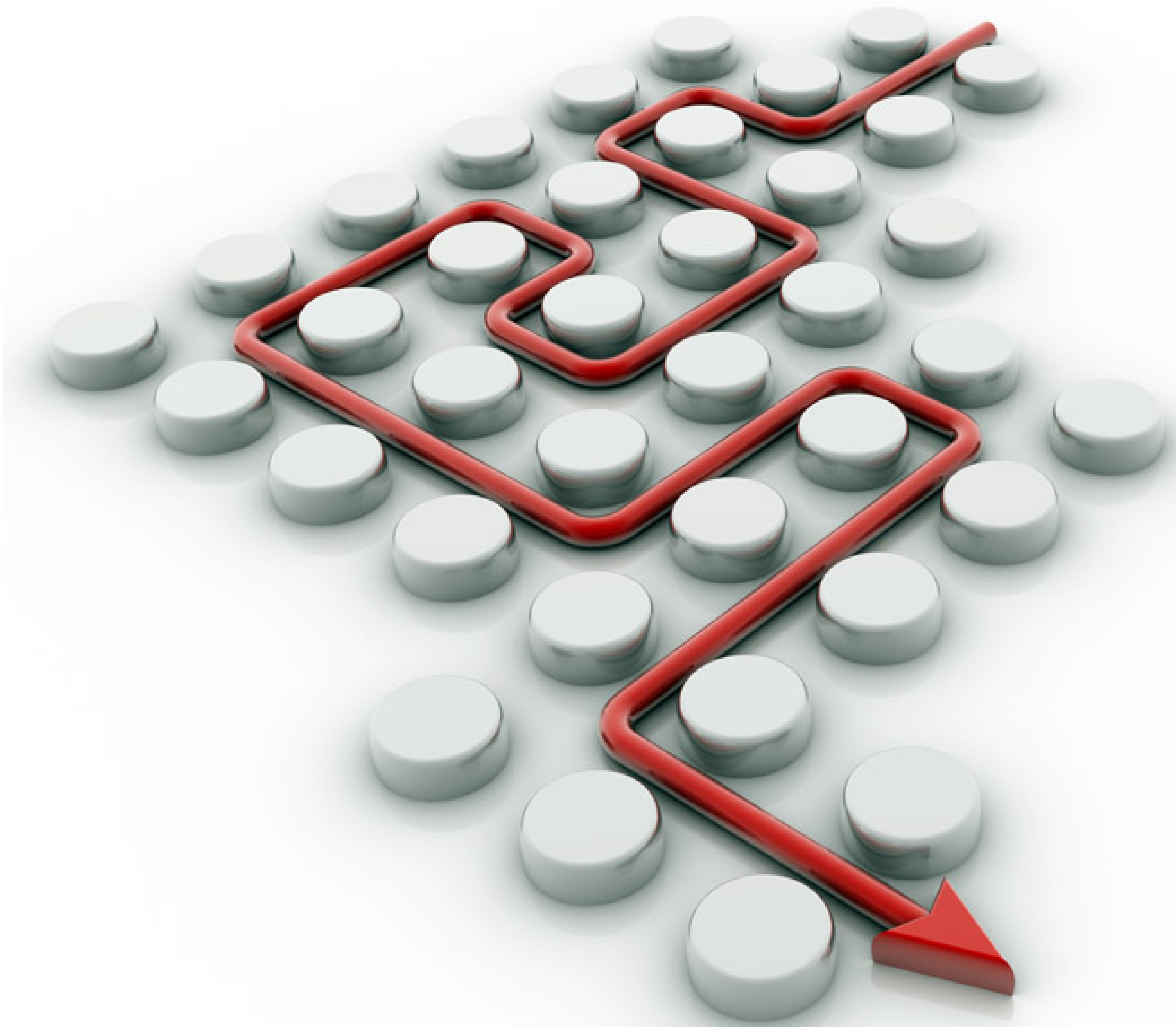
#	ProductSubcategory	Month	Year	TotalSales	_TYPE_	_FREQ_	_WAY_
34	Women's	12	2005	\$13,878.40	7	72	3
35		12	2005	\$172,663.50	8	431	1
36		12	2005	\$172,663.50	8	431	1
37		12	2005	\$172,663.50	8	431	1
38		12	2005	\$172,663.50	8	431	1
39		12	2005	\$172,663.50	8	431	1
40		12	2005	\$172,663.50	8	431	1
41		12	2005	\$172,663.50	8	431	1
42		12	2005	\$172,663.50	8	431	1
43		12	2005	\$24,540.60	10	103	2
44		12	2005	\$8,455.45	10	48	2
45		12	2005	\$3,358.00	10	26	2
46		12	2005	\$136,309.45	10	254	2
47		12	2005	\$24,540.60	10	103	2
48		12	2005	\$8,455.45	10	48	2
49		12	2005	\$3,358.00	10	26	2
50		12	2005	\$136,309.45	10	254	2
51		12	2005	\$24,540.60	10	103	2
52		12	2005	\$8,455.45	10	48	2
53		12	2005	\$3,358.00	10	26	2
54		12	2005	\$136,309.45	10	254	2

SASApp admzce4@BBLDTL as Richard Carson sas-metadata.bbld...

Summary Statistics in DI Studio:

1. Turn of Propagation
2. Add work table for output
3. Propagate class, id, by and analysis (var) variables from source to output
4. Add `_TYPE_`, `_FREQ_` and `_WAY_` as required to output table columns
5. Under Options, Assign columns: add as required for VAR, CLASS, ID and BY variables
6. Statistics: Basic – **delete all**
7. Results: specify NOPRINT and SUM= (in the OUTPUT options)
8. Other options: delete print specs and **autoname** option from OUTPUT options
9. Create a note containing the TYPES specification, and copy into the TYPES statement options

You're now good to go



Richard Carson
Senior SAS Designer



Bridging the Gap

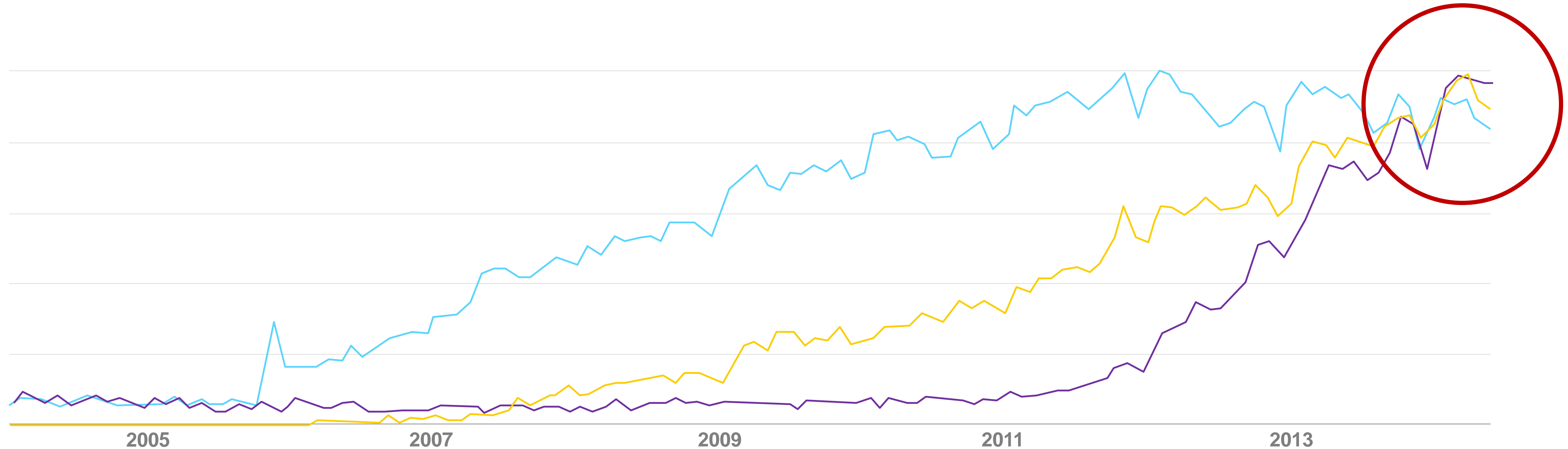
Between SAS Credit Risk Management for Banking
and APRA regulatory reporting



SMUG meeting – Big Data Analytics

Felix Liao
Business Solutions Manager
SAS Australia/New Zealand





BIG DATA
Lots of data

HADOOP
Processing Power

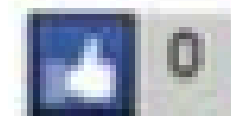
ANALYTICS
Accurate /Decisions



British Airways' BI lead: 'If you don't adopt Hadoop at least in part, you won't exist in a few years' time'

British Airways cuts memory costs and sees ROI within one year after deploying Hadoop

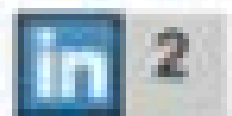
By [Margi Murphy](#) | [Computerworld UK](#) | Published 15:55, 15 April 15



0



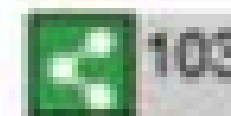
14



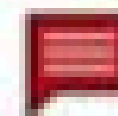
2



2



103



Enterprise business intelligence departments will 'not exist in a few years' time' if they do not adopt [Hadoop](#) in at least one instance, [British Airways'](#) data exploitation manager warned during the Hadoop Summit in Brussels today.

Also in this channel

[News](#)

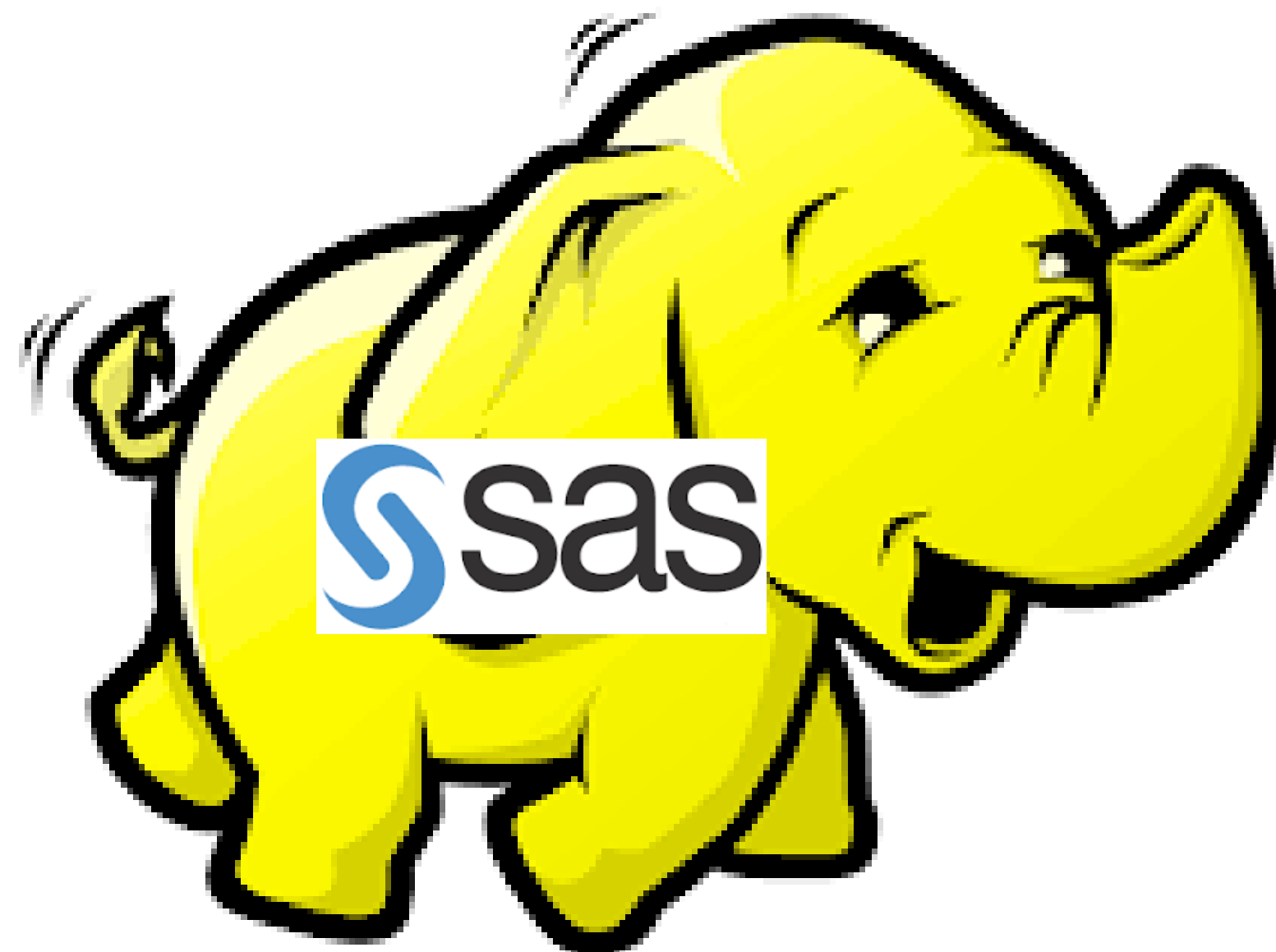
[In Depth](#)

SAS and Hadoop – The big Picture

SAS and Hadoop are made for each other

SAS and Hadoop will help you answer bigger questions

SAS is making it easy for our users to leverage Hadoop



What is Hadoop ?



- An Apache Software Foundation project
 - Open-source
 - Origins in early 2000s with contributions from Google, Yahoo! and Facebook
- Framework/Ecosystem of tools for processing Big Data
 1. Hadoop Common: The common utilities that support the other Hadoop subprojects
 2. Hadoop Distributed File System (HDFS): A distributed file system that provides high –throughput access to application data
 3. Hadoop MapReduce/YARN: A software framework for distributed processing of large data sets on compute clusters

Idea #1 - HDFS. Never forgets!

Head Node	Node 1	Node 2	Node 3	Node 4...
MYFILE.TXT				
..block1 ->	block1copy1			
..block2 ->		block2copy2		
..block3 ->			block3copy3	

Idea #1 - HDFS. Never forgets!

Head Node	Node 1	Node 2	Node 3	Node 4...
MYFILE.TXT				
..block1 ->	block1copy1		block1copy2	
..block2 ->		block2copy2		block2copy2
..block3 ->	block3copy2		block3copy3	

Idea #1 - HDFS. Never forgets!

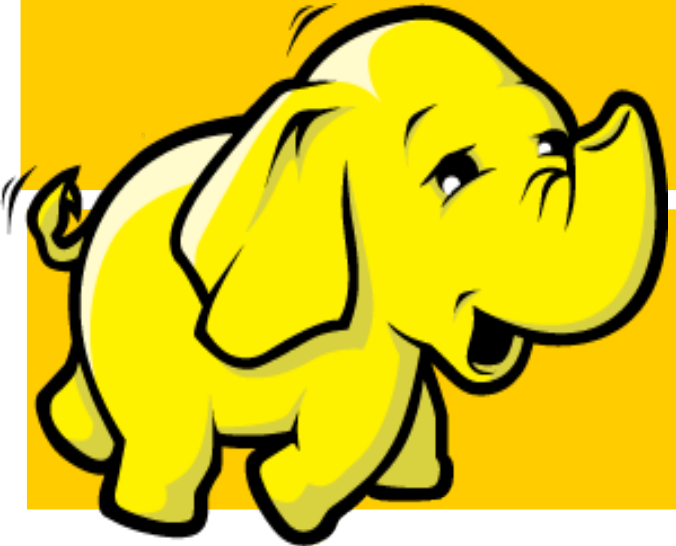
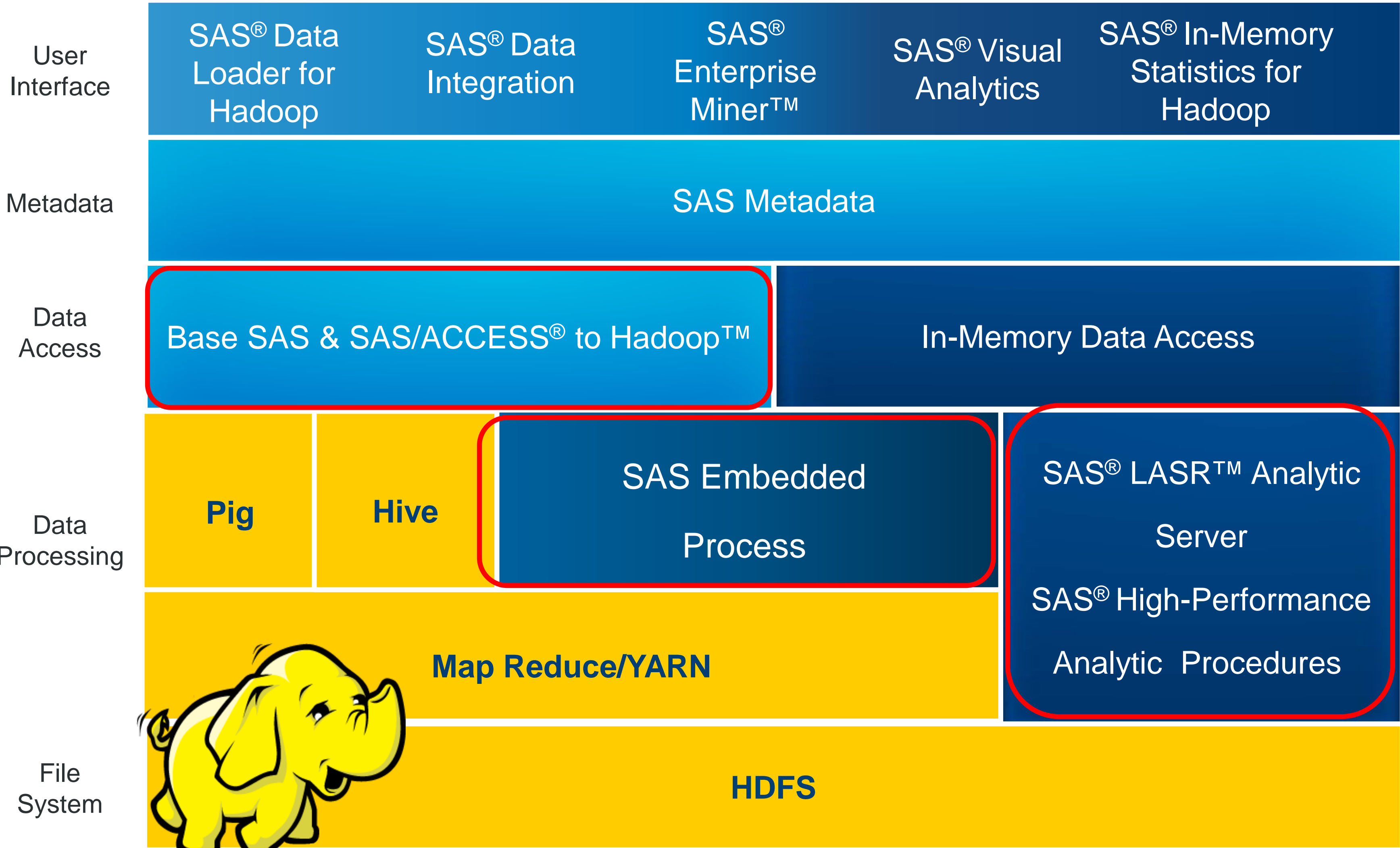
Head Node	Node 1	Node 2	Node 3	Node 4...
MYFILE.TXT				
..block1 ->	block1copy1		block1copy2	
..block2 ->		block2copy2		block2copy2
..block3 ->	block3copy2		block3copy3	



Idea #2 - MapReduce

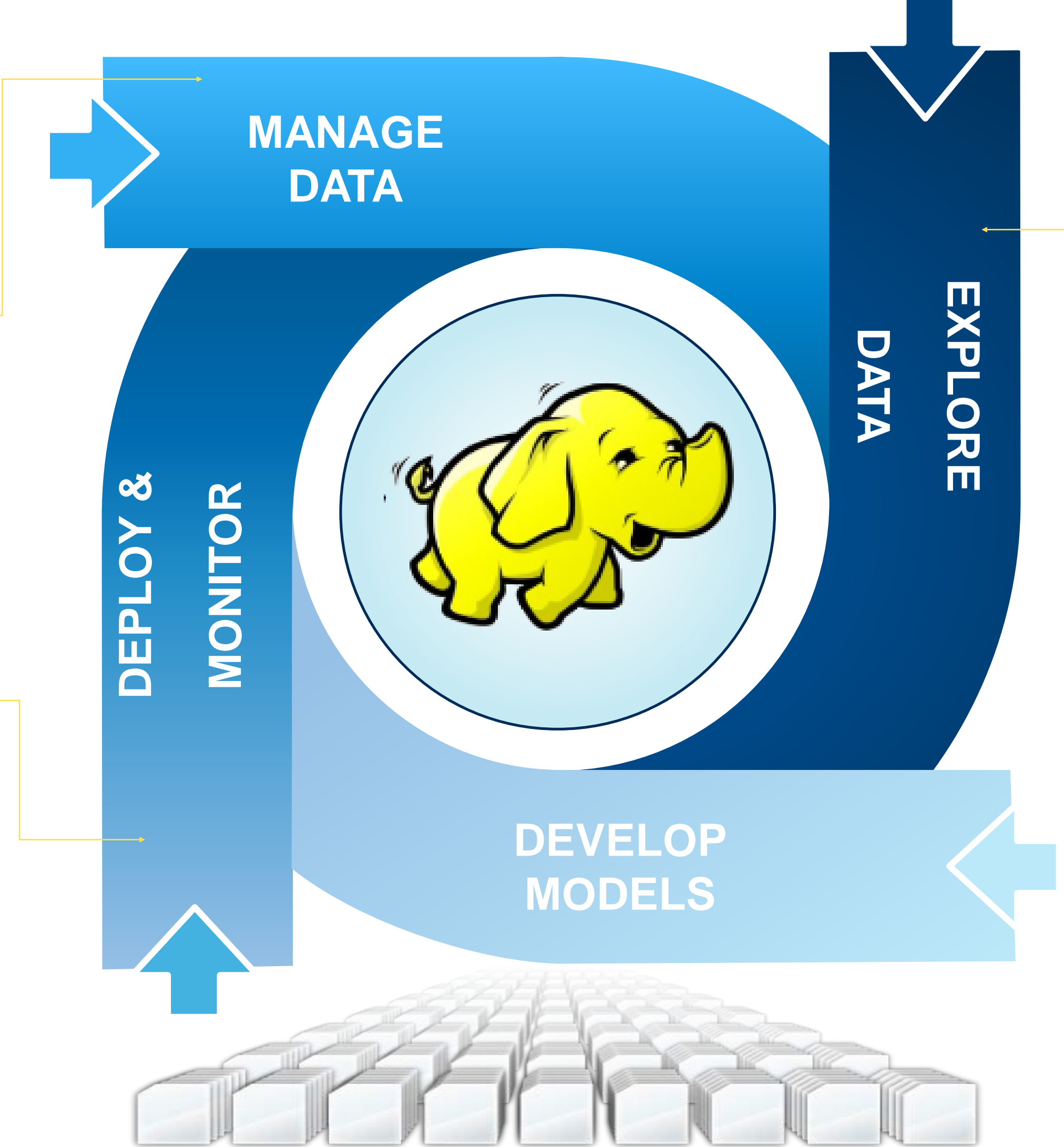
- We Want the Minimum Age in the Room
- Each Row in the audience is a data node
- I'll be the coordinator
 - From outside to center, accumulate MIN
 - Sweep from back to front. Youngest Advances

SAS® Within the HADOOP ECOSYSTEM



DATA TO DECISION LIFECYCLE on Hadoop

- SAS/ACCESS (Hadoop/Impala)
 - SAS Data Management
 - SAS Federation Server
- SAS Data Quality Accelerator for Hadoop
- SAS Code Accelerator for Hadoop
 - SAS Data Loader for Hadoop

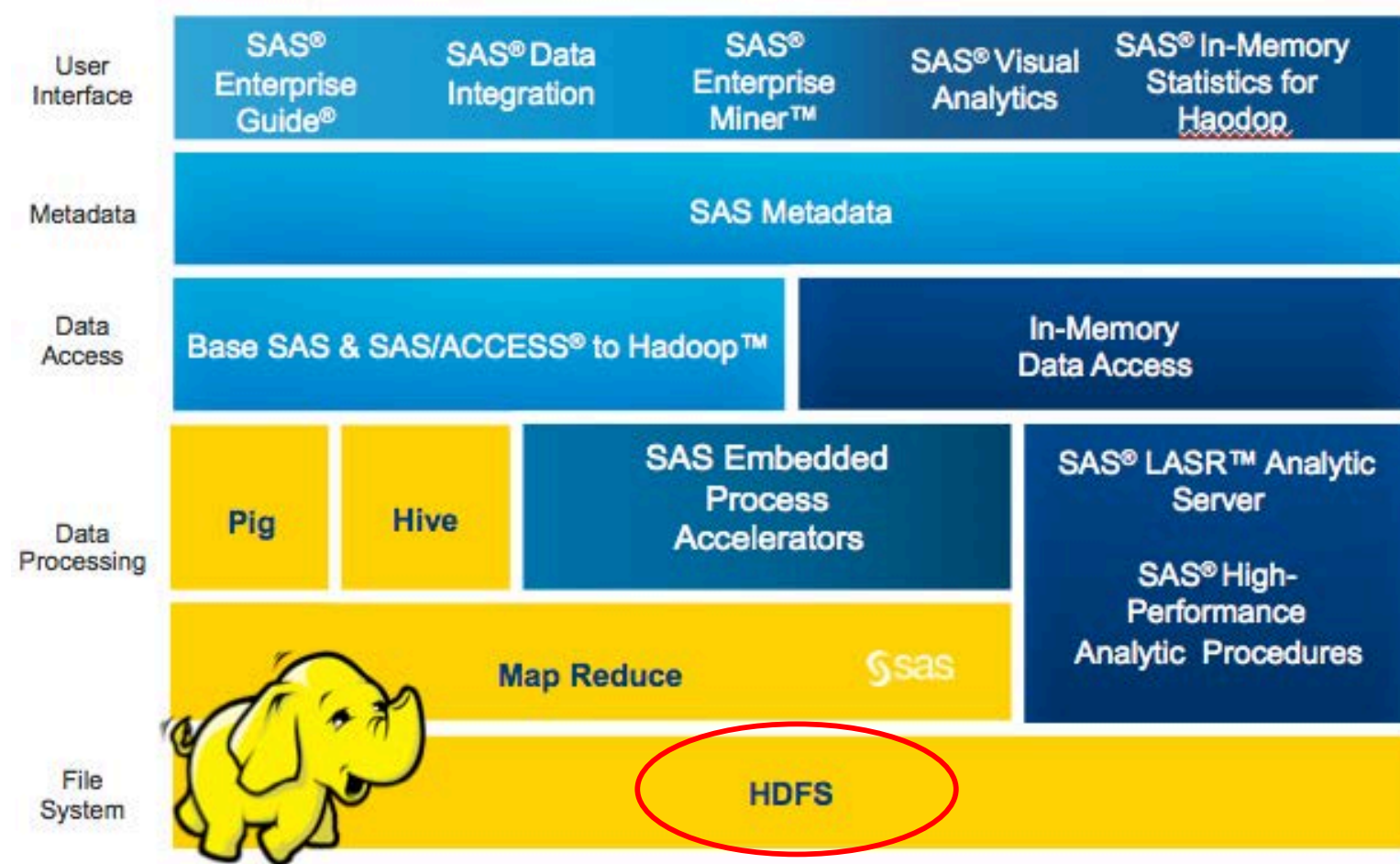


- SAS Visual Analytics
- SAS In-memory Statistics for Hadoop

- Model Manager
- SAS Scoring Accelerator for Hadoop

- SAS HPA Products
- SAS Visual Statistics
- SAS In-memory Statistics for Hadoop

MANAGE DATA READ/WRITE TO HDFS



```

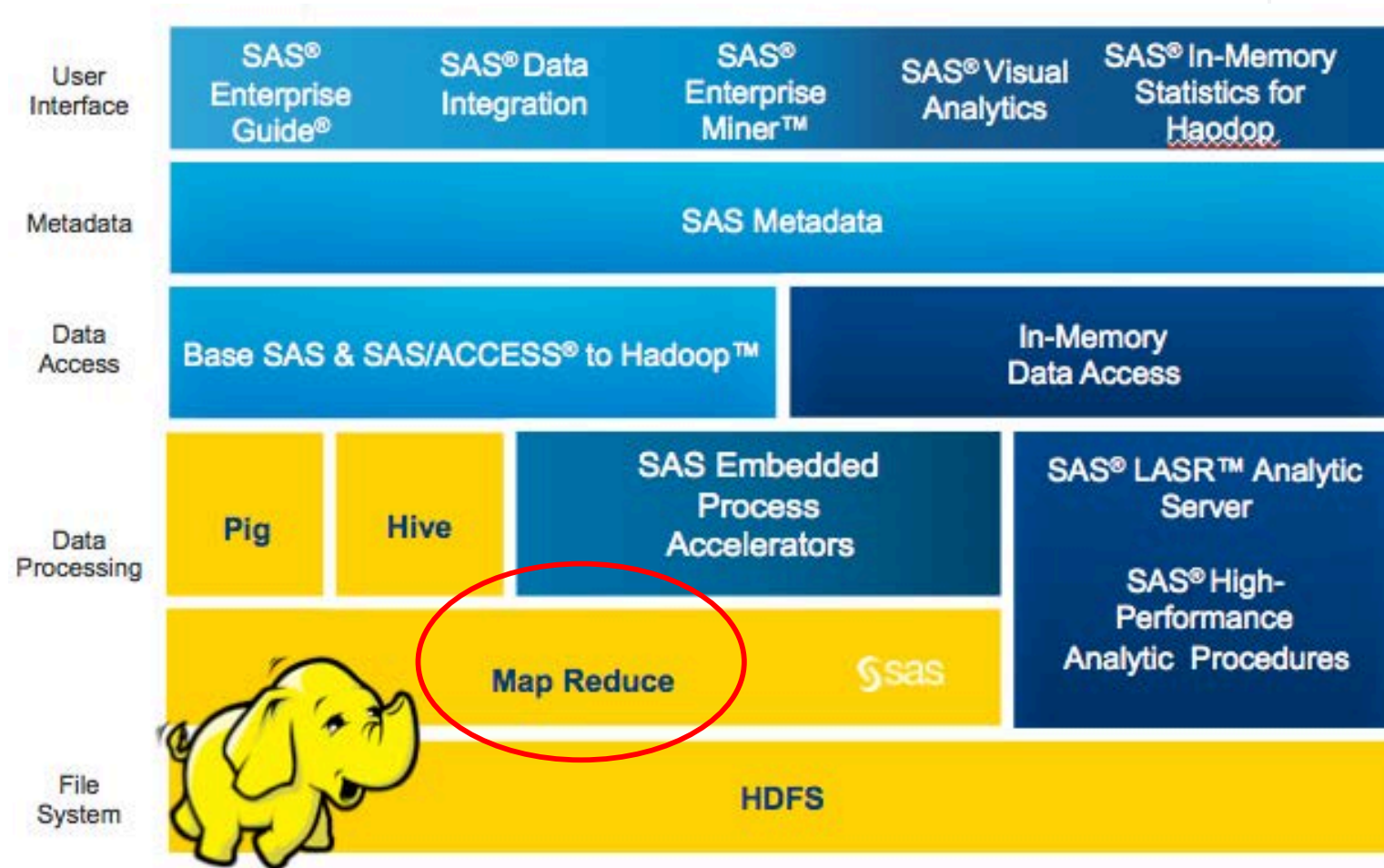
/*
Create directory on HDFS
*/
filename cfg "C:\Sample_Data\hadoop_config.xml";
proc hadoop options=cfg username="hadoop" password="hadoop";
    hdfs mkdir="/user/hadoop/testfolder" ;
run;

/*
Copy file from local SAS to HDFS
*/
filename cfg "C:\Sample_Data\hadoop_config.xml";
proc hadoop options=cfg username="hadoop" password="hadoop";
    hdfs copyfromlocal="C:\Sample_data\dept.txt" out="/user/hadoop/testfolder/";
run;

/*
Copy file from HDFS to local SAS
*/
filename cfg "C:\Sample_Data\hadoop_config.xml";
proc hadoop options=cfg username="hadoop" password="hadoop";
    hdfs copytolocal="/user/hadoop/testfolder" out="C:\Sample_data\" ;
run;

```


MANAGE DATA MAPREDUCE

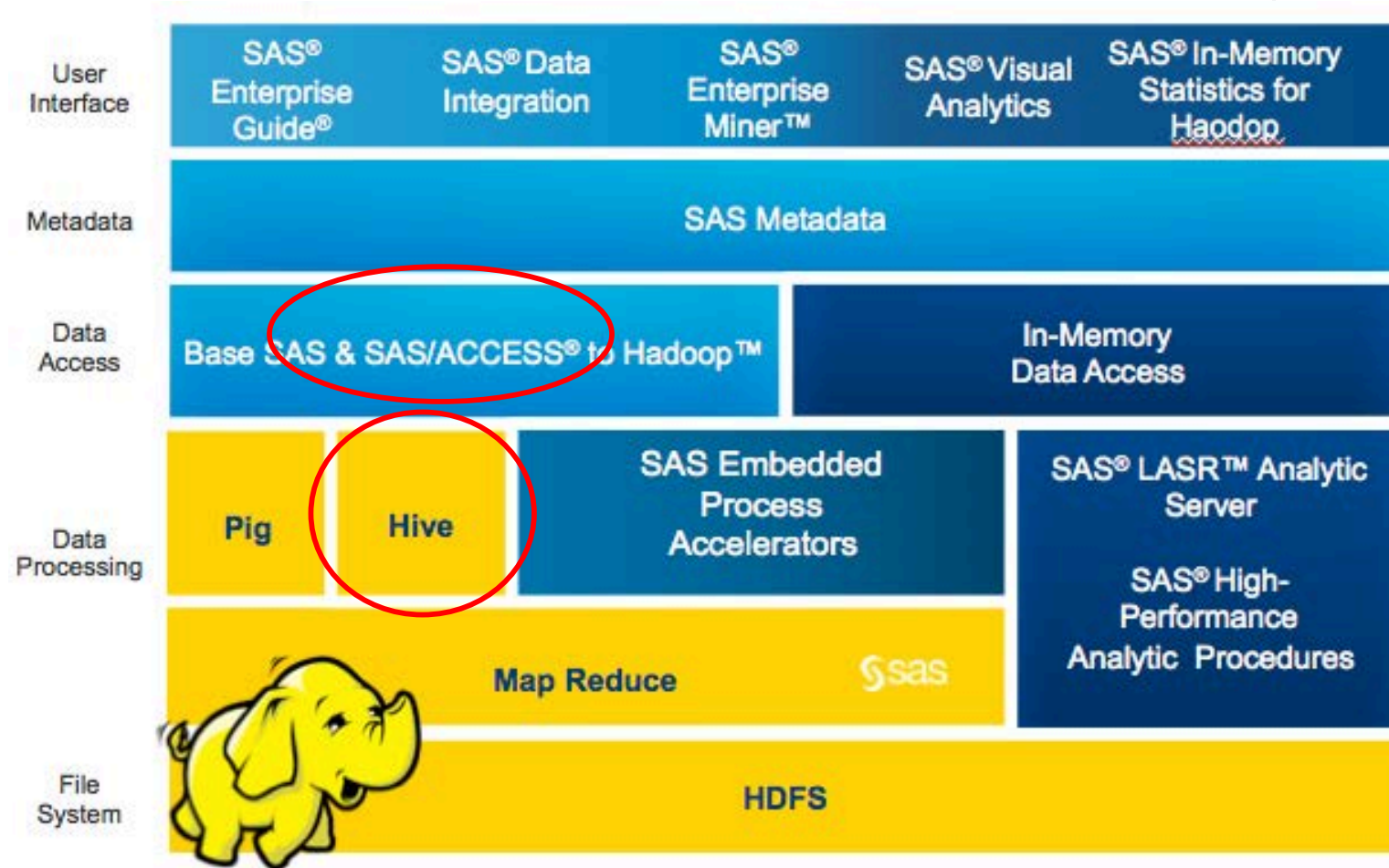


```

/*
Invoke MapReduce Word Count program
*/
filename cfg "C:\Sample_Data\hadoop_config.xml";
proc hadoop options=cfg username="hadoop" password="hadoop" verbose;
hdfs delete="/user/hadoop/output_MR1";
mapreduce
input="/user/hadoop/gutenberg" output="/user/hadoop/output_MR1"
jar="C:\Sample_data\hadoop-examples-2.0.0-mr1-cdh4.1.2.jar"
outputkey="org.apache.hadoop.io.Text"
outputvalue="org.apache.hadoop.io.IntWritable"
reduce="org.apache.hadoop.examples.WordCount$IntSumReducer"
combine="org.apache.hadoop.examples.WordCount$IntSumReducer"
map="org.apache.hadoop.examples.WordCount$TokenizerMapper"
reducetasks=0
;
run;

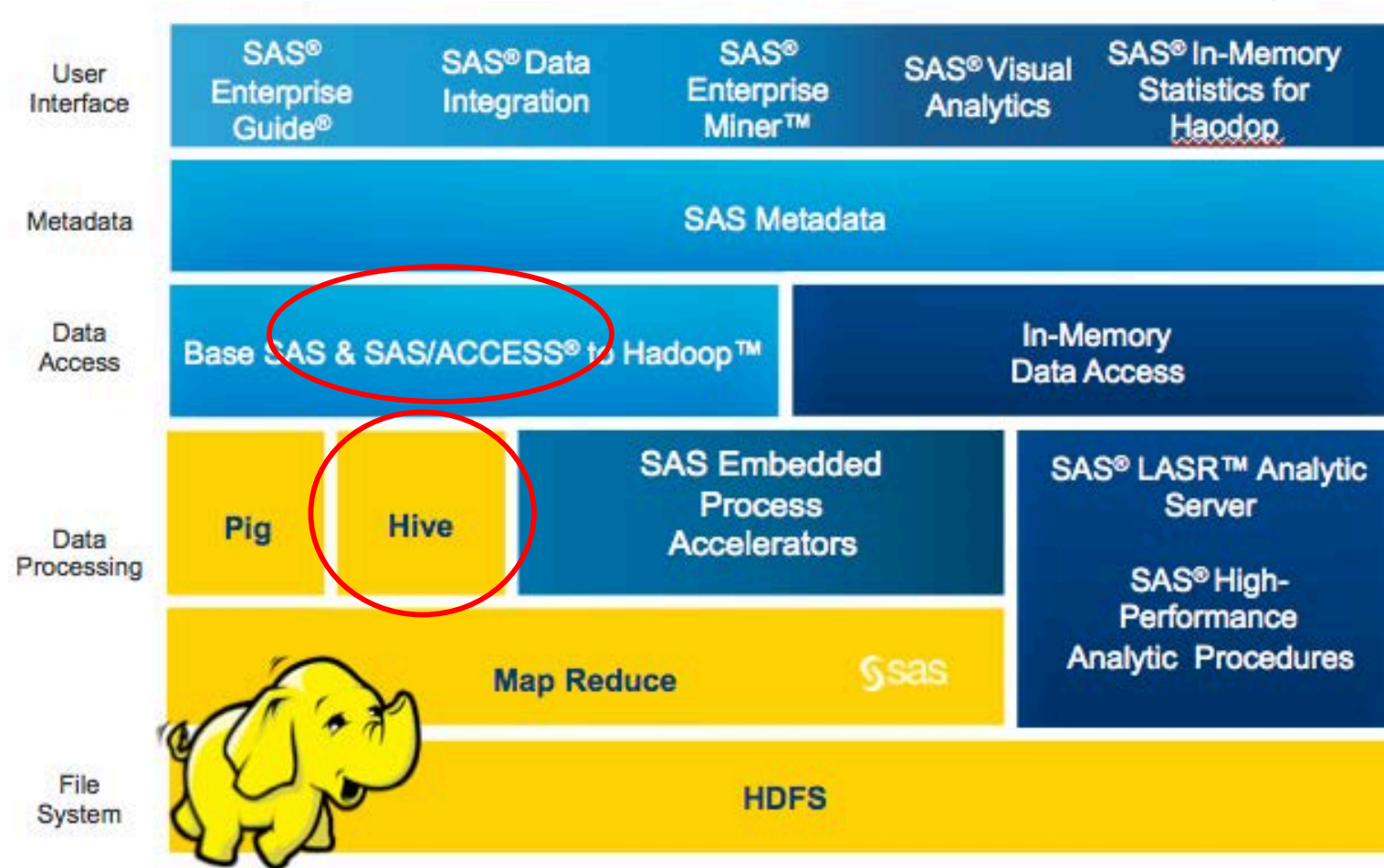
```

MANAGE DATA SAS/ACCESS



- Base SAS Procedures executed in-database for Hadoop
 - `FREQ`, `REPORT`, `SORT`, `SUMMARY/MEANS`, `TABULATE`
- Supported Hadoop distributions & combinations*
 - Cloudera CDH 5.0 running Hive/Hive2
 - Hortonworks HDP 2.0 running HiveServer2
 - IBM InfoSphere BigInsights 2.1 running Hive
 - MapR M5 2.0.1 running Hive
 - Pivotal/Greenplum HD running Hive
 - Pivotal/Greenplum MR 2.0.1 running Hive

MANAGE DATA SAS/ACCESS



```
LIBNAME cdh_hdp HADOOP PORT=10000 SERVER=sascldserv02 user=hadoop  
password=hadoop ;
```

```
/*  
Create new table  
*/
```

```
proc sql;  
connect to hadoop(PORT=10000 SERVER=sascldserv02 USER=hadoop  
PASSWORD="hadoop");  
exec( create table cars_prc (make string, model string, msrp double) ) by  
hadoop;  
quit;
```

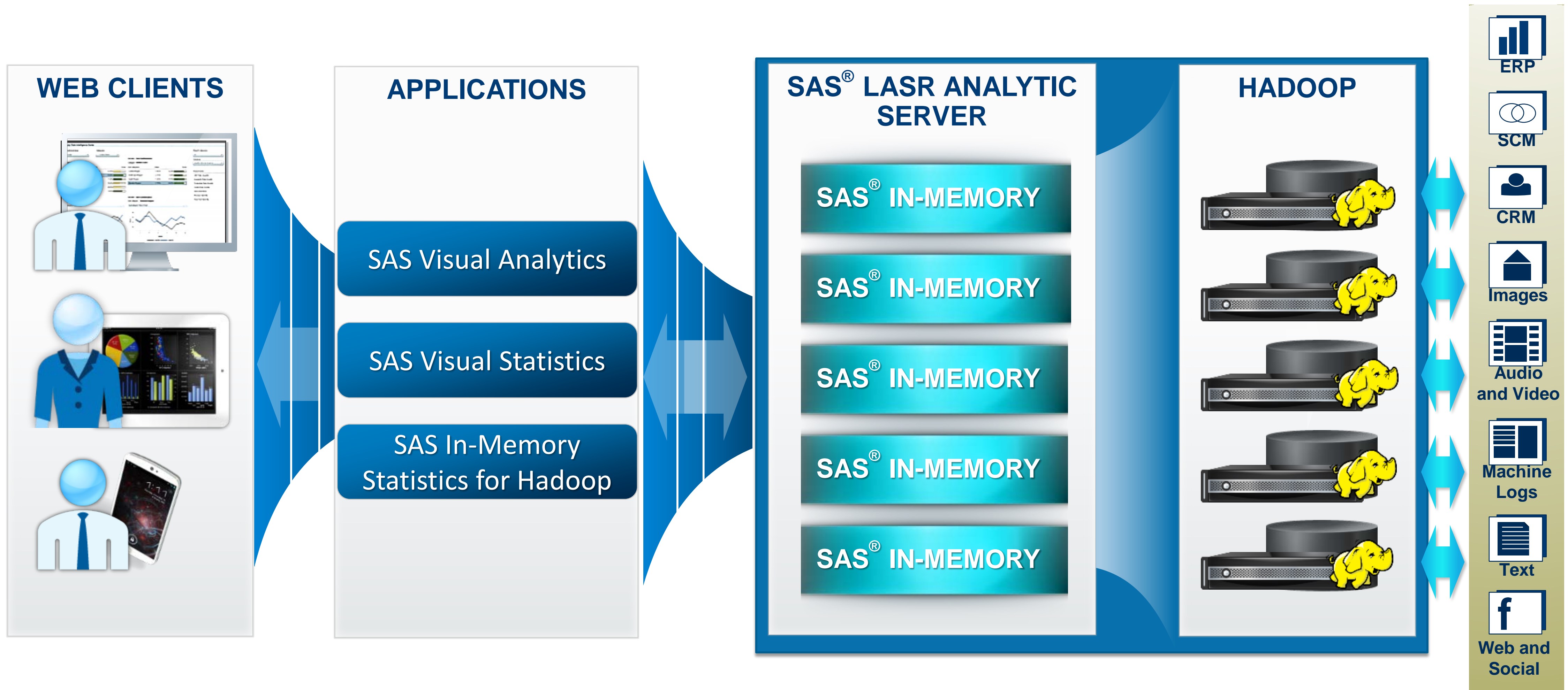
```
/*  
Copy from another table  
*/  
proc sql;  
insert into cdh_hdp.cars_prc  
select make, model, msrp  
from sashelp.cars ;
```

```
quit;  
/*  
List contents  
*/  
proc sql;  
select * from cdh_hdp.cars_prc;  
quit;
```

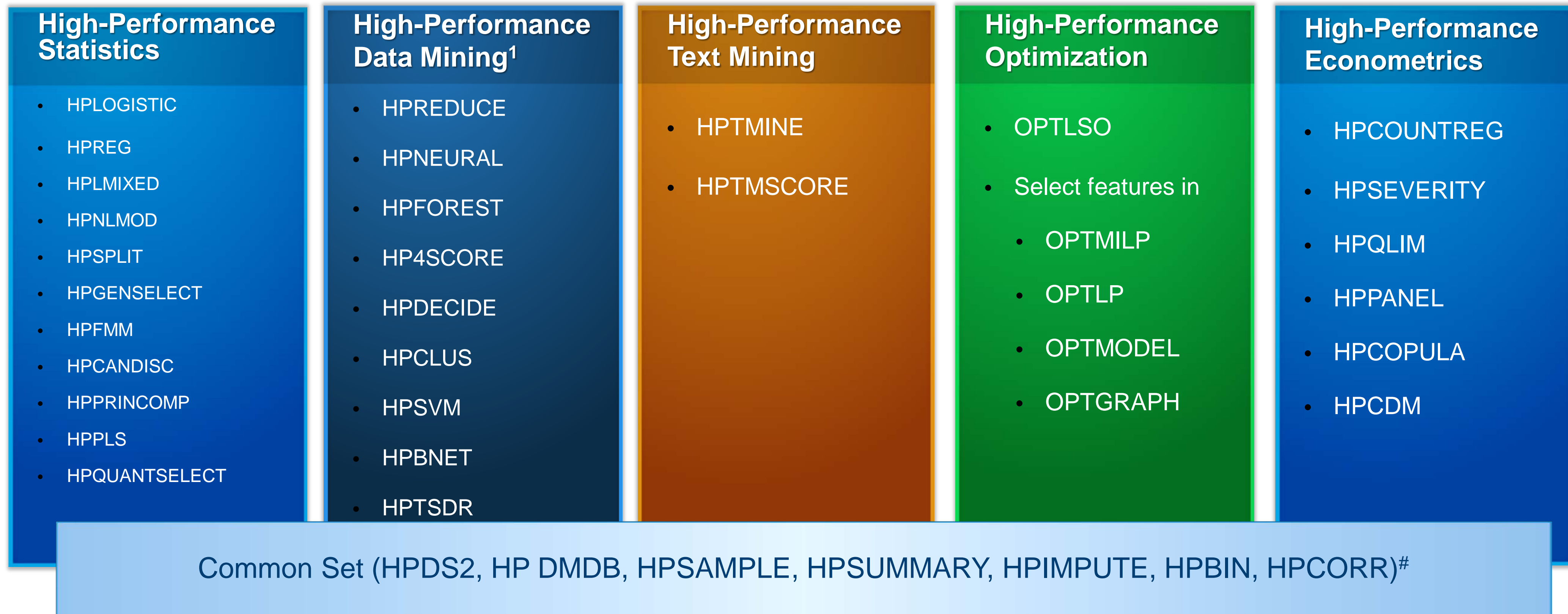
DATA DISCOVERY/VISUALIZATION

SAS IN-MEMORY ANALYTICS

In-memory processing; use Hadoop for storage persistence and commodity computing

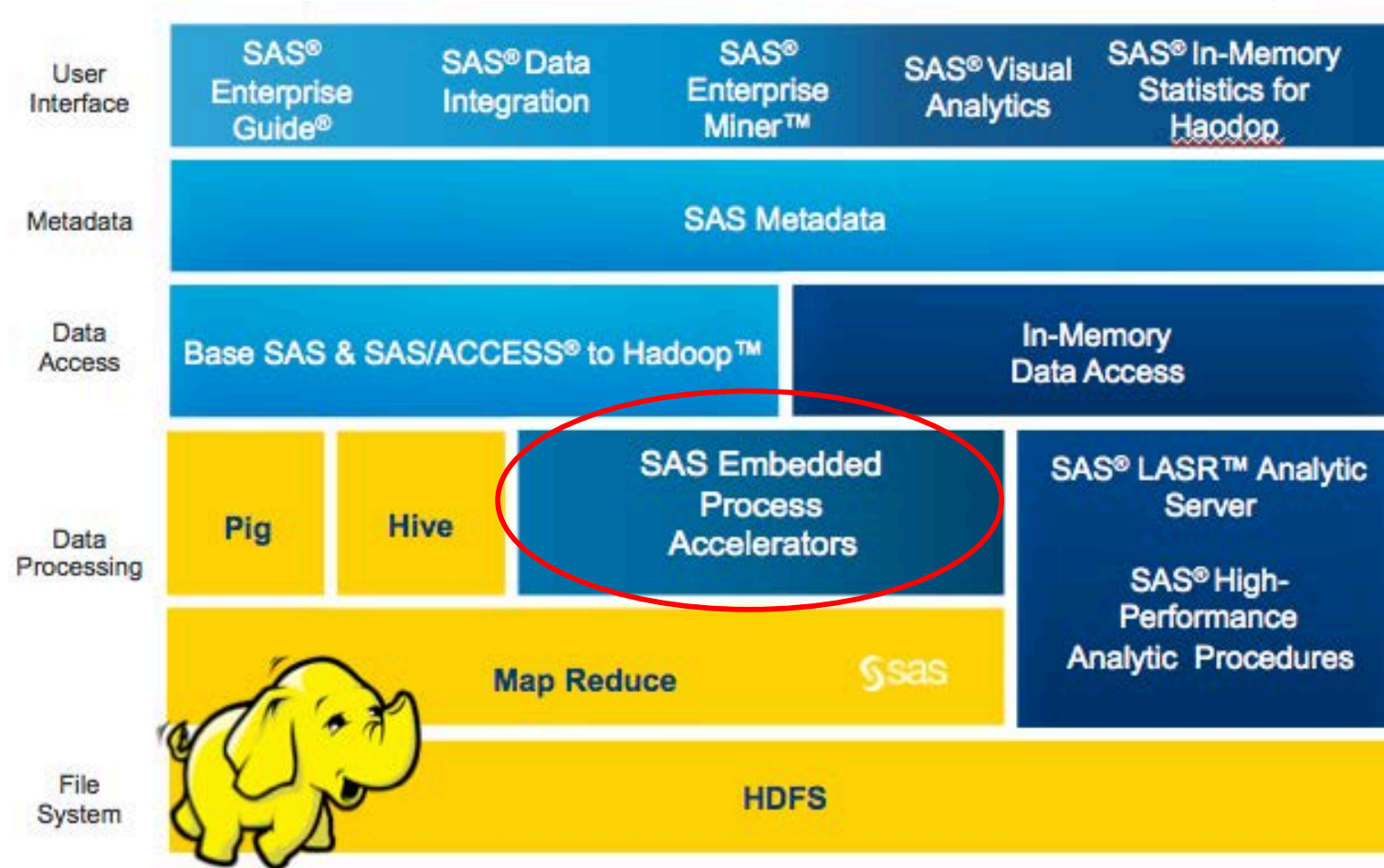


MODEL DEVELOPMENT **SAS HIGH PERFORMANCE ANALYTICS**



¹Includes SAS High-Performance Statistics

DEPLOY & MONITOR SAS SCORING ACCELERATOR FOR HADOOP



- Publish SAS® Enterprise Miner™ models or SAS/STAT linear models inside the Hadoop
- Fully integrated with SAS® Model Manager to streamline registration, validation and performance monitoring
- Reduced data movement and improve data governance by streamlining model deployment processes within Hadoop

THANK YOU!

4 short years and...

- Hadoop is on the opening session, 2 Hadoop leaders's share the panel discussion at Exec Track
 - 50% of the people attending an executive conference talk have an active Hadoop project
- SAS has built a family of applications that take full advantage of Hadoop as an analytics platform

For more Information: <http://www.sas.com/hadoop>

Felix.liao@SAS.com



@felixliao



felixliao

QUESTIONS?



SAS® FORUM
AUSTRALIA & NZ 2015

TECHNOLOGY ROADSHOW

October 2015

All major cities, across Australia & New Zealand

“The Premier Business Analytics Forum in ANZ”

ROADSHOW DATES & LOCATIONS

Dates	Locations
Fri, 16 Oct	Perth
Tue, 20 Oct	Melbourne
Thu, 22 Oct	Canberra
Tue, 27 Oct	Brisbane
Fri, 30 Oct	Sydney

DRAFT AGENDA OUTLINE

8:00AM	Registration				Discovery Area
9:00AM	Plenary				
10:30AM	Morning Tea				
Stream	Stream #1	Stream #2	Stream #3	Stream #4	
11:00AM	Presentation #1	Presentation #1	Presentation #1	HOWs #1	
11:45AM	Presentation #2	Presentation #2	Presentation #2	HOWs #2	
12:25PM	Networking Lunch				
1:35PM	Presentation #3	Presentation #3	Presentation #3	HOWs #3	
2:20PM	Presentation #4	Presentation #4	Presentation #4	HOWs #4	
3:00PM	Afternoon Tea				
3:30PM	Presentation #5	Presentation #5	Presentation #5	HOWs #5	
4:15PM	Presentation #6	Presentation #6	Presentation #6	HOWs #6	
4:55PM	Closing with Networking Drinks				

EXAMPLE TOPICS AREAS

- What's new in SAS® Visual Analytics 7.2
- Tips from the Techies
- SAS Certification preparation session
- SAS® Text Analytics - How to get the most out of your unstructured data
- R - Integrating your open source into your eco-system
- SAS® Visual Statistics - Interactive predictive modelling and visualisation
- Machine learning – It's not new to SAS
- Best practice on SAS Migration
- Introducing SAS Studio – New programmers interface
- What's New in SAS® Enterprise Guide
- Big data governance
- Convergence of data prep, analytical processing

CALL FOR SPEAKERS

We're looking for experts to showcase their skills and share insights with other SAS users locally or in other cities.

Selected speakers will:

- Receive complimentary registration at the event
- Receive professional presentation coaching session with a 3rd party provider
- Have their return flights and accommodation arranged and funded by SAS (pending probity restrictions within government agencies)

INTERESTED IN CALL FOR SPEAKERS?

Make sure you contact Hanlie Myburgh on hanlie.myburgh@sas.com to find out more

WATCH THIS SPACE!

www.sasforum.com/anz

felix.liao@sas.com



@felixliao



felixliao

Thank You!

<http://www.sas.com/au/sashadoop>

SAS LIVE WEB TRAINING

BRENDAN KIELY
EDUCATION ACCOUNT MANAGER
SOUTHERN REGION



CURRENT SAS TRAINING DELIVERY METHODS

Public Classes



Private Classes



Mentoring



dreamstime.com



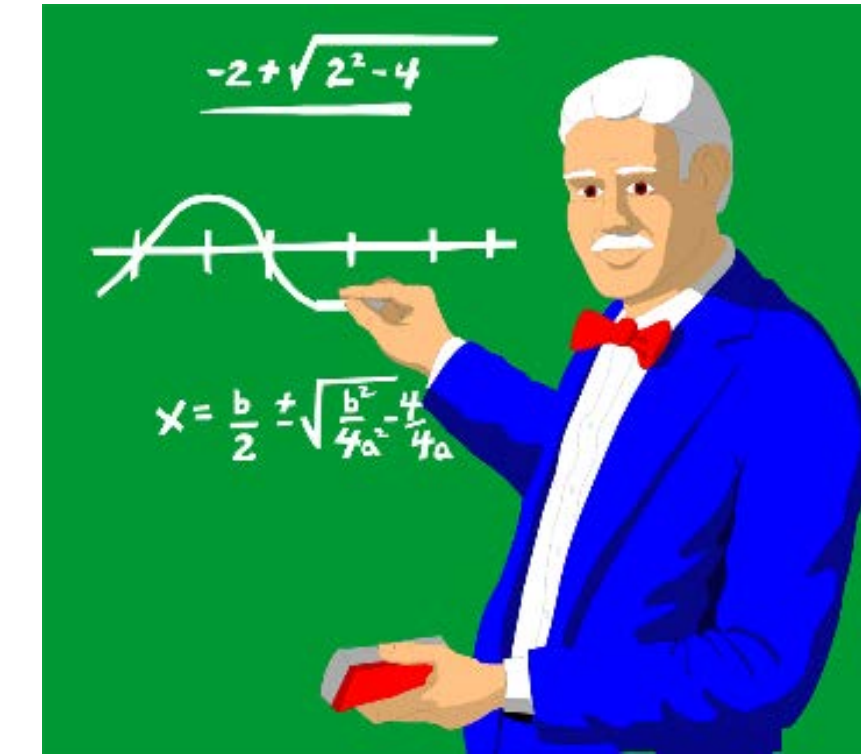
e-Learning

WHAT IF?

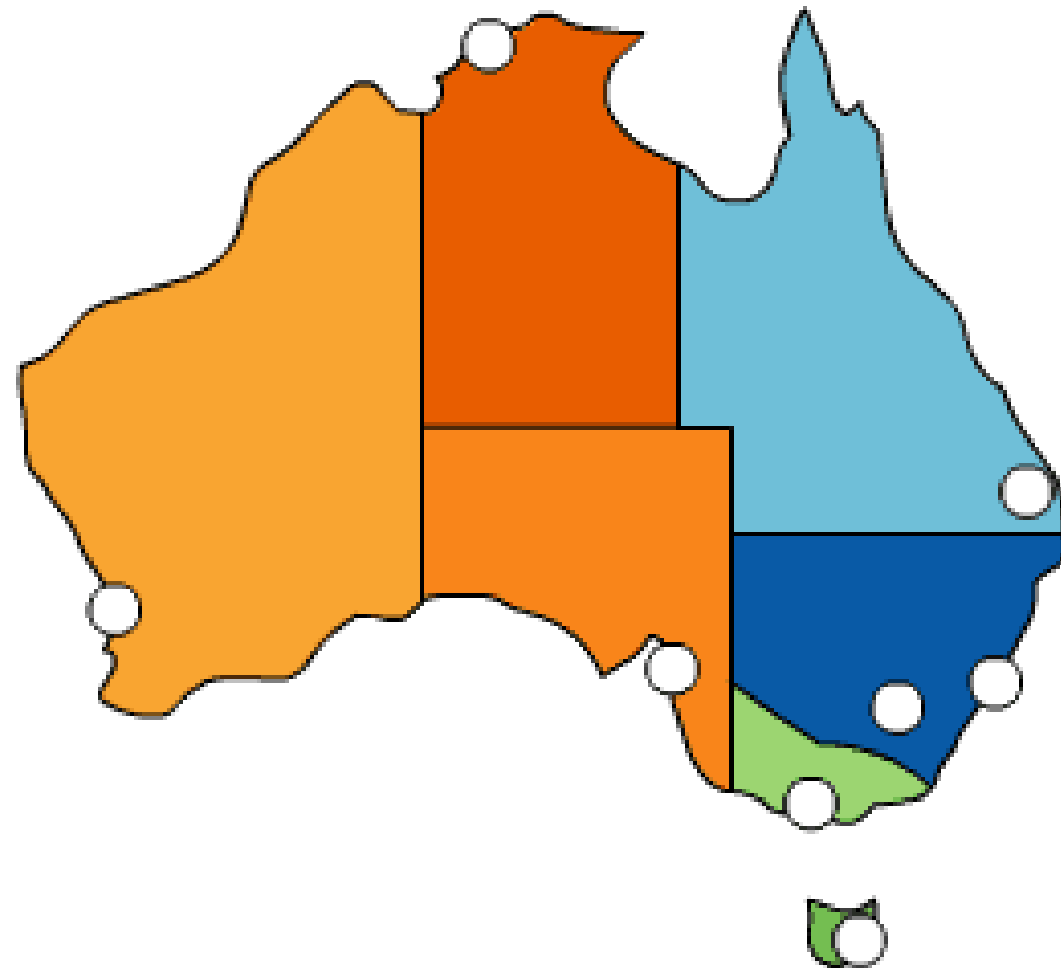
Course is not scheduled locally?



Specialist trainer located overseas?



Training is required by people in different locations?



Can't afford to spend all day out of the office?



SAS LIVE WEB TRAINING

Join the classroom right from your desktop

Ask questions and get answers in real-time

Same course content as face-to-face classes delivered in half day sessions



Led by an expert instructor who can virtually look over your shoulder

Access the latest software via a virtual lab

A SAS moderator is also present for technical assistance and chat queries

SAS LIVE WEB TRAINING



“I totally felt as if I WAS right there with you guys in a classroom”

“Best online learning experience I've ever had”

“SAS' Live Web technology is very impressive, and there is no travel!”



sas.com/au/training

ed@oz.sas.com

brendan.kiely@sas.com





QUESTIONS FOR THE PRESENTERS?

Q & A



THANK YOU
FOR YOUR ATTENDANCE

DON'T FORGET TO COMPLETE THE EVALUATION SURVEY

Please join us outside for Refreshments & Networking