

AI in Db2 for z/OS: A Brief Overview

Keziah Knopp

Data & AI on Z – Financial Services, IBM US

keziah.knopp@ibm.com



Agenda

Context

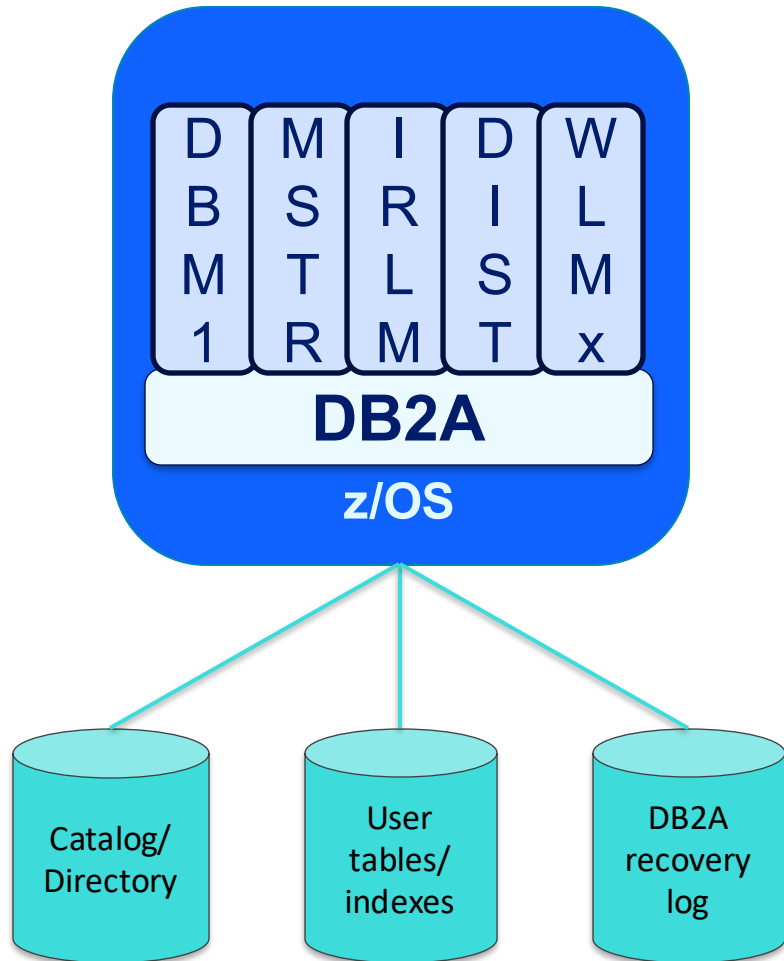
- Db2 for z/OS
- AI

Db2 AI for z/OS (Db2ZAI)

SQL Data Insights

AI built into Db2 13

Db2 for z/OS runs as a subsystem of z/OS



Db2 for z/OS comprises several resources:

Database services address space (ssnmDBM1)

- in this example, DB2ADB1
- Provides most database services

System services address space (ssnmMSTR)

- Performs variety of system functions

Internal resource lock manager (ssnmIRLM)

- Controls Db2 locking

Distributed data facility address space (ssnmDIST)

- Supports remote (network) requests

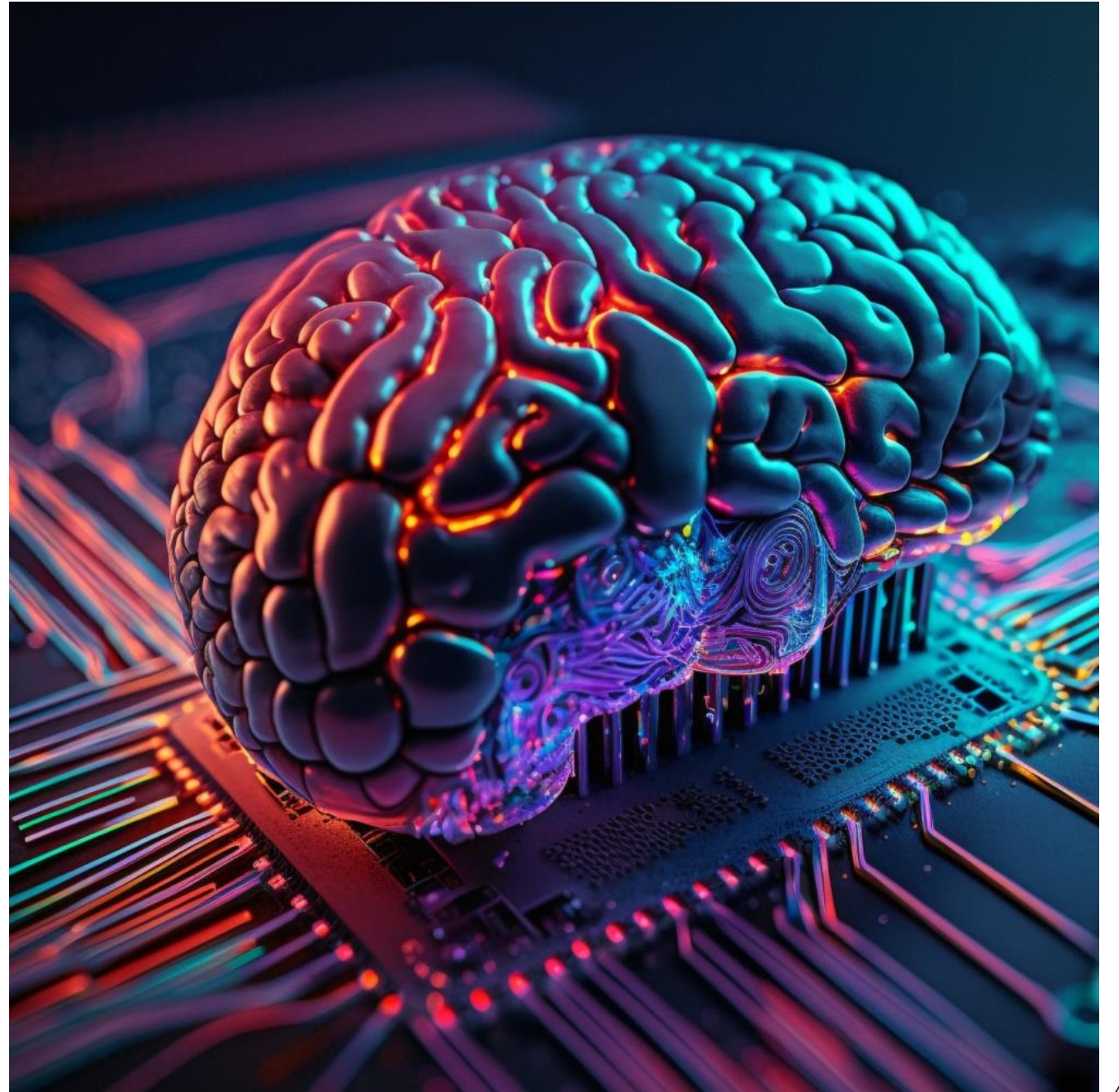
Workload manager (WLM) established address spaces (ssnmWLMx)

- Host external stored procedures

Data sets including:

- System data
- User data
- Db2 recovery log

AI: Augmented/Artificial Intelligence



What is AI?



ChatGPT



Claude

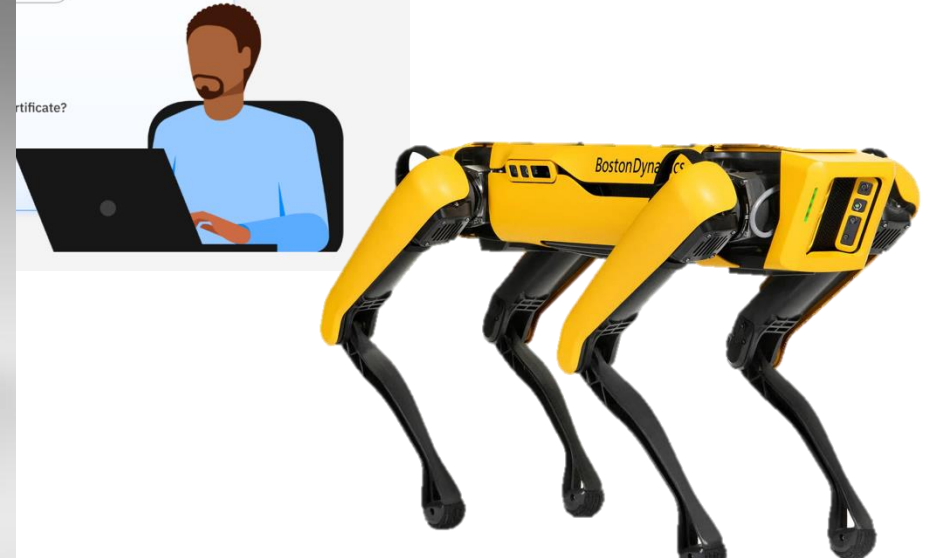
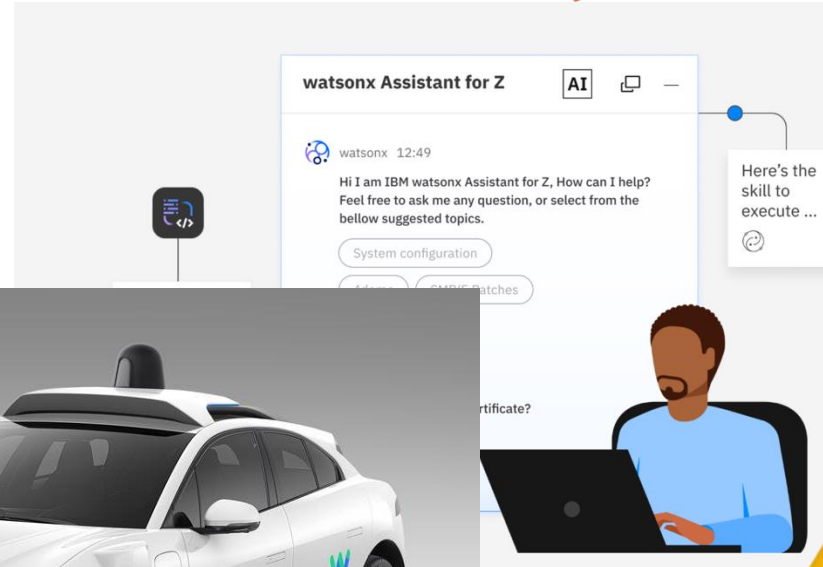
“Pop” AI

Gen AI – creates new content based on content that it knows

- watsonx Assistant for Z
- ChatGPT
- Claude
- AI image generators

Agentic AI – focuses on making decisions and taking actions based on given knowledge

- Waymo
- Boston Dynamics robotics



What is AI?

“Pop” AI

Gen AI – creates new content based on content that it knows

- ChatGPT
- Claude
- AI image generators

Agentic AI – focuses on making decisions and taking actions based on given knowledge

- Waymo
- Boston Dynamics robotics

Traditional AI

- Trains data
- Creates a model
- Identifies patterns, insights, relationships

Predictive AI

Takes traditional AI a step further and focuses on predicting the future, trends, analyses, etc.

Operational AI

Provides commercial scale AI targeted at improving efficiency, automating tasks, and enhancing workflows

- Real-time data processing
- AIOps

Why Operational AI in Db2 z/OS?



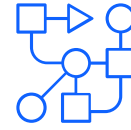
Do More With Less

- Better optimize operational cost with growing transactional workloads
- Help relieve pressure being felt by managers, DBAs, and System Admins



Reduce the high learning curve

- Minimize the need to learn the tuning knobs and instrumentations



Simplify and Protect Complex Environments

- Easily gain insights into remote application behavior and connections to Db2 z/OS
- Protect Db2 z/OS from distributed application mismanaging resources

Db2 AI for z/OS

Collect the operational data and leverage AI to improve Db2 performance and resiliency

Data Collection

Db2 for z/OS collects SQL executions, Db2 instrumentation, and operational data



Learning and optimization

Db2ZAI learns, predicts and optimizes or recommends updates on Db2 execution behavior

SQL Optimization

Reduce CPU consumption

CPU saving with better query access path



System Assessment

Improve Productivity

Less time spent analyzing data, more on finding solutions



Distributed Connection Control

Protect critical resources

Act quickly, uncover issues early, prevent a flood of connections



Bring operational AI to Db2 for z/OS to improve your database system health, efficiency and performance!

Db2 AI for z/OS - SQL Optimization

Create better access paths based on your execution environment

Data Collection

Db2 for z/OS collects SQL executions, Db2 instrumentation, and operational data



Learning and optimization

Db2ZAI learns, predicts and optimizes or recommends updates on Db2 execution behavior

SQL Optimization

Reduce CPU consumption

CPU saving with better query access paths



What It Does

- Choose optimal access paths to improve query performance with little-to-no user interference
 - Predict likely host variables
 - Predict rows to be fetched
 - Runtime optimization
 - Detect and resolve the regression

Primary Benefits

- Reduce cost by improving performance
- Reducing tuning activity work
- Stabilize access path

Db2 AI for z/OS – System Assessment

Automated performance analysis with years of experience embedded augmented by Performance Insights

Data Collection

Db2 for z/OS collects SQL executions, Db2 instrumentation, and operational data



Learning and optimization

Db2ZAI learns, predicts and optimizes or recommends updates on Db2 execution behavior

System Assessment

Improve Productivity

Less time spent analyzing data, more on finding solutions and tuning your system



What It Does

- **Identify** Db2 system tuning opportunities
- **Guided analysis** in Performance Insights without deep knowledge

Primary Benefits

- **Learn** what is normal and build the system assessment report with:
 - **Actionable** recommendations
 - Correlation on critical metrics to **help root cause analysis**

Db2 AI for z/OS – Distributed Connection Control

Greater insights and control on remote connections

Data Collection

Db2 for z/OS collects SQL executions, Db2 instrumentation, and operational data



Learning and optimization

Db2ZAI learns, predicts and optimizes or recommends updates on Db2 execution behavior

Distributed Connection Control

Protect Critical Resources

Uncover issues early, prevent a flood of connections, and provide greater insights



What It Does

- Establish remote connection and thread control through profiles
- Create real time alerts when remote connections exceed the threshold
- Analyze and visualize distributed connection and thread behavior

Primary Benefits

- Avoid the system impact from the large number of connections to the system
- Provide best practices

Db2ZAI User Interface: GBP_WRITE_FAILED_PRIM

Priority	Category	Metric name	Subsystem	Extra info	Work period name	Record count	Excp count	Excp pct	Excp min tstamp	Excp max tstamp	Metric min	Metric max	Metric mean	Metric stddev	Metric50 pctl	Metric90 pctl
🔴	Buffer pool	GBP_WRITE_FAILED_PRIM	DC1N	BP11	WORKDAY-AM-PEAK	1,080	2	0.185	2022-06-08 19:16:30	2022-06-08 19:34:50	0	3	0.004	0.109	0	0

Recommendations

- A number of write failures to the primary GBP have occurred. This indicates that the data pages in the coupling facility are being consumed faster by changed pages than the Db2 castout process can write back to the disk. Db2 will retry the affected write after the castout, but this will impact the application commit time. If retry continues to fail, the pages will be added to the logical page list (LPL) to prevent further access.
- Review the graph for GBP write, castout I/O, and write failures to see whether this is caused by a surge in update (I/U/D) activity from batch or utility. If this is caused by the surge in update activity but an infrequent occurrence, continue to monitor the exception and make sure castout is catching up.
- If this problem occurs frequently, consider reducing or spreading the overall system activity, or decrease the GBP castout thresholds (CLASST, GBPOOLT, GBPCHKPT). Currently castout class threshold (CLASST) is set as (5, 0), group buffer pool threshold (GBPOOLT) is set as 30, group checkpoint interval (GBPCHKPT) is set as 4 minutes.
- If castout activity is happening frequently, then it is possible that disk performance is not catching up with castout activities. Consider increasing the number of data entries in the (primary) group buffer pool by adjusting the directory/data ratio, and/or increase the total size of the group bufferpool.

Db2ZAI User Interface: LOCK_DEADLOCKS

!	Contention	LOCK_DEADLOCKS	DC1N	WORKDAY- PM-PEAK2	720	8	1.11	2022-06-09 00:20:00	2022-06-09 00:23:30
---	------------	----------------	------	----------------------	-----	---	------	------------------------	------------------------

Recommendations

- A higher-than-normal number of deadlocks have occurred.
- Review the deadlock messages from the master or console log (messages DSNT375I and DSNT376I) for detailed deadlock information.
- IFCID 172 (deadlock record) also contains details about deadlocks, such as contention resources, packages, statements involved with deadlocks, holders, and so on.
- The IRLM deadlock detection value is currently set to 1,000.
- If DDL operations are one of the deadlock holders, review and adjust the DDLTOX subsystem parameter value which is currently set to 1.

Db2ZAI User Interface: DDF_AVG_QUEUE_TIME

Application	DDF_AVG_QUEUE_TIME	DC1N	WORKDAY-NONPEAK	3,960	5	0.126	2022-06-08 08:44:50	2022-06-08 13:06:40	0	0.003	0	0	0
-------------	--------------------	------	-----------------	-------	---	-------	---------------------	---------------------	---	-------	---	---	---

Recommendations

- The average queue time for type 2 inactive or new connections that were queued while waiting for a database access thread (DBAT) is higher than normal time. This can be either due to workload increase or DBATs are running slower than expected.
- Review the correlation graph to see how many active DBATs are, or if there are increased DDF activities.
- If you suspect the slow down, review accounting class 1 and 2 elapsed time to see if there are long-running DRDA transactions or an unexpected wait.
- Also check DDF Enclave Settings in workload manager to ensure that the DDF work currently in the system is receiving the expected priority level, if there is high CPU utilization.
- If DDF_T2_INACT_PER_SEC is increasing, this may be due to a workload increase.
- If the increased queued time did not impact the elapsed time significantly and the response time still meets the SLA, then no action is needed. You may want to adjust MAXCONQW value based on the wait time.
- Otherwise, check to see whether MAXDBAT which currently set as 2,000 is set high enough to handle the concurrent distributed requests. Note that setting a larger MAXDBAT will affect the size of virtual and real storage, and you should investigate the impact of increasing the size before making any changes. Check the STG_REALAVAIL metric in the workload profile section, which can help you determine the amount of storage to use.

SQL Data Insights: AI-infused SQL functions

An industry-leading relational database with embedded AI capabilities



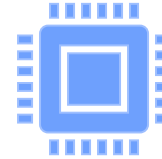
Infuses AI directly into your database on existing data to **discover hidden information**



Minimizes complexity of deploying AI into your applications



Single model used for a range of inferencing tasks over multiple fields



Exploits IBM **zIIP** and **SIMD**

SQL Data Insights: Built-in Db2 functions

FL 500

Function	Description
AI_SIMILARITY	Returns the entities that are most similar to (or dissimilar to) a particular entity
AI_SEMANTIC_CLUSTER	Returns the entities that are most similar to (or dissimilar to) a given set of up to three entities
AI_ANALOGY	Consider the relationship between value X in COL1 and value Y in COL2, and return the most analogous COL2 values if the COL1 value is Z
AI_COMMONALITY	Returns the values of a column that are outliers with respect to all rows in a table

FL 504

SQL Data Insights: Examples of functions

Built-in function - example	What the query does
<pre>SELECT V.VENDOR_NAME, AI_SIMILARITY(VENDOR_NAME, 'IBM CORPORATION') FROM VENDORS V ORDER BY 2 DESC FETCH FIRST 10 ROWS ONLY;</pre>	Returns the 10 vendors that are most similar to IBM Corporation
<pre>SELECT V.VENDOR_NAME, AI_SEMANTIC_CLUSTER(VENDOR_NAME, 'IBM CORPORATION', 'AMAZON', 'MICROSOFT') FROM VENDORS V ORDER BY 2 DESC FETCH FIRST 10 ROWS ONLY;</pre>	Returns the top 10 vendors that are most similar to the cluster of vendors comprised of IBM Corporation, Amazon and Microsoft
<pre>SELECT V.SERVICE_COUNTRY, AI_ANALOGY('IBM CORPORATION' USING MODEL COLUMN VENDOR_NAME, 'USA' USING MODEL COLUMN SERVICE_COUNTRY, 'SAMSUNG' USING MODEL COLUMN VENDOR_NAME, SERVICE_COUNTRY) FROM VENDORS V ORDER BY 2 DESC FETCH FIRST 10 ROWS ONLY;</pre>	Returns the 10 service countries that, when paired with the vendor Samsung, are most analogous to the service country USA as it is paired with IBM Corporation
<pre>SELECT AI_COMMONALITY(C.DRIVERS_LICENSE_NUMBER) AS SCORE, C.* FROM SDILAB.INSURANCE C ORDER BY SCORE ASC FETCH FIRST 15 ROWS ONLY ;</pre>	Returns the top 15 driver license numbers that are outliers with respect to all rows in the INSURANCE table

SQL Data Insights: Examples of functions

Built-in function - example	What the query does
<pre>SELECT V.VENDOR_NAME, AI_SIMILARITY(VENDOR_NAME, 'IBM CORPORATION') FROM VENDORS V ORDER BY 2 DESC FETCH FIRST 10 ROWS ONLY;</pre>	Returns the 10 vendors that are most similar to IBM Corporation
<pre>SELECT V.VENDOR_NAME, AI_SEMANTIC_CLUSTER(VENDOR_NAME, 'IBM CORPORATION', 'AMAZON', 'MICROSOFT') FROM VENDORS V ORDER BY 2 DESC FETCH FIRST 10 ROWS ONLY;</pre>	Returns the top 10 vendors that are most similar to the cluster of vendors comprised of IBM Corporation, Amazon and Microsoft
<pre>SELECT V.SERVICE_COUNTRY, AI_ANALOGY('IBM CORPORATION' USING MODEL COLUMN VENDOR_NAME, 'USA' USING MODEL COLUMN SERVICE_COUNTRY, 'SAMSUNG' USING MODEL COLUMN VENDOR_NAME, SERVICE_COUNTRY) FROM VENDORS V ORDER BY 2 DESC FETCH FIRST 10 ROWS ONLY;</pre>	Returns the 10 service countries that, when paired with the vendor Samsung, are most analogous to the service country USA as it is paired with IBM Corporation
<pre>SELECT AI_COMMONALITY(C.DRIVERS_LICENSE_NUMBER) AS SCORE, C.* FROM SDILAB.INSURANCE C ORDER BY SCORE ASC FETCH FIRST 15 ROWS ONLY ;</pre>	Returns the top 15 driver license numbers that are outliers with respect to all rows in the INSURANCE table

SQL Data Insights: Examples of functions

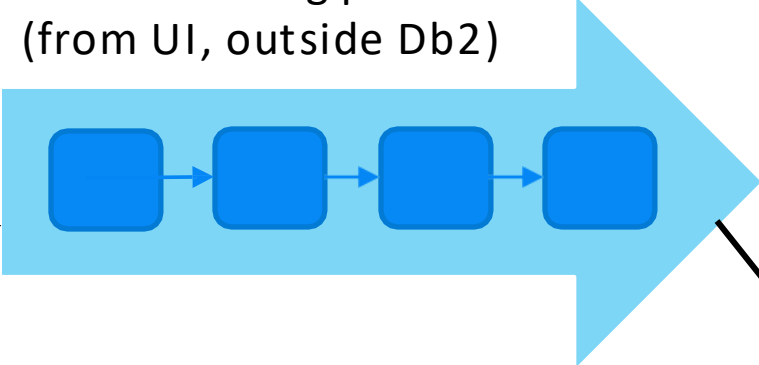
Built-in function - example	What the query does
<pre>SELECT V.VENDOR_NAME, AI_SIMILARITY(VENDOR_NAME, 'IBM CORPORATION') FROM VENDORS V ORDER BY 2 DESC FETCH FIRST 10 ROWS ONLY;</pre>	Returns the 10 vendors that are most similar to IBM Corporation
<pre>SELECT V.VENDOR_NAME, AI_SEMANTIC_CLUSTER(VENDOR_NAME, 'IBM CORPORATION', 'AMAZON', 'MICROSOFT') FROM VENDORS V ORDER BY 2 DESC FETCH FIRST 10 ROWS ONLY;</pre>	Returns the top 10 vendors that are most similar to the cluster of vendors comprised of IBM Corporation, Amazon and Microsoft
<pre>SELECT V.SERVICE_COUNTRY, AI_ANALOGY('IBM CORPORATION' USING MODEL COLUMN VENDOR_NAME, 'USA' USING MODEL COLUMN SERVICE_COUNTRY, 'SAMSUNG' USING MODEL COLUMN VENDOR_NAME, SERVICE_COUNTRY) FROM VENDORS V ORDER BY 2 DESC FETCH FIRST 10 ROWS ONLY;</pre>	Returns the 10 service countries that, when paired with the vendor Samsung, are most analogous to the service country USA as it is paired with IBM Corporation
<pre>SELECT AI_COMMONALITY(C.DRIVERS_LICENSE_NUMBER) AS SCORE, C.* FROM SDILAB.INSURANCE C ORDER BY SCORE ASC FETCH FIRST 15 ROWS ONLY ;</pre>	Returns the top 15 driver license numbers that are outliers with respect to all rows in the INSURANCE table

SQL Data Insights: Examples of functions

Built-in function - example	What the query does
<pre>SELECT V.VENDOR_NAME, AI_SIMILARITY(VENDOR_NAME, 'IBM CORPORATION') FROM VENDORS V ORDER BY 2 DESC FETCH FIRST 10 ROWS ONLY;</pre>	Returns the 10 vendors that are most similar to IBM Corporation
<pre>SELECT V.VENDOR_NAME, AI_SEMANTIC_CLUSTER(VENDOR_NAME, 'IBM CORPORATION', 'AMAZON', 'MICROSOFT') FROM VENDORS V ORDER BY 2 DESC FETCH FIRST 10 ROWS ONLY;</pre>	Returns the top 10 vendors that are most similar to the cluster of vendors comprised of IBM Corporation, Amazon and Microsoft
<pre>SELECT V.SERVICE_COUNTRY, AI_ANALOGY('IBM CORPORATION' USING MODEL COLUMN VENDOR_NAME, 'USA' USING MODEL COLUMN SERVICE_COUNTRY, 'SAMSUNG' USING MODEL COLUMN VENDOR_NAME, SERVICE_COUNTRY) FROM VENDORS V ORDER BY 2 DESC FETCH FIRST 10 ROWS ONLY;</pre>	Returns the 10 service countries that, when paired with the vendor Samsung, are most analogous to the service country USA as it is paired with IBM Corporation
<pre>SELECT AI_COMMONALITY(C.DRIVERS_LICENSE_NUMBER) AS SCORE, C.* FROM SDILAB.INSURANCE C ORDER BY SCORE ASC FETCH FIRST 15 ROWS ONLY ;</pre>	Returns the top 15 driver license numbers that are outliers with respect to all rows in the INSURANCE table

SQL Data Insights: Model representation

Model training process
(from UI, outside Db2)



A model is a Db2 table containing encoded vectors for each distinct entity in the source table

SQL semantic functions retrieve the vectors to calculate their results

USER.DATA_TABLE

CUSTOMER_ID	GENDER	BILLING	...
3668-QPYBK	F	auto	...
...



Data Engineer



DBA

DSNAIDB.<generated vector table name>

Column	Value	vector
CUSTOMER_ID	3668-QPYBK	<1280 byte vector>
GENDER	F	<1280 byte vector>
...

```
SQL: SELECT CustomerID,
AI_SIMILARITY(CUSTOMER_ID, '3668-QPYBK')
FROM USER.DATA_TABLE
WHERE ...
```



Business Analyst



App Developer

SQL Data Insights ... use cases

Uncover hidden relationships by using complex interactions between AI and critical business systems

Credit Card



Help identify if credit card rewards participants are gaming the system

Government



Search for tax cheats by finding similarity to known cheaters

Sales



Look for customers based upon similar buying patterns

Analyze customer churn

Predict sales of new products to existing customers

Insurance



Help identify customers with similar characteristics

Aid in finding appropriate insurance rates

Discover trends in policies and sales and claims

SQL Data Insights: Software & hardware requirements

Software:

- Db2 for z/OS
 - Db2 13 as built-in AI functions
 - Db2 12 technical preview provided as UDF AI functions
- SQL Data Insight UI and training services
 - separately orderable, no-charge feature of Db2 13 (FMID HDBDD18)
- z/OS maintenance
 - z/OS 2.4 and above and
 - Install 3 IBM neural network libraries

Separate install steps needed to enable SQL Data Insights
– refer to [IBM Documentation](#) and [IBM Redbooks](#)

Hardware:

- zEC12 and above
- Z14 + leverages OpenBLAS library exploitation for AI with SIMD
- Both training and SQL execution is zIIP eligible

Finding hidden information

I want to find other customers like this one

Customer ID	Gender	Senior Citizen	Dependents	Tenure	Phone Service	Multiple Lines	Internet service	Contract	Paperless billing	Payment method	Charges	Churn
001	Male	0	NO	2	YES	NO	DSL	Month-to-Month	YES	Mailed Check	120.22	YES

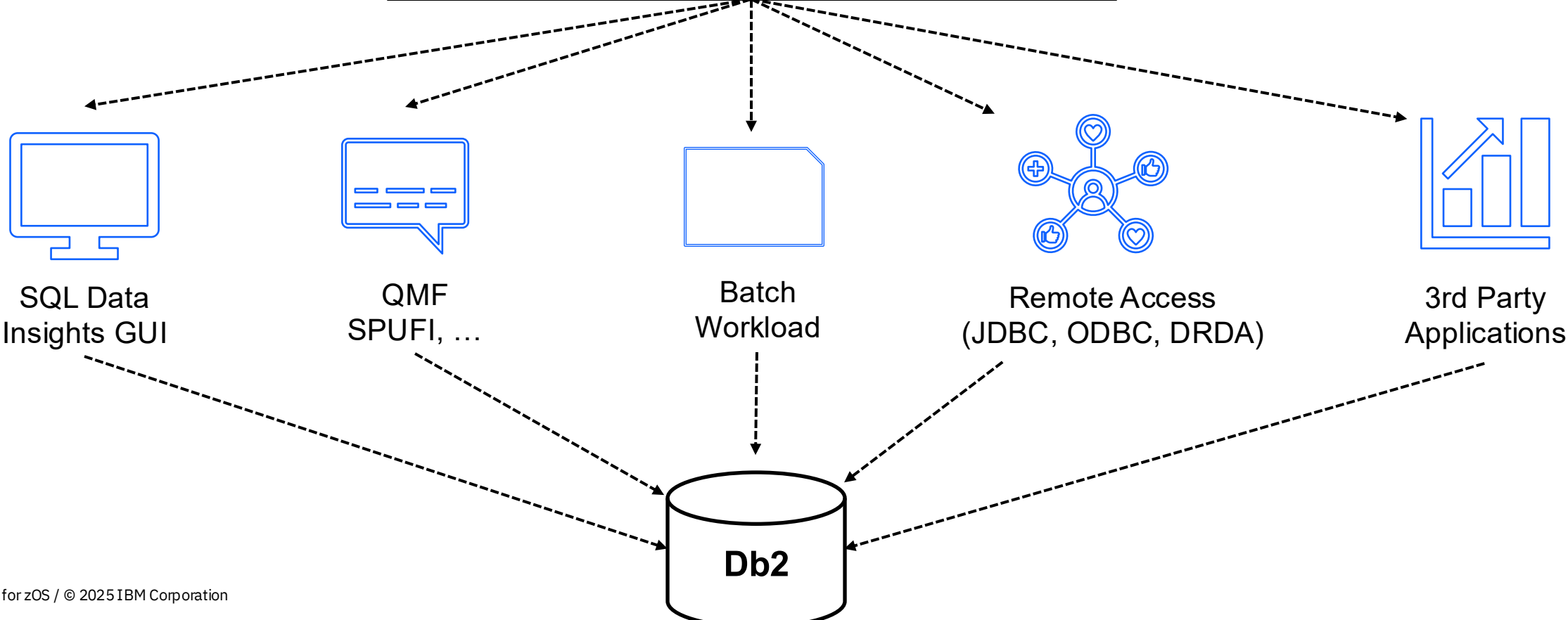
```
SELECT *,
AI_SIMILARITY (Customer_ID, '001')
AS SimScore
FROM TABLE WHERE ...
ORDER BY SimScore DESC
```

Ranked Similarity Results (most to least similar)

Sim Score	Customer ID	Gender	Senior Citizen	Dependents	Tenure	Phone Service	Multiple Lines	Internet service	Contract	Paperless billing	Payment method	Charges	Churn
0.80	004	Male	0	NO	1	YES	NO	DSL	Month-to-Month	YES	Mailed Check	48.55	YES
0.75	002	Male	0	NO	7	NO	NO	DSL	Month-to-Month	YES	Mailed Check	51.00	YES
0.70	006	Male	0	NO	3	NO	NO	DSL	Month-to-Month	YES	Mailed Check	49.80	YES
0.55	003	Female	0	NO	4	NO	NO	DSL	Month-to-Month	YES	Mailed Check	60.40	YES
0.35	005	Female	1	NO	1	NO	NO	DSL	Month-to-Month	YES	Credit Card	55.10	YES

Power any Db2 for z/OS application with AI enhanced SQL

```
SELECT CustomerID,  
       AI_SIMILARITY('3668-QPYBK',  
                    CustomerID) AS Report1  
FROM USER.DATA_TABLE  
ORDER BY 2 DESC  
FETCH FIRST 5 ROWS ONLY;
```

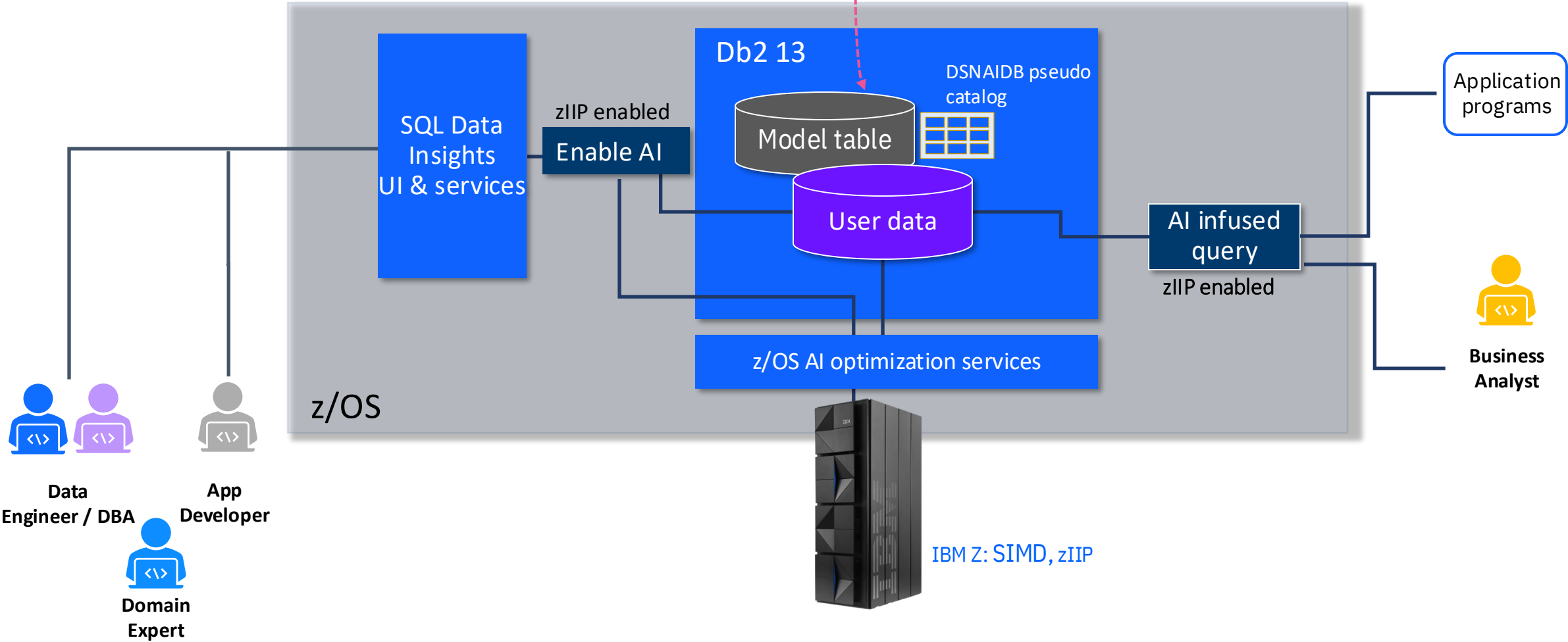


SQL Data Insights: Usage scenario

Vector information

```

8879-zZna
-0.141558 -0.346767 -0.453296 0.052447 0.476916 -
0.338483 0.000035 0.517277 0.191573 0.076891
-0.149729 1.036879 0.127160 -0.329846 -0.157252
-0.288485 0.243588 0.038326 -0.338862 0.173571
0.231060 0.149021 -0.328546 -0.058121 0.025713 ...
    
```



Experiencing Db2 SQL Data Insights

Choosing the base table for training

SQL Data Insights ⚙️ ? 👤

[Connections](#) /

Add object

bordawconnection
Choose one or more schemas to list associated Db2 objects. From the list, select the Db2 objects to add as new AI objects or enable for AI query.

Schema: 1 × Select schema ^ 🔍

Name	Schema	Last activity
<input type="checkbox"/> DSN81210		
<input type="checkbox"/> DSN81310		
<input checked="" type="checkbox"/> DSN81310		
<input type="checkbox"/> DSNRGCOL		
<input type="checkbox"/> ENCDB2		
<input type="checkbox"/> ENCDB2		

No object available. Select one or more schemas to list available objects.

Items per page: 10 ▾ 0-0 of 0 items 1 ▾ of 1 page ◀ ▶

[Cancel](#) [Add object](#) [Enable AI query](#)

Experiencing Db2 SQL Data Insights

Choosing the base table for training

The screenshot shows the 'SQL Data Insights' interface. At the top, there is a search bar with 'Schema: 1 x Select schema' and a search icon. Below this is a table with columns 'Name', 'Schema', and 'Last activity'. The table lists several objects, with 'CHURN' selected (checked checkbox). At the bottom, there are pagination controls showing 'Items per page: 10', '21-30 of 35 items', and '3 of 4 pages'. A footer bar contains 'Cancel', 'Add object', and 'Enable AI query' buttons.

Name	Schema	Last activity
<input type="checkbox"/> AIDB_DSNAIDB_OPENFOOD_MARCH13_ATT3	DSNAIDB	i
<input checked="" type="checkbox"/> CHURN	DSNAIDB	i
<input type="checkbox"/> OPENFOOD22	DSNAIDB	i
<input type="checkbox"/> OPENFOOD23	DSNAIDB	i
<input type="checkbox"/> OPENFOODFACTS1	DSNAIDB	i
<input type="checkbox"/> OPENFOODFACTS2	DSNAIDB	i
<input type="checkbox"/> OPENFOODFACTS3	DSNAIDB	i
<input type="checkbox"/> OPENFOODMARCH10ATT7ALL	DSNAIDB	i
<input type="checkbox"/> OPENFOODMARCH13ATT1	DSNAIDB	i
<input type="checkbox"/> OPENFOODMARCH13ATT4LIMITED	DSNAIDB	i

Items per page: 10 | 21-30 of 35 items | 3 of 4 pages

Cancel | Add object | Enable AI query

Experiencing Db2 SQL Data Insights

Configuring database table for training: Set up column types

SQL Data Insights

Enabling an object for AI query requires column configuration and model training. To enable DSNAIDB.CHURN for AI query, select columns, assign SQL DI data types, and specify as NULL for model training. You can also [import](#) the column configuration of the AI object from a JSON file ⓘ
. Make sure the columns defined in the file are consistent with those in DSNAIDB.CHURN.

Select columns and assign SQL DI dat... Specify column values as NULL for mo...

You can select one or more columns and assign each column a SQL DI categorical, numeric, or key data type. SQL DI uses your selections to create a column configuration and train a machine learning model for the object.

Select columns for training

21 items selected | Cancel

<input checked="" type="checkbox"/> Column name	SQL DI data type ⓘ
<input checked="" type="checkbox"/> CUSTOMERID	Key ▾
<input checked="" type="checkbox"/> GENDER	Categorical ▾
<input checked="" type="checkbox"/> SENIORCITIZEN	Categorical ▾
<input checked="" type="checkbox"/> PARTNER	Categorical ▾
<input checked="" type="checkbox"/> DEPENDENTS	Categorical ▾

Cancel Next

Experiencing Db2 SQL Data Insights

Configuring database table for training: Specify user-specified NULL values (e.g. NA)

SQL Data Insights

Enabling an object for AI query requires column configuration and model training. To enable DSNAIDB.CHURN for AI query, select columns, assign SQL DI data types, and specify as NULL for model training. You can also [import](#) the column configuration of the AI object from a JSON file ⓘ
. Make sure the columns defined in the file are consistent with those in DSNAIDB.CHURN.

Select columns and assign SQL DI dat... Specify column values as NULL for mo...

You can specify column values of your choice, such as N/A, n/a, na, NR, invalid, and empty, as NULL values. SQL DI ignores these user-specified NULL values during model training.

Specify NULL values that apply to all columns

Use semicolon to separate multiple values. Example: N/A;n/a;na;NR;invalid;empty

Column name	Specify NULL values that apply to a specif...
CUSTOMERID	Use semicolon to separate multiple values. E
GENDER	Use semicolon to separate multiple values. E
SENIORCITIZEN	Use semicolon to separate multiple values. E
PARTNER	Use semicolon to separate multiple values. E
DEPENDENTS	Use semicolon to separate multiple values. E

Back Enable

Experiencing Db2 SQL Data Insights

Initiate the training

SQL Data Insights

Enabling an object for AI query requires column configuration and model training. To enable DSNAIDB.CHURN for AI query, select columns, assign SQL DI data types, and specify as NULL for model training. You can also **import** the column configuration of the AI object from a JSON file ⓘ. Make sure the columns defined in the file are consistent with those in DSNAIDB.CHURN.

Select columns and assign SQL DI dat... Specify column values as NULL for mo...

You can specify column values of... values during model training.

Specify NULL values that apply to

Use semicolon to separate multi...

Column name	
CUSTOMERID	
GENDER	Use semicolon to separate multiple values. E
SENIORCITIZEN	Use semicolon to separate multiple values. E
PARTNER	Use semicolon to separate multiple values. E
DEPENDENTS	Use semicolon to separate multiple values. E

Cancel Enable

Back Enable

Experiencing Db2 SQL Data Insights

Initiate the training

SQL Data Insights ⚙️ ? 👤

✔️ **Enabling AI query.** AI query enablement for DSNAIDB.CHURN is in progress. Refresh this page for status updates. ✕

AI objects

bordaw2-connection

🔍 Find object Add object + Run query 📄

Name	Schema	Type	Status	Last updated	
⌵ CLAIMSDEMOSUBSET2	DSNAIDB	Table	🚨 Failed	Mar 27, 2023 5:29 PM	⋮
⌵ CLAIMSDEMOSUBSET3	DSNAIDB	Table	✔️ Enabled	Mar 27, 2023 5:59 PM	⋮
⌵ CLAIMSDEMOSUBSET4FULLSET	DSNAIDB	Table	✔️ Enabled	Mar 29, 2023 5:35 PM	⋮
⌵ CLAIMSDEMOSUBSET5FULLSET	DSNAIDB	Table	✔️ Enabled	Mar 28, 2023 3:44 PM	⋮
⌶ CHURN	DSNAIDB	Table	🔄 Enabling	Apr 3, 2023 3:44 PM	⋮

Created on Apr 3, 2023

Current status
Initializing

Last activity
🔄 Apr 3, 2023 3:44 PM
Enabling

Resources per page: 10 ▾ 11-15 of 15 items 2 ▾ of 2 pages ⏪ ⏩

Experiencing Db2 SQL Data Insights

Analyze the trained model

SQL Data Insights ⚙️ ? 👤

[Connections](#) / [AI objects](#) /

Analyze data

DSNAIDB.CHURN Last updated: Apr 3, 2023 3:50 PM ↻

Object details **Data statistics** Column influence Column discriminator >

Column ...	Db2 dat...	# of uniq...	Most co...	# of mos...	Mean va...	Standar...	Min value	Max value
MONTHLYCHARGE	DECIMAL	1585	20.05	61	64.76	30.09	18.25	118.75
TENURE	INTEGER	73	1	613	32.00	24.56	0.0	72.0
TOTALCHARGES	DECIMAL	6531	.00	11	2279.73	2266.63	0.0	8684.8
ONLINEBACKUP	VARCHAR	3	No	3088	-	-	-	-
DEPENDENTS	VARCHAR	2	No	4933	-	-	-	-
PARTNER	VARCHAR	2	No	3641	-	-	-	-
SENIORCITIZEN	VARCHAR	2	0	5901	-	-	-	-
GENDER	VARCHAR	2	Male	3555	-	-	-	-
PHONESERVICE	VARCHAR	2	Yes	6361	-	-	-	-

Intelligent performance enhancements in Db2 V13

Internal block fetch enhancement

Historically, Db2 data manager (DM – aka stage 1) would pass a query’s result set **one row at a time** to the Db2 relational data system (RDS – aka stage 2)

Db2 12 performance enhancement – *internal block fetch*: for certain types of query, DM can pass **a set of result set rows** to a buffer for retrieval by RDS

With Db2 12, buffer used to hold a block of rows will not be allocated until number of rows fetched from result set has exceeded an internal threshold

– This has the effect of delaying efficiency gain provided by internal block fetch

Db2 13 enhancement: **“intelligent” internal block fetch**

– Db2 records number of rows fetched for a query, and uses that information to **learn** that a given query is likely to fetch enough rows for internal block fetch to be helpful

– When that query is subsequently executed, internal block fetch will be **immediately enabled**

– Use of internal block fetch “from the start” improves efficiency of query execution

AI in Db2: Insert/update/delete performance (index)

Previous behavior:

Decision on use of index look-aside based on catalog statistics, which may be stale

Db2 13 behavior:

- Use of index look-aside based on execution-time information (cache of previous index probe)
- Can be used with or without FTBs
- 5-25% CPU savings for batch processing
- 2-3% CPU savings for OLTP

Previous behavior:

- If any column is a DECFLOAT column, sort does not generate machine code to help move data into sort tree
- Sort enhancements for ORDER BY and GROUP BY are not applied to other areas of sort

Db2 13 behavior:

Improved query performance

- Machine language code generated to boost efficiency of DECFLOAT sorts
- Performance enhancements for DISTINCT, GROUPING SET and PERCENTILE sorts
- More space-efficient sorting for LISTAGG within SUBSTR

Add support to reduce the length of long VARCHAR values if at the end of keys to be sorted (ML enabled)

AI in Db2: SORTL expansion using learning from execution

FL 100

Previous behavior:

- Db2 12 APAR PH31684 added support for SORTL instruction introduced with z15, but Db2 12 use of SORTL was limited
 - Length of sort key could not exceed 136 bytes
 - Length of records to be sorted could not exceed 256 bytes

Db2 13 behavior:

- Db2 13 goes beyond Db2 12 limits, makes SORTL usable in more situations
- Expanded use of SORTL driven by learned behavior of queries:
 - number of rows processed
 - size of row
 - sort pool size

Summary

Db2 AI on z/OS

- SQL Optimization
- System Assessment
- Distributed Connection Control

SQL Data Insights

AI performance enhancements

Thank you!

Keziah Knopp
Data & AI on Z – Financial Services Market
keziah.knopp@ibm.com

IBM