



Data Intelligence + Human Brilliance

Cataloging for Mere Mortals:

How to Find, Understand and Trust Your Data

Christopher Johnson | Alation Solutions Engineering

March 15, 2023

Hello!



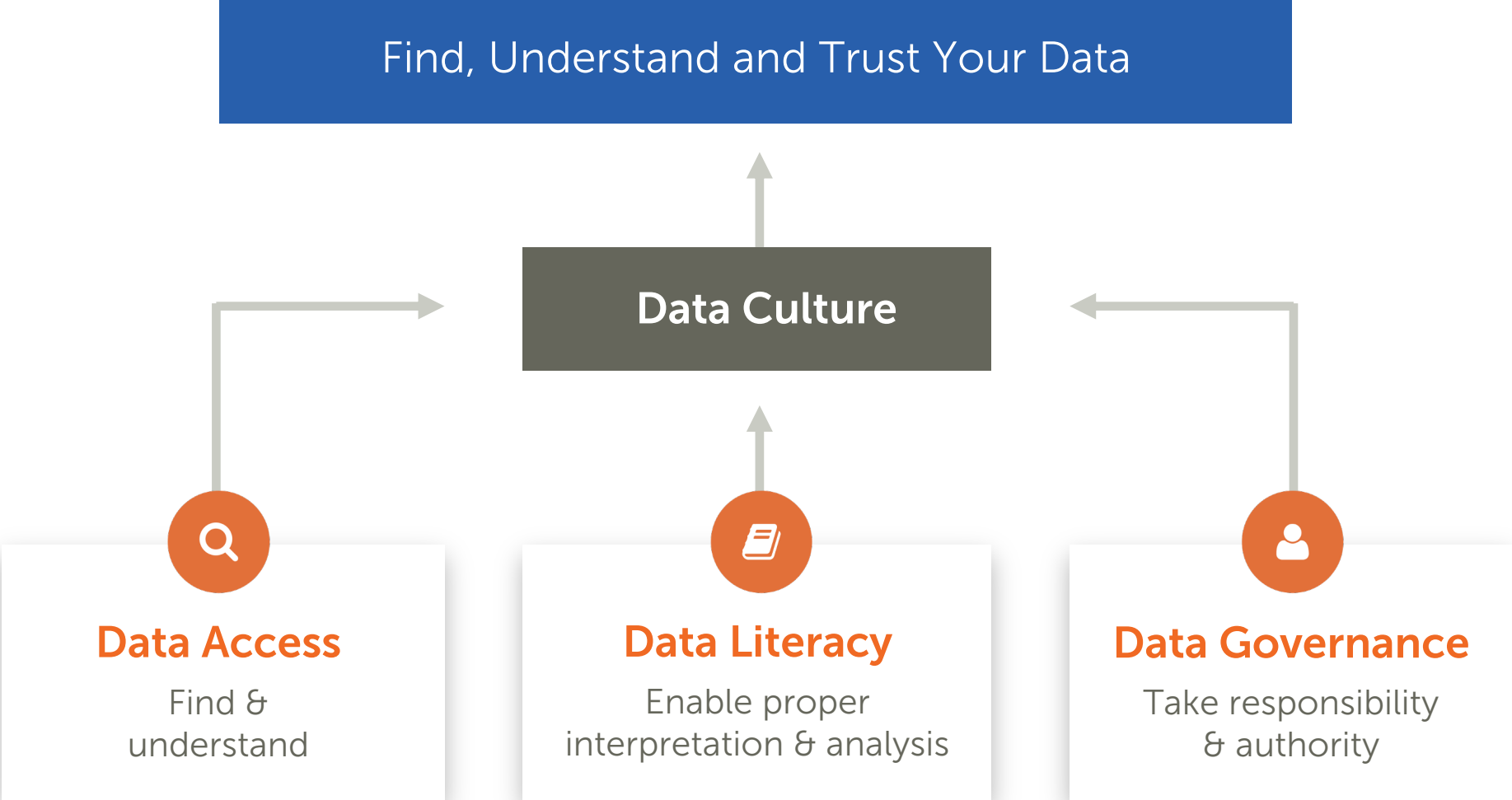
Chris Johnson

Joined Alation in 2018
25+ years in Data
DMBOK II Contributor

christopher.johnson@alation.com

[linkedin.com/in/betterdata](https://www.linkedin.com/in/betterdata)

Alation's Vision



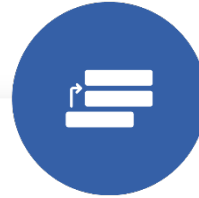
Organizations Are Increasingly Data Rich



Data Explosion

Data Growth (IDC) brings increasing regulations.

"The sum of all data created, captured, or replicated at core, edge, and endpoint locations - will expand from 33 ZB in 2018 to 175 ZB by 2025."



Data Impact

"Data and analytics (D&A) remains a top priority for organizations — both during and after the COVID-19 crisis."

[\(Gartner\)](#)



Adoption

"By 2023, overall analytics adoption will increase from 35% to 50%, driven by vertical- and domain-specific augmented analytics solutions."
(Gartner)

Data Intelligence Drives Human Brilliance...



Businesses achieve **agility** by pushing decision making down the organization



An ever-growing audience needs **current data**



Catalogs **reduce** the **cycle time** for decision making and preventing IT bottlenecks

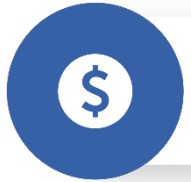


Outcome: moving from **data overload** to **data driven**

Resulting in Better Business Outcomes



Shorter time to value



Lower business costs



Business agility



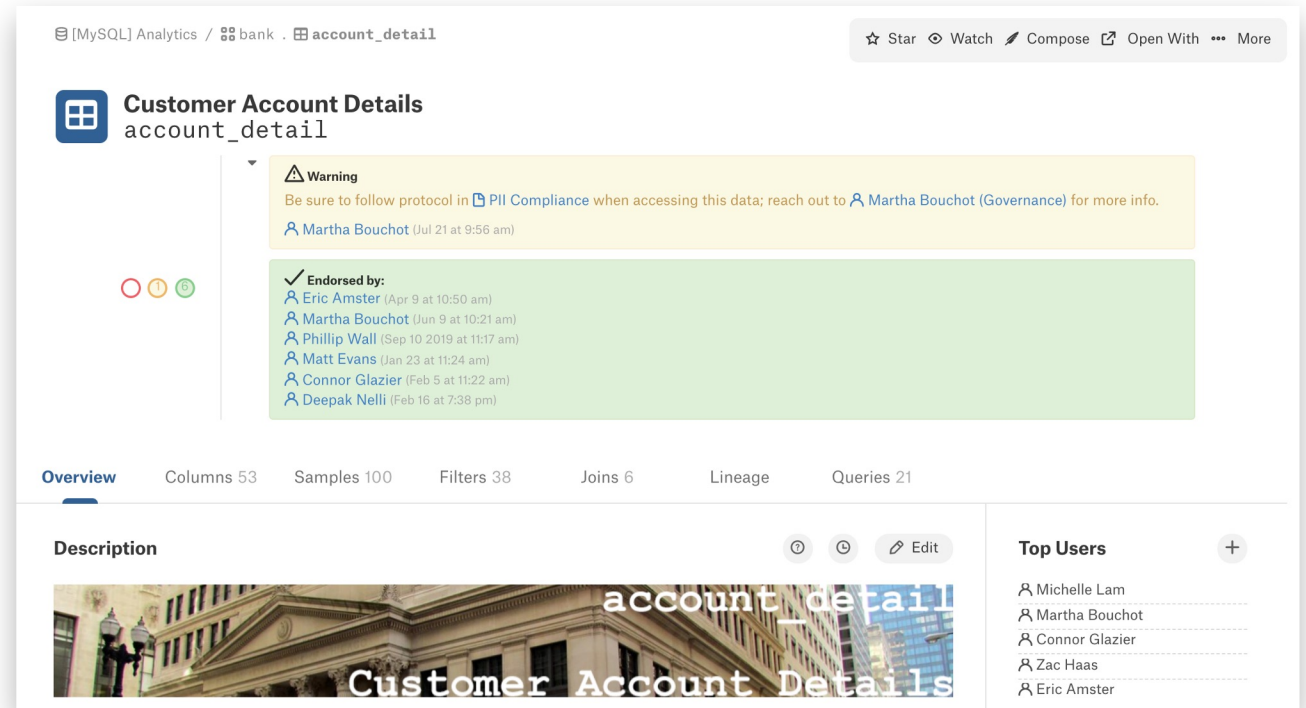
Faster/better decision making + innovation



Data driven process increases Transformation at scale

What is an Enterprise Data Catalog?

- A repository of metadata on information sources across an organization
 - Search & discovery
 - Collaboration & analysis
 - Curation & governance
- Indexes a broad range of information assets
 - Data sets, tables, articles, reports, queries, tables, conversations



Answers core questions, such as:

How to find information?

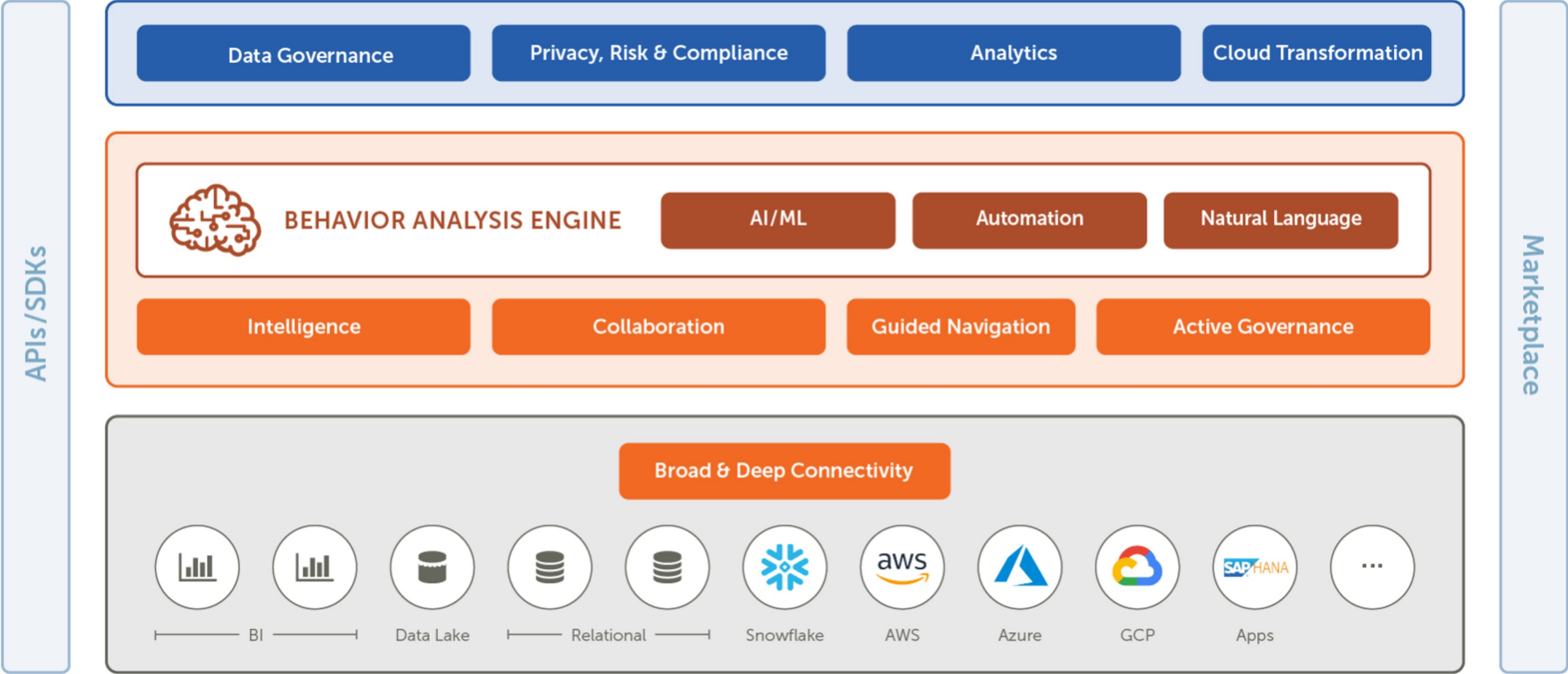
Can it be used?

Should it be used?

How should it be used?

The Alation Platform

For enterprise metadata management



Alation's Intelligence

A 'Virtual Assistant' for Your Data Producers and Consumers

Alation's Intelligence

A 'Virtual Assistant' for Your Data Producers and Consumers

- What data assets do we have?

Enterprise Data Platform / FIN_SVS.INV_BANK

☆ Star Watch Compose More

Investment Banking
FIN_SVS.INV_BANK

Overview Queries 6

Source Comments
Investment Bank Data

Description 🔔 🕒 [Suggest Change](#)

Snowflake **fin_svs.inv_bank (Investment Banking)** was migrated from MySQL **Retail Warehouse**. Please refer to [Cloud Migration Project](#) for details and please consult with [Iris Thomas \(IT Department\)](#) if there are additional needs that have not been addressed.

Tables Show 25 rows Filter

★	✓	Table	Title	📄	Popularity ↓	Columns	Rows
☆	+2	DEALFACT	Broker Dealer Revenue			27	1931
☆		BNKR_TB	Banker			7	10
☆		IND_DIM	Industry			12	777
☆	+2	CUSTMRTBL	Customer	📄		14	1201
☆	+1	CSTMR	Customer	📄		10	999
☆		REG_DIM	Region Dimension	📄		12	240
☆		CL_VOL	Call Volume	📄		6	8
☆		DEALFACT_VW	Deal Fact View	📄		19	0
☆		DDA_DET	Demand Deposit Account Details	📄		16	49
☆	+1	CSTMR_DEMS	Customer Demographics	📄		5	1998

Top Users +

- Alation SE
- Deepak Nelli
- David Ahn
- Bob Welshmer
- Aaron B

Domains +

- Investment

Data Owners 🕒 +

- Govin Na (Head of Custome...)

Stewards 🕒 +

- Sally Stewart (Data Govern...)

Tags +

- IBanking

Alation's Intelligence

A 'Virtual Assistant' for Your Data Producers and Consumers

- What data assets do we have?
- Where are they located?

Enterprise Data Platform / FIN_SVS.INV_BANK

Investment Banking
FIN_SVS.INV_BANK

Overview Queries 6

Source Comments
Investment Bank Data

Description
Snowflake fin_svs.inv_bank (Investment Banking) was migrated from MySQL Retail Warehouse. Please refer to Cloud Migration Project for details and please consult with Iris Thomas (IT Department) if there are additional needs that have not been addressed.

Tables Show 25 rows Filter

★	✓	Table	Title	Popularity ↓	Columns	Rows
☆	+2	DEALFACT	Broker Dealer Revenue		27	1931
☆		BNKR_TB	Banker		7	10
☆		IND_DIM	Industry		12	777
☆	+2	CUSTOMRTBL	Customer		14	1201
☆	+1	CSTMR	Customer		10	999
☆		REG_DIM	Region Dimension		12	240
☆		CL_VOL	Call Volume		6	8
☆		DEALFACT_VW	Deal Fact View		19	0
☆		DDA_DET	Demand Deposit Account Details		16	49
☆	+1	CSTMR_DEMS	Customer Demographics		5	1998

Top Users

- Alation SE
- Deepak Nelli
- David Ahn
- Bob Welshmer
- Aaron B

Domains

- Investment

Data Owners

- Govin Na (Head of Custome...)

Stewards

- Sally Stewart (Data Govern...)

Tags

- IBanking

Alation's Intelligence

A 'Virtual Assistant' for Your Data Producers and Consumers

- What data assets do we have?
- Where are they located?
- What do they contain?

Enterprise Data Platform / FIN_SVS.INV_BANK

Star Watch Compose More

Investment Banking

FIN_SVS.INV_BANK

Overview Queries 6

Source Comments

Investment Bank Data

Description

Snowflake [fin_svs.inv_bank](#) (Investment Banking) was migrated from MySQL [Retail Warehouse](#). Please refer to [Cloud Migration Project](#) for details and please consult with [Iris Thomas](#) (IT Department) if there are additional needs that have not been addressed.

Tables

Show 25 rows Filter

★	✓	Table	Title	Popularity ↓	Columns	Rows
☆	+2	DEALFACT	Broker Dealer Revenue		27	1931
☆		BNKR_TB	Banker		7	10
☆		IND_DIM	Industry		12	777
☆	+2	CUSTMRTBL	Customer		14	1201
☆	+1	CSTMTR	Customer		10	999
☆		REG_DIM	Region Dimension		12	240
☆		CL_VOL	Call Volume		6	8
☆		DEALFACT_VW	Deal Fact View		19	0
☆		DDA_DET	Demand Deposit Account Details		16	49
☆	+1	CSTMTR_DEMS	Customer Demographics		5	1998

Top Users

- Alation SE
- Deepak Nelli
- David Ahn
- Bob Welshmer
- Aaron B

Domains

- Investment

Data Owners

- Govin Na (Head of Custome...)

Stewards

- Sally Stewart (Data Govern...)

Tags

- IBanking

Alation's Intelligence

A 'Virtual Assistant' for Your Data Producers and Consumers

- What data assets do we have?
- Where are they located?
- What do they contain?
- How often are they used?

Enterprise Data Platform / FIN_SVS.INV_BANK

Star Watch Compose More

Investment Banking

FIN_SVS.INV_BANK

Overview Queries 6

Source Comments

Investment Bank Data

Description

Suggest Change

Snowflake fin_svs.inv_bank (Investment Banking) was migrated from MySQL Retail Warehouse. Please refer to Cloud Migration Project for details and please consult with Iris Thomas (IT Department) if there are additional needs that have not been addressed.

Tables

★	✓	Table	Title	Popularity ↓	Columns	Rows
☆	+2	DEALFACT	Broker Dealer Revenue		27	1931
☆		BNKR_TB	Banker		7	10
☆		IND_DIM	Industry		12	777
☆	+2	CUSTMRTBL	Customer		14	1201
☆	+1	CSTMR	Customer		10	999
☆		REG_DIM	Region Dimension		12	240
☆		CL_VOL	Call Volume		6	8
☆		DEALFACT_VW	Deal Fact View		19	0
☆		DDA_DET	Demand Deposit Account Details		16	49
☆	+1	CSTMRTBL	Customer Demographics		5	1998

Show 25 rows Filter

Top Users

- Alation SE
- Deepak Nelli
- David Ahn
- Bob Welshmer
- Aaron B

Domains

- Investment

Data Owners

- Govin Na (Head of Custome...)

Stewards

- Sally Stewart (Data Govern...)

Tags

- IBanking

Alation's Intelligence

A 'Virtual Assistant' for Your Data Producers and Consumers

- What data assets do we have?
- Where are they located?
- What do they contain?
- How often are they used?
- Who uses them?

Enterprise Data Platform / FIN_SVS.INV_BANK

Star Watch Compose More

Investment Banking

FIN_SVS.INV_BANK

Overview Queries 6

Source Comments ⓘ
Investment Bank Data

Description ⓘ ⓘ [Suggest Change](#)

Snowflake [fin_svs.inv_bank](#) (Investment Banking) was migrated from MySQL [Retail Warehouse](#). Please refer to [Cloud Migration Project](#) for details and please consult with [Iris Thomas \(IT Department\)](#) if there are additional needs that have not been addressed.

Tables Show 25 rows ▾ Filter

★	✓	Table	Title	Popularity ↓	Columns	Rows
☆	+2	DEALFACT	Broker Dealer Revenue		27	1931
☆		BNKR_TB	Banker		7	10
☆		IND_DIM	Industry		12	777
☆	+2	CUSTMRTBL	Customer		14	1201
☆	+1	CSTMR	Customer		10	999
☆		REG_DIM	Region Dimension		12	240
☆		CL_VOL	Call Volume		6	8
☆		DEALFACT_VW	Deal Fact View		19	0
☆		DDA_DET	Demand Deposit Account Details		16	49
☆	+1	CSTMR_DEMS	Customer Demographics		5	1998

Top Users +

- Alation SE
- Deepak Nelli
- David Ahn
- Bob Welshmer
- Aaron B

Domains +

- Investment

Data Owners ⓘ +

- Govin Na (Head of Custome...)

Stewards ⓘ +

- Sally Stewart (Data Govern...)

Tags +

- IBanking

Alation's Intelligence

A 'Virtual Assistant' for Your Data Producers and Consumers

- What data assets do we have?
- Where are they located?
- What do they contain?
- How often are they used?
- Who uses them?
- How are these data assets used?

The screenshot displays the Alation interface for a data asset named 'Broker Dealer Revenue' (DEALFACT). The interface includes a navigation bar with 'Overview', 'Columns 27', 'Samples 100', 'Filters 8', 'Joins 3', 'Lineage', and 'Queries 5'. A 'Published History' section is visible on the left. The main content area shows a list of queries. The first query, 'Product Revenue Growth by Quarter', is highlighted with an orange border. It includes a description: 'Breakdown of the growth (or not) of each product offered, quarter over quarter. This is intended to look at FY07 and provide general sense of what happened in FY08'. The SQL query is as follows:

```
select product_description, case when revenue_date between '2007-01-01' and '2007-03-31' then 'Q1' when revenue_date between '2007-04-01' and '2007-06-30' then 'Q2' when revenue_date between '2007-07-01' and '2007-09-30' then 'Q3' when revenue_date between '2007-10-01' and '2007-12-31' then 'Q4' else 'FY08' end as "Sales Quarter", sum(institutional_rev) from FIN_SVS_INV_BANK_DEALFACT where 'SALES_QUARTER' <> 'FY08' group by 1,2 order by 1,2 DESC
```

The query is attributed to Nick Haithcock, last updated on Aug 17 at 1:12pm, cloned 2 times, and executed 12 times. It was last visited on Last Tue at 2:32pm. The second query, 'Asset Under Management - per employee (AuM)', is also visible below it.

Alation's Intelligence

A 'Virtual Assistant' for Your Data Producers and Consumers

- What data assets do we have?
- Where are they located?
- What do they contain?
- How often are they used?
- Who uses them?
- How are these data assets used?
- How are these assets related?

Enterprise Data Platform / FIN_SVS.INV_BANK . DEALFACT

☆ Star ○ Watch ✍ Compose 📄 Open With ⋮ More

Broker Dealer Revenue

DEALFACT

○ ⚠️ ✓ | ▶ 1 Warning 2 Endorsements

Overview Columns 27 Samples 100 Filters 8 **Joins 3** Lineage Queries 5

Joins and Their Conditions

Search Items/Page: 10

JOIN `fin_svs.inv_bank.ind_dim`

* ON `fin_svs.inv_bank.dealfact.sub_industry_cd = fin_svs.inv_bank.ind_dim.sub_unique_cd`

Untitled ○ ○ ○ ☆ **frequency 129**

1 Endorsement

JOIN `fin_svs.inv_bank.bnkr_tb`

* ON `fin_svs.inv_bank.bnkr_tb.banker_id = fin_svs.inv_bank.dealfact.banker_id`

Untitled ○ ○ ○ ☆ **frequency 137**

JOIN `fin_svs.inv_bank.reg_dim`

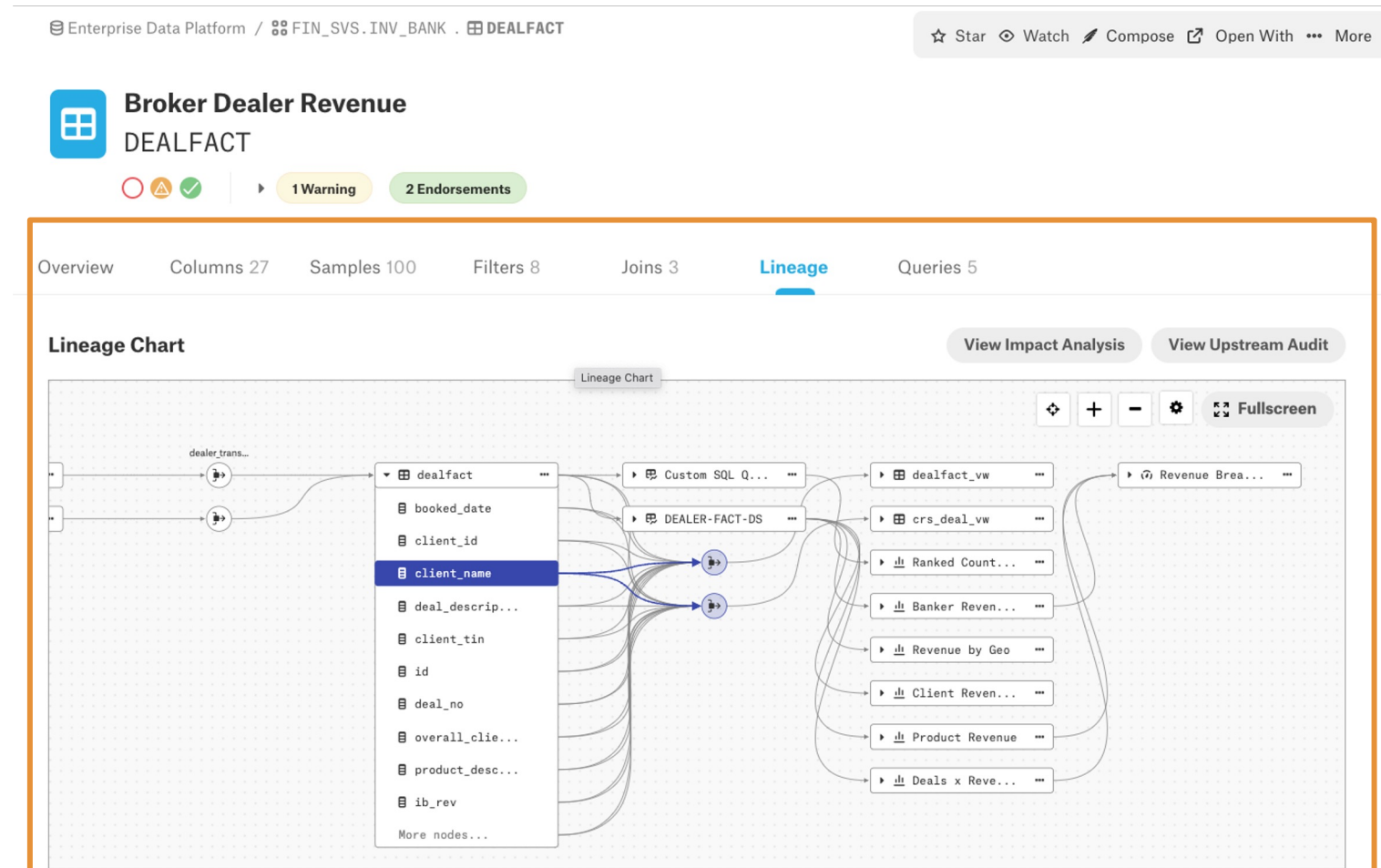
* ON `fin_svs.inv_bank.dealfact.country_cd = fin_svs.inv_bank.reg_dim.country_cd`

Untitled ○ ○ ○ ☆ **frequency 31**

Alation's Intelligence

A 'Virtual Assistant' for Your Data Producers and Consumers

- What data assets do we have?
- Where are they located?
- What do they contain?
- How often are they used?
- Who uses them?
- How are these data assets used?
- How are these assets related?



Alation's Intelligence

A 'Virtual Assistant' for Your Data Producers and Consumers

- What data assets do we have?
- Where are they located?
- What do they contain?
- How often are they used?
- Who uses them?
- How are these data assets used?
- What does a sample look like?

Enterprise Data Platform / FIN_SVS.INV_BANK . DEALFACT

Star Watch Compose Open With More

Broker Dealer Revenue
DEALFACT

1 Warning 2 Endorsements

Overview Columns 27 Samples 100 Filters 8 Joins 3 Lineage Queries 5

100 rows randomly sampled as of Sep 8 at 12:05am

Run Sample

ID	DEAL_NO	CLIENT_ID	CLIENT_NAME	BANKER_ID	COUNTRY_CD	SUB_INDUSTRY_CD	PRODUCT_DETAIL_CODE	PR
NUMBER	NUMBER	NUMBER	VARCHAR	NUMBER	VARCHAR	NUMBER	VARCHAR	VAR
Identifier	Deal Number	Client Identifier	Client Name	Banker Identifier	Country Code		Product Detail Code	Prod
1065	106985	0	*****	4	BD	27	IRD	IN
3	60552	0	*****	2	AV	24	VF	PR
1206	107764	0	*****	9	US	145	GA	GOV
496	94710	0	*****	5	VQ	25	DNP	HIC
1000	97206	0	*****	7	DR	49	RLF	REI
1783	144030	0	*****	2	SA	99	PC	COP
688	97296	0	*****	6	VQ	27	RLF	REI
271	92827	0	*****	10	RQ	81	DNP	HIC
1318	108196	0	*****	1	HA	99	IRD	IN
1481	109020	0	*****	4	AC	25	IRD	IN
095	07070	0	*****	0	MT	26	DI	DE

Alation's Intelligence

A 'Virtual Assistant' for Your Data Producers and Consumers

- What data assets do we have?
- Where are they located?
- What do they contain?
- How often are they used?
- Who uses them?
- How are these data assets used?
- How are these assets related?
- What policies and tags have been applied?

The screenshot shows the Alation data catalog interface for a table named 'Broker Dealer Revenue' (DEALFACT). The interface includes a top navigation bar with 'Enterprise Data Platform / FIN_SVS.INV_BANK . DEALFACT', a utility bar with 'Star', 'Watch', 'Compose', 'Open With', and 'More' options, and a status bar showing '1 Warning' and '2 Endorsements'. The main content area is divided into several sections: 'Overview' (Columns 27, Samples 100, Filters 8, Joins 3, Lineage, Queries 5), 'Description' (This table holds information regarding on-going investment banking deals.), 'Sample Columns' (3 of 27), 'Sample Content' (3 of 1,931), and 'Published Queries' (3 of 5). The 'Sample Content' section contains a table with columns: ID, DEAL_NO, CLIENT_ID, CLIENT_NAME, BANKER_ID, COUNTRY_CD, SUB_INDUSTRY_CD, PRODUCT_DETAIL_CODE, and PRODUCT_DETAIL_CODE. The 'CLIENT_NAME' column is highlighted with an orange box. The 'Data Policies' section on the right is also highlighted with an orange box, showing 'Region Access Policy'. Other sections include 'Top Users' (Sally Stewart, Alation SE, Deepak Nelli, Nick Haithcock, David Ahn), 'Domains' (Investment), 'Stewards' (Sally Stewart (Data Govern...)), 'Data Product O...' (Govin Na (Head of Custome...)), and 'Data Quality Info' (Profile - Last Run).

Column	Title	Type	Popularity
1 ID	Identifier	NUMBER (10, 0)	
2 DEAL_NO	Deal Number	NUMBER (10, 0)	
3 CLIENT_ID	Client Identifier	NUMBER (10, 0)	

ID	DEAL_NO	CLIENT_ID	CLIENT_NAME	BANKER_ID	COUNTRY_CD	SUB_INDUSTRY_CD	PRODUCT_DETAIL_CODE	PRODUCT_DETAIL_CODE
1065	106985	0	*****	4	BD	27	IRD	INTER
3	60552	0	*****	2	AV	24	VF	PRIVA
1206	107764	0	*****	9	US	145	GA	GOVER

Let's see it in action!



Principles For Enterprise Data Management



**Improve End to
End Data
Value Chain**



**Make Data
Easier
to Acquire**



**Quicken
Determination
of Data
Trustworthiness**



**Focus Analysts
on Creating New
Data Assets**

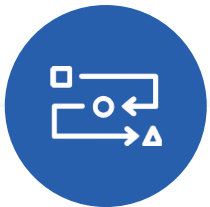
Why Customers Buy Alation



Intelligence



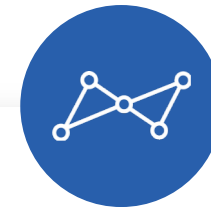
Collaboration



Guided Navigation



Active Governance



Broad Connectivity

Leading Customers From Many Industries & Use Cases

Analytics

FINNAIR

Pfizer



Zurich

Governance



Cargill



ancestry

Cloud



parexel

MassMutual

Risk & Compliance



DAIMLER



Wells Fargo

Data Privacy/GDPR



Riot Games



Digital Transformation



Genentech

Technip

About Alation

“Alation started the ML Data Catalog trend.”

-Forrester Wave, Machine Learning Data Catalogs, 2018

Gartner

**4x Gartner
MQ Leader**

Magic Quadrant
for Metadata
Management
2017–2020

FORRESTER

**2x Forrester
Wave Leader**

Machine
Learning Data
Catalogs 2018,
2020



**5x Dresner
#1 Data Catalog**

Dresner Advisory
Services Data
Catalog Market
Study 2017–2021



**Technology
Pioneer**

2016 World
Economic
Forum

Pioneer of the
data catalog market
since 2012

\$110M

in funding by

Blackstone
Costanoa
Data Collective
Dell Technologies
Icon
ISAI Cap
Riverwood
Salesforce
Sanabil
Sapphire
Snowflake Ventures

HQ in Redwood City, CA



Thank you!

Questions?