ATSCALE Technical Overview

Market

ta Availa

TABLE OF CONTENTS

Why AtScale?	1
Where does AtScale Fit in the Analytics Stack?	2
AtScale System Component Overview	3
Consumption Integration	4
Semantic Modeling	5
Data Virtualization	6
Performance Optimization	7
Analytics Governance	8
Data Integration	9
AtScale Software Component Overview	10
AtScale Design Center	10
AtScale Query Engine	13
Deployment	22
Supported Business Intelligence Tools	24
Directory Services	25
Operating System Support	25
Frequently Asked Questions	26
About AtScale	27

Why AtScale?

AtScale offers a modern approach to business intelligence and analytics in the cloud. AtScale's semantic layer platform enables analysts to perform sub-second, multidimensional analysis with popular BI ands AI tools. Enterprises rely on AtScale to overcome data and analytics challenges including: accelerating data-driven decisions at scale, creating one compliant view of business metrics and definitions, controlling the complexity and costs of analytics and reducing the risk of analytics.

AtScale helps enterprises:

SEAMLESSLY MIGRATE TO THE CLOUD

Enterprises can avoid business disruption and port analytical workloads with minimal disruptions to end users.

SIMPLIFY THE ANALYTICS INFRASTRUCTURE

Enterprises can use the best tool and platform for the job without moving data or adding new data stores.

MODERNIZE AND FUTURE PROOF THE ANALYTICS STACK

Enterprises can take advantage of data lakes and cloud data warehouses while preparing for future platforms.

SECURE AND GOVERN DATA IN ONE PLACE

With a live, governed connection to all data in a virtual model, enterprises can eliminate data copies and the security risks associated with them.

TURBOCHARGE ANALYTICS AND MACHINE LEARNING INITIATIVES

Enterprises can Instantly integrate new data sources without manual data engineering because AtScale delivers a single, super-fast, businessfriendly semantic interface for all data.

SEE ALL DATA IN A SINGLE, UNIFIED VIEW

No matter where it is stored or how it is formatted.

CONDUCT INTERACTIVE AND MULTIDIMENSIONAL ANALYSES

Using business users' preferred BI tools, whether that is Excel, Power BI, Tableau, or something else.

GET CONSISTENT ANSWERS ACROSS DEPARTMENTS AND BUSINESS UNITS

Via AtScale's Semantic Layer that standardizes queries regardless of BI tool or query language.

Where Does AtScale Fit in the Analytics Stack?

The AtScale semantic layer sits between your analytics consumption tools and your data platforms. By abstracting away the physical form and location of data, the AtScale semantic layer makes data stored in data lakes or data warehouses accessible with the same interface. Integration with enterprise data catalogs makes AtScale models discoverable and metadata shared seamlessly.



AtScale System Component Overview

AtScale provides a single, secured and governed workspace for distributed data. The AtScale semantic layer platform behaves like a logical data warehouse. The AtScale service intercepts client queries, translates logical queries into physical queries and passes those queries onto the underlying physical data warehouse or data lake for execution. As end users interact with the data in the AtScale model, AtScale automatically creates or modifies aggregate tables to optimize performance and manage costs. AtScale will create aggregates (think materialized views) on the source data platform and determine the optimal location to store those aggregates in a federated query scenario. AtScale's automated tuning functionality works consistently regardless of the underlying data platform (data warehouse or data lake).

The combination of AtScale's semantic model, data virtualization, performance optimization and analytics governance powers business intelligence (BI), artificial intelligence (AI) and machine learning (ML) initiatives resulting in faster, more accurate business decisions at scale.



In the following sections, we will describe the inner workings and benefits of each component.

Consumption Integration

AtScale speaks the languages of your analytics applications whether they are business intelligence (BI) tools, AI/ML platforms or custom applications. AtScale requires no custom client-side software installs so anyone using Excel, Power BI, Looker or Tableau can connect to AtScale and run queries immediately.

Benefits include:

Multimodal

- Supports tabular and multidimensional clients
- Supports BI and AI/ML workloads
- Application and integration friendly via REST, JDBC, ODBC, MDX and DAX

Lightweight Footprint

- No custom client-side
 drivers needed
- Wire compatible with Microsoft SSAS protocol (Excel, Power BI)
- Live connection to data without extracts or data imports

Client Optimized

- Optimized SQL interface for Tableau
- Native DAX support for Power BI and live pivot tables for Excel
- LookML integration for Looker

Semantic Modeling

The key to the AtScale semantic layer is the AtScale model. The best way to get everyone on the same page is to have everyone speaking the same language. This ensures that there won't be conflicting answers to the same questions. A single, centralized workspace for business metrics and definitions is key to offering one consistent, compliant view of data to both business users and data scientists alike.

AtScale's semantic model unifies semantic definitions and metrics for data and makes it available in one location for BI, AI/ML and applications. It works on data anywhere whether it's in a data lake or a data warehouse.

Benefits include:

Object Oriented

- Reuseable cubelets for conformance
- Library to promote sharing across models
- Multi-pass calculation for complex expressions

Multi-source

- Blend data from multiple sources in a model
- Smart push down for optimal performance
- Data platform optimized

Open

- XML-based markup language
- Programmable APIs for model creation
- Native SQL for full platform functionality

Query Virtualization

AtScale's data virtualization automates the sourcing, curation and modeling of data on premises or in the cloud. It blends live data from multiple data sources into virtual, logical views. Virtualization makes IT more agile with the ability to store data in the most suitable platforms while providing the flexibility to adopt new platforms in the future without rearchitecting their stack or disrupting their downstream data consumers.

AtScale's Data virtualization provides access to enterprise data by functioning as an abstraction layer on top of a variety of data platforms but without manually moving data.

Benefits include:

Built for BI

- Supports tabular and
 multi-dimensional modeling
- Supports SSAS compatible
 MDX and DAX (Excel)
- Time Intelligence, Hierarchies, Cell-level calcs

Graph-based Planner

- Node-based query planning and optimization
- Dialect and data platform
 optimized
- Scalable to thousands of dimensions

Integration First

- Query delegation and directory support
- Integration with enterprise
 data catalogs
- Dimensional sharing model (Conformed Dims)

Performance Optimization

Gathering live data from multiple sources across the organization can be a long, manual process. Data engineers should be creating new value for the business rather than simply preparing and moving data for business reporting.

AtScale's autonomous performance optimization technology identifies query patterns and creates and manages intelligent aggregates, just like the data engineering team would do. The Al-driven optimizer learns from user behavior and data relationships and takes care of data updates and changes, so business users can focus on gathering insights from data and data engineers can focus on other projects. With AtScale, the moment a model is published, data access is "live". AtScale builds aggregates in real-time in response to user activity and automatically tunes queries without additional manual intervention.

Benefits include:

Machine Learning Powered

- Uses query behavior patterns
- Leverages HLL++ to computer statistics
- Proactively optimizes at model time

Adaptive, Always Working

- Automatically prunes, splits, merges aggregates
- Automatically adapts to data and model changes
- Supports full and
 incremental updates

Platform Optimized

- Optimizes locale via
 preferred agg storage (PAS)
- Platform specific
 partitioning and ordering
- HLL++ libraries adapted for each platform

Analytics Governance

AtScale's patented security capabilities respect native data platforms security by supporting end-to-end user delegation and impersonation. AtScale's object-level security supports user and group access rules while providing discoverability for a 360-degree feedback experience with model designers. With integrations with enterprise data catalog and governance tools, AtScale can enforce data governance rules using AtScale's virtualized governance layer.

Benefits include:

Environmentally Smart

- Enterprise Directory
 Integration (AD/Octa/OAuth)
- Secure Cluster Support (Kerberos, Ranger, files)
- Secure Data Transport
 Protocols (TLS)

Just in (query) Time

- Configurable delegation
 query execution
- Impersonation throughout
 (including client)
- Query time row level
 masking for user/groups

Object Oriented

- Role-based, object level control for users/groups
- Column masking in query tools (Perpectives)
- Multi-tenancy support via organization level

Data Integration

First generation data virtualization was not designed for the large analytical workloads that are typical of today's BI and AI use cases. AtScale's deep expertise in multidimensional analytics along with federated query processing provides unparalleled support for BI and AI tools alike. AtScale speaks to data lakes and data warehouses with data platform optimized SQL so performance is as fast or faster than hand-written queries. Rather than processing data locally, the AtScale engine pushes queries down to the underlying data platform to eliminate data movement and scale performance along with the data platform, without the need of managing a separate analytics infrastructure.

Benefits include:

Environmentally Smart

- Enterprise Directory
 Integration (AD/Octa/OAuth)
- Secure Cluster Support (Kerberos, Ranger, files)
- Secure Data Transport
 Protocols (TLS)

Just in (query) Time

- Configurable delegation
 query execution
- Impersonation throughout (including client)
- Query time row level masking for user/groups

Object Oriented

- Role-based, object level control for users/groups
- Column masking in query tools (Perpectives)
- Multi-tenancy support via organization level

AtScale Software Component Overview

AtScale Design Center

AtScale Design Center is a browser-based application that subject matter experts (SMEs) use to create and publish AtScale models. AtScale Design Center is organized around the following object model:

ORGANIZATIONS: The 'top-level' structure contains sets of users, groups, roles, permissions, and more. AtScale instances will have at least one Organization. Additional Organizations can be created and they are mutually exclusive. Nothing is shared between Organizations.

PROJECTS: A Project is a collection of one or more AtScale models. Each project contains a shared Library with objects like Datasets and Dimensions that can be shared with other models in that project. There is no limit to how many Projects you can create or how many objects can be shared.

MODELS: A Model is created within a Project and can use shared objects from that Project's Library. A Model is a collection of Datasets, Dimensions, Measures, Hierarchies, and Calculations along with their Relationships that form the basis of a virtual, multidimensional view of your source data.

Models in AtScale visually look like a 'star' or 'snowflake' schema. However, there is no requirement for any particular physical data structure or layout: data can be normalized or denormalized or a little bit of both.

In the following screenshots, you can see how AtScale Design Center defines models.

	QUERIES	AGGREGAT	TES SETTING	S SEC	CURITY SUPPORT	ORCHESTRATOR				🛕 🥹 Admin 🗸
🥵 Sales Insights - Sr	owflake CUBE CANVA	S CUBE MA								
										Preview X 🔚 Dimensions X 🔐 Measures X
	Color Dimension	# ≡	factinternetsales	* =	■ Order Dimension	Bate Dimension	*=			
	A Color		oraeroece	o Storg	Order Dimension Domence	Date Month Hierarchy Vear				🕫 Internet Sales Cube 🗸
		String	DIMENSION LEVELS (4)		currencylley Decimal(38,0)	year H Duarter	DateTime			
			size		salesordernumber String currencykey Decimal(38,0)	quarter	DateTime		*	
	I Style Dimension	₩≡			Order Line Item salesorderlinenumber Decimal(38,0)	month	DateTime			• E Dimensions
	style ⇔ Style		weight MEASURES (#)	2010	currencykey Decimal(38,0)	datekey	Decimal(38,0)		Ä	Customer Attributes
	style	String	productkey	(Dec)38.0)	Eustomer Olmension	ats Date Week Hierarchy			œ	Date Attributes
	1 Size Dimension	\$∦ ≡	orderdatekey	Decist.0)	A Customer Hierarchy	year L- Wook	DateTime			Drodust Attributer
	ch Size		orderquantity	there(SR,0)	Customer Name customerkey Decimal(38,0)	week	DateTime		-	 Initiation and an and a second second
	Size size	String	unitprice	.trouma		datekay	Dectmal(38,0)			Customer Metrics
		ak -	salesamount	1200-000	Product Dimension	■ Neporting Year				Product Metrics
	A Weight	375 E	- & Calculated Tax	House	Product Line	Reporting_year	Detelime			Sales Metrics
	⊷ Weight weight	String	RELATIONSHIPS (11)	/ /	Product Category	3 reporting_holf_yea ⊶ Reporting Quarter	r DateTime			🔸 🚞 Time Relative
			salesorderlinenumber	1	productline Strin productsubcategoryke@ecimal(38.0	reporting_quarter	DateTime	Q		
			preductiley	Dec(SEO)	Product Name productkey Decimal(38,0) reparting_month	DateTime			
			orderdstekey	Dec(18.0)		reporting_week	DateTime			
						datekey	Decimal(\$8,0)			
						Custom Year	DataTime			
						Gustom Quarter	Deservice			
						reparting_quarter	DateTime			
	WARNINGS 0									<

The above screenshot shows the main AtScale Design Center canvas. In the center, you can see blue titled panels that represent datasets (or fact tables), green titled panels that represent normal dimensions and gray titled panels that represent degenerate dimensions (dimensions based on a fact table). The orange arrows represent relationships between the respective objects (datasets and dimensions). In the right hand panel, you can see a preview of the dimensions, measures and hierarchies (the logical model) that the end query consumer will see.

$\mathbf{\Lambda}$	PROJECTS	Q	UERIES	AGGREGATES	SETTINGS	SECURIT	TY SUPP	PORT OI	RCHESTRATOR					Ļ	😫 Admin 🕶
💦 Sal	es Insights - Snowl	flake	CUBE CANVAS												
😨 Inte		🚅 factin	ternetsales \times												
✓ Pac	e 1	f10 ~	Row L	imit: 1000	Apply								Colum	1.Order: Natural	
(antonionto)												- and sedan to ser		in the second	S calactator and
satesorde	nur ••• satesorde	TUDE IN	sates_reasonr •	CONVERT DATA TYPE	broduct_into or	derdatekey ••••	customerkey	orderquantity •••	unitprice (Dot •••	salesamoune	taxamt (Dour •••	Corderdate (+++	shipdatekey (i •••	currencykey (***	• Catectater ····
SO4369			5,9	Boolean	size:62,weight:15	20050701	21768		3578.270019531	28626.16	2290.1	2005-07-01 00:0	20050708		2433.2236000000
SO4369				Date	size:44,weight:21	20050701	28389		3399.9899902343	3399.99	272.0	*2005-07-01 00:0	20050708		288.99915
SO4369				DataTima	size:44,weight:21	20050701	25863		3399.9899902343	3399.99		*2005-07-01 00:0	20050708		288.99915
SO4370				Daterine	color:Black,size:6	20050701			699.09802246093			*2005-07-01 00:0			178.26965
SO4370				Decimal	size:44,weight:21	20050701	11003		3399.989990234	3399.99		*2005-07-01 00:0	20050708		288.99915
SO4370				Float	color:Red,size:44	20050702	27645		3578.270019531	3578.27	286.26	2005-07-02-00:0	20050709		304.1529500000(
, SO4370				Integer	size:62,weight:15	20050702	16624		3578.270019531	3578.27	286.26	*2005-07-02 00:0	20050709		304.15295000000
S04370					color:Black,size:4	20050702	11005		3374.989990234	6749.98		*2005-07-02 00:0	20050709		573.7483
504370			NULL	Upper Case	color:Silver.size:3	20050702	11011		3399.9899902343	3399.99	272.0	*2005-07-02 00:0	20050709		288.99915
\$04370				Lower Case	color-Red cize-48	20050703	27621		3578 270019531	3578 27	285.25	*2005-07-03 00-0	20050710		304 1529500000
304370					CotoriRed,size.46	20030103			3310.270019331		03:003	2003-07-03 00.0	20030710		304.1329300000
504370				Extract	color:Red,Size:48	20050703	27616		5578.270019531.	3578.27	286.26	2005-07-03 00:0	20050710	100	304.15295000000
S04370			NULL	Replace	size:52,weight:19	20050703	20042		699.09802246093	699.1	55.93	2005-07-03 00:0	20050710	98	59.42350000000K
S04370					color:Red,size:52	20050703			3578.270019531	3578.27	286.26	2005-07-03 00:0	20050710		304.15295000000
S04371					size:56,weight:14	20050703						2005-07-03 00:0			304.15295000000
S04371					size:56,weight:14	20050704						*2005-07-04 00:0			608.30590000000
504371					color:Red,size:44	20050704	13513		3578.270019531	10734.81	858.79	2005-07-04 00:0	20050711		912.45885
S04371					size:62,weight:15	20050705	27601		3578.270019531	7156.54		2005-07-05 00:0	20050712		608.30590000000
S04371					color:Red,size:44	20050705	13591		3578.2700195312	7156.54		*2005-07-05 00:0	20050712		608.30590000000
ERI															<

In the above screenshot, you can see the AtScale Design Center data wrangling view for the "factinternetsales" Snowflake table. You can apply transformations and create new calculations for the semantic model.

$\mathbf{\Lambda}$	PROJECTS QUE	RIES AGGRE	GATES SETTINGS	SECURITY	SUPPORT	ORCHESTRATOR			¢	e Adr	min 🔻
👧 S	ales Insights - Snowflake C	UBE CANVAS CUBE									
😥 Int			er Dimen 🗙 🔚 Geography Dime.								
								0	👬 Hierarchies 🗙		
	📰 Geography Dimension	* =	dimcustomer	* =	E Gender Dimension	* =			Filter		
	🔥 Geography City			0	👬 Gender Hierarchy				🏦 Customer Hierarchy		
	Country country	String	UNUSED (1)		Gender	String			📁 🔢 🗁 Customer Name		
	⇒ State		datefirstpurchase		gunuenney			m	📄 💿 Occupation		***
	statekey	String	DIMENSION LEVELS (2)	-	[📄 🧿 Last Name		
		String	customerkey	Dec(55,0)					😑 💿 First Name		
	city	String	fullname	String							
	GeoKey geographykey	Decimal(38,0) 🔨	RELATIONSHIPS (2)	•							
>	🚓 Geography Zip		geographykey	Dec(38,0)							
ALC: NO	⊷ Country		gender	String 🖉							
	country	String	SECONDARY ATTRIBUTES (3)								
		String	firstname								
	postatcode	string	lastname								
	geographykey	Decimal(38,0)	occupation	String							
							<u>u</u>				
8	RRORS 0 WARNINGS	0									<

This image shows a "Customer" dimension built from a base table called "dimcustomer" and connected to two other dimensions, "Geography" and Gender". By building conformed dimensions, new models are both easy to create and consistency is insured across models.

$\mathbf{\Lambda}$	PROJECTS	QUERIES	AGGR	EGATES	SETTINGS	SECURITY	SUPPORT	ORCHESTRATOR		♠	e Admin -
 ✓ Back Search 				Internet Version 1 Pub	t Sales Cub	C PUBLISHED					
DRAFT	Sales Insights - S	inowflake									
PUBLIS											
- &	Sales Insights - S	nowflake									
i	🗑 Internet Sales (le.com:11111/Sales%20insights%20·%2	20Snowflake		
						http://atscale-node-01		n:10502/xmla/default			
				ODBC & SQ							
				Connect a E	31 Tool						
				+	H H H H H H H H	E	KCEL	POWER BI			

Once a model is published in AtScale Design Center, it is ready for consumption by BI users and data scientists. In this image, you can view the connection instructions for Tableau, Excel, Power BI and connection strings for AtScale's JDBC, ODBC and MDX interfaces for custom applications.

AtScale Query Engine

The AtScale Query Engine acts as a query interface for business intelligence, AI/ML tools and applications. Tools can connect to AtScale via one of the following protocols:

- 1. ODBC/JDBC (SQL)
- 2. XMLA (MDX or DAX)
- 3. Python
- 4. REST

For the tools that speak SQL, the AtScale engine appears as a Hive SQL warehouse. For the tools that speak MDX or DAX, AtScale appears as a SQL Server Analysis Services (SSAS) cube. For applications speaking REST or Python, AtScale appears as a web service. The following screenshots show how the AtScale Query Engine appears in a variety of consumption tools.

AtScale's semantic layer provides the same logical view of business-friendly data regardless of the BI and/or AI/ML tool. Users can interact with data using the same dimensions, hierarchies, and measures defined in the Design Center. With AtScale, data is delivered as a service to all data consumers without any restrictions to share and collaborate.



In this image, you can see how the AtScale semantic layer for the "Internet Sales Cube" appears to consumers in Tableau. This image shows the results of "Order Quantity by Product Name" query.



In this image, you can see how the Power BI users can leverage the built-in SQL Server Analysis Services (SSAS) connector. This means that there is no need to install custom client-side drivers to access AtScale models.

🖫 🤌 🖓 Untitled - Power Bl Desktop	₽ Search		Sign in 🔵 🗕 🗇 🗙
File Home Insert Modeling View	Help		
A Cut Paste Copy ✓ Format painter Clipboard Cut Clipboard Cut Cut Cut Cut Cut Cut Cut Cut	Enter Dataverse Recent data sources v Queries	lew Quick Ser asure measure Calculations Se	Sitivity Publish ristivity Share
Auto recovery contains some recovered files that have	n't been opened. View	recovered files	< Visualizations > <
E SQL Server Server Server ○ Inttp://atscale-nod Database (optional) ○ Import ● Connect live ▶ MDX or DAX quer	Analysis Services database e-01.docker.infra.atscale.com:10502/xmla/default	Cancel	Fields Fields
Page 1 of 1			

In this image, you can see how Power BI connects in "Live" query mode to the AtScale Query Engine.



In this image, you can see how the AtScale semantic layer for the "Internet Sales Cube" is inherited automatically in Power BI, requiring no additional models for analytics consumers.



In this image, you can see how the AtScale semantic layer for the "Internet Sales Cube" appears to consumers in Power BI. This image shows the results of "Order Quantity by Product Name" query.

. 5-		Book1 - Excel			D	avid Mariani	<u> </u>	m –	a x	ć
File H	ome Insert	Page Layout Formulas Data Review View Add-ins Help	o Team	Q -	Tell me what	t you want to	do		🖻 Share	
Get External	New Do	w Queries Connections 2↓ ZA Clear Data Connection Wizard ?	×	What	at-If Forecas	현를 Group 현를 Ungro st 때 Subtr	p ~ + oup ~ -			
Data *	Query ~ Lo Re	Connect to Database Server		Analy	/sis ∽ Sheet		Judi	-		
	Get & Ira	Enter the information required to connect to the database server.	10015		Forecast	Out	line	121		^
A1	• E 🗙		_							٣
A A	В	1. Server name: de-01.docker.infra.atscale.com:10502/xmla/default	к	L	M	N	0	р	Q	-
2	_	2. Log on credentials	-							
4		Use Windows Authentication								
5		O Use the following User Name and Password								
6		User Name:								
7		Password:								
8										
9										
10										
11		Cancel < <u>B</u> ack <u>N</u> ext > Finish	e							
12			-							
13										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										Ŧ
	Sheet1	(+)	: (4)						•	
Ready							<u> </u>			0%

In this image, you can see how the Excel users can leverage the built-in SQL Server Analysis Services (SSAS) connector. This means that there is no need to install custom client-side drivers to access AtScale models. Also, notice that we are connecting using Window Authentication (Active Directory).

	⊟ 5× ਟੋ - ∓				В	ook1 - Excel				Davi	id Mariani 🛛	i 1	ক্র —	o x
F	ile Home Insert Page Layo	ut Formulas	Data	Review	View	Add-ins	Help	Team	PivotTable	e Analyze	Design	Q	Tell me	🖻 Share
Piv	Active Field: Product Line Field Settings Down Active Field	+∃ Group ↓	Insert Slice Insert Tim Filter Coni Filter	er eline nections	Refresh C	Change Data Source ~ Data	Actions	Image: Fields, Image:	ltems, & Set Tools ~ onships Iculations	is * Pivo	tChart Recor Pive Tools	inmendec otTables	Show	~
А	13 • : × ✓ fx	S												¥
1	A	В	с	D	E	F	G	н	1					
1	Row Labels 🔹	Order Quantity								PivotT	able Fie	elds		* X
2	BM									Chaorafie	de to add to	reports		- 25 v
3	⊞ M -1	7861.								choosene		report		- H
4	■ M -22									Search				Q
5	Women's Mountain Shorts, L	589.												
6	Women's Mountain Shorts, M	542.								4	✓ Product I	lierarchy	1	
1	women's Mountain Shorts, S	520.									Product Line			
8	■ IVI -27	404.									Product Cate	gory		
10	■ M -20	3250.									Product S	ubcatego	ID VID	
11	= M -37	10280										ubcolcyc		
12	⊞ B	24459								🔺 📃 Shij	Date Dimen	sion		
13	±s	36740								Þ 💼	Date Attribute	es		
14	®T	7254								A Size	Dimension			*
15	Grand Total	95167.												
16										Drag field	s between are	eas below	r:	
17										W. 1994		1.0	III Coloma	
18										1 Filters	.	-	Columns	
19														
20														
21														
22										= Paur			N Values	
23										= NOWS	1.0			-
24										Product	Hierarchy		Urder Quantity	
25														
26														
27									-					
	Sheet1 +			ł	4				Þ	Defer	Layout Updat	e		Update
Rea	ady									B		四 -		- + 100%

In this image, you can see how the AtScale semantic layer for the "Internet Sales Cube" appears to consumers in Excel. This image shows the results of "Order Quantity by Product Name" query.

💭 Jupyter	Al-Li	nk Last Checkp	oint: 3 minutes a	ago (unsaved	changes)								2		Logo	ut
File Edit	View	Insert Cell	Kernel	Widgets	Help							Trusted	ø	Pyth	non 3	0
8 + % 2	6	↑ ↓ ► R	un 🔳 C 🕨	Code	~											
									7							
In [5]:	impor from	t pandas <mark>as</mark> atscale impo	pd rt AtScale													
In [6]:	proje	ct = AtScale serve usern organ proje model)	(r='http://c ame='johnly ization='bb ct_id='0681 _id='f99aaf	ovid19.ats nch', be3ab7-a28 a511-23c0- e3-66ac-4d	cale.cc 8-4cb4- 442a-43 f6-5096	-4c70- 370-d 5-130)-63b9(11dcf2))5cef9)	da33dc 841ac1 2aa1'	c0e', L',							
In [8]:	df = df	ord:	data([' <mark>stat</mark>	e', 'categ	ory', '	'tota	al_sal	.es'])								
Out[8]:	sta	te category	total_sales													
	7 (A FOODS	27248769.81													
	4 0	A HOBBIES	18339506.75													
	3 (A HOUSEHOLD	34823077.10													
	0	FOODS	21499035.25													
	5 1	TX HOBBIES	13762165.83													
	6	TX HOUSEHOLD	26337989.50													
	2 \	VI FOODS	20728269.17													
	8 \	VI HOBBIES	13926504.88													
	1 \	WI HOUSEHOLD	26230027.60													

In this image, you can see how the AtScale semantic layer appears to a data scientist using a notebook and AtScale's Python interface.

Deployment

AtScale micro-services install on a Linux server or virtual machine within any environment, either on-premise or in the cloud using a RPM. The AtScale instance serves both as a query endpoint for BI/AI tools and modeling endpoint for AtScale Design Center, a browser-based design environment for creating and managing models.

For the query federation option, AtScale also installs and manages a Spark service to manage heterogeneous joins across multiple data sources.



This image shows a graphical representation of the AtScale service components.

$\boldsymbol{\wedge}$	PROJECTS	QUERIES	AGGREGATES	SETTINGS	SECURITY	SUPPORT	ORCHESTRATOR	<u> </u>	8 Admin -
	Luster Mane	agement	LIVE MODE C						
								LIVE DETAILS	
VIR	TUALIZATION		111 - 77 - 77 - 77 - 77 - 77 - 77 - 77			2	INACTIVE	DETAILS HOSTNAME: otscole-node-01.docker/infra otscole.com	
Ĩ	SUPERVISOR atscale-node-01.dock	er.infra.atscale.com						✓ SERVICES ● agent: Running	
•	LISTENER atscale-node-01.dock	er.infra.atscale.com						coordinator: Running database: Running	
*	WORKER atscale-node-01.dock	er.infra.atscale.com						 directory: Running egress: Running 	
ø	DATA AGENT							engline: Runnling ingress: Runnling	
								modeler: Running orchestrator: Running	
								service_registry: Running servicecontrol: Running	

This image shows the AtScale Design Center Orchestrator console for managing AtScale components and services.

The AtScale platform consists of multiple services:

1. Agent

Installed on a virtualization fabric node to communicate with the virtualization service.

2. Balancer

Internal load balancers for routing traffic for High Availability (HA).

3. Coordinator

Installed at least 3 nodes for managing High Availability (HA).

4. Database

AtScale internal database for storing application and configuration data.

5. Directory

Internal LDAP directory used if an external directory is not defined.

6. Engine

AtScale query engine service.

7. Health

Health check service.

8. Ingress Bridge service for virtualization configuration services.

- 9. Modeler AtScale Designer Center service.
- 10. Servicecontrol AtScale services manager.
- **11. Virtualization_listener** Virtualization fabric listener service.
- 12. Virtualization_supervisor Virtualization fabric supervisor service.
- 13. Virtualization_worker Virtualization fabric worker.

Supported Business Intelligence Tools

Excel and Power BI contain the required drivers. SQL Tools such as Tableau must use one of the supported Hive drivers listed below. AtScale requires a Hive driver because it emulates the Hive SQL interface.

ΤοοΙ	Version(s)	Connection Type
Tableau Desktop and Server	2021.1.0, 2020.4.5, 2020.3.1, 2020.2.2, 2020.1.5, 2019.4.0	Thrift SQL
Looker	21	\$32.00
Excel	2016, 2013, 2010	XMLA (MDA)
Power Bl	NA	XMLA (DAX)

Unsupported Tools: The following BI tools have basic connection and query support, however they are not fully supported: Microstrategy, Business Objects, Cognos, Saiku, Spotfire.

Experimental Tools: AtScale has experimental support for Qlik. To use Qlik with AtScale, you must have a special AtScale license. Ask your sales representative about obtaining a Qlik-enabled license.

Directory Services

- Windows Active Directory 2012 (Supports Kerberos authentication protocol)
- Google Directory Service (NOTE: Does not support Kerberos, which is required for Tableau Single Sign-on)
- LDAP 2 compatible directory service providers

Operating System Support

- CentOS 6
- CentOS 7
- Ubuntu 16.04
- Ubuntu 18.04
- Ubuntu 20.04

Frequently Asked Questions

What do I need to deploy AtScale?

You need to configure AtScale to point to a supported data platform as listed in the Integrations section of this document. While not required, you will also want to configure AtScale to access your directory service (AD/LDAP) and your external load balancer for High Availability (HA) configurations. For AtScale installation, at least one Linux server or virtual machine is required with some basic prerequisites to install the AtScale software. For client tool access, you may need the appropriate JDBC/ODBC drivers if they aren't already installed. No additional driver is necessary for Excel, Power BI or tools that use the XMLA (MDX, DAX) protocol.

Is there a trial and/or open-source version of AtScale?

You need to configure AtScale to point to a supported data platform as listed in the AtScale supports a proof-of-concept trial. Please contact us to discuss your use case and/or project needs to determine if a proof-of-concept trial would be appropriate.

How does AtScale interact with my data platform?

AtScale acts as a client to your data platform(s) and will generate optimized, platform specific SQL based on the AtScale model defined in the AtScale Design Center.

Once a cube is published, it is immediately available for BI and/or AI/ML activity. There is no pre-processing or data movement required when publishing a model. Data consumers can connect to the AtScale engine via ODBC/JDBC (SQL), XMLA (MDX, DAX), REST or Python interfaces and begin querying the model.

AtScale intercepts inbound queries from end user's query tools and rewrites queries for execution on a data platform, leveraging any available AtScale managed aggregates that would be beneficial to the user's query.

Simultaneously, AtScale's machine learning algorithms are monitoring user activity and managing its aggregations to automatically optimize query performance. AtScale creates, manages and stores aggregate tables in a schema in the underlying data data platform(s).

What are the options for aggregate creation?

Aggregates may be triggered in 3 ways:

- 1. Demand-based Aggregates are generated heuristically based on user query behavior.
- 2. **Predictive Aggregates** are generated proactively based on model design. For example, dimensional aggregates may be generated to facilitate fast lookups for building reports.
- 3. User-defined Aggregates are defined by the AtScale Design Center modeler and are stored inside the AtScale model. Users can specify combinations of dimensions and measures to design an aggregate manually and these aggregates will automatically build when the model is published.

In addition to these types, settings are available for adjusting behavior and thresholds for creating demand and prediction based aggregates.

How are the acceleration structures managed and kept current?

There are three methods of controlling how and when the acceleration structures are refreshed.

- 1. Aggregates may be refreshed on a time or calendar basis using AtScale's built-in scheduler.
- 2. Aggregates may be refreshed on a file trigger basis by using AtScale's file watcher utility. This method is often used in conjunction with an ETL pipeline to trigger a refresh upon completion of an ETL flow.
- 3. Aggregates may be refreshed using AtScale's REST API. As with the file trigger option, this method is often used in conjunction with an ETL pipeline workflow.

Aggregates can be updated either incrementally or in full refresh mode. Incremental updates allow for the appending of new or changed data whereas a full refresh will rebuild the aggregates from scratch.

ABOUT ATSCALE

AtScale enables smarter decision-making by accelerating the flow of data-driven insights. The company's semantic layer platform simplifies, accelerates, and extends business intelligence and data science capabilities for enterprise customers across all industries. With AtScale, customers are empowered to democratize data, implement self-service BI and build a more agile analytics infrastructure for better, more impactful decision making. For more information, please visit www.atscale.com and follow us on LinkedIn, Twitter or Facebook.

atscale.com