

A photograph of three people sitting around a table, smiling and laughing while eating breakfast. A woman with short blonde hair is on the left, a man with glasses and a beard is in the center, and a woman with curly dark hair is on the right. They are all holding coffee cups.

Data for Breakfast



WELCOME!

Sign up for your
Snowflake free trial
(Enterprise edition)



<https://trial.snowflake.com>

Slides are available
online



Please place cell
phones on silent



AGENDA

- | | |
|----------------------|-----------------------------------|
| 9:00 - 9:25 | Snowflake the Cloud Data Platform |
| 9:25 - 10:05 | Technical Deep Dive |
| 10:05 - 10:25 | Unlock AI Without Limits |
| 10:25 - 10:45 | Q&A With Speakers |
| 11:00 - 1:00 | Zero to Snowflake Workshop |



SNOWFLAKE

THE CLOUD DATA PLATFORM



DATA...THE NEW FRONTIER



NEW TECHNOLOGY CHANGES HOW WE USE DATA

Diversification of Analytics



Analytics is growing in importance, everywhere, and for everyone

Explosion of Data



IoT, mobile, and social open up new opportunities for insight

Rise of the Cloud



Cloud gives us the ability to scale and centralize data

TAKING A PLATFORM APPROACH LEADS TO BIG BUSINESS IMPACT

Make Better, Quicker Business Decisions



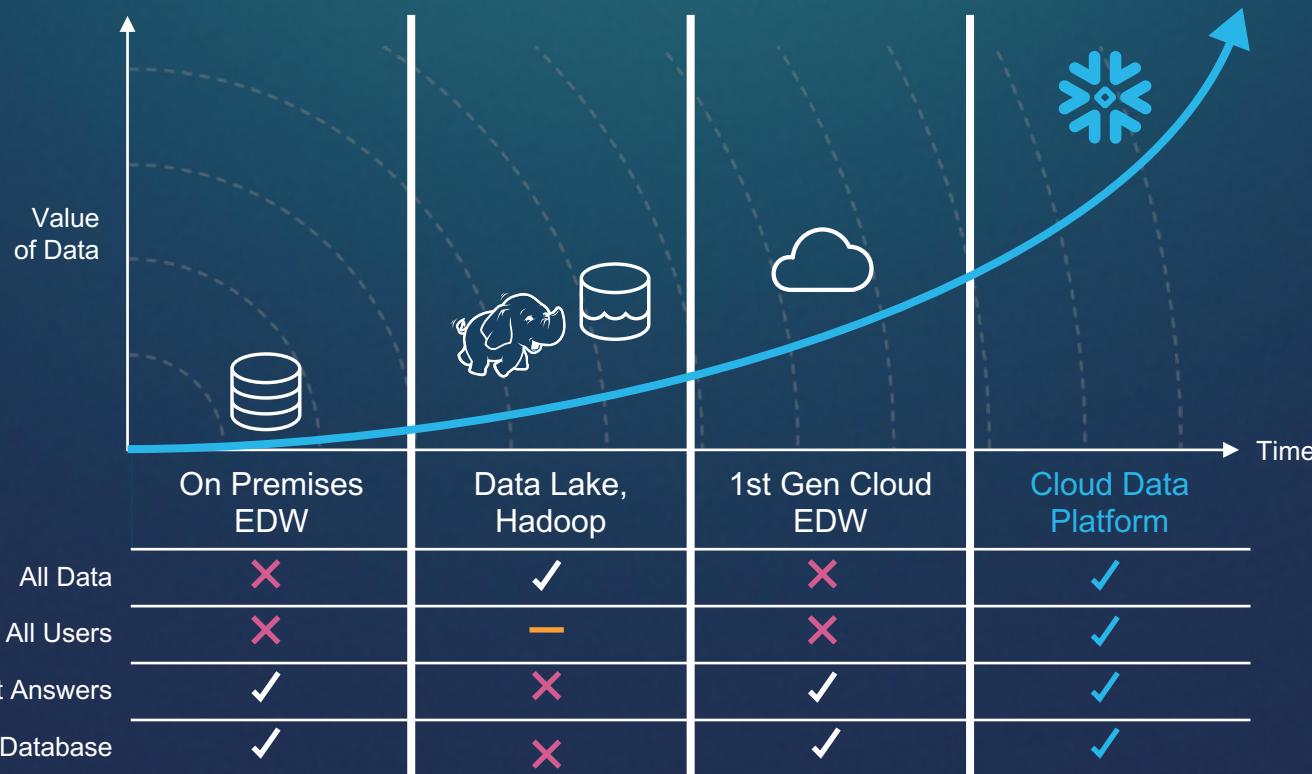
Reduce the Cost of Scaling Data Management and Analytics



Create a Great Customer Experience with Data

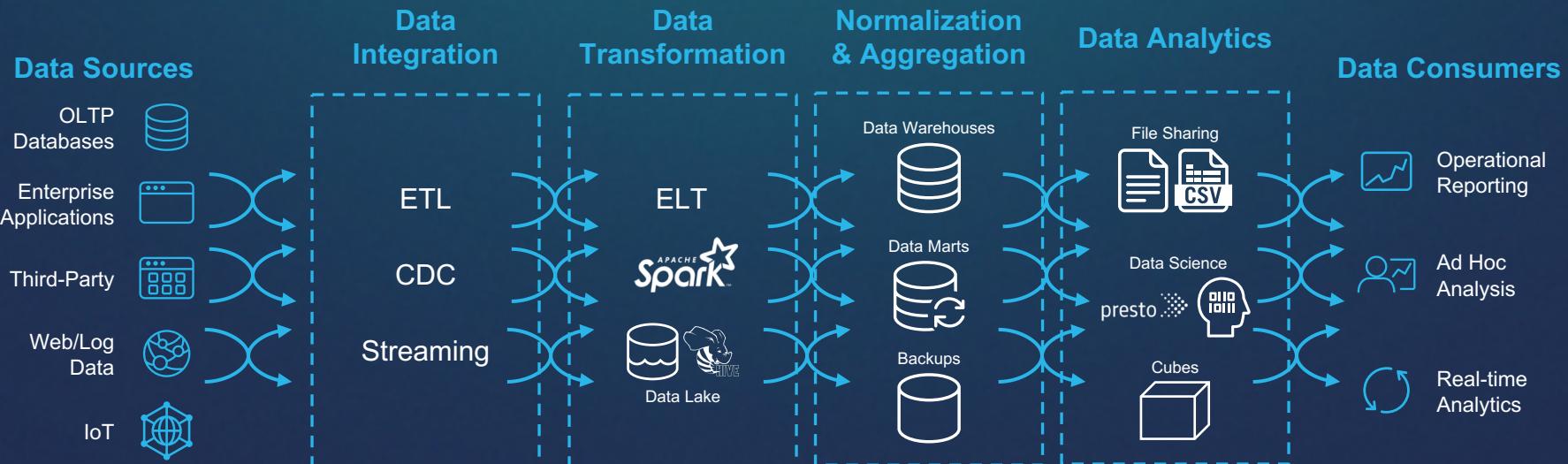


JOURNEY TO A CLOUD DATA PLATFORM



TRADITIONAL DATA ARCHITECTURE

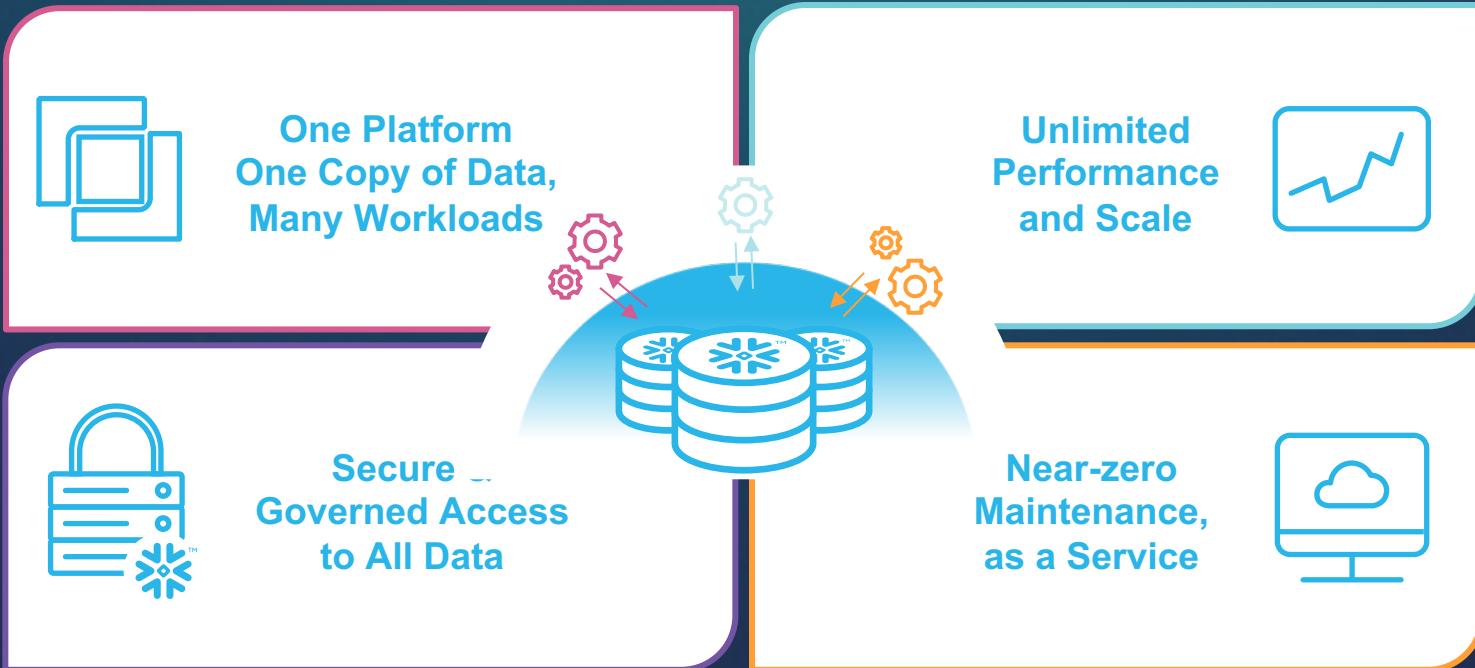
Complex and Costly with Multiple Copies of Data



MODERN DATA ARCHITECTURE WITH SNOWFLAKE



THE VALUE OF A CLOUD DATA PLATFORM



ONE PLATFORM, ONE COPY OF DATA, MANY WORKLOADS



Data Warehouse

Modernize data warehousing to deliver faster analytics at scale



Data Engineering

Rethink transformation with robust and integrated data pipelines



Data Lake

Simplify and accelerate your data lake with one platform for all your data



Data Applications

Develop apps with fast and scalable analytics that delight customers



Data Exchanges

Empower your ecosystem with secure, governed access to all data



Data Science

Simplify and accelerate machine learning and artificial intelligence



THE IMPACT OF A CLOUD DATA PLATFORM



SITUATION

- Shift to the cloud with focus on speed and value
- Managing infrastructure
- Significant governance, regulatory requirements

VALUE

- 60+ analytics teams served by single cloud data platform
- Migrated in < 90 days
- Load speeds improving 86%
- 5x faster complex queries
- Improved governance and democratized insights

SITUATION

- Multiple data warehouses, some end-of-life
- Data silos caused ambiguity and reporting disparities
- Poor consumer insight

VALUE

- Diverse data ingested for analytics, data science
- Improved agility, scalability and analytics performance
- In depth customer knowledge and improved services



SITUATION / PAIN

- Painfully slow analytics cycles
- Limited ability to answer complex questions
- Inability to provide business continuity and ensure security



SOLUTION / VALUE

- Hundreds of newly empowered analysts
- Improved scalability and faster query times
- Guaranteed security and data availability, 24/7/365





PROVEN BY OVER 3000 CUSTOMERS



Platform



BI/Analytics



ETL



EVER EXPANDING ECOSYSTEM

Data Science



Services



TECHNICAL DEEP DIVE



HOW IS SNOWFLAKE UNIQUE?



ARCHITECTURE



A CLOSER LOOK

Traditional Architectures



Shared-disk

Additional capacity requires forklift upgrade

Reads/Writes at the same time cripples the system

Replication requires additional hardware



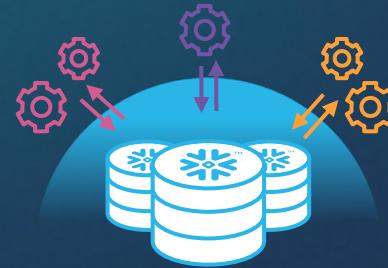
Shared-nothing

Resizing cluster requires redistributing data. Shut down requires unloading.

Each cluster requires its own copy of data (ex: test/dev, HA)

Vacuuming processes needed to maintain sort and distribution for performance

Snowflake



Multi-cluster, shared data

- Centralized, scale-out storage that expands and contracts automatically
- Independent compute clusters can read/write at the same time and resize instantly
- Automated backup across multiple availability zones/regions
- AWS, Azure, GCP



SNOWFLAKE ARCHITECTURE



MODERN DATA ARCHITECTURE WITH SNOWFLAKE CLOUD DATA PLATFORM



ONE PLATFORM, ONE COPY OF DATA, MANY WORKLOADS



Data Warehouse

Modernize data warehousing to deliver faster analytics at scale



Data Engineering

Rethink transformation with robust and integrated data pipelines



Data Lake

Simplify and accelerate your data lake with one platform for all your data



Data Applications

Develop apps with fast and scalable analytics that delight customers



Data Exchanges

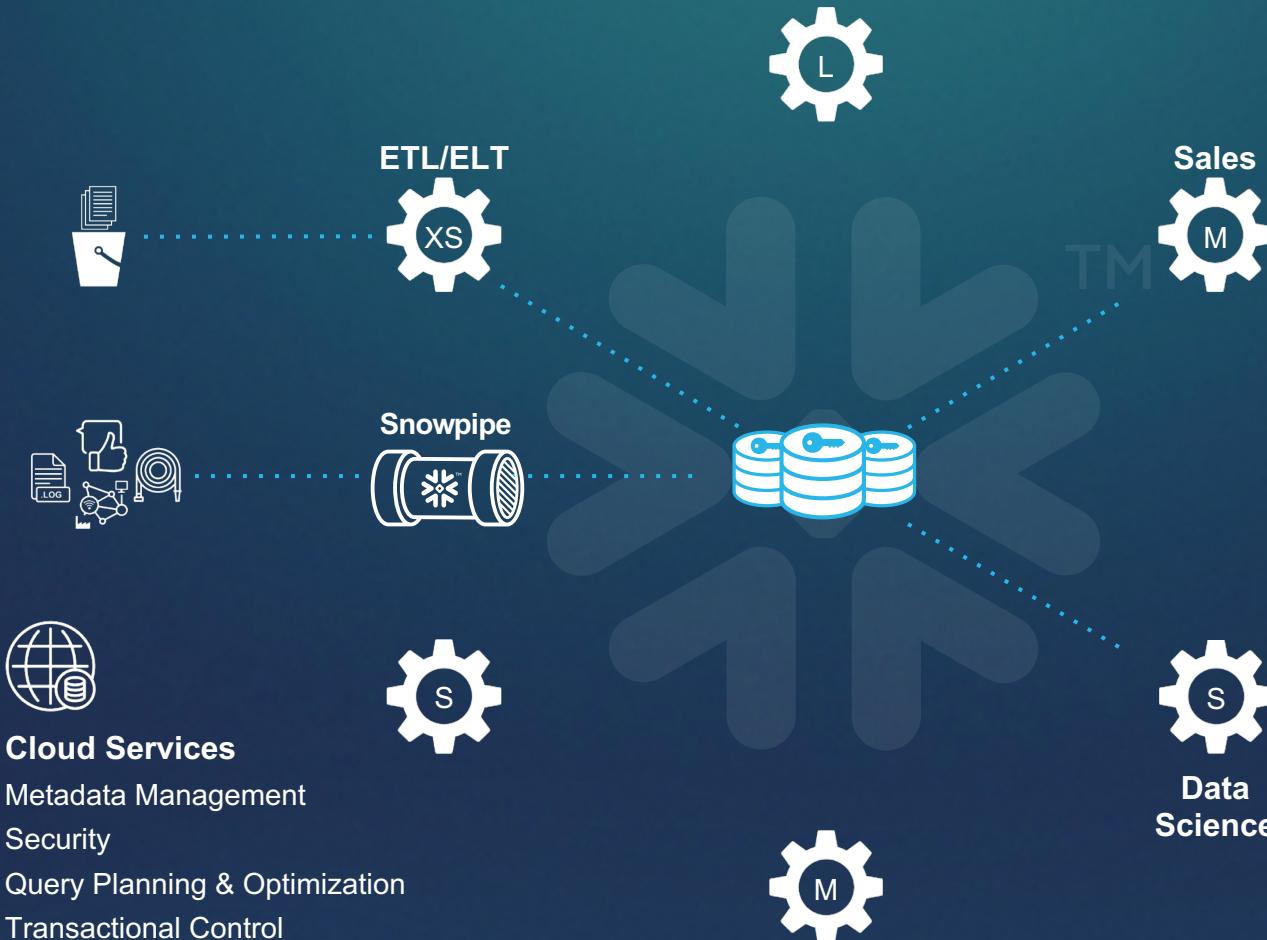
Empower your ecosystem with secure, governed access to all data

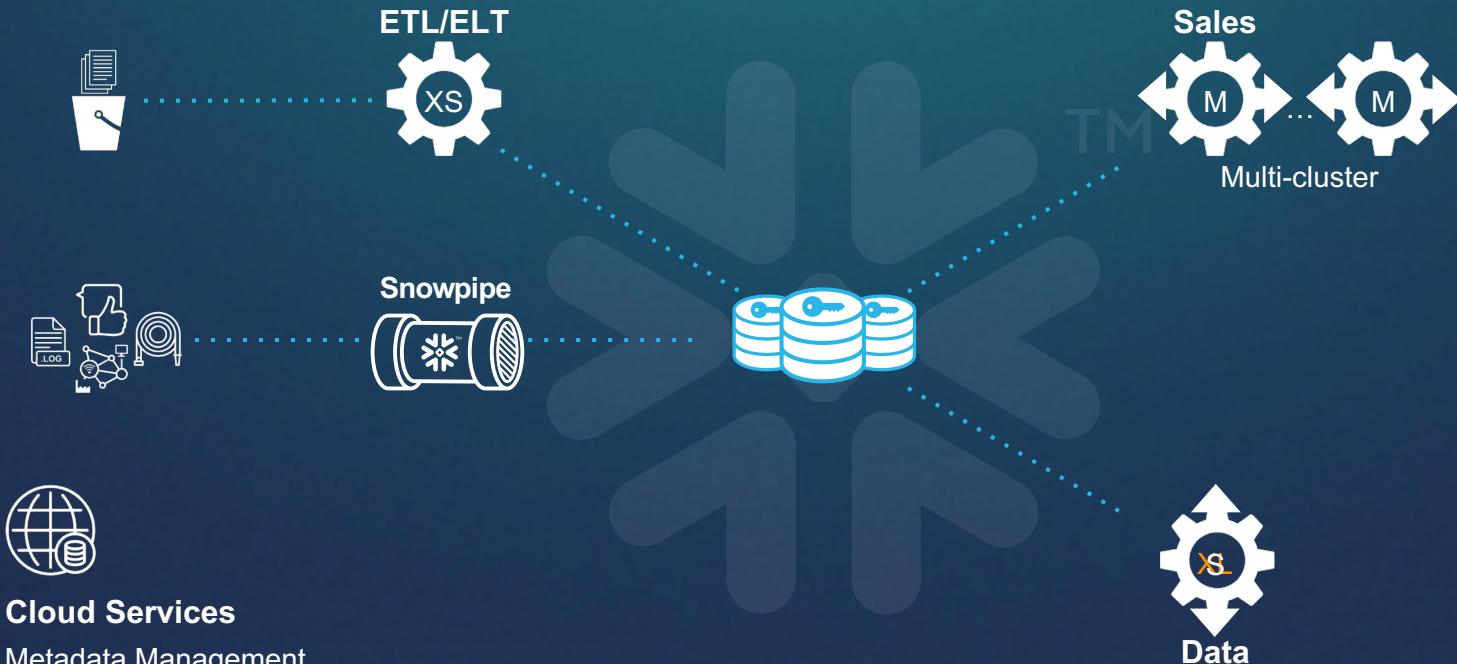


Data Science

Simplify and accelerate machine learning and artificial intelligence







Cloud Services

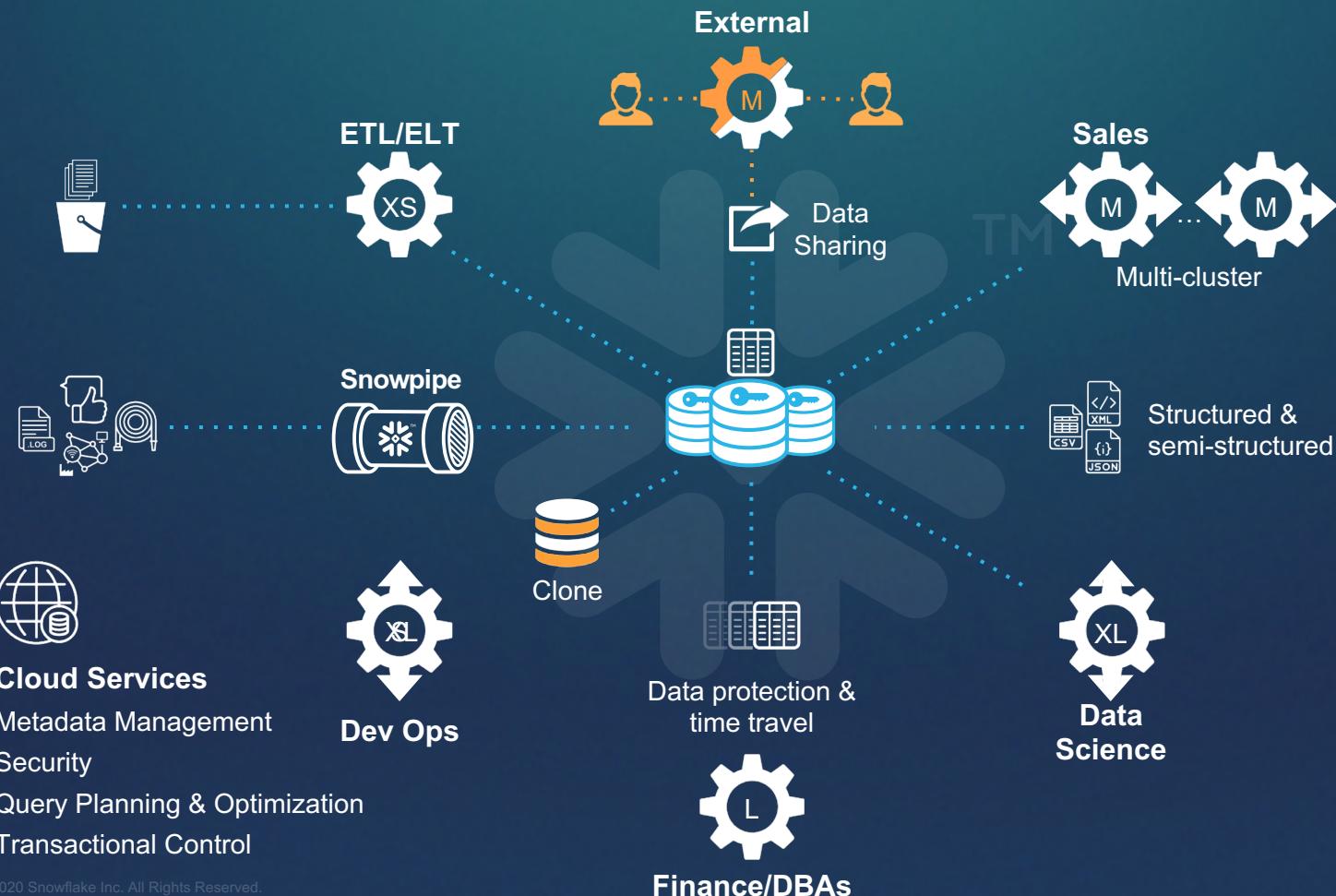
Metadata Management

Security

Query Planning & Optimization

Transactional Control





SECURE BY DESIGN, DATA AVAILABILITY

Authentication



- Embedded multi-factor authentication
- Key Pair authentication
- Federated authentication / SSO supported

Access Control



- IP whitelisting
- Roles-based access control model
- Granular privileges on all objects & actions

Data Encryption



- All data encrypted, always, end-to-end
- Encryption keys managed automatically

External Validation



- Certified against enterprise-class requirements
- PCI and HIPAA available

Data Availability



- User error: Time Travel, Failsafe, Cloning
- Zone failure: Data replicated to multiple zones in a region
- Region/Provider failure: Data replication & failover

MORE ON WORKLOADS



Data
Lake



Data
Exchange

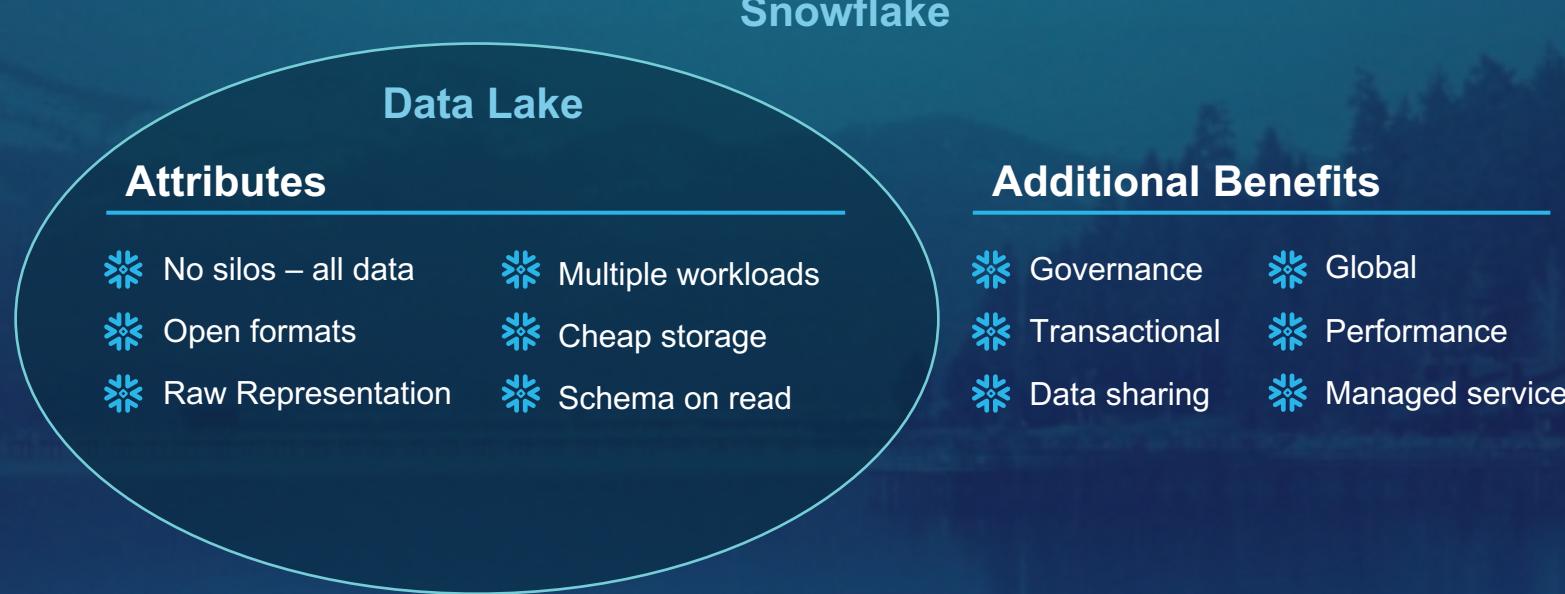


Data
Engineering

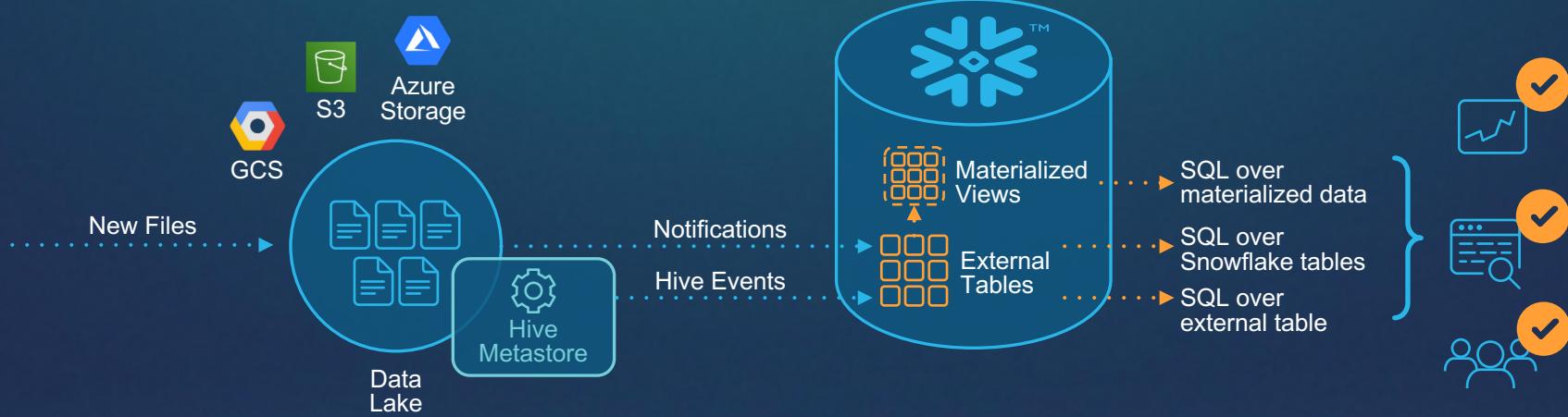


Data
Science

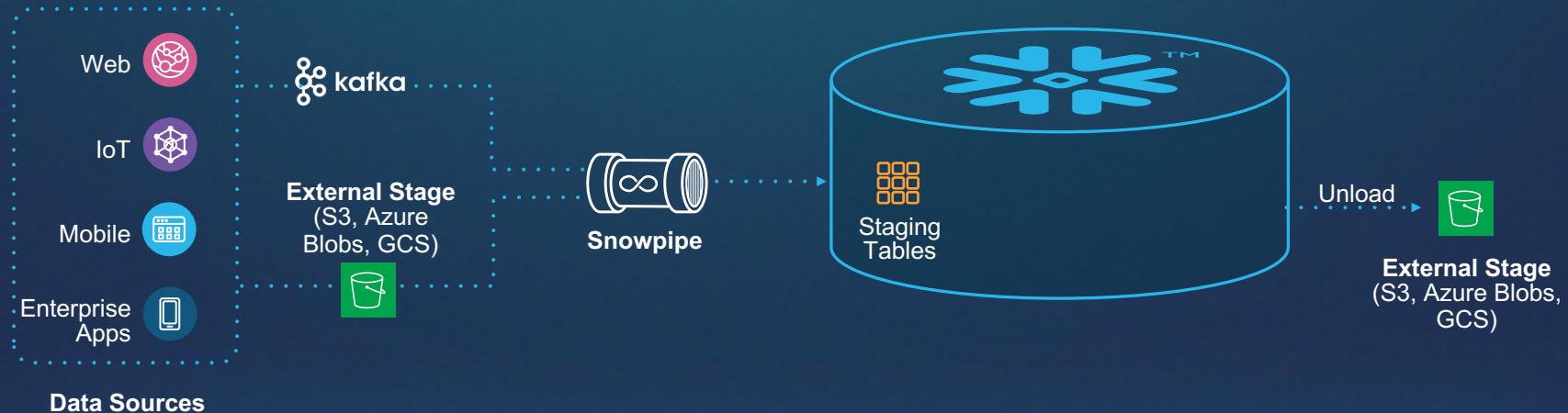
HOW DOES SCALABLE CLOUD DATA PLATFORM ENABLE DATA LAKES?



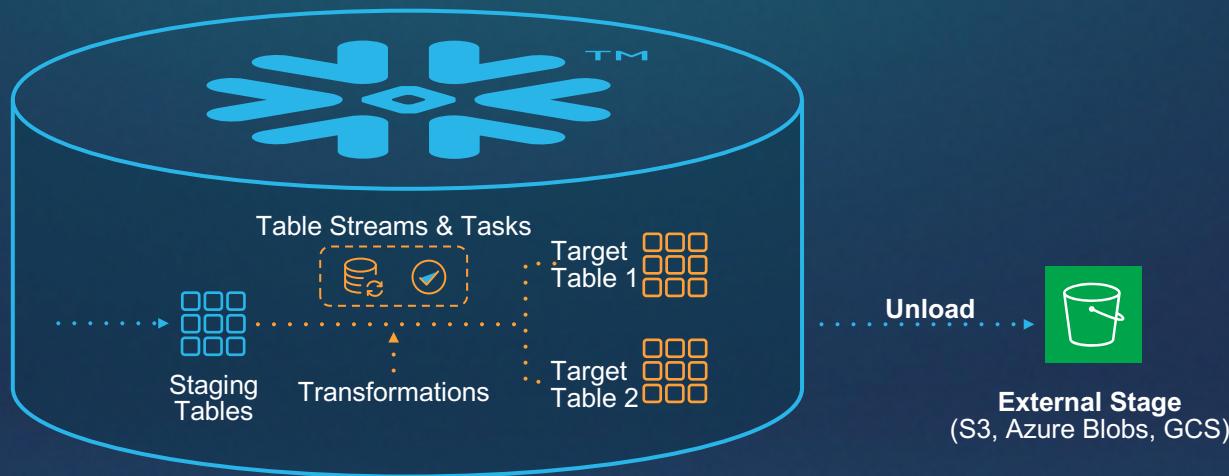
AUGMENTING EXISTING DATA LAKES



LOW LATENCY INGEST

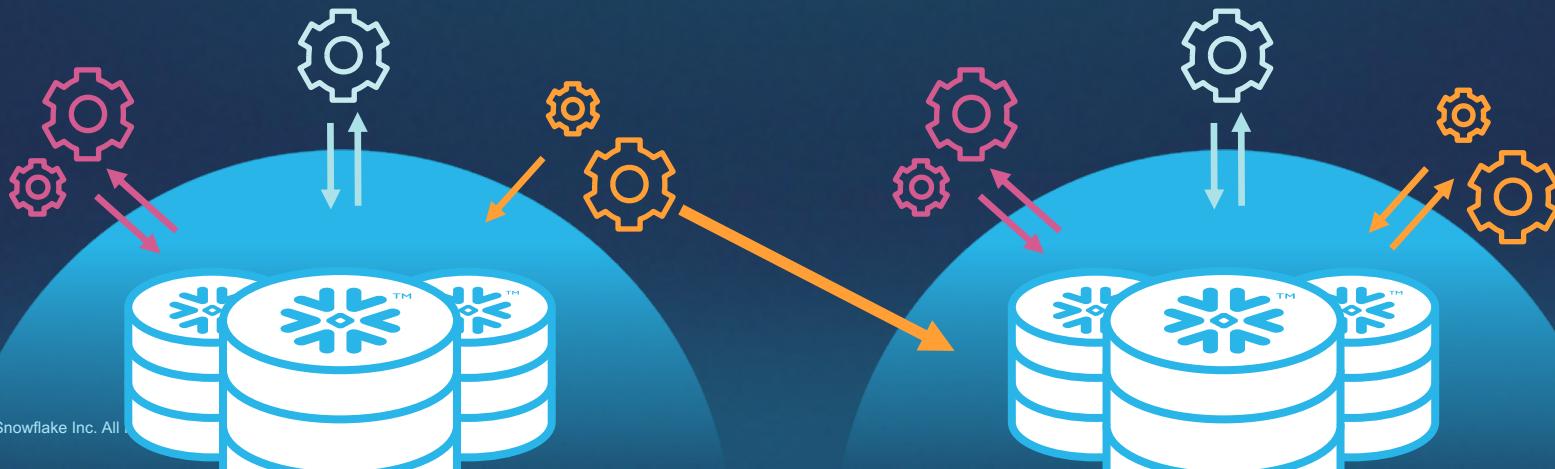


SCALABLE TRANSACTIONAL TRANSFORMATIONS



SECURE DATA SHARING

Secure Live Frictionless Personalized Global



ENABLING AI, ML, AND DATA SCIENCE

- Improving data science speed and efficiency with centralized source of high performance data
- Accelerating data exploration and preparation by 10-100x
- Connectors to leading and emerging technologies
- First class ecosystem of partners



Amazon SageMaker



UNLOCK AI WITHOUT LIMITS

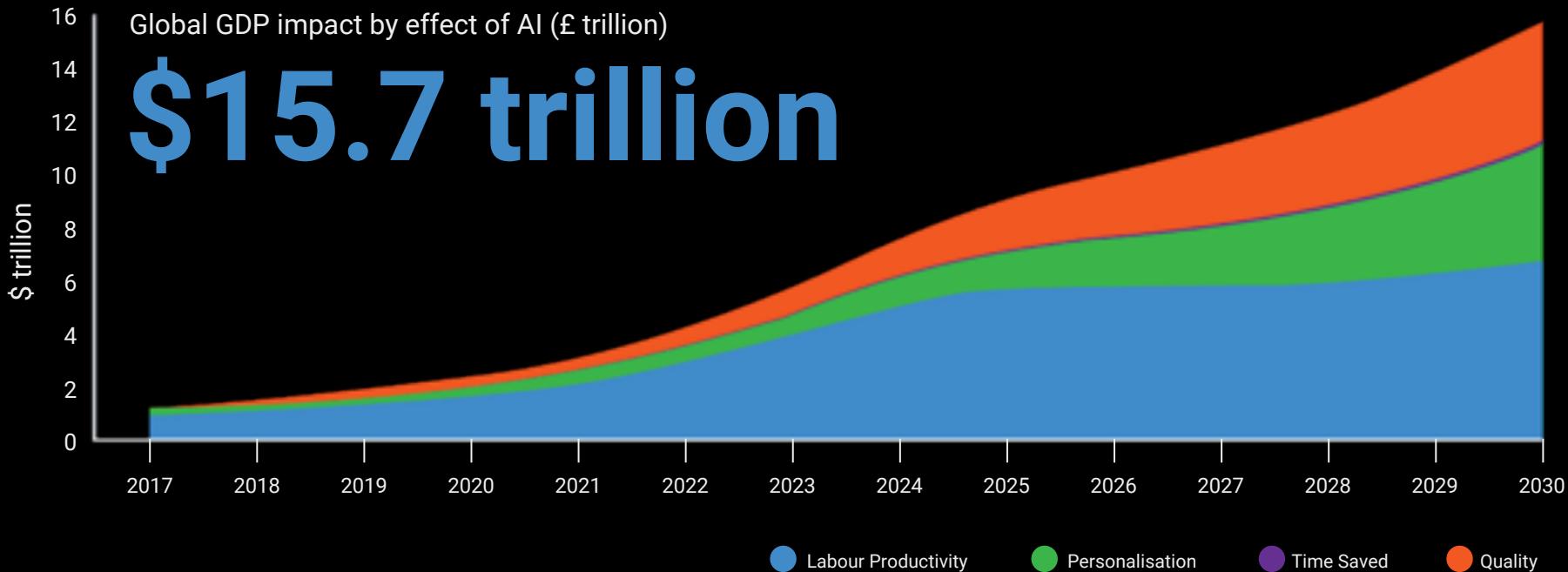


snowflake[®]



DataRobot

AI is transforming the world

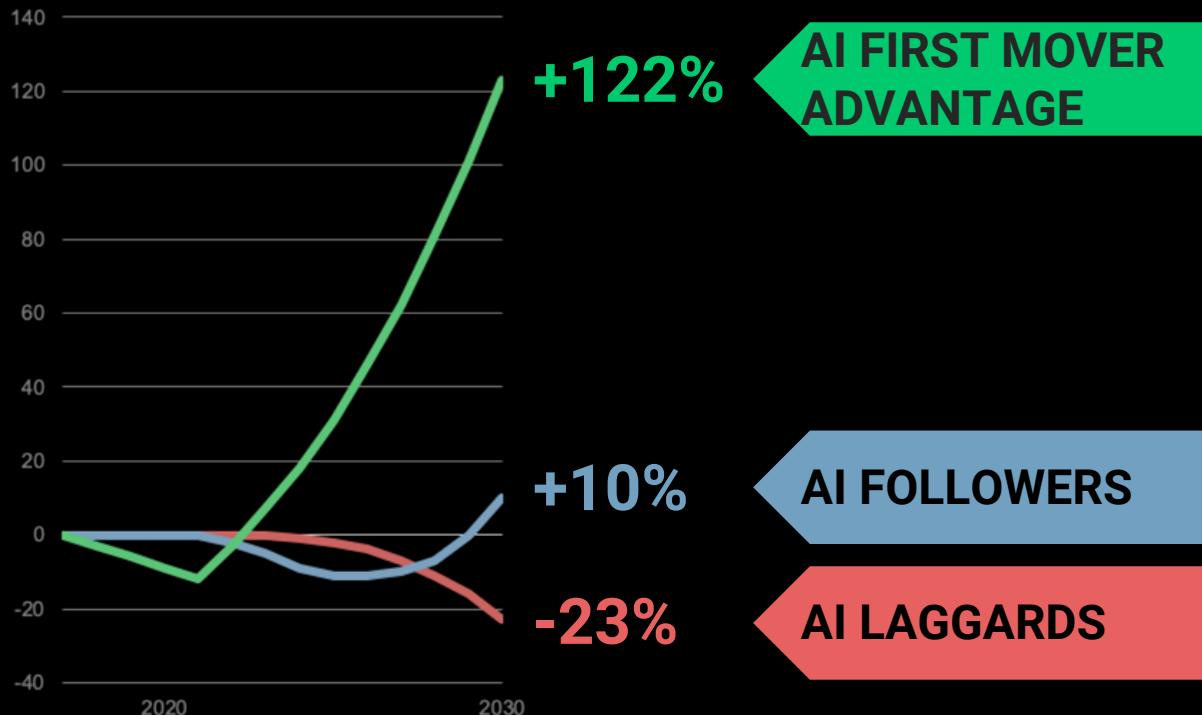


Source: PWC

Sizing the prize: What's the real value of AI for your business and how can you capitalise?

AI is a massive opportunity for your organization

Changes In Cash Flow by AI Adoption

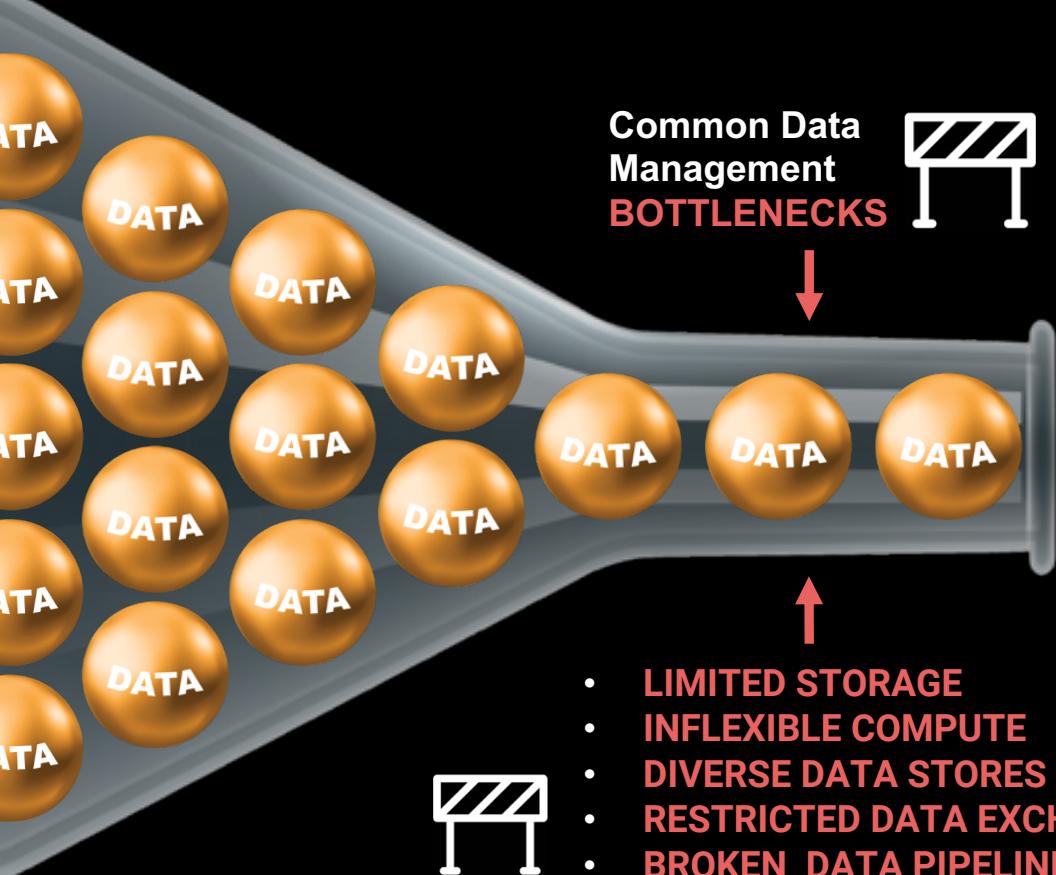


"Relative Changes In Cash Flow by AI Adoption Cohort" [Notes from the AI frontier: Modeling the impact of AI on the world economy](#) by McKinsey & Company

"The [AI] winners may take all and late adopters may never catch up."

Harvard Business Review
November 2018

You need help to overcome the barriers to adoption



BUILD

96%

Experience problems with model development

DEPLOY

90%

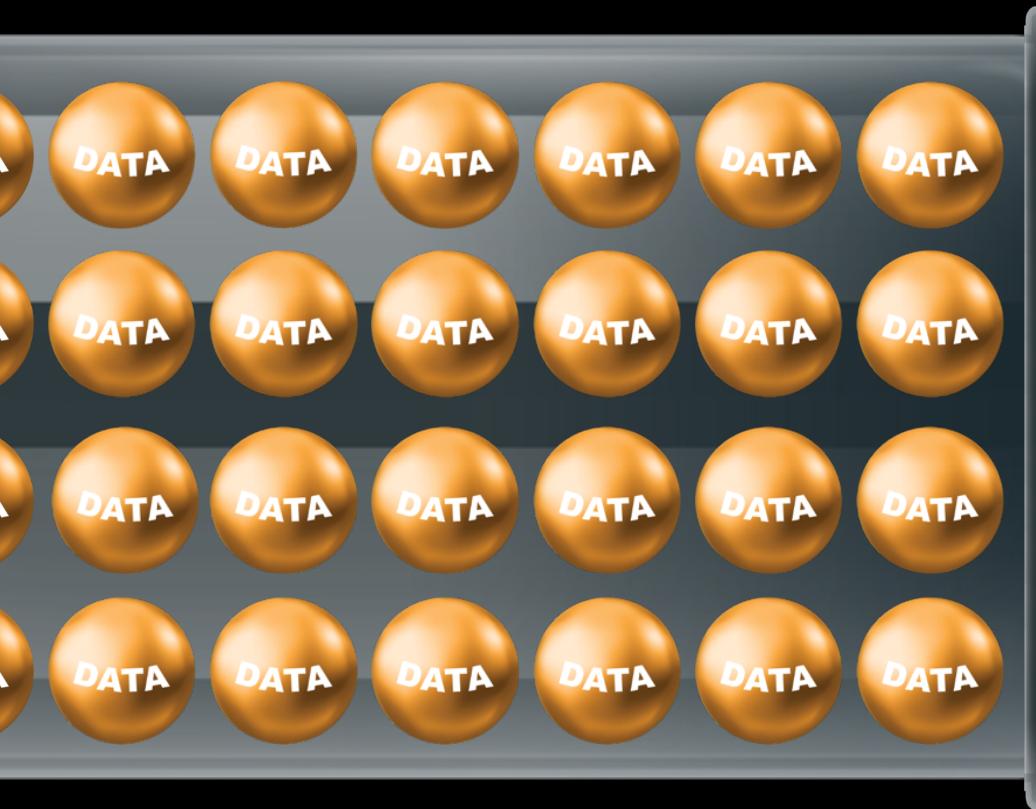
Cite obstacles to moving AI models into production

MONITOR

1%

Monitor their AI and ML assets once in production

We have joined forces to unlock AI without limits



Simple, scalable & speedy
Unlimited storage & compute
Seamlessly connected
Mutually cloud native
Leader in Data Management
Leader in Enterprise AI



MILES DAVIS

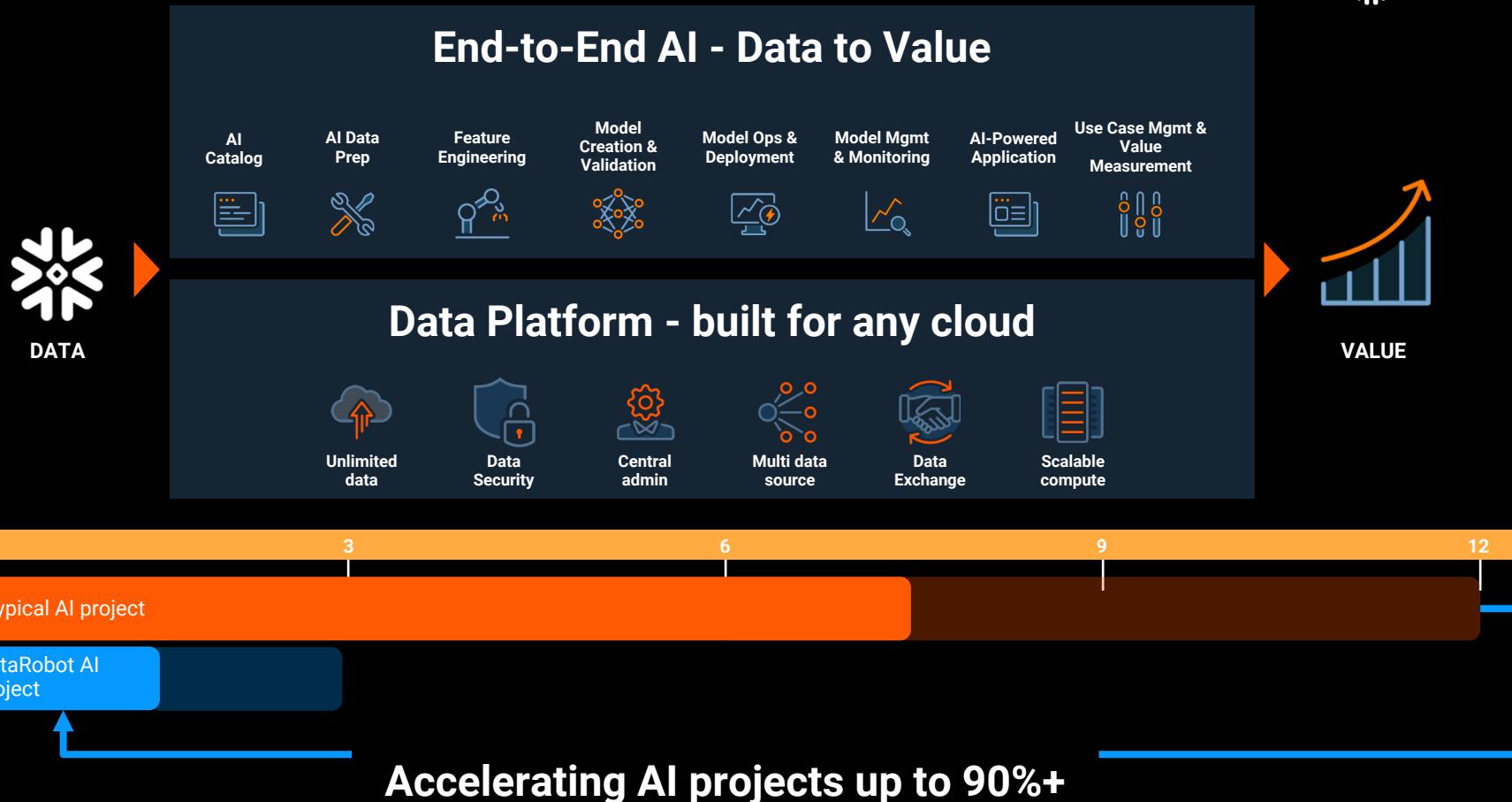
DATA & ANALYTICS MANAGER | HARMONY

Trusted Enterprise
AI Platform



Trusted AI-Native
Strategic Partner

Trusted Enterprise AI Platform to accelerate ROI



You can start with many tried and tested AI use cases



1 Real-time optimization

Enhancing routing and predicting delays for logistics and time sensitive processes



2 Strategic optimization

Planning store/site locations to maximize sales/support based on local demand and competition



3 Predictive analytics

Identifying customer churn risks following behavior changes or an interaction, for proactive retention



4 Predictive maintenance

Anticipating failures of critical machinery and equipment, enabling proactive repairs to be scheduled



5 Interpreting unstructured data

Reading customer data/ emails/tickets and routing the query based on type, priority and urgency



6 Forecasting

Estimating product, service and staffing demand at a hyper local level based on trends and events



7 Trend/anomaly detection

Identifying and preventing new cases of fraud based on suspicious activities and behavior changes



8 Radical personalization

Customizing offers for every individual customer based on their history and personal profile

Trusted by customers across all industries globally



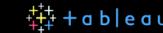
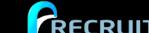
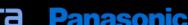
Transportation & Manufacturing



Finance



Technology



Healthcare



Insurance



Retail



PERSONAL LOANS

SIMPLE 100% ONLINE PROCESS

Personalised interest rates
from 6.99% – 29.99% p.a.*

[Apply now](#)[Get a quote](#)

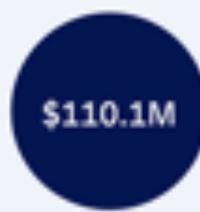
Won't affect your credit score



Total Application (\$)

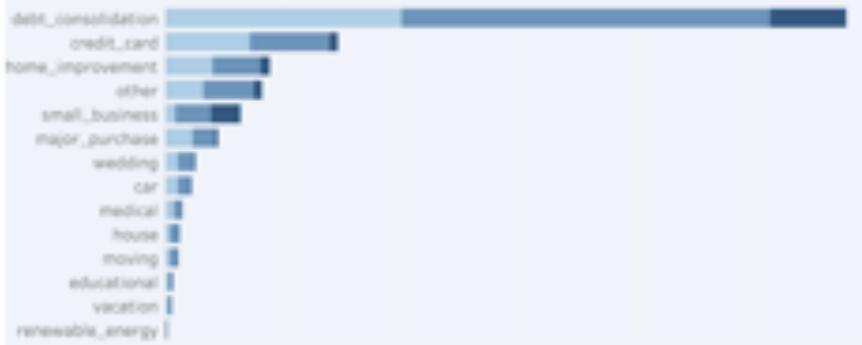
Total Funded (\$)

Risk Level

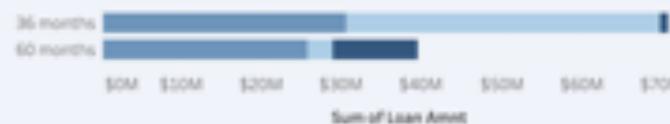


DataRobot

Paid
(\$107.4M)

Loan Purposes**Executive Summary**

	Sum of Loan Amnt	Funded Amnt	Number of Applications
High Risk	\$11.2M	\$11.2M	782
Moderate Risk	\$54.7M	\$54.7M	4,888
Low Risk	\$42.5M	\$41.6M	4,310
Grand Total	\$110.3M	\$107.4M	9,980

Term**Avg. Interest Rates**

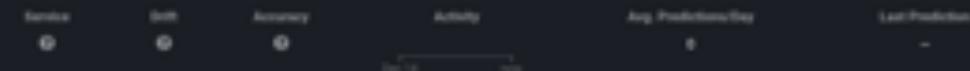
36-months

60 months

Loan (\$) vs. Fund (\$)**Grade & Risk**

Grade

Loan default model
(Deployment Prediction Server | Loans scoring)



Overview Service Health Data drift Accuracy Integrations Settings

Integrations



Snowflake
Not Configured



Scoring Code

DataRobot

app.datarobot.com/new

DataRobot Data Models Deployments Insights Jupyter Repository AI Catalog

Create project from an AI Catalog item

Search Cancel

Sort by Creation date ▾ 1-11 of 11 Select a dataset to see details...

B

- Snowflake_10k_HistoricLoansPerformance / Dataset**
Modified Jan 29, 2020 by James Lawson
STATIC
- E-commerce_churn_train / dataset**
Modified Jan 29, 2020 by James Lawson
STATIC
- Pred_Main_Bin / Dataset**
Modified Jan 29, 2020 by James Lawson
STATIC
- Store_Sales_Forecast_Train.xlsx / Dataset**
Modified Jan 29, 2020 by James Lawson
STATIC

DataRobot

app.datarobot.com/projects/5e31c6545d5f03045451ec9/eda

DataRobot Data Models Deployments Insights Jupyter Repository AI Catalog

Snowflake_10k_HistoricLoansPerformance

What would you like to predict?

is_bad

Number of rows

is_bad

Start

Modeling Mode: Autopilot

Feature list: Informative Features

Optimization Metric: LogLoss

Time-Aware Modeling

DataRobot has detected 1 time feature in your data. Would you like to use time series modeling or out-of-time validation?

Set up time-aware modeling

WORKERS

Using 0 of 20 total workers across all projects.

STATUS

- ✓ 1. Uploading Data
- ✓ 2. Reading raw data (Quick)
- ✓ 3. Exploratory Data Analysis (2.742 sec.)

Leaderboard Learning Curves Speed vs Accuracy Model Comparison Prediction Apps

☰ Menu Q Search Add New Model Filter Models Export

Metric LogLoss

 Model Name & Description

Feature List & Sample Size T

Validation

Cross Validation

Holdout

WORKERS

Using 0 of 20 total workers
across all projects 28

STATUS

Autopilot has finished

ACTIONS

Rerun Autopilot

Unlock project Holdout for
all models

eXtreme Gradient Boosted Trees Classifier with Early Stopping and Unsupervised Learning Features

Ordinal encoding of categorical variables | Converter for Text Mining | Auto-Tuned Word n-Gram Text Modeler using token occurrences | Missing values Imputed | Search for differences | Standardize | One-hot Encoding | Partial Principal Components Analysis | K-Means Clustering | eXtreme Gradient Boosted Trees Classifier with Early Stopping and Unsupervised Learning Features.

M129 BPN 64.8%

DR Reduced Features M73		80.0%		0.3634		0.3551		0.3630
-------------------------	--	-------	--	--------	--	--------	--	--------

ENET Blender

M134 M79+72+75+76+72+...

Multiple Feature Lists		64.0%	
------------------------	--	-------	--

0.3659

0.3564

ENET Blender

M135 M72+11+9+73

Multiple Feature Lists		64.0%	
------------------------	--	-------	--

0.3660

0.3569

AVG Blender

M132 M72+11+9+73

Multiple Feature Lists		64.0%	
------------------------	--	-------	--

0.3660

0.3569

Advanced AVG Blender

M133 M79+72+75+76+72+...

Multiple Feature Lists		64.0%	
------------------------	--	-------	--

0.3671

0.3572

DataRobot Data Models Deployments Insights Jupyter Repository AI Catalog Snowflake_10k_HistoricLoansPerformance

Leaderboard Learning Curves Speed vs Accuracy Model Comparison Prediction Apps

Menu Q Search Add New Model Filter Models Export Metric LogLoss

Model Name & Description Feature List & Sample Size Validation Cross Validation Holdout

eXtreme Gradient Boosted Trees Classifier with Early Stopping and Unsupervised Learning Features

Ordinal encoding of categorical variables | Converter for Text Mining | Auto-Tuned Word N-Gram Text Modeler using token occurrences | Missing Values Imputed | Search for differences | Standardize | One-Hot Encoding | Partial Principal Components Analysis | K-Means Clustering | eXtreme Gradient Boosted Trees Classifier with Early Stopping and Unsupervised Learning Features

M129 RPII # 64.0% RECOMMENDED FOR DEPLOYMENT

Evaluate Understand Describe Predict Compliance

Feature Impact Feature Effects Prediction Explanations Word Cloud

Feature Impact

Feature	Impact (%)
desc	~98%
annual_inc	~75%
term	~65%
HI_rate	~55%
title	~45%
inq_last_6months	~35%
grade	~25%
revolUtil	~20%
emp_occupation	~15%
loan_amnt	~10%
installment	~8%
zip_code	~6%
open_acc	~5%
total_acc	~4%
sub_grade	~3%
funded_lamount	~2%
purpose	~1%

Using 0 of 20 total workers across all projects 20

STATUS Autopilot has finished

ACTIONS Run Autopilot Unlock project Holdout for all models

DataRobot Data Models Deployments Insights Jupyter Repository AI Catalog Snowflake_10k_HistoricLoansPerformance

Leaderboard Learning Curves Speed vs Accuracy Model Comparison Prediction Apps

Menu Search Add New Model Filter Models Export Metric LogLoss

Model Name & Description Feature List & Sample Size Validation Cross Validation Holdout

Get top 3 explanations for low high predictions: Compute & Download

Validation Predictions

Preview of explanations for predictions from validation data within highlighted area(s):

ID	PREDICTION	EXPLANATIONS
6257	0.718	term = "60 months" desc = "The time is right to ... title = "Small Business Loan"
7336	0.623	desc = "Borrower added on ... title = "Small Business Loan!" term = "60 months"
3500	0.582	desc = "Borrower added on ... term = "60 months" title = "Business Loan"
910	0.622	int_rate = 5.79 desc = "Borrower added on ... annual_inc = 83000
618	0.621	int_rate = 5.79 desc = "Borrower added on ... credit = 70"

workers: Using 0 of 20 total workers across all projects. STATUS: Autopilot has finished. ACTIONS: Run Autopilot, Unlock project Holdout for all models.

Worksheet 1

+ -

 All Queries | Saved 6 minutes ago

Context: ACCOUNTADMIN COMPUTE_WH (X) SNOWFLAKE_SAMPLE_DATA PUBLIC ...

Results Data Preview

← Open History

Table: LENDINGTREE.PUBLIC.LendingTree_Scored

Data

Details

Filter result...

Columns ▾

SLNO	revol_bal	revolutil	total_rev	initial_list_status	mths_since_last	policy_code	Prediction	Positive Probabi	Reason 1 Strength	Reason 1 Strength	Reason 1 Feature
1214	9.3	20	f	NA	1	1	FALSE	0.047275	---	-0.632484	int_rate
1466	86.2	5	f	NA	1	1	FALSE	0.112501	***	0.26099	annual_inc
29763	56.26	21	f	NA	1	1	FALSE	0.0603	---	-0.619301	int_rate
4787	10.8	20	f	NA	1	1	FALSE	0.083169	**	0.315783	annual_inc
56006	84.7	38	f	NA	1	1	FALSE	0.289554	***	0.46436	int_rate
2296	38.3	29	f	NA	1	1	FALSE	0.104191	---	-0.608982	annual_inc
3634	371	22	f	NA	1	1	FALSE	0.061383	---	-0.512833	int_rate
11832	56.3	34	f	NA	1	1	FALSE	0.097507	***	0.349549	term
4062	18.1	21	f	NA	1	1	FALSE	0.080543	---	-0.20443	annual_inc
10581	32.8	18	f	NA	1	1	FALSE	0.187204	***	0.452096	emp_title
2899	39.7	8	f	NA	1	1	FALSE	0.165861	***	0.31219	annual_inc
6839	80.4	20	f	NA	1	1	FALSE	0.083224	---	-0.238318	desc

PERSONAL LOANS

SIMPLE 100% ONLINE PROCESS

Personalised interest rates
from 6.99% – 29.99% p.a.*

[Apply now](#)[Get a quote](#)

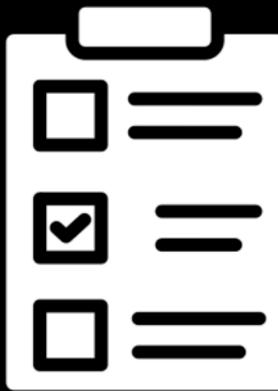
Won't affect your credit score



Take your next step to unlock AI without limits



Learn
More



Explore Your
Use Case(s)



Starter
pack



FIND OUT MORE

www.DataRobot.com/partner/snowflake

SPEAKER Q&A



JOIN US AFTER THE BREAK

FROM ZERO TO
SNOWFLAKE

in 90 minutes



© 2020 Snowflake Inc. All Rights Reserved.

THANK YOU

Data for Breakfast

