

Snowflake vs. Molecula Cheat Sheet

		
Data Readiness & Access		
Feature Extraction	 No access to features	 Automates feature extraction from raw data
No Pre-Aggregation	 Only data that has been aggregated and moved into Snowflake is available to users	 Tap directly into data sources
Distributed Data Access	 Access to data from multiple sources all housed within one Snowflake cloud instance	 Can ingest streaming, structured raw and unstructured raw data with ease
Secure for Sharing	 Full data security compliance	 Data is stored in anonymized, binarized format
SQL Compatible	 Full library of SQL available	 SQL is available for relevant queries, but not full library of SQL
Single Infrastructure Across All Projects	 Cloud data warehouse that is often used in conjunction with data lakes and other data infrastructure tools	 Overlay feature store allows access to distributed data sources to power advanced analytics from one location
Performance		
Performance with Complex Analytics	 Query performance	 Query performance

	degrades with complexity	improves with complexity
Query Performance with High Volumes of Data	 Query performance deteriorates as you scale	 Homomorphic compression reduces data footprint up to 90% while maintaining query speeds
Query Performance with High User Concurrency	 Up to 10 concurrent users	 Unlimited concurrent
Query Performance with Distributed Data Sources	 Ability to query seamlessly across disparate data once it has been aggregated within Snowflake	 Ability to query seamlessly across disparate data (across clouds, on prem, in edge)
Real-Time Data Availability to Query	 Some data can be accessed in real-time via partnered replication tools	 Real-time data is available
No Federation	 Performance requirements may force users to federate	 No need to federate -- all projects can be run on features within feature store
No Caching	 Uses a Result Cache as part of its internal architecture to attempt to boost performance	 No need to cache -- feature-first format continuously updates without causing large footprint growth
Feature Serving		
Feature Re-Use	 No feature access	 Feature re-use without limitation
Feature Vectors	 No built-in feature engineering capabilities	 Continuous feature extraction related to entities

Analytics Types		
Ad-Hoc	 <p>Ability to do analysis on the fly is present with Snowflake. Performance varies based on complexity of queries and volume of data</p>	 <p>Ability to perform ad-hoc analysis is very minimally affected by the volume of data or complexity of queries</p>
Real-Time	 <p>Snowflake has minimal real-time capabilities for advanced analytics use cases</p>	 <p>Real-time feature updates and millisecond query performance enable real-time analytics</p>
Free Text/NLP	 <p>Snowflake has some free text analytics capabilities that enable NLP</p>	 <p>Molecula does have some free text analysis capabilities but we would not recommend using Molecula for this use case</p>
BI (Diagnostic/Descriptive)	 <p>Snowflake is great for business intelligence initiatives</p>	 <p>Visualization tool compatibility is limited but this functionality is available upon request</p>
Customer 360/Segmentation	 <p>Snowflake is great for customer segmentation efforts</p>	 <p>Molecula is excellent for customer 360 and segmentation initiatives, especially when real-time or predictive requirements are necessary</p>
ML (Predictive/Proactive)	 <p>Snowflake can help perform some machine learning initiatives but isn't best suited for most cases</p>	 <p>Feature-first data format enables millisecond query performance at scale and continuous analysis</p>
Image Analytics	 <p>Snowflake has some image analytics capabilities</p>	 <p>Molecula is not suited for image analytics</p>

Costs		
Low Total Cost	<p style="text-align: center;">●</p> <p style="text-align: center;">High costs due to storage + compute</p>	<p style="text-align: center;">●</p> <p style="text-align: center;">Low costs due to 60-90% reduced footprint</p>
Low Entry Cost	<p style="text-align: center;">●</p> <p style="text-align: center;">Free trial available</p>	<p style="text-align: center;">●</p> <p style="text-align: center;">Free trial available upon request</p>
No Copy	<p style="text-align: center;">●</p> <p style="text-align: center;">Copies data from sources and moves it into cloud database</p>	<p style="text-align: center;">●</p> <p style="text-align: center;">Since only features are extracted from data and all are housed in one feature store, there is no need to copy data</p>
Minimal Additional Cost per Project	<p style="text-align: center;">●</p> <p style="text-align: center;">Pay for storage + time of compute for each project as well as the additional technologies and data copies required for advanced analytics</p>	<p style="text-align: center;">●</p> <p style="text-align: center;">Single feature store instance powers all projects</p>