

[thegeneralist.substack.com /p/the-s-1-club-snowflake-and-the-data](https://thegeneralist.substack.com/p/the-s-1-club-snowflake-and-the-data)

## The S-1 Club | Snowflake and the Data Blizzard

MDA Gabriele

31-39 minutes

---

# THE DATA BLIZZARD



Since its founding in 2012, Snowflake has grown into a leader in the cloud data warehousing space. The company's intuitive solution makes it easy to run analytical queries with a customer-aligned billing model. Intriguingly, Snowflake has established itself without the founder-led culture that dominates most of the tech industry. Instead, three external CEOs have guided the company to a ~\$500M annual revenue run rate, with 132% year-over-year growth. That's attracted pre-IPO investment from Warren Buffett — a rare foray into tech for the Oracle of Omaha. Big winners of the listing include the secretive Sutter Hill Ventures and Altimeter Partners.

Snowflake's toughest task may be in weathering the coming storm. Amazon, Microsoft, and Google all have products in the space, and with limited enduring defensibility, the company could see its position pressured.

As a member of The Generalist, you have access to The S-1 Club's full report. To learn about Snowflake's kooky founders, mercurial CEO Frank Slootman, and the mission to win the "data cloud," keep reading. If you enjoy this analysis, we'd appreciate you sharing it.

[Share](#)

Tomorrow, September 12th at 7 pm ET, Danielle Morrill, Sam Parr, Annika Lewis, Aashay Sanghvi, David Wei, and I will discuss our research and answer questions. If your idea of a fun Saturday evening

is talking about enterprise software and one of the fastest-growing companies of all-time, we'd love to see you there.

[Join the event](#)

## Analysts

[David Wei](#)

[Justin Gage](#)

[Annika Lewis](#)

[Aashay Sanghvi](#)

[Karine Hsu](#)

[Mario Gabriele](#)

## Introduction

One last job.

It's a scene you've watched before. A grizzled pro approached by an old friend in need of help. Perhaps they meet in a dingy basement, under the veil of secrecy, or share words in a banal local diner or coffee shop.

The friend pleads with the pro. A daughter is sick, a partner in danger, a brother in a bad spot with some worse people. Or sometimes it's a question of upside — the money is just too good to pass up.

*One last job*, the friend beseeches.

It always works. It has to — otherwise, there would be no film.

The "One Last Job" premise is so popular it's attained the status of a trope, serving as the basis for iconic films like *Heat*, *Inception*, *Gone in Sixty Seconds*, and countless others.

After a period of reluctance, the pro succumbs. The job is on and the game afoot.

In 2018, Frank Slootman was retired and comfortable. The bolshy Dutchman had fulfilled his destiny in the corporate world, transforming Data Domain from a 20-person company into an "800-pound gorilla" doing \$1B in sales, before taking the reins at ServiceNow. Starting in 2011, Slootman supercharged the business, increasing revenue from \$75M to \$1.5B over his six-year spell.

Now it was time to relax. To sit on a few boards and enjoy life on the water. Slootman's sailboat, *The Invisible Hand* (really), had won the Transpac Honolulu Race the year prior. A man like Slootman would surely have seen such success not as a culmination but the start of sustained dominance.



#### Renegade Sailing

But then Slootman got a call.

We can only speculate as to who made first contact. Before long, though, Slootman was talking to [a familiar face](#): Mike Spieser. The two men had gotten to know each other starting in 2014 as both served on the board of Pure Storage, a fast-growing data storage company. Mutual respect grew.

If anyone knew Snowflake and what the company needed, it was Spieser. He'd been there since the company's early days, both as an investor through his firm Sutter Hill Ventures (SHV), and as an operator — he served as Snowflake's first CEO before passing the torch to Bob Muglia. Formerly of Microsoft and Juniper Networks, Muglia had performed admirably, overseeing a period of breakout growth, but when it became clear Slootman might be available, Snowflake's board pounced.

The message the former ServiceNow chief received [was clear](#): "We want you."

[According to Slootman](#), he didn't consider any other roles, and if the opportunity had come a year earlier, he would have turned it down. But, "Snowflake [was] a special company...they just don't come along often."

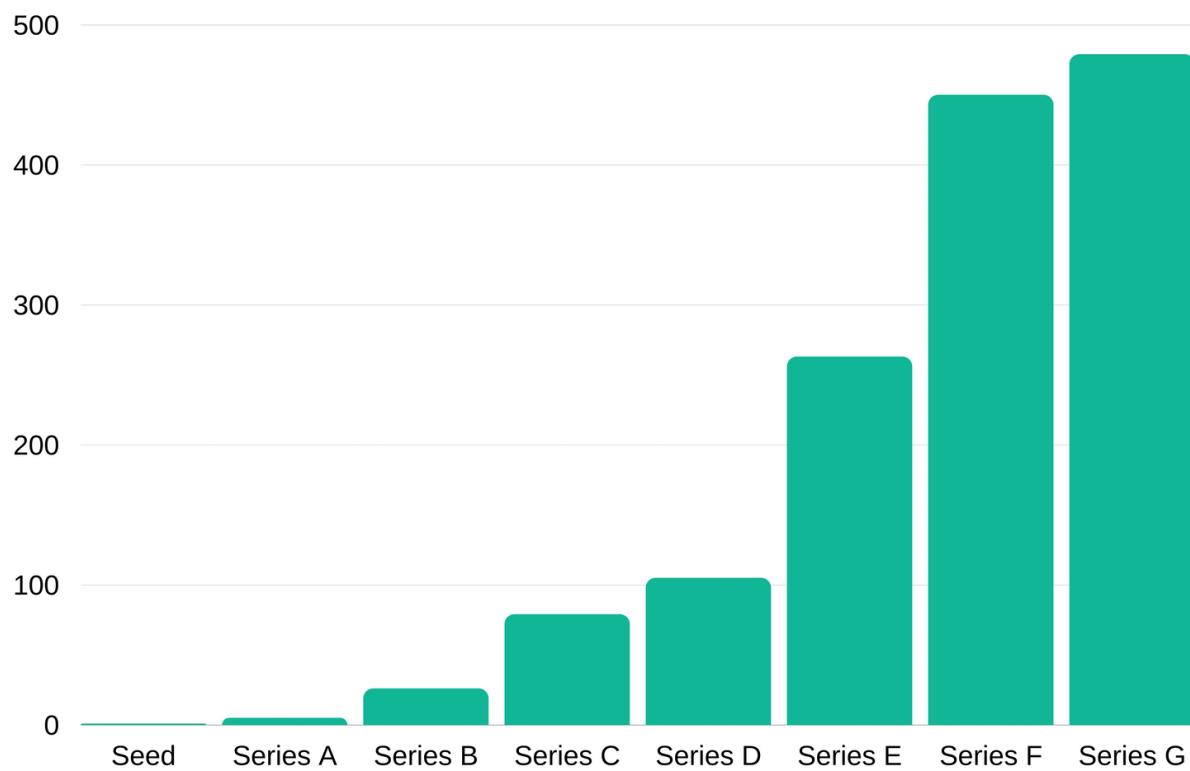
For Slootman, it was time for one last job. This is his story, in part, but more importantly, the tale of a company too good to pass up.

Instead, as alluded to above, Snowflake has been helmed by three outside CEOs over the last six years. First Spieser, then Muglia, and finally Slootman.

That pattern rarely screams "wild success," but Snowflake is the exception to the rule: the company has strategically leveled-up over time, identifying the right leader at the right time to maintain their startling trajectory.

While every company has its trials and tribulations under-the-hood, from the outside, Snowflake has executed, dare we say it, perfectly. From January 31, 2019, to January 31, 2020, the company grew revenue by over 173% to \$264M, with the following six months seeing Snowflake hit an Annual Run Rate (ARR) of \$483M. Though Snowflake is still operating at a loss — losing \$348M as of January 2020 — that growth has allowed them to access capital at a jaw-dropping clip. Snowflake has raised \$1.4B with \$1.1B arriving in the past two years.

## Snowflake's funding history



Crunchbase

Safe to say: it's snowing in data land.

[Share](#)

## Number of mentions in the S-1

- Data: 697
- Capital One: 31
- Data Cloud: 26
- Cloud Data Platform: 28
- Silo: 15
- Sutter Hill Ventures: 12

- Scarcity: 1

## Market

Wintery tailwinds are expected to power growth in the world of data. While Snowflake pegs their market at ~\$81B it may be even more extensive than that. As of 2018, the International Data Corporation (IDC) found global data storage generated [\\$88B in revenue](#) with a capacity of 700 exabytes added. By 2023, that figure is expected to nearly double, reaching ~\$176B in revenue and 11.7 zettabytes in storage.

The data warehousing market was believed to be [\\$13B in 2018](#), rising to \$30B by 2025, a 12% CAGR. The APAC region is expected to grow more quickly, at a 15% CAGR.

Encouragingly for Snowflake, incremental capacity and renewals are expected to be driven by the cloud. Gartner predicts [75% of databases](#) will be in the cloud by 2022, with just 5% "repatriated" on-premises. Other sources suggest that hybrid strategies — the marriage of cloud and on-premise data warehousing — may expand rapidly. Just as encouragingly, "unstructured data" — a significant driver of Snowflake's rise — is growing between [40-50%](#) per year.

It's worth noting that Snowflake is already well-established in the space. Per Datanyze, the company is the third-largest with an estimated [10.08%](#) market share, just behind SAP Business Warehouse and Apache Hive.

DATA WAREHOUSING MARKET SHARE TABLE

Ranking	Technology	Domains	Market Share
1	<a href="#">SAP Business Warehouse</a>	2,044	15.43%
2	<a href="#">Apache Hive</a>	1,462	11.03%
3	<a href="#">Snowflake</a> 	1,335	10.08%
4	<a href="#">Amazon Redshift</a>	1,281	9.67%
5	<a href="#">IBM Netezza DataWarehouse Appliances</a>	1,034	7.80%
6	<a href="#">Oracle Autonomous Data Warehouse Cloud</a>	989	7.46%
7	<a href="#">Omniture Data Warehouse</a>	903	6.82%
8	<a href="#">IBM Cognos Impromptu</a>	825	6.23%
9	<a href="#">Pivotal Greenplum</a>	689	5.20%
10	<a href="#">IBM Data Warehouse Systems</a>	502	3.79%

Mordor Intelligence

Gartner recognized Snowflake as a "Leader" though the company did trail more familiar names like Oracle, Microsoft, and AWS in terms of "completeness of vision" and "ability to execute."

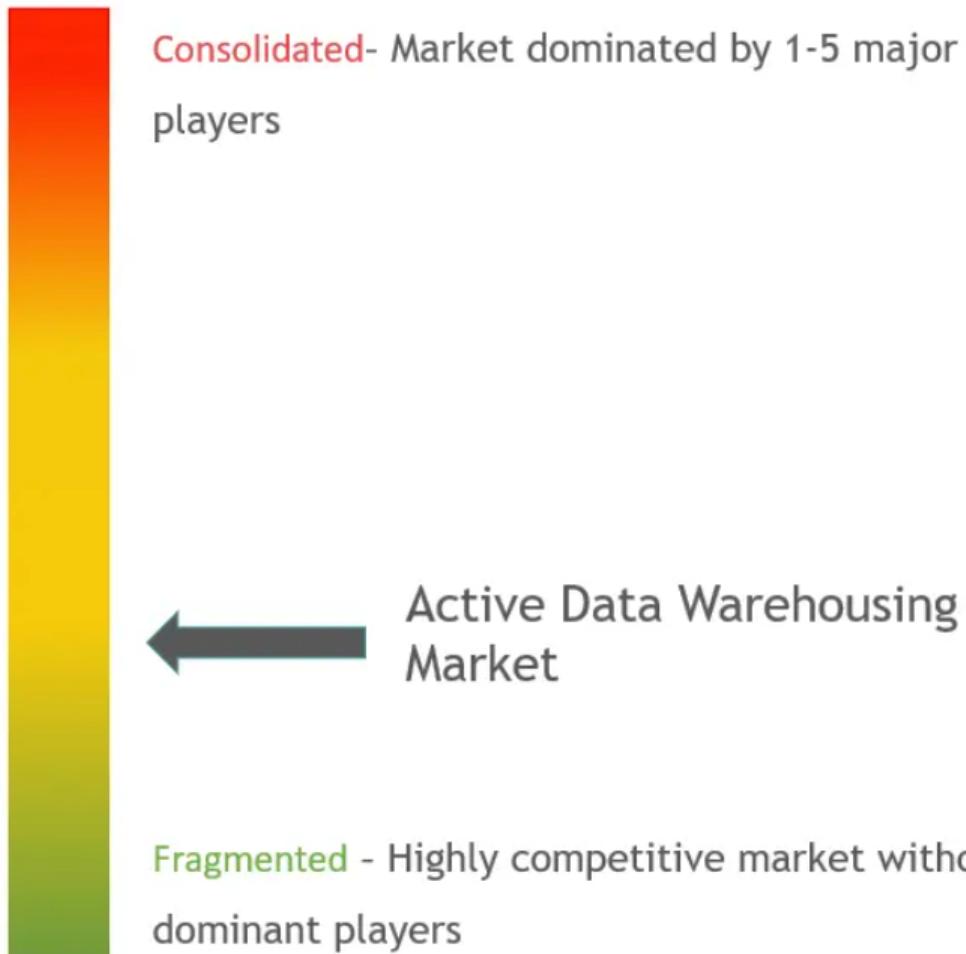
**Figure 1. Magic Quadrant for Data Management Solutions for Analytics**



#### Gartner's Magic Quadrant

Finally, it's worth noting that the market seems to be up for grabs with considerable fragmentation in the space. That's created a fluid dynamic between businesses as rivals compete and collaborate. That structure has aided Snowflake's rise over the past eight years but could make data domination a considerable challenge. In "Competition," we'll dive into one method to track Snowflake's market position, focused on "developer mindshare."

# Market Concentration



Mordor Intelligence

## What *is* a database?

If you've gotten this far, you probably know that Snowflake is a database.

But what does that mean? How does Snowflake compare to multi-billion dollar incumbents like Oracle and MongoDB?

Fundamentally, there are two types of databases: transactional and analytical. Snowflake is printing money by excelling at the latter.

Transactional databases are all around us. With few exceptions, every app is built on a database. That's what allows developers to store user emails, preferences, account information, and anything else that gets customized. These are called transactional databases because the data is generated through the little transactions that happen every time a user takes an action:

- A user signs up → add a row to the users' table
- A user changes their password → update an existing row
- A user loads their profile page → read from the users' table

If your app has many users, these transactions *constantly* occur (we're talking about thousands per second). As a result, early databases were optimized for this use case: to add and update rows quickly and ensure data wasn't corrupted. They're also well-suited to filter specific sets of data.

Over time, teams realized that all of the data they collected could be pretty valuable. By analyzing it, they might be able to answer questions critical to their business. How often did users log in? Where were they based? What was the most popular pricing plan?

The problem was that the kinds of queries required to unearth this information are *markedly different* from those used in transactional databases. They're typically longer, more resource-intensive, and involve combining data sources from multiple places. Transactional databases weren't designed to handle them.

Transactional query	Analytical query
"Get me this particular user's address"	"Show me the number of unique users per month that visited our site and then signed up for our Pro plan with a Paypal account"

Enter the data warehouse.

To better serve these types of queries, teams began designing databases specifically for analytics, copying data from their transactional database into this new analytical database. These new structures were optimized for gnarly, analytical queries, as was the underlying infrastructure (servers).

Thanks to their more fine-toothed functionality, teams have ended up storing orders of magnitude *more* data in analytical warehouses, tracking events like website page views and product interactions.

This is what Snowflake is: an analytical database. More precisely, it's an analytical database that requires *no infrastructure management* from its users, often referred to as a "managed data warehouse." Users don't need to deal with deploying servers, scaling, networking, or managing permissions - they just dump data into Snowflake and query it (extremely) quickly. Snowflake's even [built an in-product query engine](#) to accelerate queries through low-level optimizations.

## The Product

Snowflake isn't the only managed data warehouse on the market — [Amazon Web Service's \(AWS\) Redshift](#) and [Google Cloud Platform's \(GCP\) BigQuery](#) are popular options.

What is it that has made Snowflake such a success? Above all, it's a combination of flexibility, service, and UI.

### Splitting storage and compute

It's worth noting that with a database like Snowflake, two pieces of infrastructure keep the lights on: storage and compute. Snowflake takes care of efficiently storing your petabytes of data *and* making sure your queries run quickly. This idea — separating storage and compute in a data warehouse — was relatively novel when Snowflake started. Today, [entire query engines like Presto](#) exist just to run queries efficiently, with no storage included.

There are some clear advantages to splitting storage and compute: stored data is located remotely on the cloud, saving local resources for the compute load. Moving storage to the cloud delivers lower cost, has higher availability, and provides greater scalability.

### No vendor lock-in

CIOs — and increasingly, developers — don't want to be tied to a specific cloud provider. The majority of enterprises have adopted a [multi-cloud](#) approach (and that can mean many things). As such, there's a natural reluctance to choose solutions like BigQuery that are native to a single cloud (Google, in this case).

Snowflake provides a different level of flexibility. It runs on AWS, Azure, or GCP, satisfying the "multi-cloud checkbox" that enterprise buyers have. Often such a strategy is theoretical rather than concrete.

Amidst tech's giants battling for control of the cloud, Snowflake is a sort of warehousing Switzerland.

Of course, this isn't to say that there's *no* lock-in. By signing up to Snowflake, users are committed to a vendor: Snowflake, itself.

## Fully managed

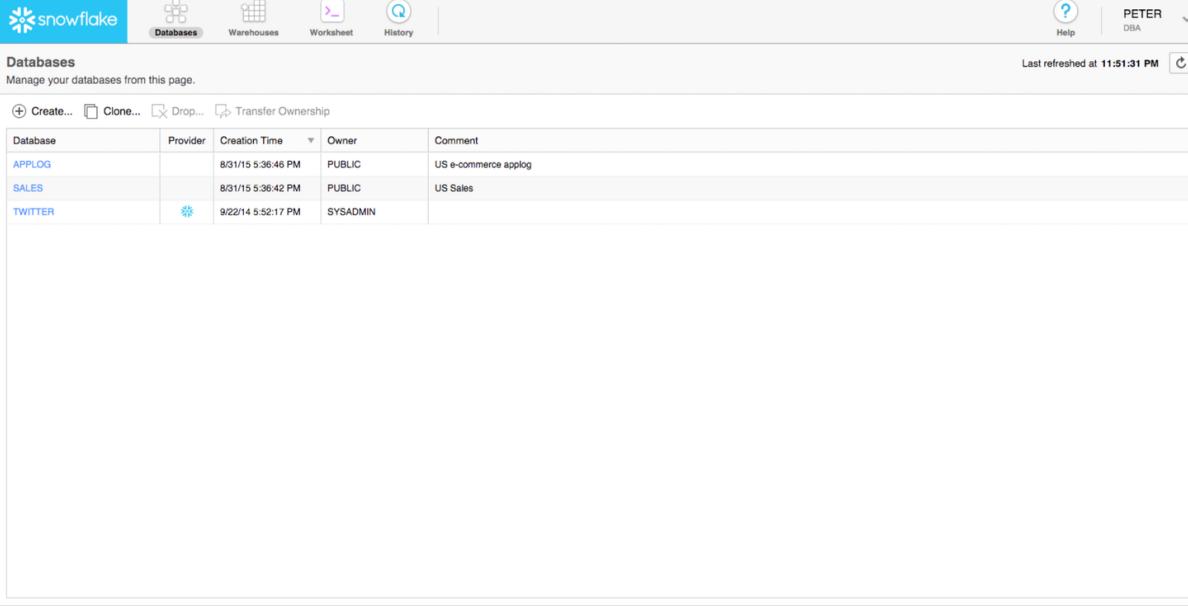
If you want to build a data warehouse from scratch, there's a lot of infrastructure to manage. Even if you're outsourcing your servers to a cloud provider, you still need to deal with sizing the right instance (4GB RAM? 48GB RAM?), scaling as your workloads grow, and networking components together. If you use a managed service like AWS's Redshift, you need to worry about handling sizing and scaling.

As mentioned above, Snowflake is a fully managed service — users don't need to worry about any infrastructure at all. You just put your data into the system and query it.

While fully managed services sound great, no good deed goes unpunished, which means there's a tradeoff: price. Snowflake customers need to be intentional about storing and querying their data because fully managed services are expensive. When teams decide whether to build or buy, they need to forecast cost savings by comparing Snowflake ownership's total cost to building something themselves.

## Impressive UI and SQL ergonomics

Snowflake's UI for querying and exploring tables is surprisingly pleasant to look at and intuitive to use.



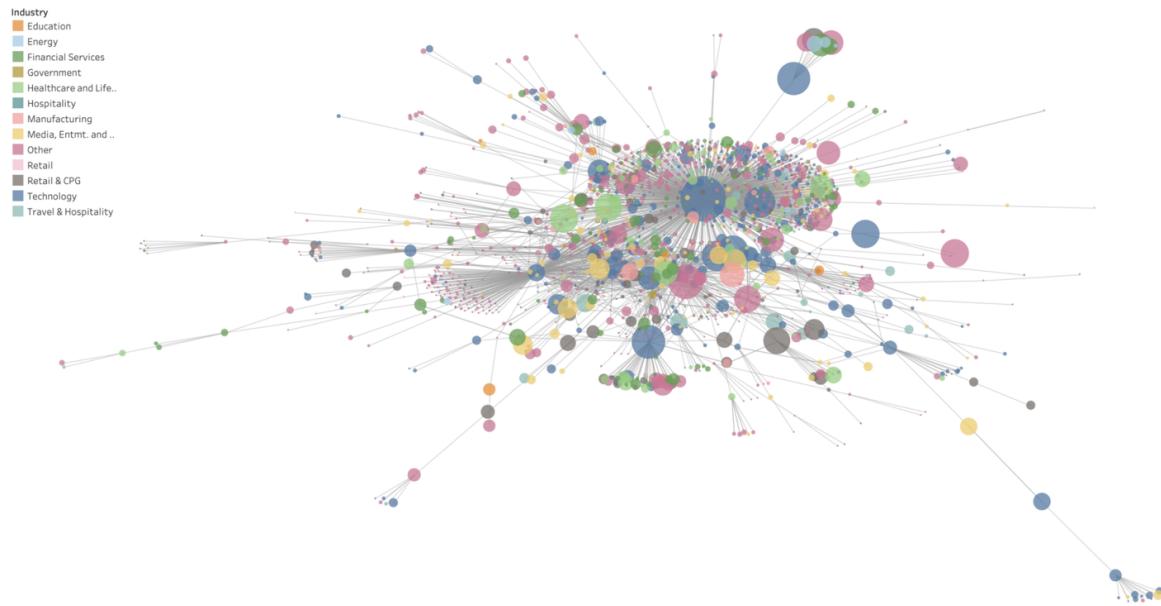
The screenshot shows the Snowflake web interface. At the top, there's a navigation bar with icons for Databases, Warehouses, Worksheet, and History, along with a user profile for 'PETER DBA'. Below the header, a banner says 'Manage your databases from this page.' and 'Last refreshed at 11:51:31 PM'. A toolbar below the banner includes buttons for '+ Create...', 'Clone...', 'Drop...', and 'Transfer Ownership'. The main area is a table titled 'Databases' with columns: Database, Provider, Creation Time, Owner, and Comment. It lists three databases: APLOG (Provider: PUBLIC, Owner: PUBLIC, Comment: US e-commerce applog), SALES (Provider: PUBLIC, Owner: PUBLIC, Comment: US Sales), and TWITTER (Provider: SYSADMIN, Owner: SYSADMIN, Comment: null). The table has a light blue header and white rows.

Snowflake's product

Another nice touch from Snowflake is their SQL ergonomics. SQL (Structured Query Language) is the programming language that developers and data scientists use to query their databases. Each database (PostgreSQL, MySQL, etc.) has its own SQL flavor, with slightly different details, wording, and structure. Snowflake's SQL seems to have gathered the best from many SQL dialects, adding useful functions. A great example is the [PIVOT function](#), which regular SQL tends not to have but saves analysts a bunch of time (if you've ever tried to create a pivot table from scratch, you'll understand).

## Data sharing

In a bid to establish its network effect credentials, Snowflake outlines its "data sharing" functionality in the S-1. As it stands, this appears to be a relatively immature part of Snowflake's product, though it does seem to have had some use during the COVID outbreak. The S-1 notes that one client shared epidemiological data via the company's "Data Marketplace," subsequently viewed by "hundreds" of customers. A representation of this interaction can be found below.



**Snowflake's S-1 filing**

## Business Model

"I...[do] not like SaaS that much as a business model, felt it not equitable for customers," Slootman [wrote in 2019](#), discussing his decision to join Snowflake.

The concept of equity and alignment is a critical part of the company's business model. Rather than charging per seat on a monthly or annual basis, Snowflake bills customers based on usage, offering on-demand and upfront billing. Discounts are given for the latter.

Furthermore, Snowflake bills its three core functions — storage, compute, and cloud services — separately. Compute is the larger financial commitment. For [example](#), purchasing 4TBs of data storage per year and 955 in compute credits costs roughly \$23K per year. Approximately \$22K would be spent on compute, with just \$1K allocated to storage.

## PRICING EXAMPLE 1

Customer has purchased the Snowflake Capacity Standard Service with Premier Support in the United States. They load data every night and have 8 users using the database during business hours. This customer:

- Stores an average of 4 TB's of data.
- Runs a batch loading task every night for 2.5 hours, using a Small Standard Virtual Warehouse.
- Enables eight users to work 5 days a week from 8am until 6pm using a Medium Standard Virtual Warehouse.

### DATA LOADING REQUIREMENTS

Parameter	Customer Requirements	Proposed Setup
Data Load Frequency	Daily	Small Standard Virtual Warehouse (2 Credits per hour)
Loading Window	2.5 hours	(2 Credits x 2.5 hours per day x 31 days per month)
Credits Required per month	155	

### USER REQUIREMENTS

Parameter	Customer Requirements	Proposed Setup
No. of Users	8 Users	Medium Virtual Warehouse (4 Credits per hour)
Time Slot	8AM-6PM (10 hours)	
Credits Required per month	800	(4 Credits x 10 hours per day x 20 days per month)
Credits Required per month	955	(800 Credits for Users + 155 for Data Loading)

### STORAGE REQUIREMENTS

Data Set size (per month)	4TB (After Compression)	
Data storage price (per year)	\$1,104	4TB x \$23 (per TB per month, based on Storage table) x 12 months
Virtual Warehouse (per year)	\$21,774	955 (Credits per month) x \$2 (per credit) x 0.95 (5% sample discount) x 12 months
Total Price (per year)	\$22,878	Data Set (per year) + Virtual Warehouse (per year)

### Public Comps, Snowflake

This model is suited to the large enterprise contracts on which Snowflake focuses. The company boasts recognizable logos across sectors, including media, financial services, healthcare, manufacturing, and technology. Sample customers include sleepier businesses like Office Depot, McKesson, and Nielsen, alongside startups DoorDash, Instacart, and Rent the Runway.

Advertising, Media & Entertainment	Financial Services	Healthcare, Wellness & Life Sciences
2K Games Accordant Media KIXEYE Nielsen PLAYSTUDIOS Sharethrough	AXA Bankrate Capital One CapSpecialty Chime Experian	Amino Asics HC1 McKesson Strava
Manufacturing & Retail	Online Services & Marketplaces	Technology
Logitech Madison Reed Office Depot Sainsbury's Sony US Foods	Ask Blackboard DoorDash Instacart OfferUP Rent the Runway	Adobe Akamai DocuSign Dropbox Keboola Micron

### Snowflake's S-1 filing

Snowflake primarily leverages a top-down sales motion to secure these deals, relying on robust sales development, inside sales, and other related teams. Indeed, of [all public SaaS](#) companies, Snowflake

spends most on Sales & Market as a percentage of revenue at 70%. As a comparison, Zoom spends just 37%.

Capital One is a particularly crucial relationship. In 2017, Snowflake signed the financial giant and has successfully deepened the relationship since. For the fiscal year ending January 31, 2020, Capital One represented [11% of Snowflake's revenue](#) — both a feat and a concern. That's significant customer concentration, a worry slightly mitigated when looking at the rest of Snowflake's customer base: the company has 3,117 customers. Of the Fortune 500, Snowflake counts 146 as clients, and they contribute 26% of revenue. Reasonable enough.

The signing of Capital One is also a sign of the trust Snowflake has built. The company has earned HIPAA and FedRAMP compliance certifications, improving security. That's important for all companies, but particularly those that handle sensitive information. The fact that Snowflake has signed Capital One, Experian, and DocuSign is indicative of the reliable reputation they have earned in the market.

## Management team

As discussed in "Snowflake's History," co-founders Dageville, Cruanes, and Zukowski are all still with the company. But while Snowflake's success began with the three co-founders' vision and technical expertise, Snowflake has blossomed under the leadership of external CEOs, chiefly Bob Muglia (2014-2019) and Frank Slootman (2019-present).

Slootman most recently served as CEO of ServiceNow from 2011 to 2017 and is an operations guru known for being something of an IPO specialist after ushering ServiceNow, Data Domain, and now Snowflake into the public markets. That's been achieved by focusing on boosting sales and trimming corporate expenditures. Slootman's track record at his previous postings is impeccable. ServiceNow grew revenue from \$100M to \$1.4B in revenue during his tenure, while Data Domain was guided to a \$2.4B acquisition from EMC.

From an incentive perspective, executive compensation is mostly cash salary with both an equity and non-equity component primarily determined by the compensation committee. Specific KPIs are largely lacking:

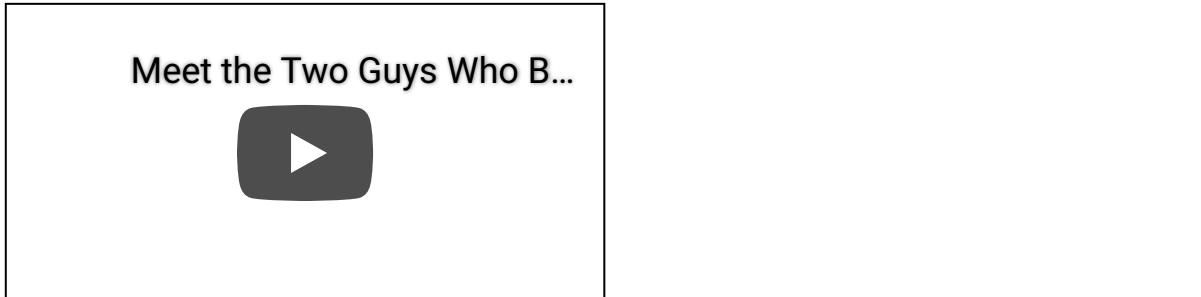
Name and Principal Position	Salary	Option Awards <sup>(1)</sup>	Stock Awards	Non-Equity Incentive Plan Compensation <sup>(2)</sup>	All Other Compensation	Total
Frank Slootman <sup>(3)</sup> <i>Chief Executive Officer</i>	\$ 287,990	\$ 59,874,582	\$ —	\$ 307,886	\$ 233 <sup>(4)</sup>	\$ 60,470,691
Robert L. Muglia <sup>(5)</sup> <i>Former Chief Executive Officer</i>	72,885	—	16,652,753 <sup>(6)</sup>	—	315,772 <sup>(7)</sup>	17,041,410
Michael P. Scarpelli <sup>(8)</sup> <i>Chief Financial Officer</i>	138,461	20,157,901	—	150,730	146 <sup>(4)</sup>	20,447,238
Benoit Dageville <i>President of Products</i>	292,707 <sup>(9)</sup>	1,836,642	—	70,308	887 <sup>(4)</sup>	2,200,544

[Snowflake's S-1 filing](#)

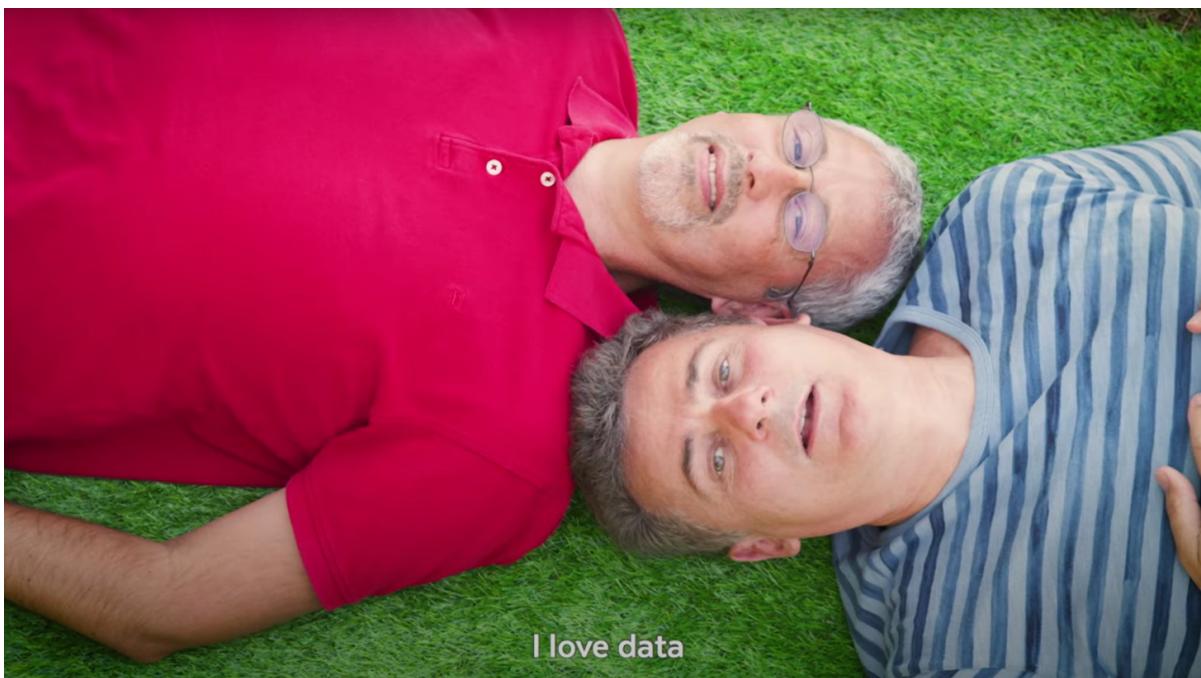
It's interesting to see that Slootman (5.9% holds more of the company than founder Dageville (3.9%). Along with Muglia (3.3%), those three men make up the largest affiliated equity holders. SHV leads external investors with 20.3%.

## Snowflake's history

Though Slootman and Spieser are central to this filing's tale, they are not where the adventure begins. To understand Snowflake's inception and its culture, there's only one place to start: [the first 0:25 of this video](#).



If this screenshot of founders Benoit Dageville and Thierry Cruanes doesn't entice you, we're not sure what will.



Benoit and Thierry are kooky, playful, anything but boring. In that respect, they aren't what you'd picture when asked to imagine two long-time data architects.

Back in the summer of 2012, these jesters left their presumably-high-paying-and-cushy Oracle jobs. They teamed up with Marcin Zukowski, the former founder of Vetcwise, a database management company. The trio had a big idea: build a cloud-native data warehouse from the ground up.

To be clear, this is probably the *least* casual of ambitions one could have. Building a successful company as an entrepreneur is a daring aspiration in and of itself, but the complexity and scope of Benoit and Thierry's challenge is hard to overstate. For those that aren't intimately familiar with enterprise data infrastructure, this isn't your mother's mobile app. This is even far from standard B2B SaaS. Data warehouses are, well, absolute *behemoths*.

But the trio knew what they were doing. Benoit and Thierry had worked on legacy architectures at Oracle, understood the drawbacks of the then-newly-hyped Hadoop, and knew there had to be a better way. Between them, they already held over 120 patents — this was not a team light on experience. Still, the task was formidable. The team built, often beneath a veil of secrecy, emerging from stealth with something unique, one-of-a-kind. A Snowflake, if you will.

Interestingly, none of the founders have ever served as CEO, though they remain key leaders: Benoit acts as President of Products, while Thierry is CTO. Marcin may have taken a step back — [his LinkedIn](#) identifies him as a "co-founder."

Name of Beneficial Owner	Shares	%	% Total Voting Power Before this Offering <sup>(1)</sup>
<b>Named Executive Officers and Directors:</b>			
Frank Slootman <sup>(2)</sup>	15,213,149	5.9 %	5.9 %
Robert Muglia <sup>(3)</sup>	8,084,086	3.3 %	3.3 %
Michael P. Scarpelli <sup>(4)</sup>	4,597,266	1.9 %	1.9 %
Benoit Dageville <sup>(5)</sup>	8,360,000	3.4 %	3.4 %
Jeremy Burton <sup>(6)</sup>	477,546	*	*
Teresa Briggs <sup>(7)</sup>	50,000	*	*
Carl M. Eschenbach <sup>(8)</sup>	—	—	—
Mark S. Garrett <sup>(9)</sup>	901,204	*	*
Kelly A. Kramer <sup>(10)</sup>	8,333	*	*
John D. McMahon <sup>(11)</sup>	1,237,110	*	*
Michael L. Speiser <sup>(12)</sup>	49,564,848	20.3 %	20.3 %
Jayshree V. Ullal <sup>(13)</sup>	50,000	*	*
All directors and officers as a group (13 persons) <sup>(14)</sup>	90,198,644	33.8 %	33.8 %
<b>Other 5% Stockholders:</b>			
Entities affiliated with Altimeter Partners Fund, L.P. <sup>(15)</sup>	36,286,307	14.8 %	14.8 %
Entities affiliated with ICONIQ Strategic Partners III, L.P. <sup>(16)</sup>	33,752,048	13.8 %	13.8 %
Entities affiliated with Redpoint Ventures V, L.P. <sup>(17)</sup>	21,928,585	9.0 %	9.0 %
Entities affiliated with Sequoia Capital Growth Fund III, LP <sup>(18)</sup>	20,619,156	8.4 %	8.4 %
Entities affiliated with Sutter Hill Ventures <sup>(19)</sup>	49,564,848	20.3 %	20.3 %

Snowflake's S-1 filing

## Investors

When unpacking Snowflake's investor base, there's only one place to begin: SHV, venture capital's silent assassins.

It's hard to find much information about the firm, despite having been in business since 1962. When you navigate to [Sutter's website](#), you're met by a wall of blue, a logo, and an address. That's it. In a funding environment in which firms pay escalating amounts to brand themselves and establish an online presence, SHV's minimalism is a consummate flex. (See also: [Benchmark](#)).

# SHV

ESTABLISHED 1962

Sutter Hill Ventures  
Bldg A 200, 755 Page Mill Rd  
Palo Alto, California 94304

© 2020. All rights reserved.

[Sutter Hill's website](#)

Whatever game the market, SHV isn't bothered. They know what they're doing, and firms are unlikely to attain such staying power without success. According to Crunchbase, the firm has made [287 investments](#) with 81 exits along the way. That includes familiar names like Sumo Logic (another impending IPO), GlassDoor, Smartsheet, and Yext.

But perhaps more interesting than SHV's portfolio is their approach. While we've seen new venture funds come to market in the last couple of years outlining a "studio" approach, SHV has played the game for at least a decade. In an interview with Strictly VC, Managing Partner Sam Pullara describes the firm's practice of "origination" — building companies in-house with external entrepreneurs. This is a strategy that runs counter to the narrative that funds can't incubate enduring businesses.

"The best example of what we do is Pure Storage, which my [Sutter Hill] partner Mike Speiser helped start when he was leaving Yahoo to come here. He worked with [founder and CTO] John Colgrove for eight months, trying to figure out the best company to start in flash storage. John started it, and Mike joined as interim CEO as John built out the team. Then Aneel Bhusri of Greylock [Partners] did the Series B, and the rest of that is going amazing."

Though details are scant, Snowflake seems to have been built with the same approach. Indeed, it's hard to overstate just how involved SHV appears to have been in its success. From what we understand, Speiser helped the original architects [design the commercial organization](#) around the product, with Snowflake's sales and go-to-market teams running through SHV until recently.

That hard work has paid off. At IPO, SHV is the largest external shareholder with a 20.3% stake in Snowflake. They're followed by Altimeter Partners (14.8%), ICONIQ Partners (13.8%), Redpoint Ventures (9.0%), and, inevitably, Sequoia Capital (8.4%).

As mentioned earlier, Snowflake has raised a total of \$1.4B on its path to IPO, illustrating that while costs are relatively low to *start* software businesses, scaling is expensive. Critical rounds include the \$5M Series A led by SHV in 2012, a \$26M Series B round led by Redpoint in 2014, and the monster \$450M Series F in late 2018, spearheaded by Sequoia. As part of that round, enterprise software veteran Carl Eschenbach — formerly of VMWare and involved with the firm's investments in UiPath and Gong — took a board seat.

There's one final investor worth noting: Warren Buffett. The Oracle of Omaha is expected to plow [\\$570M into Snowflake](#), investing \$250M in a private placement at \$80 per share. He's agreed to purchase a

further 4M shares from former CEO Muglia at the IPO price, which could cost between \$303 and \$344M. Salesforce Ventures is also grabbing \$250M as part of the private placement.

Given Berkshire Hathaway's reticence to invest in tech companies, particularly those operating at a loss like Snowflake, this show of confidence represents something of a coup.

## Financial highlights

Ask an investor to wave a wand and create a business, and it might end up looking a lot like Snowflake. There's plenty to admire in the company's performance to date. With that in mind, we're going to focus on three fundamental pieces of the business: revenue growth, net revenue retention, and the margin profile.

### Revenue growth

There are no two ways of putting it: Snowflake is growing *fast*. As noted under "Snowflake's History," revenue expanded 173% over the last two fiscal years and 132% when comparing the six months leading up to July 31, 2019, with the same period in 2020. Some sources have suggested that Snowflake's trajectory makes them [one of the fastest-growing](#) software companies at IPO ever.

This has been achieved at a significant scale. As noted earlier, Snowflake logged \$264M in its 2020 fiscal year, up from \$96.7M.

### Net revenue retention

Customers are using Snowflake more and more. Thanks to the usage-based pricing model, Snowflake has had success "landing and expanding" with their clients. Their most recent six month period showed a net retention rate of 158%, meaning that even accounting for churn, customers spent 58% more than they had compared to the same period in 2019. Remarkably, Snowflake's net revenue retention stood at 223% the year prior.

This [compares favorably](#) to several public software companies, including Datadog (146%), Slack (138%), and Slootman's old haunt, ServiceNow ([~130%](#)).

### Moderate margins

If there's one weakness in Snowflake's financial performance beyond the decision to eschew profitability, it's the gross margin profile.

Unlike many public software companies, Snowflake can't brag about 75-90% gross margins. Because of where the company sits in the stack, they're reliant on public cloud providers like AWS, Azure, and GCP. That's demonstrated graphically in the S-1 filing.



As a result, gross margins are comparatively modest, around 60%. This represents an improvement in performance in prior years, which saw margins of [44-45%](#).

Part of that is due to increased negotiating power with those public cloud providers. From the S-1:

"Our improved gross margin during the last four quarters presented is primarily attributable to higher volume-based discounts for purchases of third-party cloud infrastructure, increased scale across our cloud infrastructure regions, and improved platform pricing discipline."

## Valuation

What is Snowflake worth? This is the (multi)billion-dollar question for Slootman, SHV, and Buffett himself.

As background, Snowflake last raised capital in February of this year at a \$12.4B valuation. There are whispers \$ SNOW's debut may see the business priced north of \$20B, with \$30B within the realm of possibility given the scalding hot IPO-market. That's quite a price for a company yet to reach an annual recurring revenue rate (ARRR) of \$1B.

As referenced in "Market," Snowflake considers its TAM to be \$81B, with additional upside possible should data sharing become a more significant part of its business. The S-1 notes that the "data sharing opportunity" has "not been defined or quantified by any research institutions" but considers it to be "substantial and largely untapped." Putting that to the side, should Snowflake IPO at a \$25B valuation, it would represent one of the industry's highest multiples relative to sales: 95x for the 2020 financial year and 52x for the six months ending July 31.

Does Snowflake deserve this premium?

Ultimately, the company is a best-in-class provider, growing quickly and operating in a rapidly expanding market. If we assume Snowflake's trajectory continuing with revenues reaching \$850M by the 2022 financial year, a sales multiple looks more reasonable at 29x. That's in-line with other public SaaS companies, including Datadog, Zoom, and Coupa, all who see Enterprise Value (EV) to Next Twelve Months (NTM) sales at roughly 30x. More broadly, high-growth SaaS sees EV/NTM sales between 20-25x.

## Competition

The competitive landscape is one of the most exciting parts of the Snowflake story. As noted, the "frenemies" dynamic seems unique in the enterprise software landscape.

### The Big Three and beyond

To put it simply: Snowflake relies on the big cloud providers (AWS, GCP, Azure) for raw materials. Snowflake's product, in effect, is storage and compute. Without the big cloud providers, they don't have storage. Access is mission-critical.

Simultaneously, Amazon, Google, and Microsoft each have a directly competing product to what Snowflake offers. Redshift, BigQuery, and Synapse are going after the same customers that Snowflake is. Given those are three of the most successful, highly-capitalized businesses of all-time that's cause for concern. To date, Snowflake seems to have benefited by being the "Switzerland" of the space — staying neutral as public cloud providers engage in a slugging match. They've also capitalized on their first-mover advantage. But looking forward, it's not clear what barriers prohibit this "Big 3" from catching up.

In addition to the big cloud providers' SQL databases, the OG on-prem SQL data players remain. Oracle, SAP, and Teradata are striving to compete in this brave new (cloud) world.

Finally, it's worth considering a competitor of similar vintage to Snowflake. Databricks, founded in 2013, started as an extension of Apache Spark, initially focusing on the data science part of the analytics stack. Over time, they've encroached on Snowflake's domain, with their new Lakehouse product providing storage. While Snowflake's been squarely focused on storage (and compute) to date, the company has also suggested an interest in data science workflows. Databricks is a small company relative to the giants listed above, last valued at \$6B. But in five years down the line, we may see more robust competition as feature sets converge.

## Tracking market share

Given the space's competitive dynamics, tracking Snowflake's market share becomes increasingly important for those interested.

Though we've mentioned other KPIs above — notably revenue growth, margin profile, and net revenue retention, all relative to competitors — alternative metrics may provide a unique vantage. In particular, we think tracking "developer mindshare" may give a differentiated view on how Snowflake is trending relative to its rivals and is a potential leading indicator of future market share.

Leveraging publicly available data on Github, we were able to identify the frequency and velocity of code and repositories on specific technologies. The results are as of August 2020, summarized in the tables below:

Repositories	Snowflake	Hadoop	Cloudera	BigQuery	Synapse	Apache Hive	Yellowbrick	RedShift
7/31/2020	3,333	34,801	1,776	2,825	2,030	431	61	2,490
6/30/2020	3,247	34,427	1,749	2,727	1,967	424	51	2,412
3/31/2020	2,947	33,199	1,681	2,460	1,775	397	46	2,152
12/31/2019	2,631	32,192	1,626	2,237	1,671	380	42	2,006
9/30/2019	2,372	31,020	1,553	1,991	1,575	352	36	1,855
6/30/2019	2,186	29,975	1,494	1,777	1,475	335	35	1,640
3/31/2019	1,992	28,679	1,421	1,596	1,374	319	31	1,482
12/31/2018	1,774	27,418	1,346	1,441	1,292	297	28	1,397
9/30/2018	1,610	25,981	1,240	1,299	1,200	282	26	1,298
6/30/2018	1,490	24,448	1,172	1,146	1,129	258	24	1,197
3/31/2018	1,356	22,739	1,067	1,019	1,039	231	18	1,091
12/31/2017	1,210	21,006	985	912	968	209	17	1,009
9/30/2017	1,114	18,826	879	782	868	202	16	901
6/30/2017	1,034	17,081	784	694	806	185	14	828
3/31/2017	930	15,276	693	610	736	170	14	743
12/31/2016	813	13,814	607	548	663	150	14	663
9/30/2016	716	12,459	541	480	604	133	12	583
6/30/2016	644	11,253	468	427	548	118	12	504
3/31/2016	585	9,920	413	365	495	100	11	432
12/31/2015	484	8,817	361	307	426	88	11	384
9/30/2015	428	7,867	324	254	382	80	11	321
6/30/2015	397	7,056	292	215	340	72	9	266
3/31/2015	362	6,082	252	159	291	67	8	219
12/31/2014	290	5,388	211	130	267	60	7	166

YoY Chg	Snowflake	Hadoop	Cloudera	BigQuery	Synapse	Apache Hive	Yellowbrick	RedShift
7/31/2020								
6/30/2020	49%	15%	17%	53%	33%	27%	46%	47%
3/31/2020	48%	16%	18%	54%	29%	24%	48%	45%
12/31/2019	48%	17%	21%	55%	29%	28%	50%	44%
9/30/2019	47%	19%	25%	53%	31%	25%	38%	43%
6/30/2019	47%	23%	27%	55%	31%	30%	46%	37%
3/31/2019	47%	26%	33%	57%	32%	38%	72%	36%
12/31/2018	47%	31%	37%	58%	33%	42%	65%	38%
9/30/2018	45%	38%	41%	66%	38%	40%	63%	44%
6/30/2018	44%	43%	49%	65%	40%	39%	71%	45%
3/31/2018	46%	49%	54%	67%	41%	36%	29%	47%
12/31/2017	49%	52%	62%	66%	46%	39%	21%	52%
9/30/2017	56%	51%	62%	63%	44%	52%	33%	55%
6/30/2017	61%	52%	68%	63%	47%	57%	17%	64%
3/31/2017	59%	54%	68%	67%	49%	70%	27%	72%
12/31/2016	68%	57%	68%	79%	56%	70%	27%	73%
9/30/2016	67%	58%	67%	89%	58%	66%	9%	82%
6/30/2016	62%	59%	60%	99%	61%	64%	33%	89%
3/31/2016	62%	63%	64%	130%	70%	49%	38%	97%
12/31/2015	67%	64%	71%	136%	60%	47%	57%	131%

Taken together, Snowflake shows steady growth in developer mindshare, though BigQuery is accelerating more rapidly. RedShift shows relatively similar growth, while Hadoop, Cloudera, Synapse, and Apache Hive are in relative deceleration. Those interested in the stock may want to consider tracking this information going forward.

## The cautionary tale of Hadoop and Cloudera

While all looks rosy for Snowflake at the moment, observers of the space may recall the (relative) demise of similarly hyped business: Cloudera, MapR, and Hortonworks. Swift technological advancements have the potential to benefit insurgents and deposition incumbents.

After its release in 2006, Apache's Hadoop received significant buzz through the early 2010s. Hadoop was an open-source ecosystem, and for a period, the best data warehousing solution for on-prem databases. A trio of companies emerged off the back of Hadoop, serving as vendors. Cloudera, Map R, and Hortonworks helped companies leverage Hadoop's features to create high-volume data management strategies.

That trio dominated the on-prem traditional data warehouse industry before they faced a classic Innovator's Dilemma. As they focused on servicing their existing market, they failed to devote time and resources to the next frontier: the cloud.

Over time, all three lost share to Snowflake and the offerings of the Big 3. Unable to compete architecturally, the trio faced increasing pressure exacerbated by strategic mistakes. In 2019, Cloudera and Hortonworks merged in an all-stock deal. Though Cloudera now trades on the NYSE, it's \$3.9B is a fraction of what it might have been when seeing what Snowflake has managed to achieve. The same year as the Cloudera merger, Hewlett Packard acquired MapR for [less than \\$50M](#), a far cry from the \$280M raised, and believed to be below its revenue.



&

CLOUDERA

- **Best at:** Open-source ecosystem that focuses on dynamic data set analytics
- **Weak at:** managing data that is shared across an enterprise – Hadoop is not a database, has no data safety or workload management.
- **Downfall:** Hadoop took a short-sighted approach and marketed itself as an alternative to a data warehouse (a new market) instead of focusing on its advantage: evolving big data and demonstrating its value
- **Nail in the coffin:** “object storage” innovation, which launched in the cloud and provided cheap storage and support for flexible data types (Hadoop’s area) for 3x cheaper

- **Best at:** last of the original three (MapR and Hortonworks) big data cos that built products and services based off of Hadoop
- **Weak at:** difficulty of monetizing a business based off of open-source software
- **Downfall:** inability to build a competitive cloud product
- **Nail in the coffin:** the migration of big data storage to the cloud reduced demand for on-prem, which was Cloudera’s main market. Cloudera is now focusing on machine learning applications and hybrid cloud

Snowflake was the beneficiary of the evolution of data warehousing towards the cloud, but like all things tech-oriented, it is susceptible to disintermediation via innovation. Today, no threat appears imminent. But mindful investors will remember that the backers of Cloudera, Hortonworks, and MapR may have once felt the same.

## Why now?

There's a rumor that circulated in May of last year. In an interview with [The Information](#), then-CEO Bob Muglia made an innocuous comment. When asked when Snowflake would go public, he answered that it would be at least a year, maybe more.

"We have enough cash and don't need to go public," he said.

Two weeks later, he was ousted. The man brought in to replace him — Slootman, of course — was a different profile of leader. A disciplined, aggressive CEO with experience taking companies public, having done so twice already. Snowflake is his chance to "three-peat."

All to say that Snowflake has been mulling a listing for some time, and put in place a leader well-suited to make it happen in relatively short order. In [October 2019](#), just five months into his tenure, Slootman stated that he expected Snowflake to go public early summer in 2020 or after the presidential elections.

Both the favorable market conditions and coronavirus tailwinds seem to have altered that calculation slightly. Rather than waiting for late November or December, Snowflake is capitalizing on the sympathetic market conditions to list this fall.

While investors' appetite for cloud stocks is undoubtedly a large part of the calculus, there may be another element. As mentioned earlier, Snowflake has executed at breakneck speed. At the back of Slootman and Co.'s mind may be a question: can we keep it up? It's hard to find fault with Snowflake's metrics at the moment, but as competition heats up, that may not always be the case.

If you enjoyed our analysis, we'd very much appreciate you sharing with a friend. We'd also love to have you join us for tomorrow's event.

Share

[Join the event](#)

