

Data-driven VC Landscape 2023



By Andre Retterath
datadrivencv.io

In partnership with



affinity



Synaptic



VAUBAN
from Carta

VESTBERRY

Why read this report?



Market overview

Get a clear and transparent overview on the state of digitization in venture capital



What leaders do

Learn more about leading data-driven VCs, size of their engineering teams, focus across the value chain, preferred tools, and more



Practical guidance

Solve your “buy versus build tradeoff” with a comprehensive tech stack overview, step-by-step guides, thought leaders and content to follow

“By 2025, more than 75% of VC investor reviews will be informed using AI and data analytics.”

- Gartner

Foreword.

“ Following the launch of my [Data-driven VC newsletter](#) in September 2022, I’m excited to share the first episode of the “Data-driven VC Landscape” as the next building block to create more transparency around the state of digitization in Venture Capital.

This report sheds light on the leading data-driven VC firms and the people behind them driving the revolution, to understand the specifics of their setups, focus across the value chain, preferred tools, and a lot more. Beyond novel insights, it contains actionable guides and content to become more data-driven yourself.

Enjoy & stay driven!”



Andre Retterath
datadrivencv.io
Earlybird VC

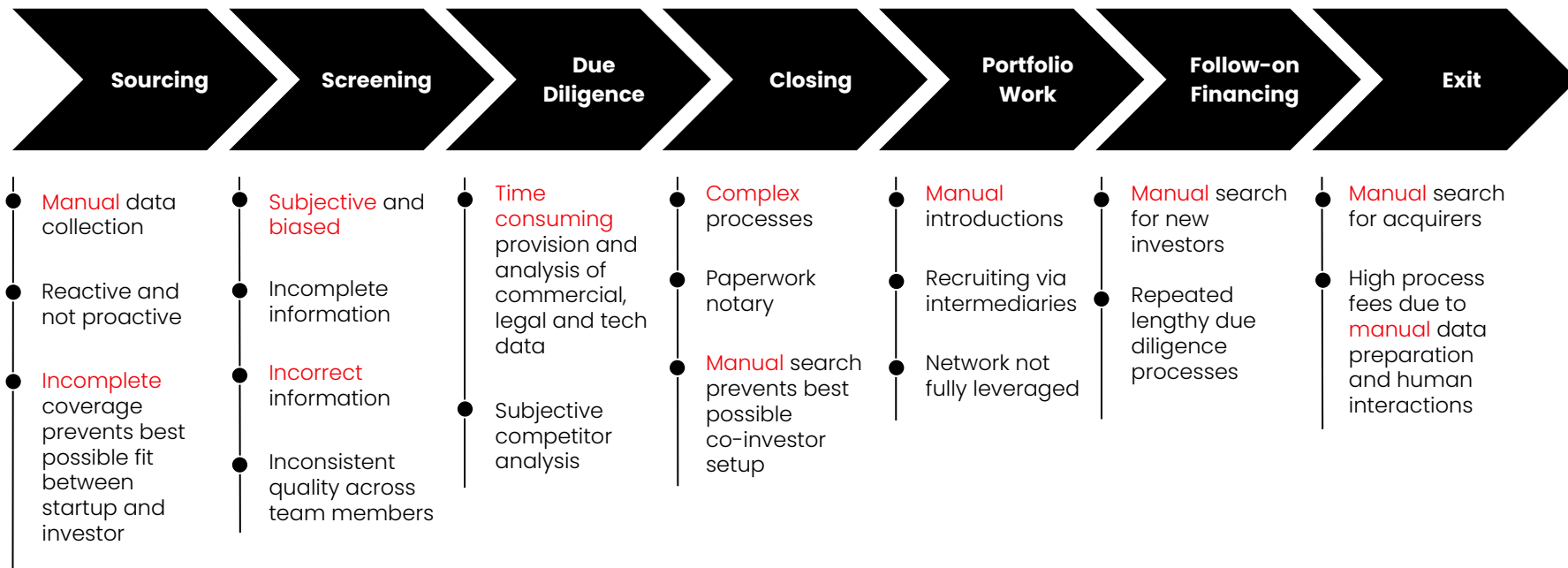
Table of Contents

- 1 | Why should VCs become more data-driven?**
- 2 | Market benchmarking**
- 3 | VC execution – how do you get there?**

1 | Why should VCs become more data-driven?

Why should VCs become more data-driven?

The VC investment process is broken.



It's manual, inefficient, non-inclusive and biased.

Why should VCs become more data-driven?

VC firms can solve three main **pain points** with data-driven approaches and AI.

1

Efficiency

Always busy, distracted and **unable to focus on actual value drivers.**

2

Effectiveness

Miss out on **outlier opportunities** due to incomplete coverage and wrong prioritization.

3

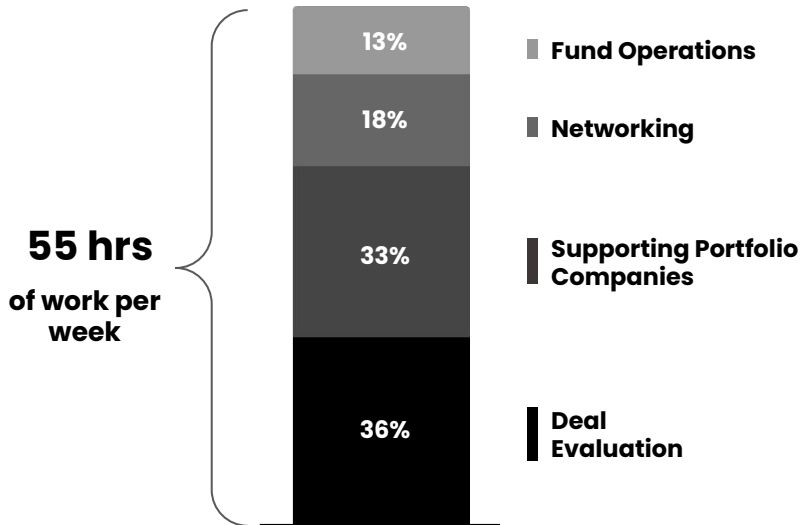
Inclusiveness

Biased and exclusive allocation of capital leads to suboptimal utilization of founder potential.

Why should VCs become more data-driven? **Efficiency**

VCs are always busy.

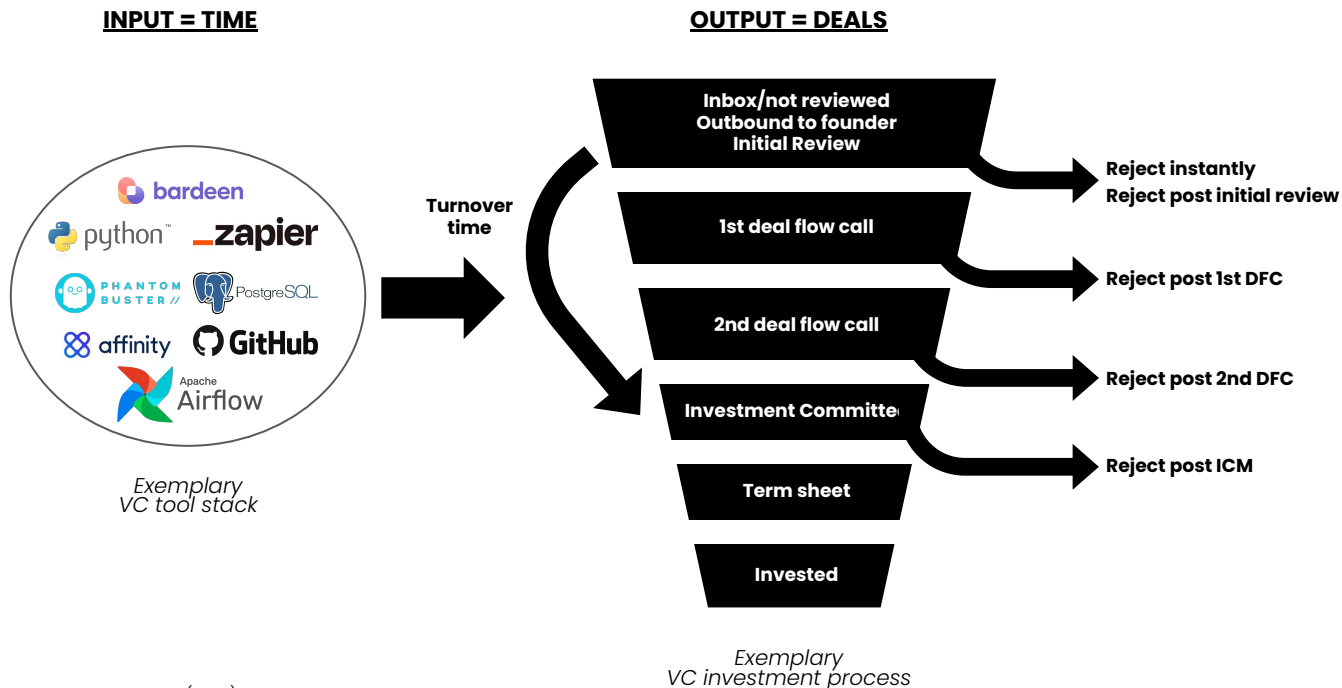
How venture capitalists spend their time
% of time spent



Source: Actual screenshot of my calendar mid of March 2023, blurred for privacy. HBR [survey on 900 VCs](#), TechCrunch [survey on 100 VCs](#).

Why should VCs become more data-driven? **Efficiency**

Leverage AI and modern tools to **increase input-output ratio.**



Productivity effect of using AI

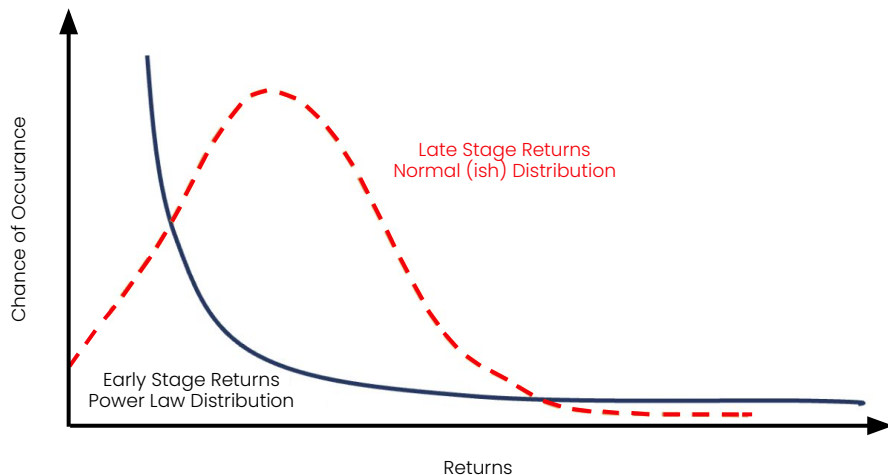
- **Time to perform tasks decreases** by 0.8 SDs
- **Output quality increases** by 0.4 SDs
- **Inequality decreases** as AI allows lower ability teams to achieve more

Why should VCs become more data-driven? **Effectiveness**

VC returns are power-law distributed, there is a large cost in missing outliers.

VC fund return distribution of early-stage vs late-stage

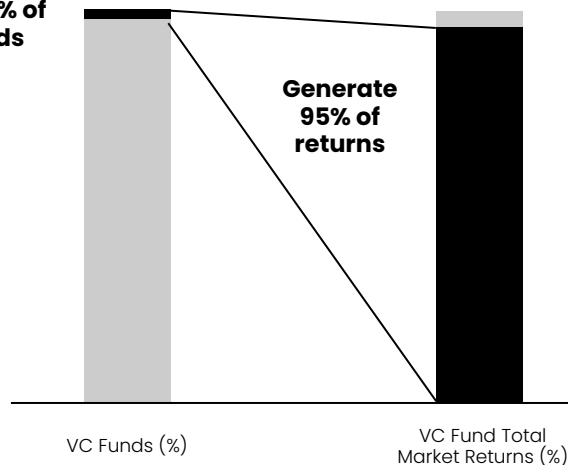
Chance of occurrence



VC fund returns

% returns

Top 2% of funds

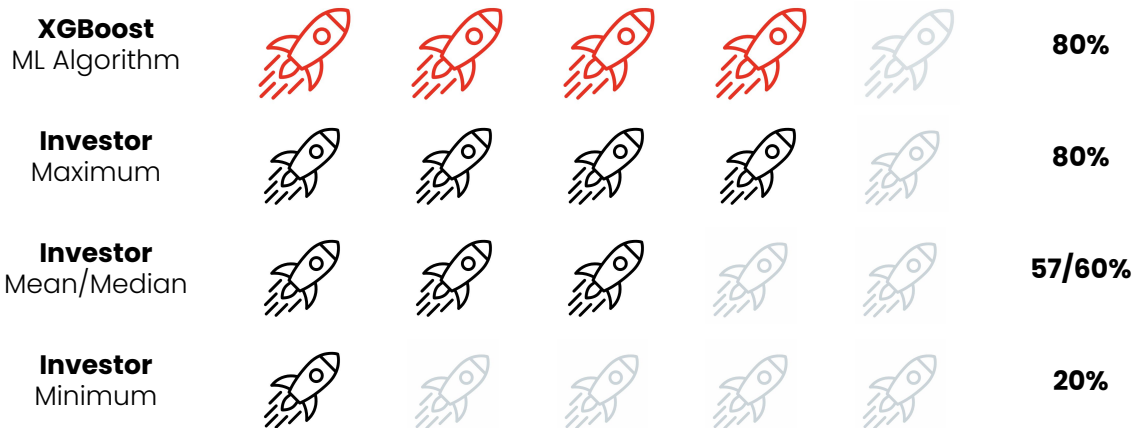


Why should VCs become more data-driven? **Effectiveness**

Machine learning models are proven to outperform human investors in screening.

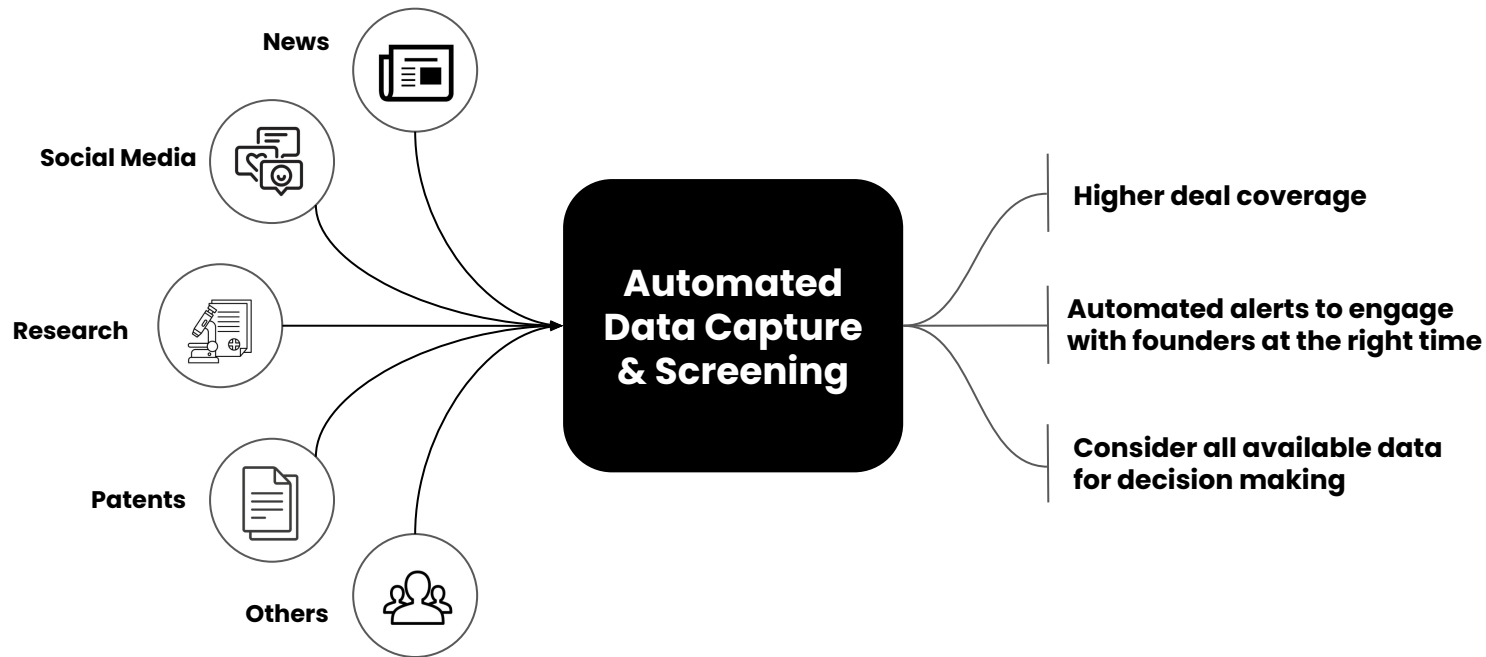
Human vs Computer

% Performance in Accuracy and Recall



Why should VCs become more data-driven? **Effectiveness**

Achieve more comprehensive deal coverage and **reduce miss rates.**



Why should VCs become more data-driven? **Inclusiveness**

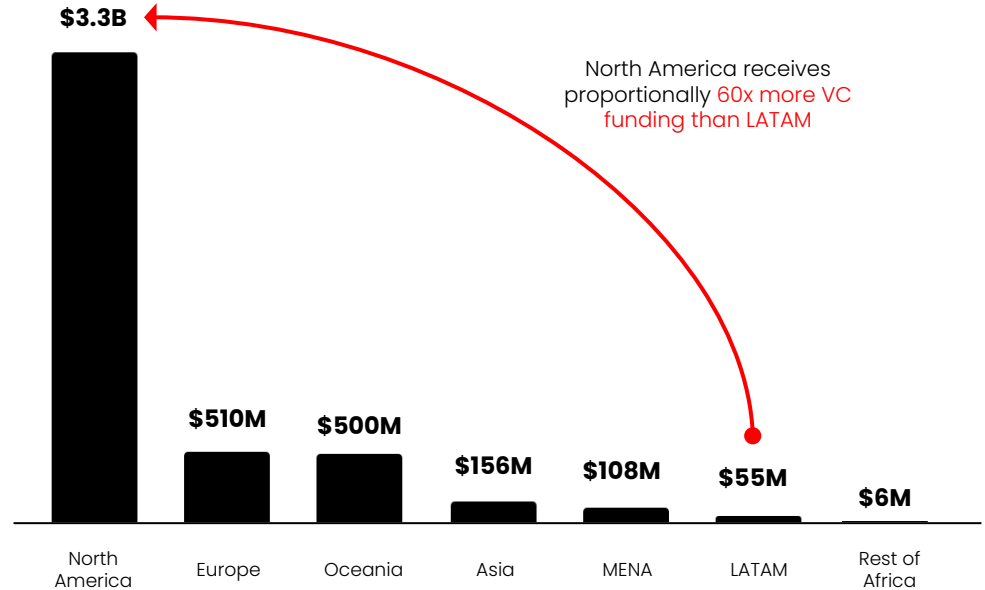
Talent is distributed equally, **capital** is not. There is a huge **disparity** in funding across the world.

Venture capital is an instrument to allow highly talented and ambitious people realize their potential.

But circumstances are the defining criteria for whether someone will become an inventor.

- White children are **3x more likely** to become inventors than black children.
- **Less than 20%** of inventors are female, and differences in ability do not explain the disparity.
- Top performing children are much more likely to become inventors, but **only if they come from high-income families.**

VC funding by region in the world
2017-2022 funding per 1M people

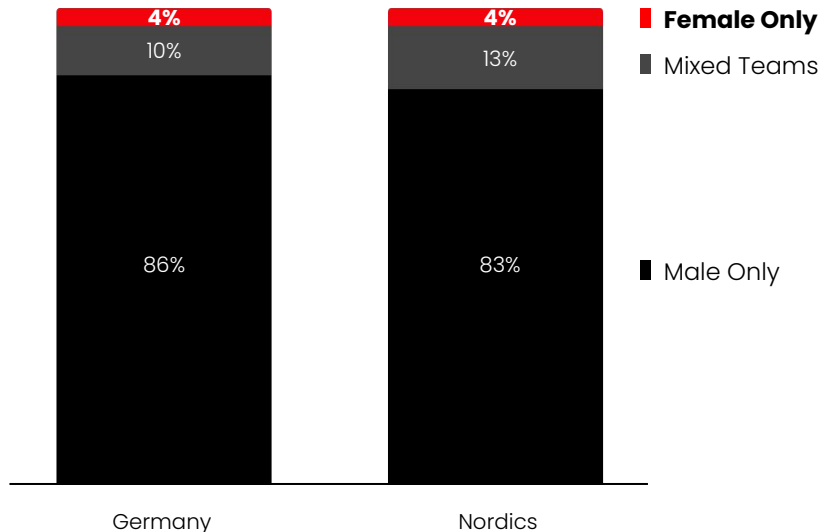


Why should VCs become more data-driven? **Inclusiveness**

Data-driven approaches can **reduce bias** and the inequality gap in startup funding.

Gender inequality in startups across Europe

% of startup founder team composition



- Venture capital **funding decisions are male-driven**, only 10% of VC decision makers are female in the industry. Opportunities for female owned companies were only 2.2% of global VC funding.
- This is despite evidence which shows that companies with **female founders and leaders outperform male-owned startups**.
- Data-driven initiatives have been shown to help venture capital firms **reduce gender bias and make better, fairer investment decisions**.

“ I think we are approaching the tipping point where the majority would need some degree of data-driven culture to become and remain competitive.

So to some extent, it is moving from a nice-to-have to a must. And you need to be data-driven because everything is larger, faster, earlier; there are so many funds around and differentiation is hard to show sometimes, and finally because quality is increasing overall, and to win the best deals and allocate your capital wisely you can't simply trust your gut anymore.”

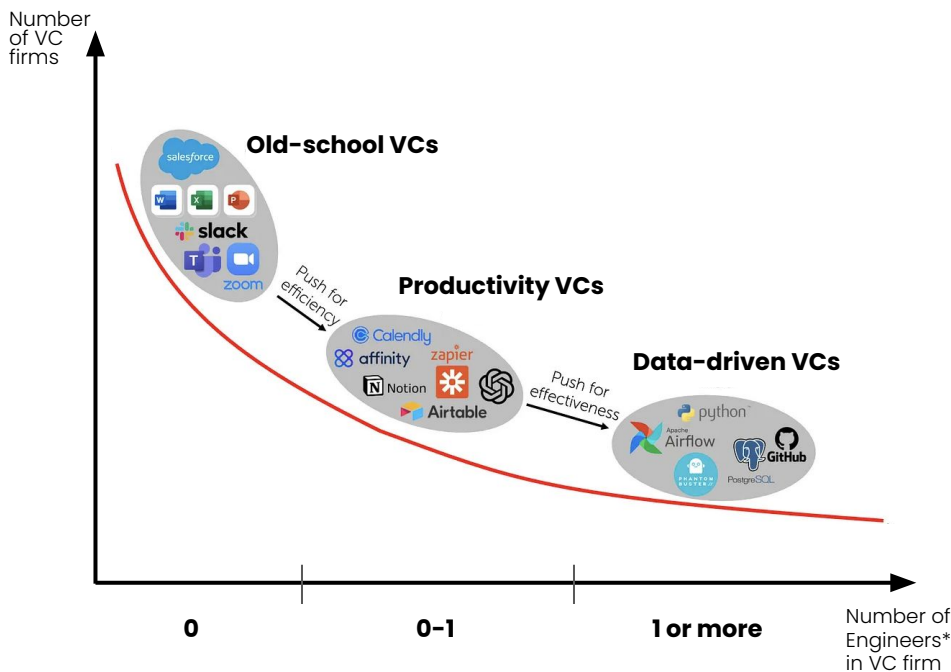


Francesco Corea
Director of Data Science



2 | **Market Benchmarking**

VCs evolve gradually from basic tools to adding their own scalable solutions.



Source: DDVC. Tech stack typically adds up and becomes more complex from left to right.
*Engineers summarizes developers and data positions of any kind, no investors included.

Old-school VCs

Focused on **manual workflows**, a simple tool stack with a CRM system like Salesforce, basic Email, Slack/WhatsApp for communication, etc.

Productivity VCs

Successfully took the first leap to **modern off-the-shelf tech stack** with VC-focused CRM systems like Affinity, automated workflows with Zapier, use Notion for knowledge sharing, etc.

Data-driven VCs

Develop their **own scalable solutions** to capture data, automate workflows and **bring the core of their business in-house**.

Data-driven VC landscape 2023.

The leading 151 data-driven venture firms are paving the way for new industry standards.

The leaders were ranked using 331 community nominations, number of engineers in the team, relative share of engineers in the firm and number of segments covered with internal tools across the value chain. Find the full list [here](#).

Data-driven VCs are defined by fulfilling all of the following conditions:

- At least one engineer in the team.
- At least one community nomination.
- Proven to develop internal tooling in at least one segment of the value chain.

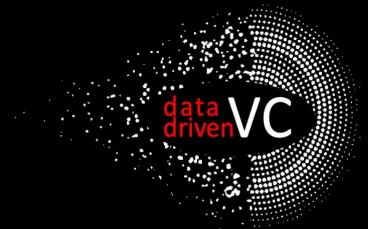
Top 20 Data-driven VC Firms

| | | | | | | | | | | |
|--|--|--|--|--|--|--|--|--|--|--|
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |

Data-driven VC Landscape 2024

Join the community and
participate in the next edition.

[Register Here](#)



Market Benchmarking

What does a data-driven VC look like?

The median data-driven VC firm:

2012

Founding Year

\$540M

AUM

43

Employees

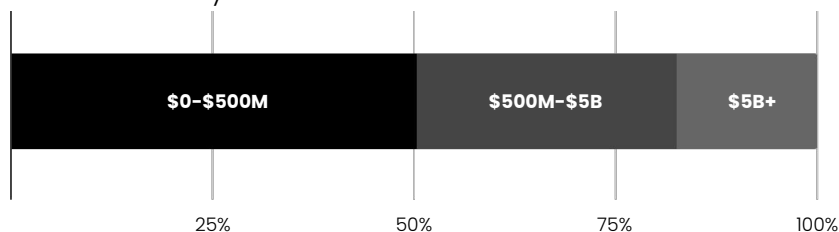
3

Engineers

Source: DDVC. Based on 151 data-driven venture firms.

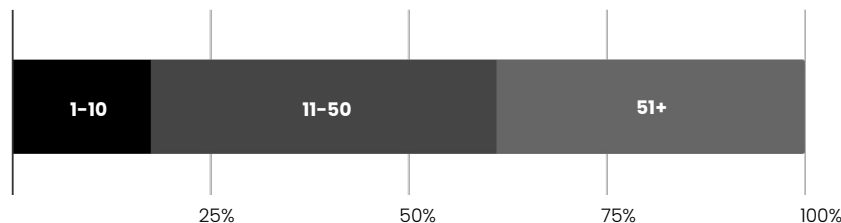
Fund proportion by AUM cohort

% of funds analyzed



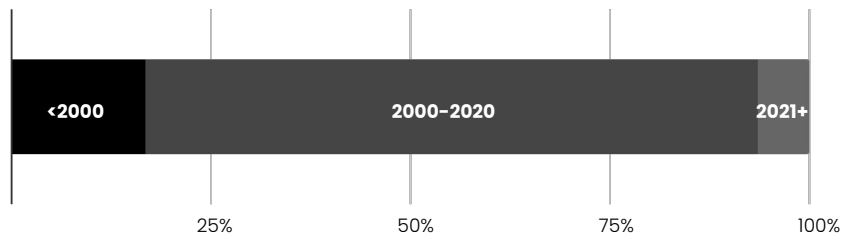
Fund proportion by # of employees cohort

% of funds analyzed



Fund proportion by founding year cohort

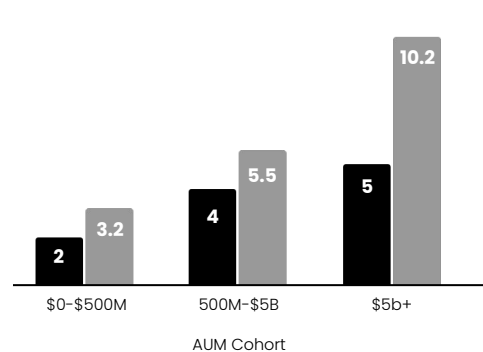
% of funds analyzed



Engineering team size by AUM

of engineers

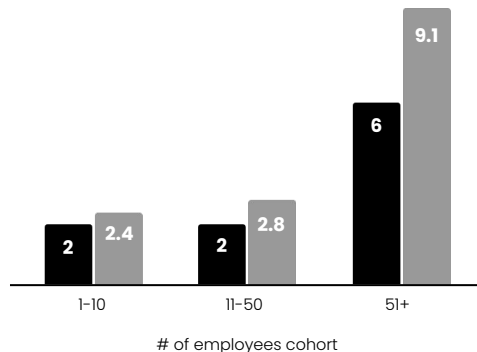
■ Median Engineers
■ Average Engineers



Engineering team size by # of employees

of engineers

■ Median Engineers
■ Average Engineers



Market Benchmarking

Efficient scalability: Engineering teams and AUM growth

VC assets under management (AUM) and number of engineers are exponentially correlated.

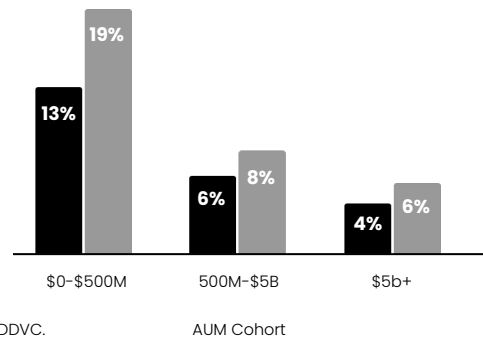
Two major learnings:

- 1) VCs need at least **one dedicated engineer to start** inhouse data-driven initiatives, independent of AUM and team size.
- 2) **Engineering activities scale well**, strongly increasing AUM per engineer in larger funds.

Engineering team size by AUM

% of engineers in relation to team

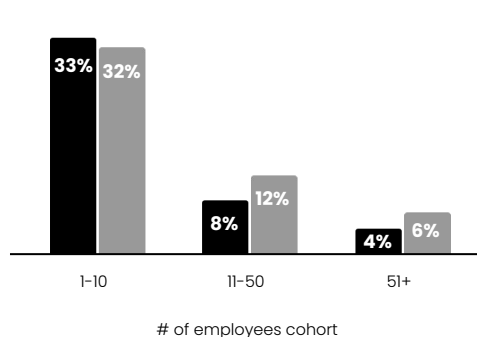
■ Median Engineers (%)
■ Average Engineers (%)



Engineering team size by # of employees

% of engineers in relation to team

■ Median Engineers (%)
■ Average Engineers (%)



“ *The data-driven VC space has matured a lot. As I meet more VC peers, I continue to be impressed with the innovations and craftsmanship of typically very small teams, sometimes even one-person bands. You see great data work ranging from extreme operational efficiency all the way to thesis matching at scale, forecasting or macro analysis of trends. That said, there is still a long way to go, not only by VCs that haven't started but, most importantly VCs that are not yet tapping their full potential.*

To truly differentiate, VCs must go deeper, and use first-principles thinking in combination with Data and AI techniques, but keep the genuine interest, personal touch and value add approach that characterizes great VCs.”

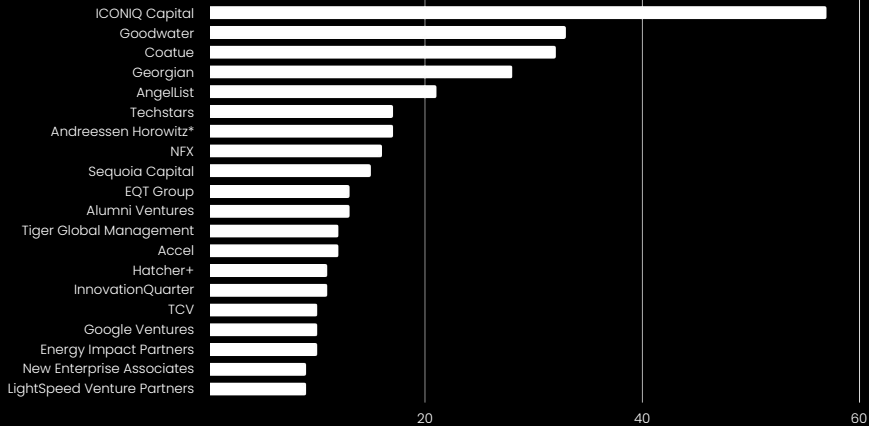


Pietro Casella
Chief Architect & MD

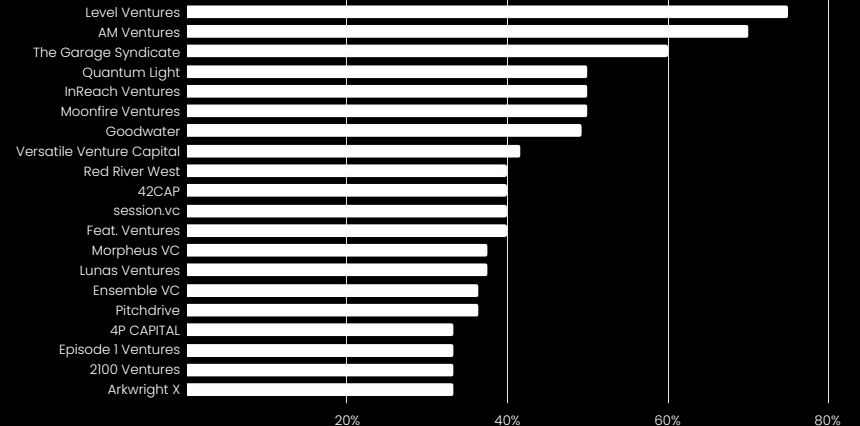


Top data-driven VCs by engineering team size and relative number of engineers in their firm.

By number of engineers

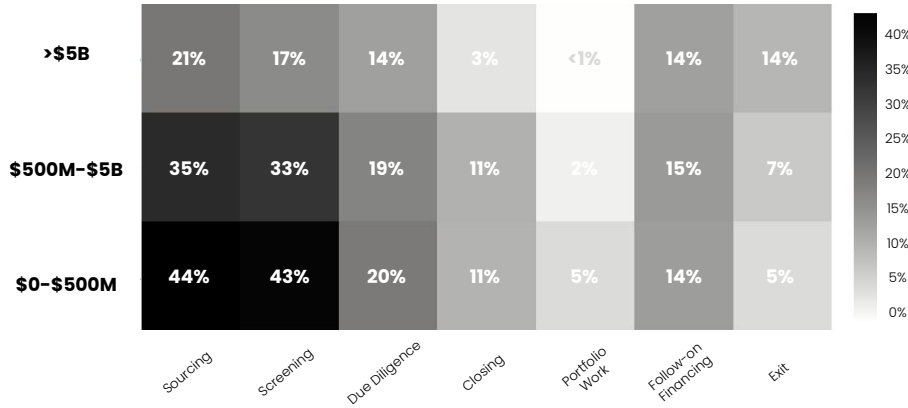


By % of engineers



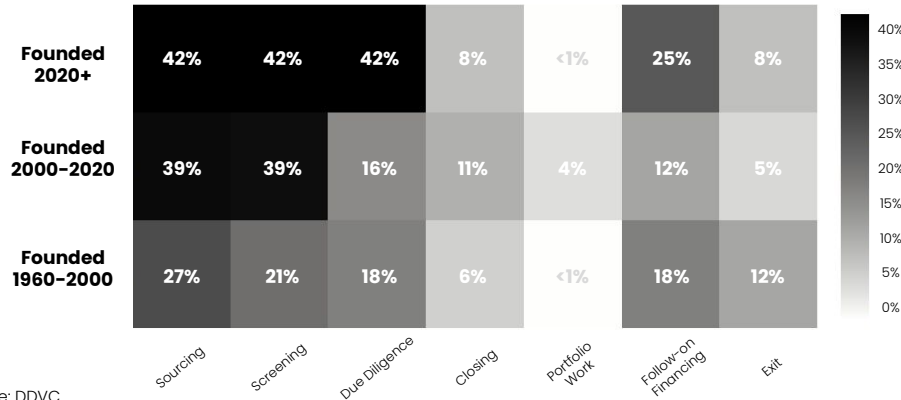
Where DDVCs act in the value chain by AUM

% of DDVC funds analyzed



Where DDVCs act in the value chain by age

% of DDVC funds analyzed



Source: DDVC.

Market Benchmarking

Data-driven VCs focus efforts on initial stages of the value chain.

Sourcing and screening are the most penetrated stages of the investment process, closely followed by portfolio value creation and due diligence.

Younger and smaller firms more broadly build tools across the value chain than their older and bigger peers. Newer firms benefit from greenfield situations, whereas established VCs face more friction due to migration requirements, process changes and cultural reluctance.

The impact of data-driven initiatives is expected to be stronger for younger and smaller firms than for more established and bigger peers.

“ *For a firm which has started and is beginning to see some added value, I'd say don't neglect the front end. We invested hundreds of hours of development at Red River West to make sure that our platform is web-based, works on desktop and mobile, is intuitive, easy to use, provides different levels of details on scoring, presents the data and insights in a visual way etc.*

By making sure that the tool is used by every single person within the organisation and not just the "IT minded ones", you create a virtuous circle where the tool is central to all processes and benefits from the input of all parties to be even better. "

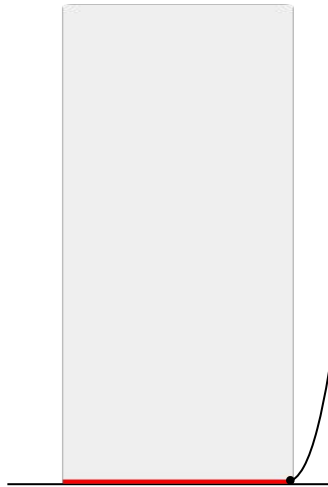


Olivier Huez
Partner



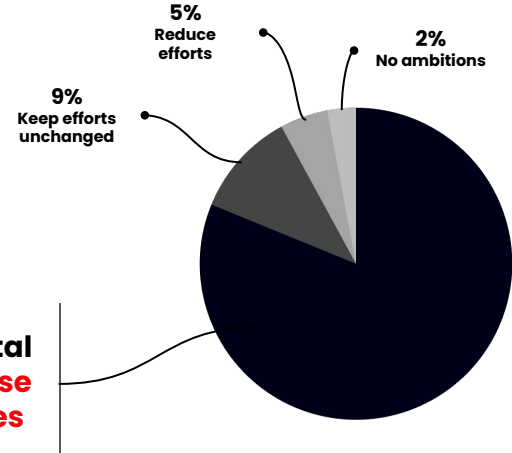
The data-driven revolution has just begun.

Data-driven VCs
% of total active VC firms



Only **1%** of venture capital firms **have internal data-driven initiatives** today.

Data-driven ambitions
% of total survey responders



84% of venture capital firms want to **increase efforts and resources**

People driving the data-driven VC revolution in 2023.

Top 20 Thought Leaders

| | | | | | | | | | |
|--|---|---|--|--|--|---|---|--|---|
|  Yohei Nakajima UNTAPPED CAPITAL |  Francesco Corea GREYCROFT |  Alexis Ohanian SEVEN SEVEN SIX |  Mike Arpaia MOONFIRE VENTURES |  Rob Kniaz HOXTON VENTURES |  Karl Lorey STARTUP RADAR |  Amir Sani TECHSTARS |  Chris Farmer SIGNALFIRE |  Dominik Dellermann VENCORTEX |  Marco Squarci SPECTER |
|  Olivier Huez RED RIVER WEST |  Pietro Casella EQT |  Stephane Nasser OPENVC |  Penny Schiffer RAIZED.AI |  Maximilian Fleitmann WIZARD VENTURES |  Vlastimil Vodicka LEADSPICKER |  Abel Samot RED RIVER WEST |  Jonas Vetterle MOONFIRE VENTURES |  Ali Tamaseb DCVC |  Henrik Landgren ARK CAPITAL |
|  Fabian Koenig PROPTTECHI |  Drake Dukess GRAVITY |  Andreas Göldi BTOVENTURE |  Eva-Valérie Gferer MORPHAIS |  Tom Wehmeier ATOMICO |  Niccolò Sanarico PRIMO VENTURES |  Daniel Sawko SHIP SHAPE |  Sarah Guemouri ATOMICO |  Anastasia Gonotskaya REDSTONE |  Ilya Kirnos SIGNALFIRE |
|  Bartosz Trocha SONARHOME |  Viet Le LA FAMILIA |  Margaux Wehr BALDERTON CAPITAL |  Atte Honkasalo NGP CAPITAL |  Marton Medveczky FLASHPOINT VC |  Abe Othman ANGELLIST |  Ali Kokaz ONE PEAK PARTNERS |  Adam Shuaib EPISODE 1 VENTURES |  Ties Boukema DAWN CAPITAL |  Cyril Parisot INDEX VENTURES |
|  Max Ruderman HARMONIC |  Gomathi Ramalingam LEFT LANE CAPITAL |  Dhiana Deva EQT |  Sina Sadegh WIZARD VENTURES |  Stefano Gurciullo REDSTONE |  Marc Felske REMUS CAPITAL |  Miguel Pinho SEEDCAMP |  Anton Ask Astrom EQT |  Mustafa Torun INVEST-NL |  João Nunes PLAYFAIR CAPITAL |
|  Jonas Sommer B2VENTURE |  Ben Smith INREACH VENTURES |  Lindsey Higgins PALE BLUE DOT |  Aaron Ford TCV |  Hem Wadhar SEQUOIA |  Nico Abanese ASCENSION |  Ernests Stals STARWATCHER.IO |  Eon Mattis INOVEBUS |  Sergei Mosunov WALE.IA |  Alessio Fanelli DECIBEL PARTNERS |
|  Stephan Wirries VENTECH |  Le Minh Nguyen KATAPULT |  Naiel Iqba TIGER GLOBAL |  Guy Conway KOBLE |  Francesco Farina MOONFIRE VENTURES | | | | | |

Source: DDVC. Thought leaders ranked using 158 community nominations, follower base and amount of relevant content in last 12 months. Data as of April 30, 2023

Data-driven VC thought leaders have extensive technical education.

3x more PhDs and **3x less MBAs** than VC industry average

Data-driven activities require a deeper and more thorough methodological skill set than the average VC investment job.

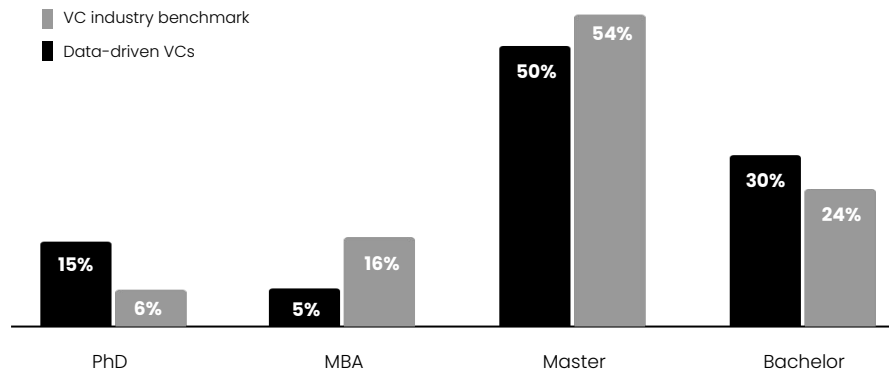
2x more STEM backgrounds than VC industry average

Many engineering, data analytics and computer science graduates have started to explore novel approaches to VC. Being at the forefront of tech innovation, they position themselves as the disrupters and thought leaders in their firms.

Source: Based on 65 thought leaders (DDVC) & 111 VC investors ([Retterath, 2020](#)).

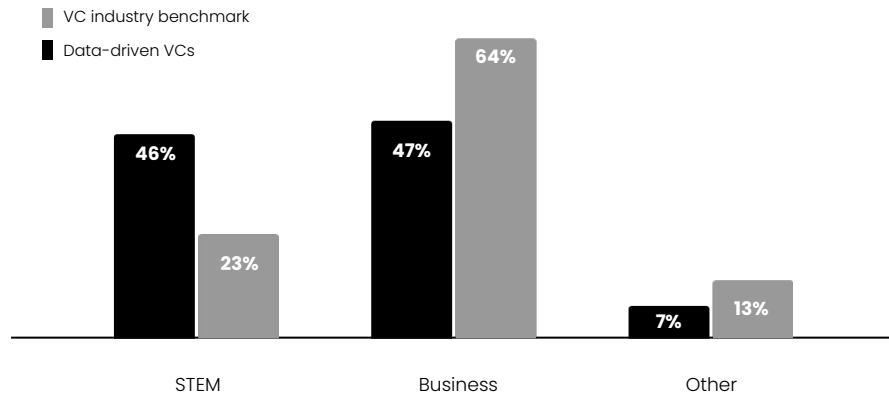
Thought leaders by academic title

% of DDVC thought leaders



Thought leaders by academic subject

% of DDVC thought leaders



“
Often times the first step that General Partners make when they decide that they want to get more quantitative is to hire someone to own this part of the firm. **Unfortunately, many people see these data roles at VC firms as an entry point to becoming a VC investor themselves.** This creates a scenario where the data people aren't spending as much of their time, effort, or focus on building large-scale, robust, data-driven infrastructure as they are on trying to network with others, talking to founders, etc.

Personally, I see myself as a **computer scientist working on venture capital as my subject matter.** I love the VC industry as subject matter but I wouldn't want to be a full-time investor. I'm a technologist.”



Mike Arpaia
Partner

MOONFIRE

3 | VC Execution – how do you get there?

VC Execution - how do you get there?

VC Tech Stack: Most used external tools by data-driven VCs.

 Synaptic

 Specter

 Proxycurl
By Nubeta

 SimilarWeb

 bright data

 PhantomBuster

**Alternative Data-Sources
and Lead Generation**

 crunchbase

 dealroom.co

 PitchBook.

 CBINSIGHTS

 S&P
CAPITAL IQ

 Tracxn

**Startup Databases
and Research**

 affinity

 VESTBERRY

 carta

 Airtable

 pipedrive

 HubSpot

**CRM, Portfolio and LP
Management**

 bardeen

 Calendly

 Notion

 SUPERHUMAN

 zapier

 OpenAI

Productivity

Find the compiled list
of all 400+ tools in the
table below:

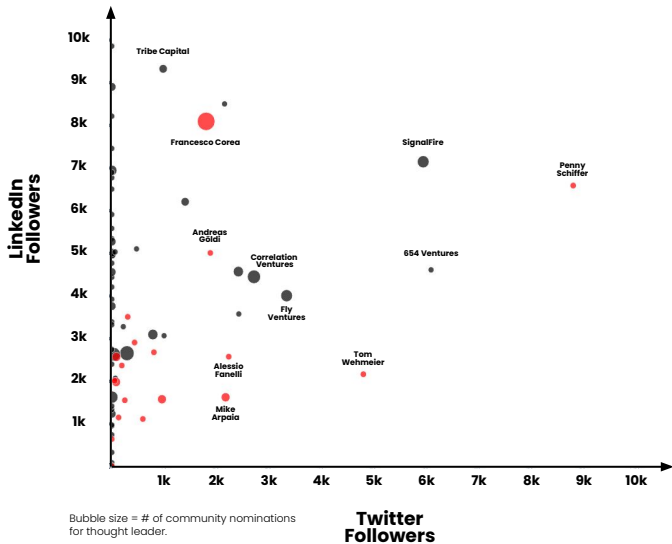
**[View Complete
VC Tech Stack](#)**

VC Execution - how do you get there?

Thought leaders share their insights with growing audiences across LinkedIn and Twitter.

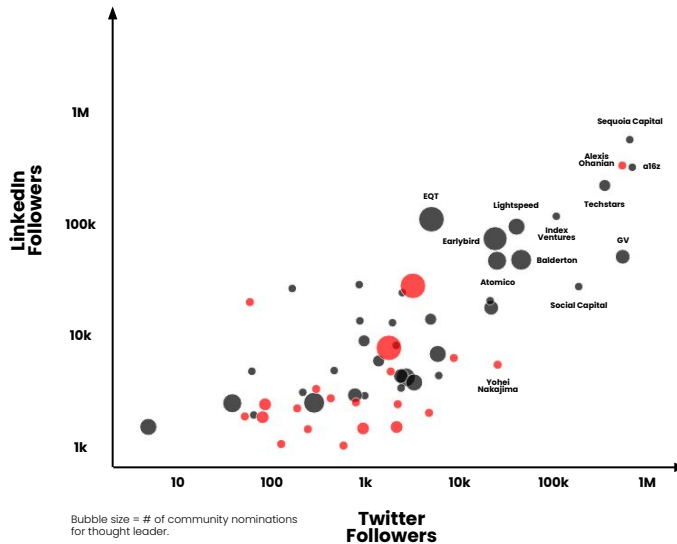
Thought leaders and data-driven VCs (Troposphere)

of Twitter and LinkedIn followers - linear scale



Thought leaders and data-driven VCs (Stratosphere)

of Twitter and LinkedIn followers - log scale



Follow the top 20 data-driven VCs and thought leaders below:

[View Top DDVC Leaders](#)

“*The industry has gone through three phases: first the initial exploration of using **data to source and evaluate startups** by some specialized VCs, more recently followed by more **VCs realizing the key value of data** but struggling to prove its added value and achieve positive signal to noise ratios, and finally, now using **data has become mainstream** as AI capabilities improve with the availability of large language models and autonomous agent technology.*”



Penny Schiffer
CEO & Co-founder



VC Execution - how do you get there?

Practical links.

List of all data-driven VC firms

| Rank | VC Firm | Location | Year Founded | Focus |
|------|---------------------|----------------|--------------|---------------------|
| 1 | Accel | Menlo Park, CA | 1999 | Enterprise Software |
| 2 | Andreessen Horowitz | Menlo Park, CA | 2009 | Enterprise Software |
| 3 | Benchmark | Menlo Park, CA | 1986 | Enterprise Software |
| 4 | First Round Capital | Menlo Park, CA | 2002 | Enterprise Software |
| 5 | GV | Menlo Park, CA | 2004 | Enterprise Software |

[Access Here](#)

List of all data-driven VC thought leaders

| Rank | Thought Leader | Location | Year Founded | Focus |
|------|--------------------------|----------------|--------------|---------------------|
| 1 | David Sacks | Menlo Park, CA | 1999 | Enterprise Software |
| 2 | Marc Andreessen | Menlo Park, CA | 2009 | Enterprise Software |
| 3 | Ben Horowitz | Menlo Park, CA | 1986 | Enterprise Software |
| 4 | David Heinemeier Hansson | Menlo Park, CA | 2002 | Enterprise Software |
| 5 | Paul Graham | Menlo Park, CA | 2004 | Enterprise Software |

[Access Here](#)

Complete VC Tech Stack

| Rank | Tech Stack Component | Location | Year Founded | Focus |
|------|----------------------|-------------------|--------------|---------------------|
| 1 | Slack | San Francisco, CA | 2009 | Enterprise Software |
| 2 | Zoom | San Jose, CA | 2011 | Enterprise Software |
| 3 | Dropbox | Foster City, CA | 2007 | Enterprise Software |
| 4 | Asana | San Francisco, CA | 2008 | Enterprise Software |
| 5 | GitHub | San Francisco, CA | 2008 | Enterprise Software |

[Access Here](#)

VC Execution - how do you get there?

Most read newsletter articles.

**10x your productivity
with ChatGPT.**

[Access
Here](#)

**An emerging VCs
Tech Stack.**

[Access
Here](#)

**The power of
GPT4 and LLMs in
Venture Capital.**

[Access
Here](#)

**The VC digitization
journey.**

[Access
Here](#)

**Top 13 Chrome
Plugins every VC
should use.**

[Access
Here](#)

**First steps as a
data-driven VC
without coding skills.**

[Access
Here](#)

Data-driven VC Newsletter

Since September 2022, the DDVC newsletter has produced over 30 editions of market analyses and data-driven tips for VCs.

Join our growing community of over 8,000 like-minded professionals!

[Join for free](#)

Some notes on our methodology.



Global VC Funds

Productivity VC Funds

Data-driven VC Funds

This report is based on several external sources and a survey conducted by **250+ VC funds** across Europe, North and South America, Asia, Oceania and MENA.

Data on usage and coverage of tech stacks and VC value chain are based on **151 data-driven VC funds** per DDVC definitions.

Rankings and inclusion are based on **fund surveys and publicly available information** - this means if firms or professionals do not actively communicate about their work, they might not have been included in the report.

About our data.

Sources of data

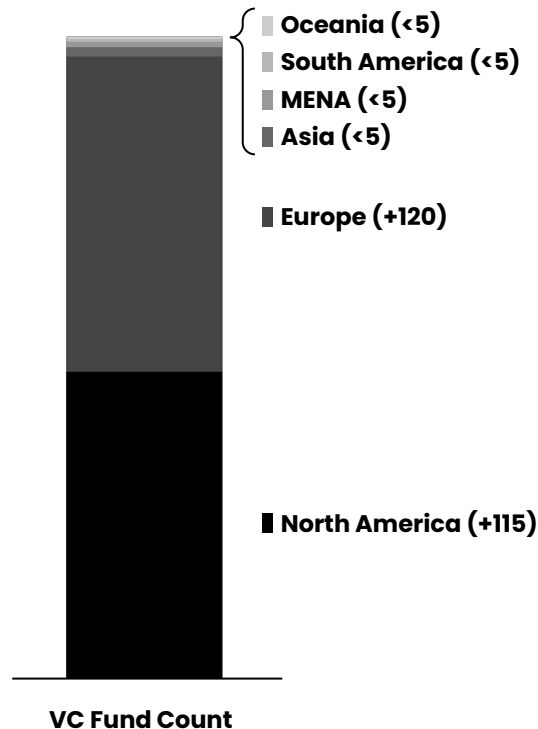
All data for this report was captured by a combination of surveys with VC funds as well as Pitchbook, Crunchbase and Dealroom.co databases. Public data in news articles, LinkedIn, Twitter and others was also used as a source.

Definitions

Venture capital (VC) is a form of private equity and a type of financing that investors provide to startup companies and small businesses.

Data-driven VC firms (DDVC) fulfill all of the following criteria: 1) at least one engineer (which summarizes developers and data roles of any kind) in the team, 2) at least one community nomination, and 3) proven to develop internal tooling in at least one of the segments across the VC value chain.

Region distribution of total count of VC funds nominated and analyzed
% of total funds



Our valuable partners.

We count ourselves lucky to have such a dedicated set of partners, aligned with our mission of championing data-driven innovation and the use of AI in Venture Capital. The Data-driven VC Landscape is a massive undertaking, bringing together a huge volume of data and voices from the leading actors in our field, and our partners are crucial to that process. We couldn't do this without them.



Affinity is a **relationship intelligence platform** that empowers dealmakers in relationship-driven industries to find, manage, and close more deals. With the most automated relationship intelligence insights and technology, Affinity enables leaders to drive deals, free themselves from data drudgery, and ensure their teams can take action with confidence, knowing the context and history of every relationship. The Affinity platform, including Affinity CRM, is used by over 2,000 relationship-driven organizations worldwide. Founded in 2014, Affinity is headquartered in San Francisco and backed by leading investors, including Menlo Ventures, Advance Venture Partners, 8VC, and MassMutual Ventures.



Synaptic helps **investors get company insights** by unifying a wide spectrum of alternative data on a single no-code platform. We cover alternative data points across various facets of business performance — such as web traffic, app reviews, employee and hiring data, product reviews, search trends, funding info, firmographics, and more. We unify data from proprietary and custom sources with a rich analytics toolkit, enabling more efficient ways to complete and complement the research workflow—from sourcing and tracking to the due diligence of companies and sectors. We work with top Venture Capital firms, Private Equity firms, and Hedge Funds across the globe.



Vauban from Carta is the easiest way to **launch & run your venture investing**. We offer SPVs and VC funds for GPs at all stages of the journey - from your first syndicate deal to operating a billion-dollar venture fund. Our end-to-end digital platform automates your back office so you can focus on what matters: finding the next unicorn & building investor relationships. Backed by Carta, the largest VC infrastructure company, we are in a unique position to provide valuable data insights for our customers on what's really happening in the world of venture capital.



Vestberry is a **portfolio management solution** tailored to the needs of VC investors seeking a data-driven advantage. Our platform consolidates vital portfolio information, enabling venture capital professionals to concentrate on extracting valuable insights from their data, rather than managing it. Leveraging a no-code approach, Vestberry empowers users to effortlessly construct their data infrastructure and seamlessly integrate with hundreds of data sources for unparalleled analytical capabilities. Trusted by top-tier fund managers from around the globe, who have come to rely on Vestberry as a partner in delivering excellence in portfolio management.

Thanks for taking the time to read this year's Data-driven VC report.

Please spread the word and share it with like-minded people to make our industry more efficient, effective, and inclusive.

If you are interested in becoming a partner for next year's report please drop me an email on andre@datadrivencv.io with the subject line 'DDVC Report 2024' and I will get back to you in due course.

We look forward to sharing the results with you next year!

*Stay driven,
Andre*

