dataiku + databricks

SURVEY REPORT

AI, Today

Insights From 400 Senior AI professionals on Generative AI, ROI, Tech Stack, & More

Introduction

With the rise of Generative AI, people around the world are excited about this new era of technology. But does the excitement hold for those leading data, analytics, and IT functions, the people who will be charged with turning Generative AI from cool technology into day-to-day reality (not to mention business value) in the workplace?

To answer this question, in June 2023, Dataiku and Databricks surveyed 400 senior AI professionals in large companies around the world.[1] In addition to attitudes and adoption of Generative AI, we also wanted to understand the larger environment in which we find ourselves today, from tech stacks to spend on AI tools and services, AI use case reach, and more.

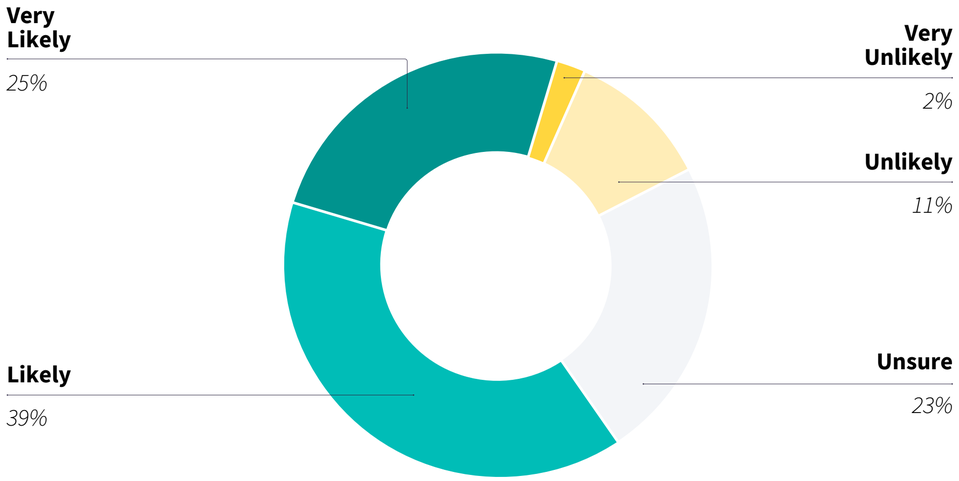
1 Respondents are manager level or higher, come from the U.S., EMEA, APAC, and work at organizations with >$3 billion in annual revenue. See p. 24 "About the Survey" for additional details about respondents and methodology.

What we found was enthusiasm and action around Generative AI. However, it's much more nuanced than pure excitement. For example, when asked whether they are more worried or excited about the future of AI, the majority of our respondents were in the former camp.

Why might senior AI professionals be worried about the future of AI?

Our survey results reveal that AI professionals are still facing some very real challenges in democratizing data, much less AI (much less Generative AI), across their organizations. The good news is that there are many signs of success and progress.

Generative AI Adoption in the Next 12 Months



Q20: Are you planning to use Generative AI or large language model (LLM) technology for your business in the coming year?

55%

of respondents say they are more worried than excited about the future of AI.

45%

of respondents said they are already experimenting with Generative AI.

3 Things AI Pioneers Do Differently to Succeed

While there is no easy button for democratizing, accelerating, and building trust in AI, there are definitely some keys to success. What makes true AI Pioneers[2] stand out against their peers?

Our survey revealed that these respondents have executive buy-in, are investing in the right infrastructure and tools in order to accelerate time to value, and are open to tackling more use cases and experimenting with techniques.

2 We defined an "AI Pioneer" as a respondent who meets at least two of the four following criteria: (1) has an advanced level of AI adoption and impact at their organization, (2) has a mature organizational structure for AI (such as a Hub-and-Spoke or Embedded model), (3) has a dedicated framework to account for the value associated with AI initiatives, and/or (4) delivers a return of > $1 for every dollar spent. See p. 24 "AI Pioneers: Data & Definition" for additional details.

1. AI PIONEERS HAVE EXECUTIVE SUPPORT

Let's face it: You won't get far with advanced analytics or AI initiatives if there's no support. That's true of grassroots, bottom-up support from frontline workers, but it's also true of executive support.

Our data shows that AI Pioneers are much more likely to have a data leader in their C-suite. They're also more likely to be confident that their executives understand the risks and benefits of AI, meaning they have partners in the C-suite that acknowledge and are invested in the success of AI initiatives (a key component of trust in AI for the rest of the company).

70%

of AI Pioneers have a data leader in the C-suite,

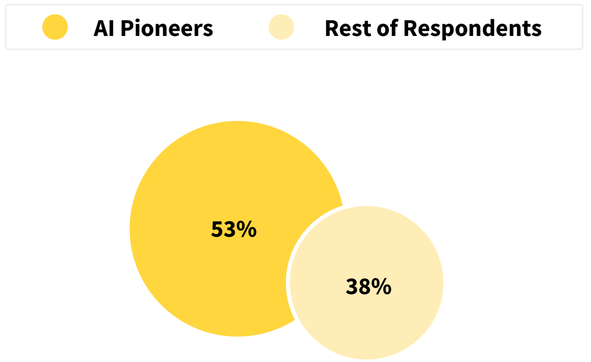
compared to only

43%

of the rest of respondents

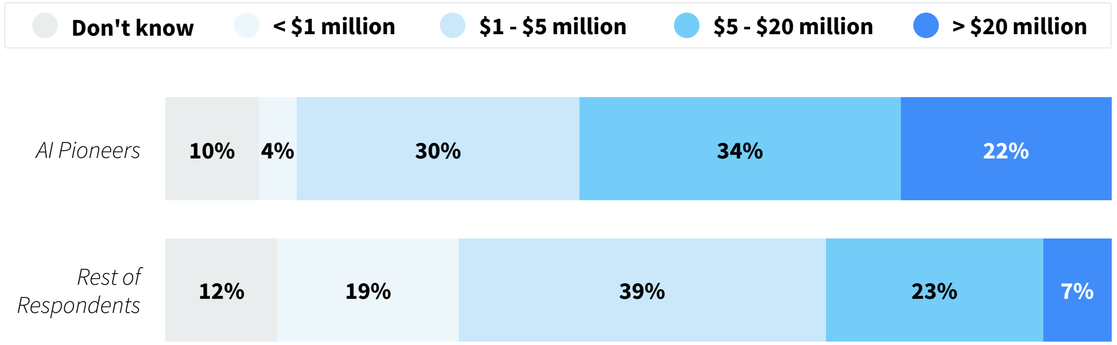
Do you think the leaders at your organization (C-suite — CEO, CMO, CFO, etc.) understand the risks and benefits of AI?

% responding with "Yes/Generally Agree."



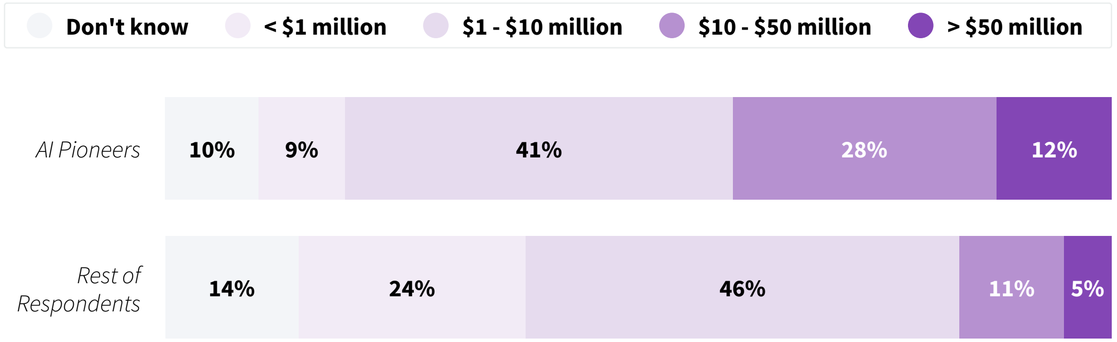
In theory, more executive support could open doors to more budget. Indeed, we see that **AI Pioneers have a larger budget** dedicated to both data and AI tools as well as services over the next 12 months.

Data and AI Technology Tools Budget — Next 12 Months



Q16: Approximately how much is your organization budgeting to spend on data and AI technology and tools in the next 12 months?

Data and AI Services Budget — Next 12 Months



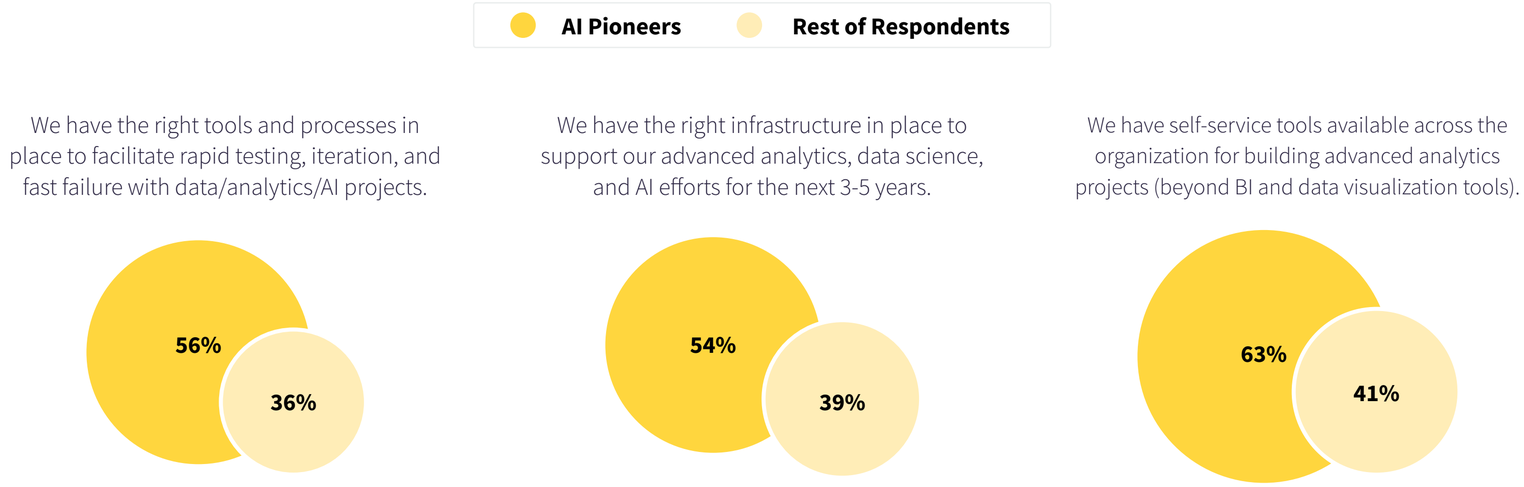
Q17: How much is your organization budgeting to spend on data and AI services (consultants, system integrators, etc.) in the next 12 months?

2. AI PIONEERS ARE INVESTING IN THE RIGHT INFRASTRUCTURE & TOOLS TO ACCELERATE

Pioneers know that AI (including Generative AI) is not a one-off initiative to build a single moonshot use case — it's about being able to handle hundreds (even thousands) of use cases with repeatability and reliability.

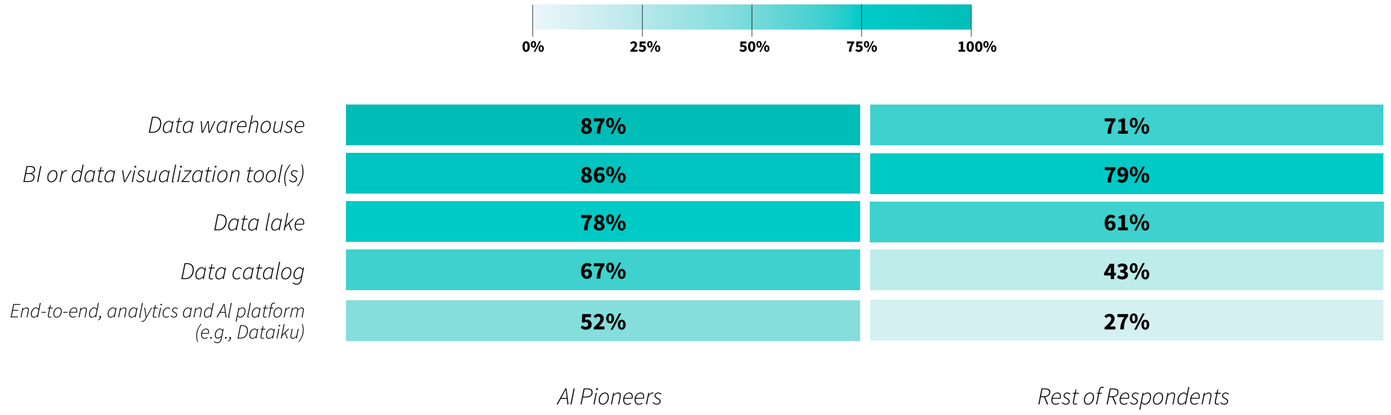
It's not about avoiding failure, but failing fast and being able to accelerate across all parts of the AI lifecycle. AI Pioneers know this and invest in tools accordingly.

% Responding With "Yes/Generally Agree"



The Modern Data & AI Stack

% selected "Yes, this is currently part of our data, analytics, & AI technology stack."



Q18: Which of the following are a part of your data, analytics, and AI technology stack?

> We as an industry should be really careful about how we talk about data science and AI projects failing.
>
> Failure is a necessary part of exploring the unknown. In general, it's absolutely part of doing good science, doing advanced development, and I feel like a lot of times in the industry, we'll say, well, 85% of data science projects never make it to production and present that like it's a fundamental flaw in the way we're doing data science.[3]

Will Benton
Principal Product Architect
NVIDIA

3 *https://blog.dataiku.com/how-to-make-your-ai-projects-successful-insights-from-nvidia*

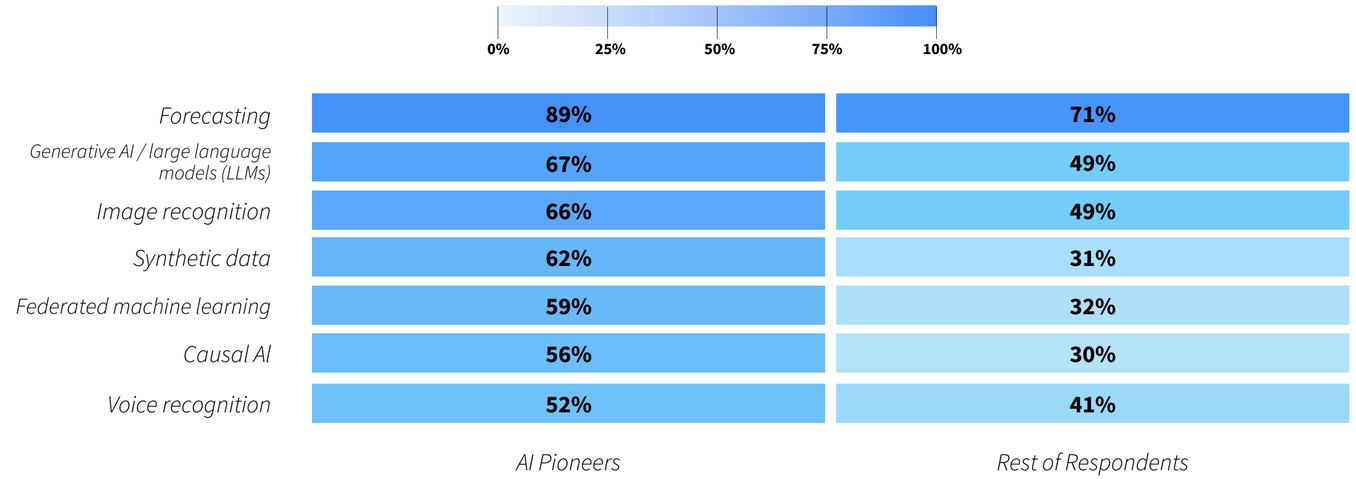
3. AI PIONEERS ARE TACKLING MORE USE CASES & TECHNIQUES

AI Pioneers also know that being successful with advanced analytics and AI is not about executing on one or two use cases, but about weaving it into every part of the business — scale matters.

We see that AI Pioneers already have (or are working on) use cases across a wider swath of business units and using a wider range of techniques and technologies.

Techniques Used for Data Science, Machine Learning, or AI Use Cases

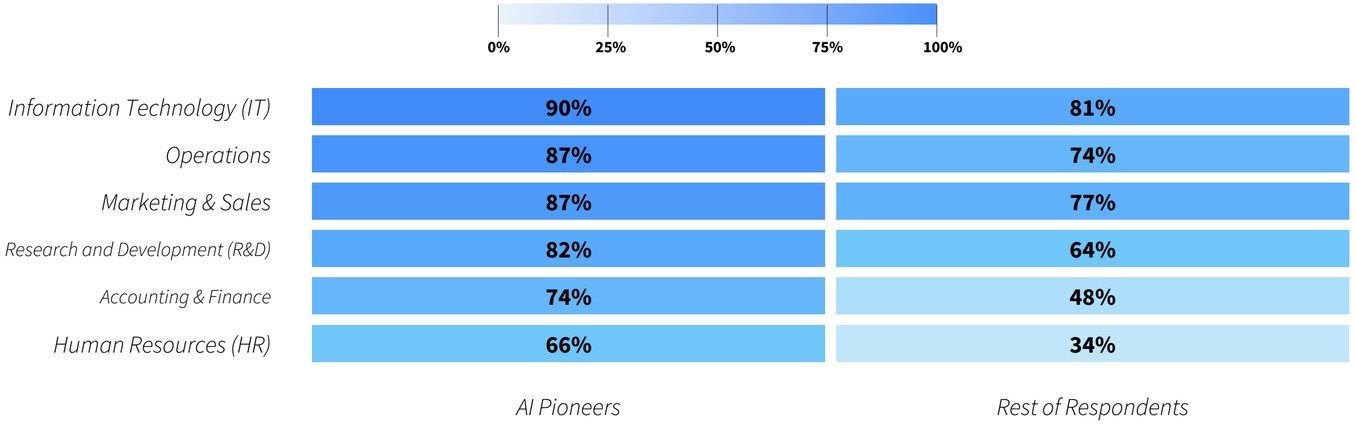
% selected "Experimenting" or "Actively using or have used"



Q15: What data science, machine learning, or AI techniques are you using for any use case in your organization?

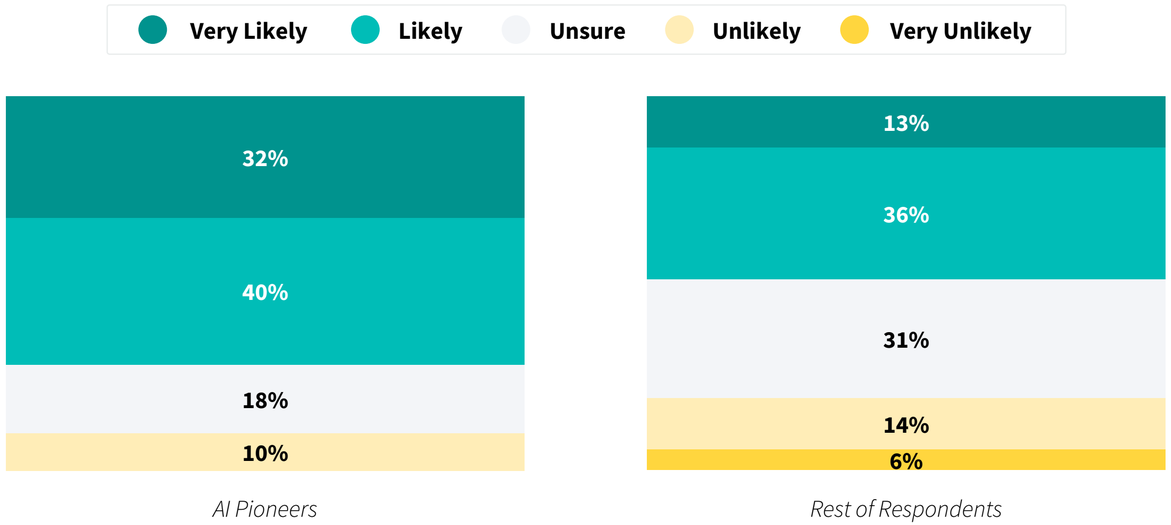
Top Use Cases for AI

% selected "We have developed or are currently developing use cases in this area" or "Use cases in this area are on our road map for the next year"



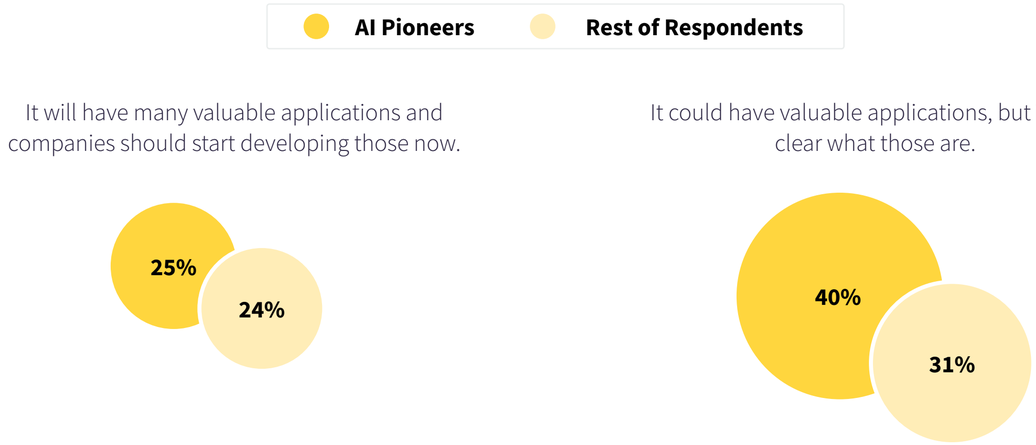
% of respondents who selected "We have developed or are currently developing use cases in this area" or "Use cases in this area are on our road map for the next year"

Plans to use Generative AI or Large Language Model (LLM) Technology Over the Next Year



Q20: Are you planning to use Generative AI or large language model (LLM) technology for your business in the coming year?

Attitudes About ChatGPT



This includes Generative AI, which is poised to give AI Pioneers an even bigger leap forward as they plan to use the technology in the next year.

Interestingly, though AI Leaders are overwhelmingly planning to use Generative AI or Large Language Model (LLM) technology, they are unsure about the applications of ChatGPT — which is based on LLM technology — for work. Perhaps a glimmer of hope for the rest of the pack to catch up on AI progress?

Deep Dive on Generative AI Insights

Given the amount of attention that Generative AI has been receiving in the popular and business media, data leaders and executives are rightly wondering whether this is yet another technology hype cycle that will fizzle in the coming months.

It is not. Given that the capabilities of these models far surpass what even their developers could have expected, it is right to see these models as ushering us into a new era. Our data supports this, showing the excitement as well as real action (mixed with, rightfully, some concern and caution) happening to move toward Generative AI use cases.

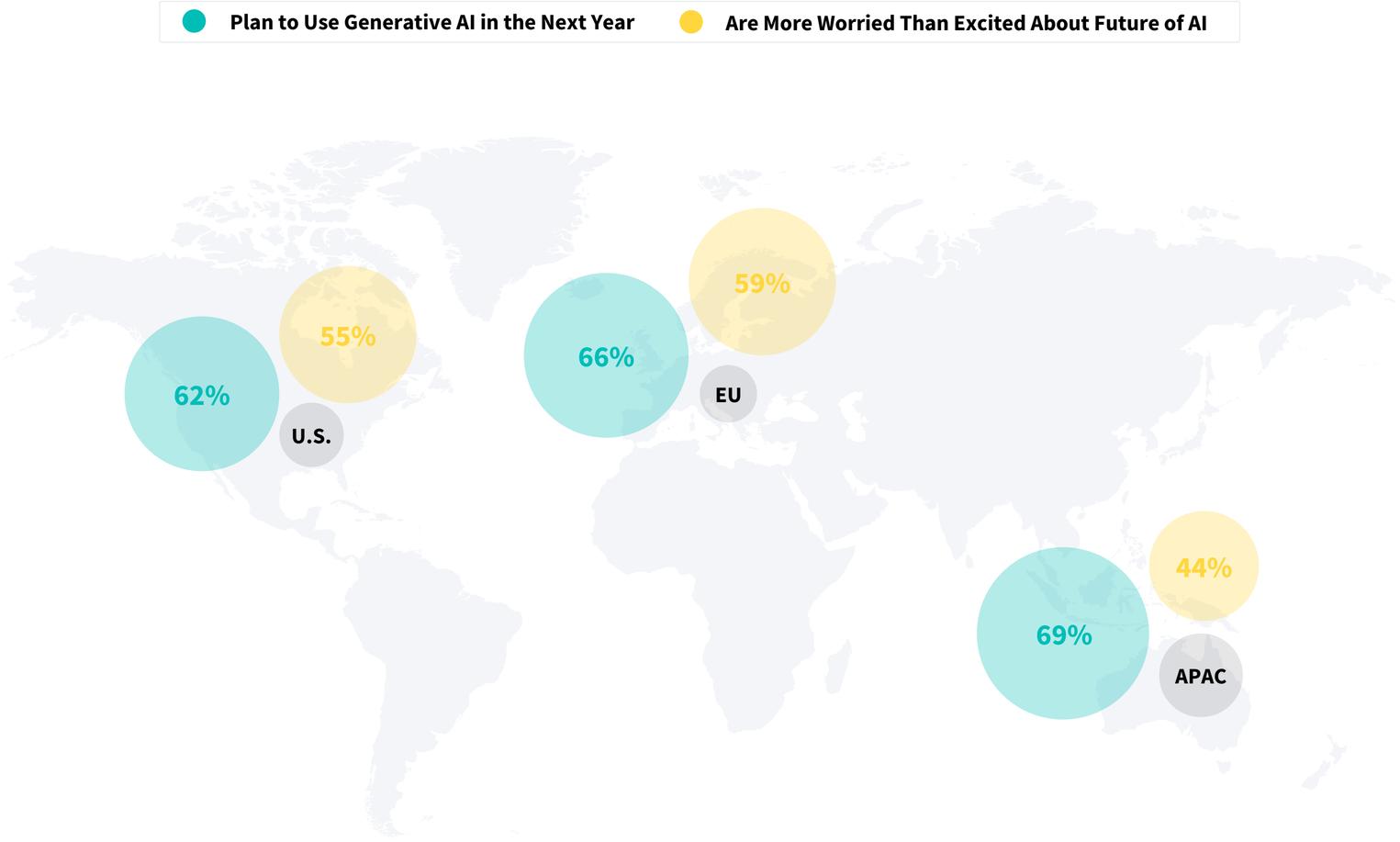
64%

of organizations will “Likely” or “Very Likely” use Generative AI technology for their business over the next year.

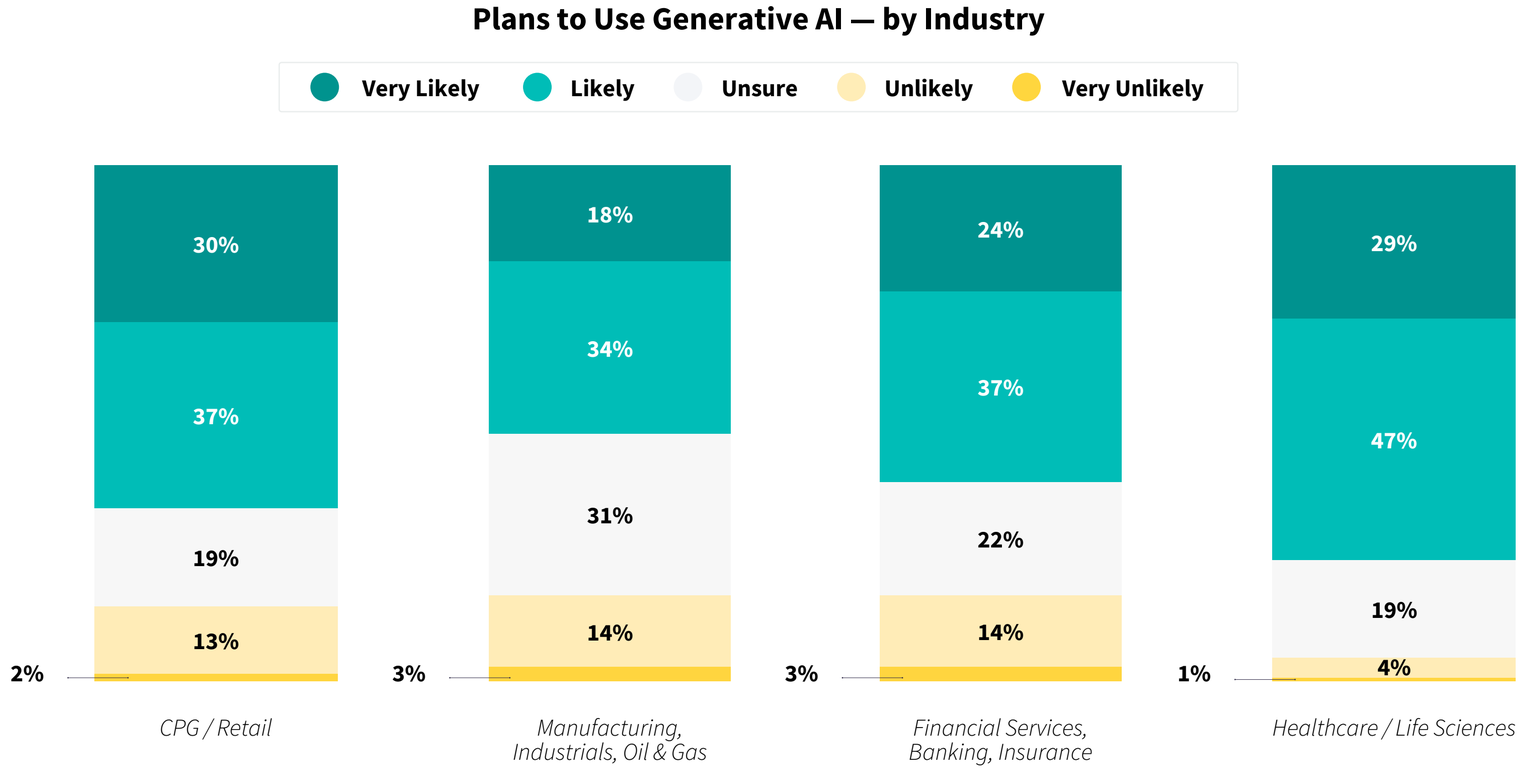
It's clear that companies that succeed in harnessing Generative AI and putting it into practice will enjoy a sustainable competitive advantage in their markets.

While more than 60% of data leaders already have Generative AI — and more specifically large language models (LLMs) — baked into their plans, APAC is even more keen on the tech than the rest of the world. This aligns with additional survey insights that reveal that APAC is generally less fearful and more excited about AI than other regions.

Plans to Use Generative AI & Fears About AI



There is some variation by industry when it comes to plans for using Generative AI or LLMs over the next year — healthcare and life sciences seems to be the most enthusiastic, along with CPG/retail. Interestingly, the C-suite was less likely to respond “unsure” than other levels of the organization and 70% of them answered “Likely” or “Very Likely.”



Q20: Are you planning to use Generative AI or large language model (LLM) technology for your business in the coming year?

70%

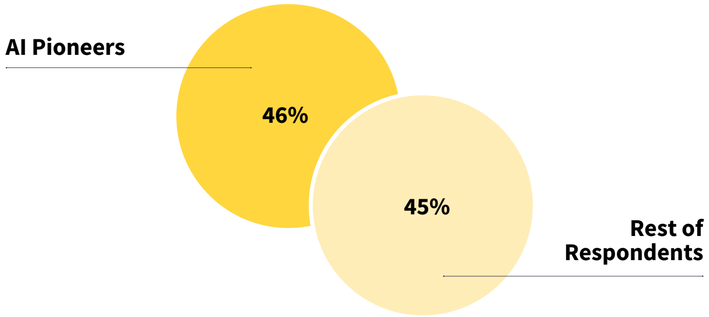
of C-suite respondents said that they are “Very Likely” or “Likely” to use Generative AI for their business in the coming year. Contrast this with only …

50%

of VP-level respondents who said they are “Very Likely” or “Likely” to use Generative AI for their business in the coming year.

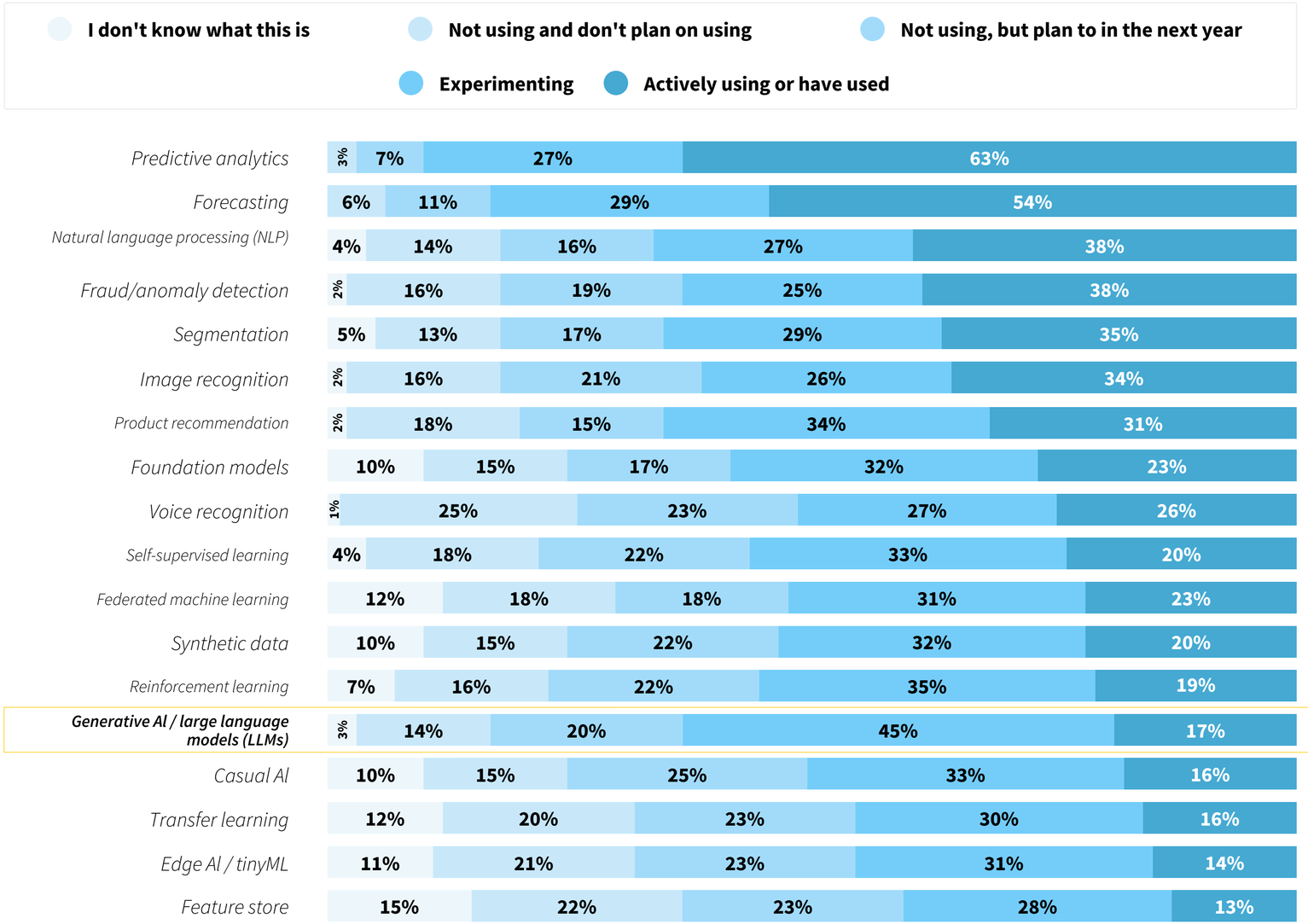
Not only is now the time to start building a Generative AI strategy, but it's the time to start actually experimenting with Generative AI use cases and LLMs — and nearly half of organizations are doing just that.

% Of Organizations That Say They Are "Experimenting" With Generative AI



Interestingly enough, for those specifically "experimenting" with the technology, this number does not shift much for AI Pioneers versus the rest of the respondents, indicating that non-Pioneers might be seizing this opportunity to iterate and test before firming up their strategy and, in doing so, catch up to the AI Pioneers.

Techniques Used for Data Science, Machine Learning, or AI Use Cases



Q15: What data science, machine learning, or AI techniques are you using for any use case in your organization?

33%

of People Are Unclear on What the Valuable Applications of ChatGPT Are

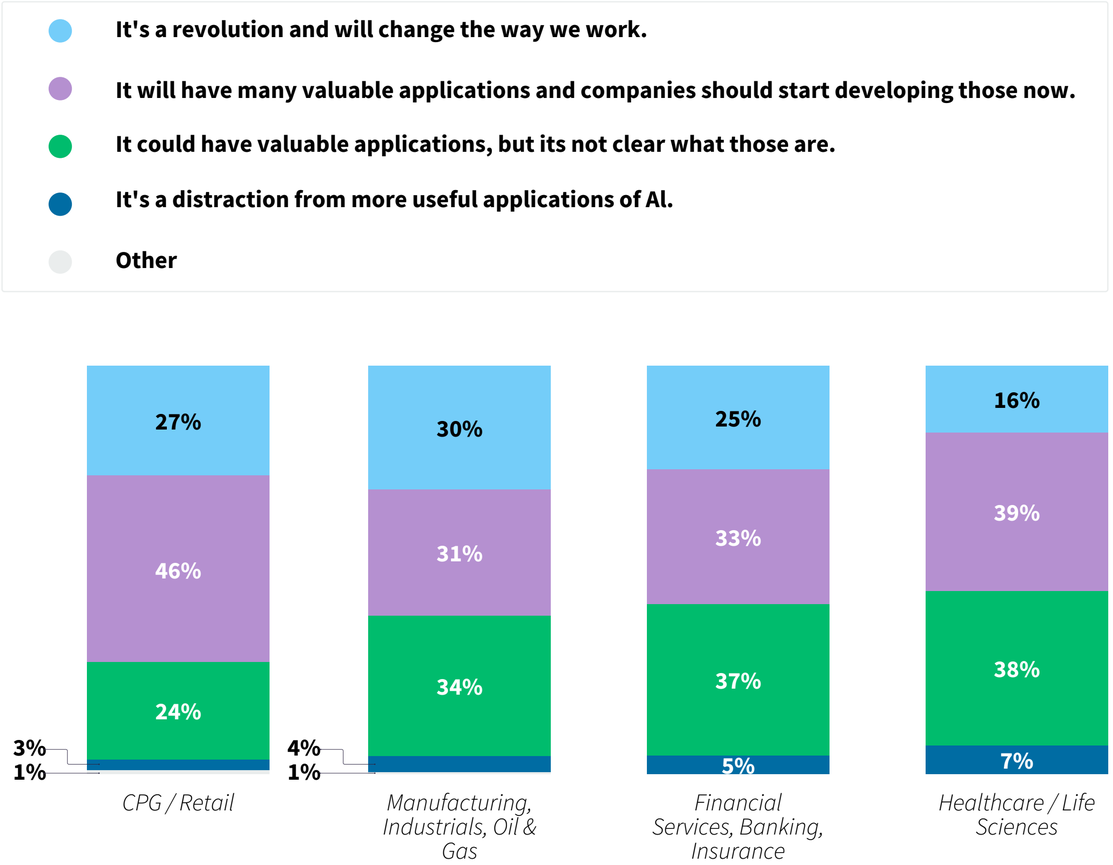
When it comes to how respondents feel about ChatGPT and similar technologies, responses are mixed. Although the most common response is that “It will have many valuable applications and companies should start developing those now,” it is closely followed by more than one-third of respondents who are unclear on the precise applications of ChatGPT and similar technologies.

While there’s been a lot of buzz about Generative AI in the news, it’s a rare sight when it comes to actual enterprise applications. For organizations looking to change that and get started on real-world enterprise applications, check out the [Dataiku Generative AI Use Case Collection.]

Which of the following most closely aligns with your personal feelings about ChatGPT and similar technology?

37%	“It will have many valuable application, and companies should start developing those now.”
33%	“It could have valuable applications, but it’s not clear what those are.”
24%	“It’s a revolution and will change the way we work.”
5%	“It’s a distraction from more useful applications of AI.”

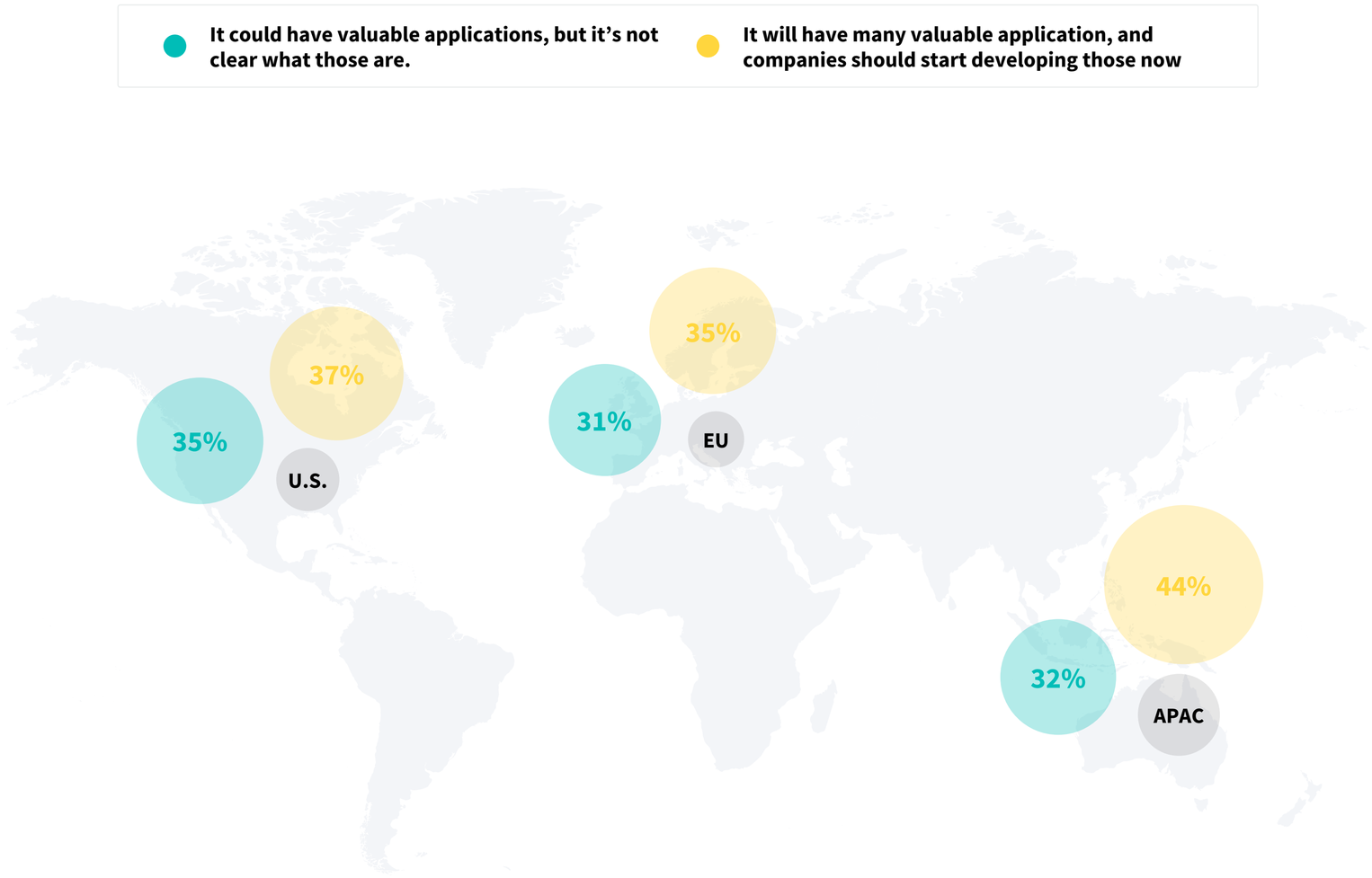
Attitudes About ChatGPT — by Industry



Q23: Which of the following most closely aligns with your personal feelings about ChatGPT and similar technology?

When it comes to regions, **APAC is the most confident** that ChatGPT and similar technologies will have many valuable applications and companies should start developing those now.

Attitudes About ChatGPT — by Region



Beyond the Buzz: Further Insights on Scaling AI

Outside of what sets AI Pioneers apart and insights on Generative AI, our survey revealed a few more interesting trends around what's happening with organizations' data and AI efforts that we'll present in this section.

DATA QUALITY & ACCESS REMAINS A BARRIER FOR ALL

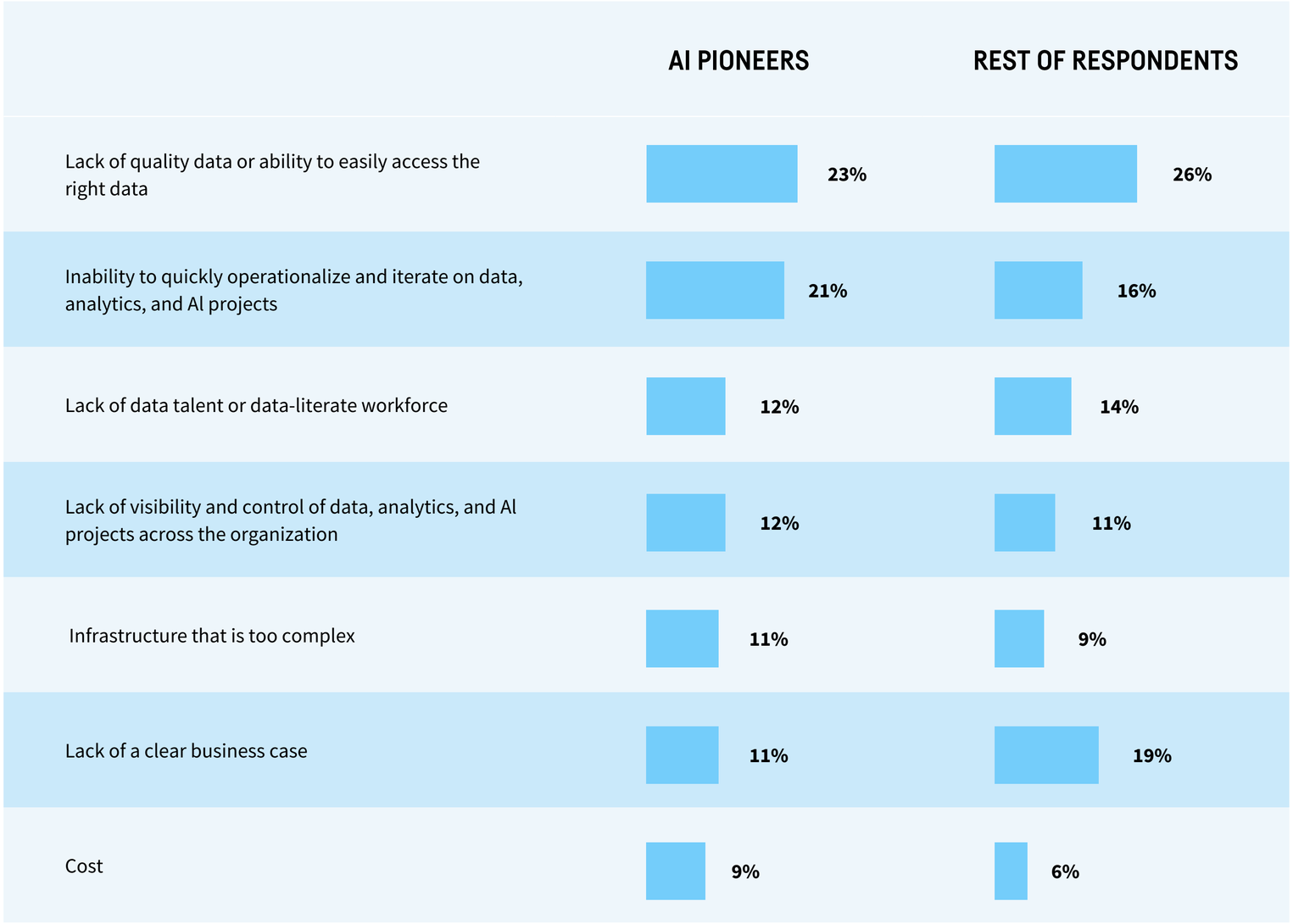
21%

of organizations get > $5 ROI on AI for every $1 spent.

So that begs the question:

What's getting in the way of generating value?

% of Respondents Who Ranked Each Challenge as Their Top Barrier to Getting More ROI From AI



The good news is that many respondents and their organizations are taking the first steps to addressing data quality challenges, including having clear owners of data quality.

55%

say “We have clear owner(s) responsible for data quality.”

45%

“Trusted data is easily accessible to frontline people when needed.”

“When people talk about some of the success that we’ve had … they point to things like next best action, and they point to the sophisticated algorithms that we’re using, and the work we’re doing around predictive analytics, and all the work we’re doing around data visualization. And to be clear, I’m exceedingly proud of that stuff …

But when I look at what ultimately is driving the real success of all of that, it’s the hard nuts and bolts work that we’ve done around data quality. It's that we have data stewards that come to work every day who are accountable for the accuracy. It’s the data quality engines that run every night. It’s the issues management processes that we have in place. It’s our data definitions that we put into our systems so people can go to and they know. It’s our monthly governance meeting.

It’s the hygiene that we put around that — and if there is a data quality problem (and we see them all the time) there is an infrastructure that can take it in, it can evaluate it, and action can be taken … I would argue that is one of the competitive advantages that we have as an organization.”[4]

Jeff McMillan
Chief Analytics and Data Officer at Morgan Stanley

4 *https://www.dataiku.com/stories/detail/morgan-stanley/*

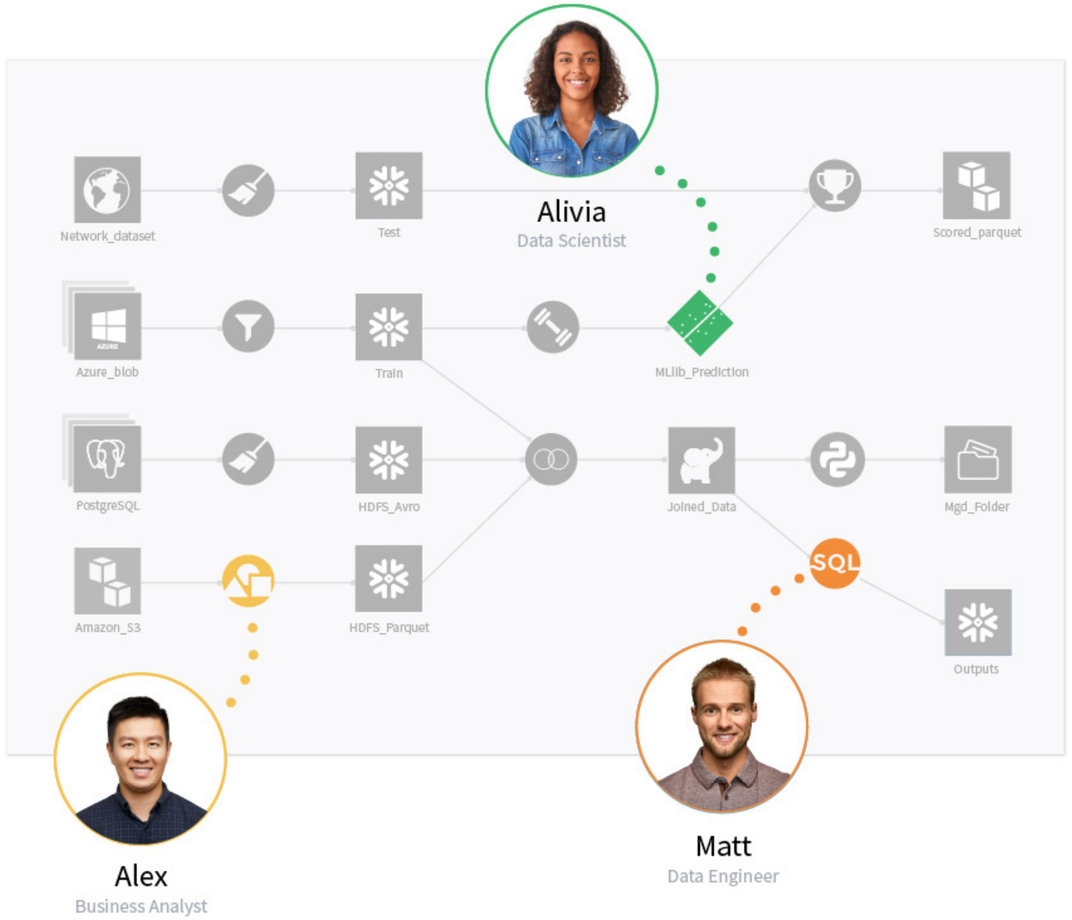
“COLLABORATIVE DATA SCIENCE” IS BECOMING THE REALITY

At Dataiku, we’ve been talking about the idea of true collaboration between data and non-data experts for more than 10 years. Democratization of data and, now, Generative AI won’t happen if all the tools and technologies stay in the hands of the few.

The good news is that interdisciplinary teams (i.e., business people working with data people) today are becoming status quo for the overwhelming majority of companies — and just in time for the next wave of AI!

77%

of organizations say their advanced analytics development teams are interdisciplinary



EFFICIENCY IS THE NAME OF THE GAME

Though there is some variation by industry, for the most part, lines of business that have lots of manual processes to optimize are the most developed in terms of advanced analytics, data science, and AI use cases.

Bottom line: Those getting ahead are not just thinking about moonshot use cases, but also the mundane that bring big efficiencies and gains to day-to-day work.

Use Cases for AI (Top 6)

% of respondents who answered they "Developed or are currently developing use cases in this area"

1. Information Technology (IT)

Overall	57%
Retail & CPG	55%
Manufacturing, Industries, Oil, & Gas	57%
Financial Services, Banking, & Insurance	61%
Healthcare & Life Sciences	56%

2. Operations

Overall	52%
Retail & CPG	54%
Manufacturing, Industries, Oil, & Gas	49%
Financial Services, Banking, & Insurance	43%
Healthcare & Life Sciences	53%

3. Marketing & Sales

Overall	49%
Retail & CPG	64%
Manufacturing, Industries, Oil, & Gas	41%
Financial Services, Banking, & Insurance	44%
Healthcare & Life Sciences	47%

4. Research & Development (R&D)

Overall	45%
Retail & CPG	41%
Manufacturing, Industries, Oil, & Gas	44%
Financial Services, Banking, & Insurance	33%
Healthcare & Life Sciences	61%

5. Customer Service

Overall	42%
Retail & CPG	42%
Manufacturing, Industries, Oil, & Gas	36%
Financial Services, Banking, & Insurance	46%
Healthcare & Life Sciences	40%

6. Accounting & Finance

Overall	29%
Retail & CPG	31%
Manufacturing, Industries, Oil, & Gas	31%
Financial Services, Banking, & Insurance	29%
Healthcare & Life Sciences	26%

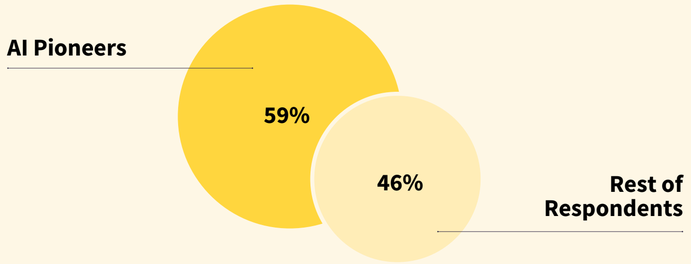
PEOPLE WANT MORE AI REGULATION

Industries move at different paces in their adoption of AI Governance and Responsible AI — some being naturally more accustomed to adapting business processes to ever-evolving regulations than others. However, with discussion of Generative AI and LLMs nearly everywhere, it's not surprising that more than half of respondents believe that AI requires more official regulation.

55%

of respondents say that they believe AI requires more official regulation (via laws, government oversight, etc.).

This figure is even higher among AI Pioneers:



Raft Framework For Responsible Generative AI Applications

55%

Of Respondents Are More Worried Than Excited About the Future of AI

Our survey shows that people across multiple industries are anxious about the future of AI. With the rise of Generative AI, we are in uncertain waters, to be sure.

Our take? It will take more stringent regulations and the development of more robust Responsible AI policies across more companies and teams to turn this worry into excitement.

In order to equip enterprises for the arrival of a wave of policy proposals seeking to protect both workers and consumers from potential harms of AI, such as the U.S. Blueprint for an AI Bill of Rights and the EU AI Act, Dataiku created its RAFT Framework (Reliable, Accountable, Fair, and Transparent) for both traditional and Generative AI systems. Databricks has also made a ***commitment to helping enterprises deploy AI responsibly.***

Get The Full RAFT Framework

About the Survey

BACKGROUND & OBJECTIVES

Dataiku and Databricks conducted this joint survey to:

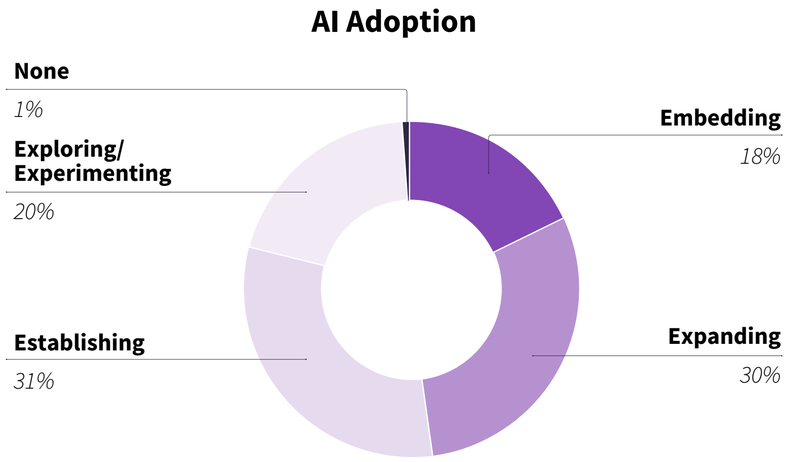
- Learn about the differing levels of adoption, sentiment, and barriers to AI in general.
- Understand the perception and level of adoption specifically around new technologies like Generative AI and Large Language Models (LLMs).
- Assess trends in tech stack and tooling in today's largest organizations.

AI PIONEERS: DATA & DEFINITION

Note that throughout this survey report, an "AI Pioneer" is a respondent who met at least two of the four following criteria. Here's a look at each of these baseline criteria along with its overall breakdown plus a breakdown by industry and region.

1. MATURITY

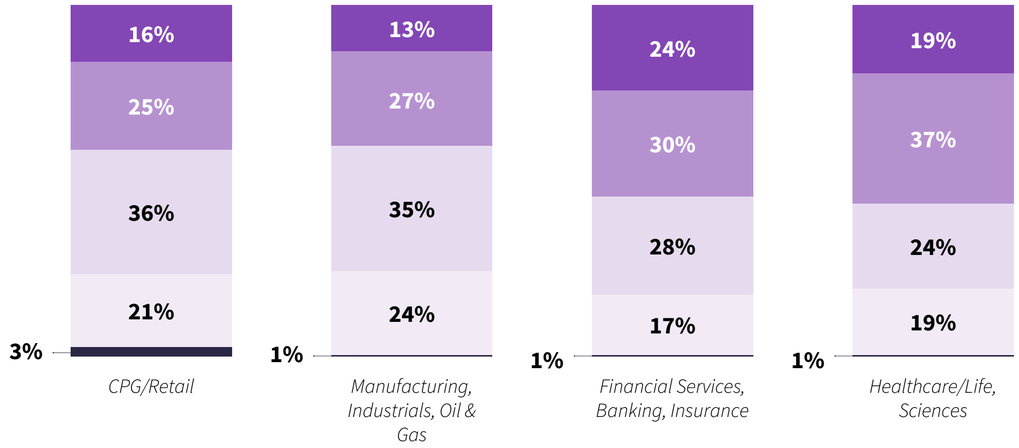
Organization has an advanced (i.e., Expanding or Embedding) level of data science and AI adoption.



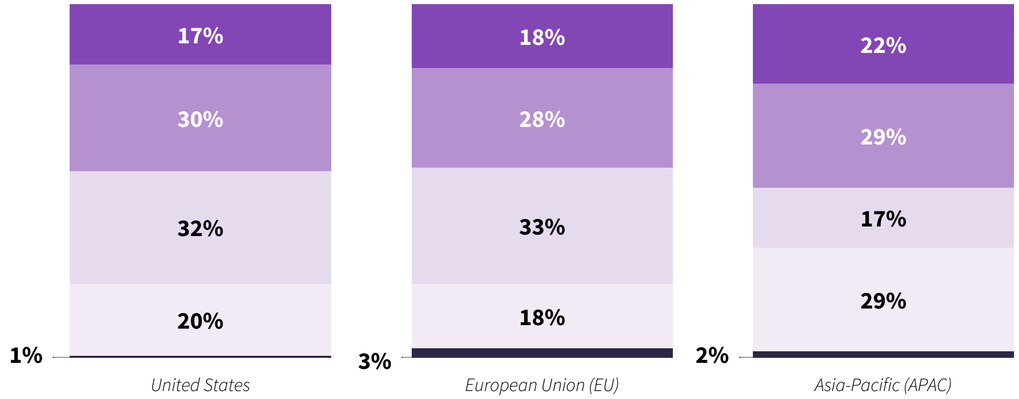
- **None:** *We don't leverage advanced analytics (data science or AI).*
- **Exploring/Experimenting:** *We're still experimenting with our first projects and use cases.*
- **Establishing:** *We've proved value with several use cases and we're working on building processes to scale out our ability to execute.*
- **Expanding:** *We're working on expanding the use of data science and AI beyond a central team and into different functions and departments across the organization.*
- **Embedding:** *We already have people across different functions and departments using data science/AI, and we're working on embedding into even more activities, making data part of the DNA of the organization.*

Q7: What best describes the level of adoption of data science and/or AI at your organization?

AI Adoption — by Industry



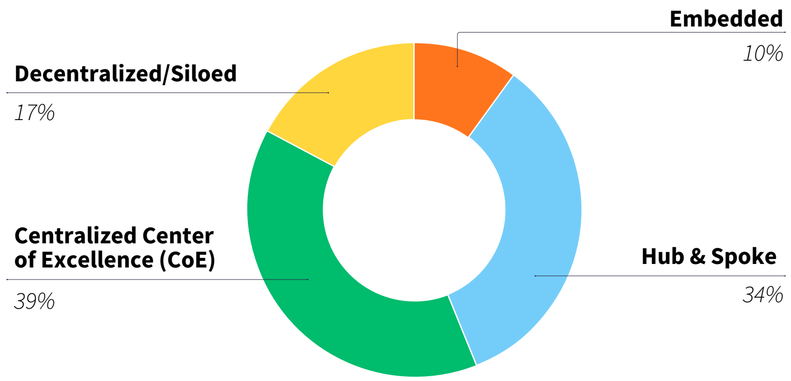
AI Adoption — by Region



2. ORGANIZATIONAL OR OPERATIONAL MODEL

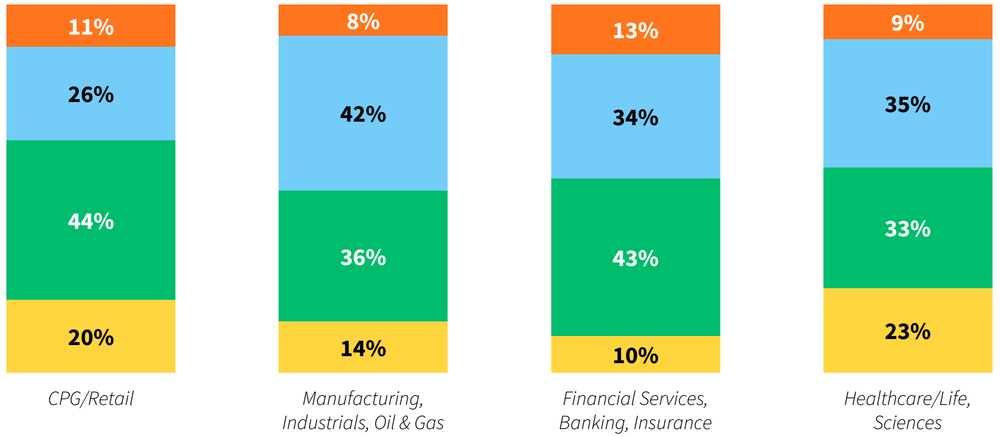
Organization has a mature (i.e., Hub & Spoke or Embedded) data science, analytics, and AI organizational structure.

AI Operational Model

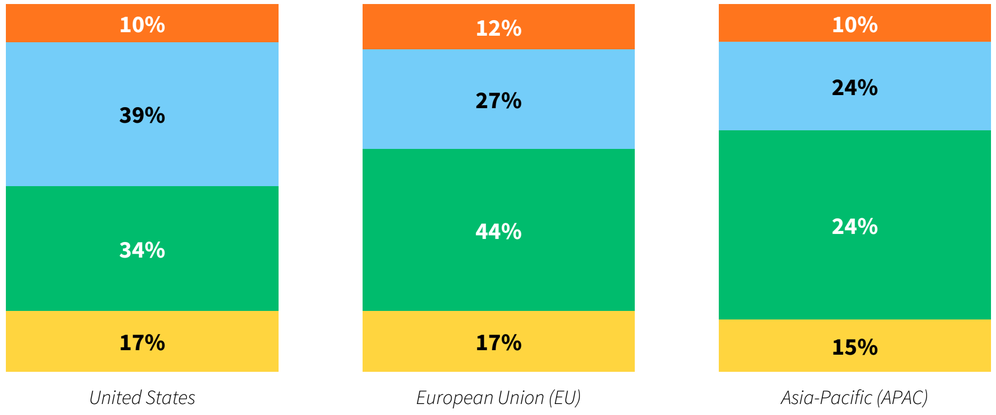


Q8: What best describes your organization's structure around data science, analytics, and AI initiatives?

AI Operational Model — by Industry



AI Operational Model — by Region



DECENTRALIZED /SILOED

Individual teams within a company do their own independent experimentation with Al.

Little to no sharing of infrastructure, data, best practices, or talent.

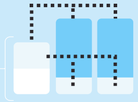
Almost always a temporary operating model, where the goal is to determine if there's enough Al value to further invest.

CENTRALIZED CENTER OF EXCELLENCE

Known as a CoE, this operating model is designed to jumpstart the adoption of Al within an organization.

A centralized team develops and maintains Al products for many business units and functions.

HUB & SPOKE

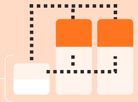


CoE functions get distributed around an organization: Al experts (advanced, graduate-level data scientists) are in the hub, business units and functions are in the spokes, and they collaborate on product development.

Like in a CoE, the hub is responsible for infrastructure, standards, and tracking industry innovation.

However, ownership of Al products shifts to the spokes.

EMBEDED



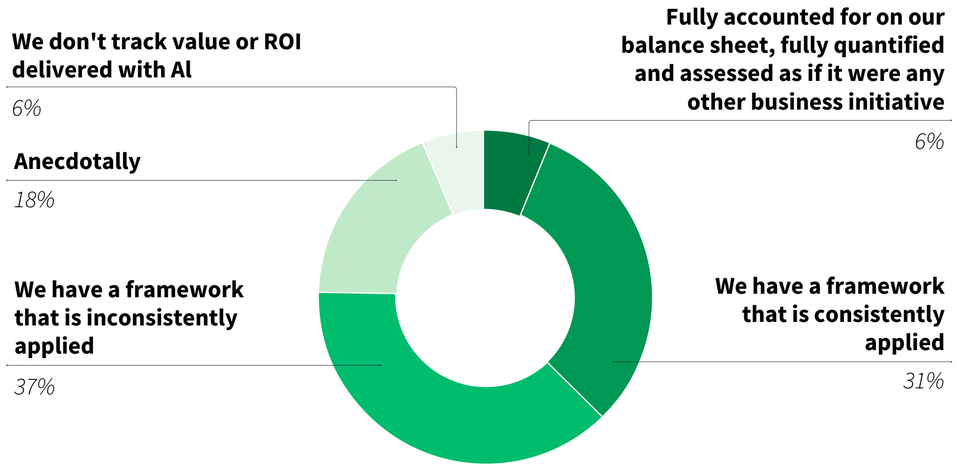
Very few central, shared resources and rules such as Responsible Al guidelines, infrastructure, and a few common, curated datasets.

The most decentralized, agile, and innovative structure since many business units and functions are involved, and they are loosely jonnected by rules and resources.

3. RETURN ON INVESTMENT (ROI) METHODOLOGY

Organization has a framework for tracking ROI that is consistently applied or AI initiatives are fully accounted for on our balance sheet, fully quantified and assessed as if it were any other business initiative.

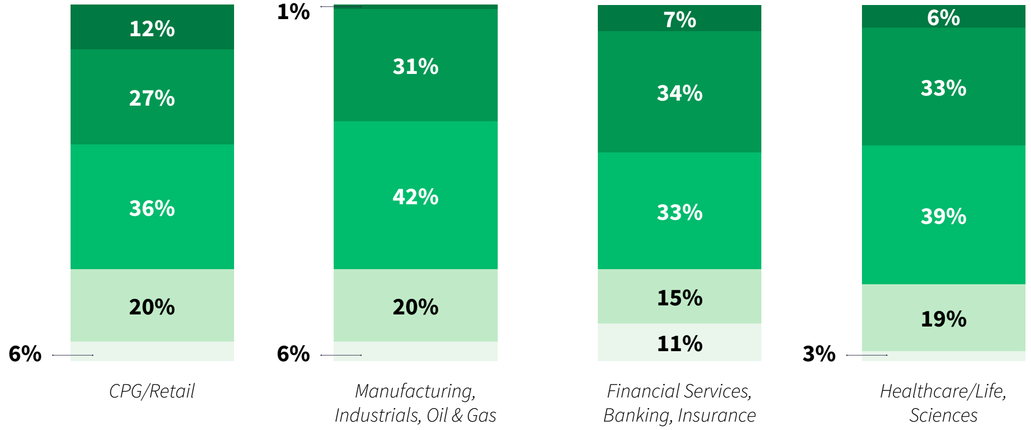
AI ROI Methodology



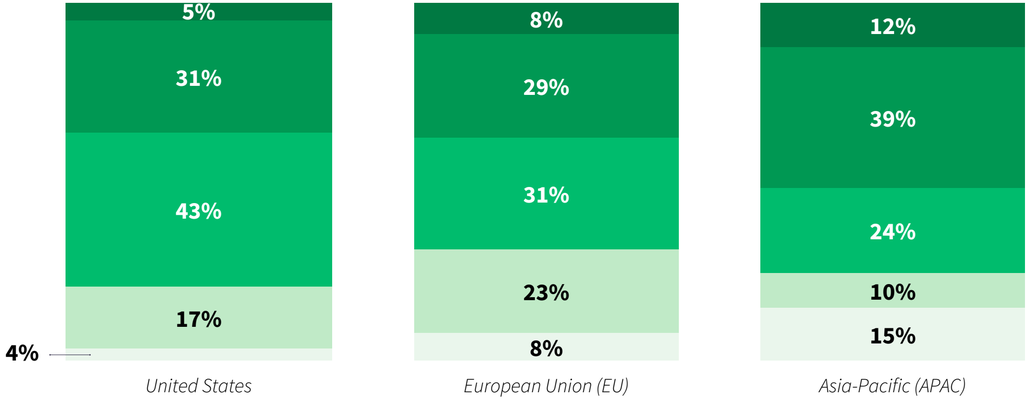
Q9: How does your organization account for the value delivered with data, analytics, and AI initiatives?

- **Fully accounted for on our balance sheet, fully quantified and assessed as if it were any other business initiative**
- **We have a framework that is consistently applied**
- **We have a framework that is inconsistently applied**
- **Anecdotally**
- **We don't track value or ROI delivered with AI**

AI ROI Methodology — by Industry



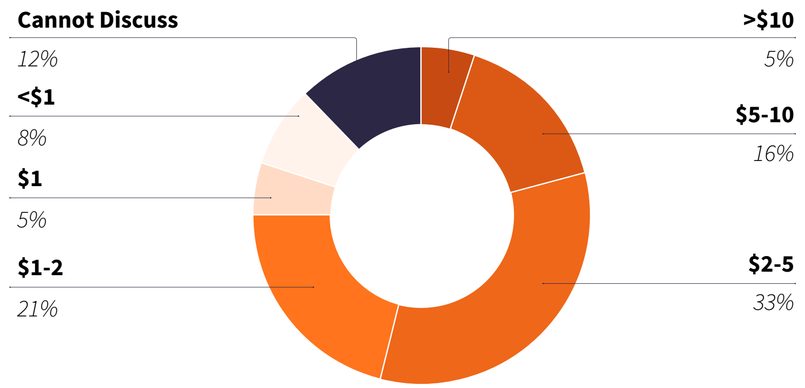
AI ROI Methodology — by Region



4. ACTUAL ROI

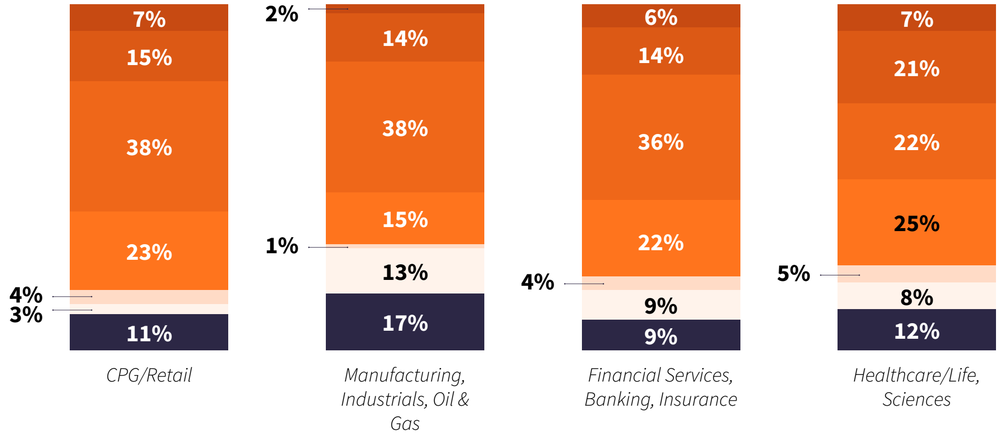
Organization delivers a return of > $1 for every $1 spent on analytics and AI initiatives.

ROI on AI per $1 Spent

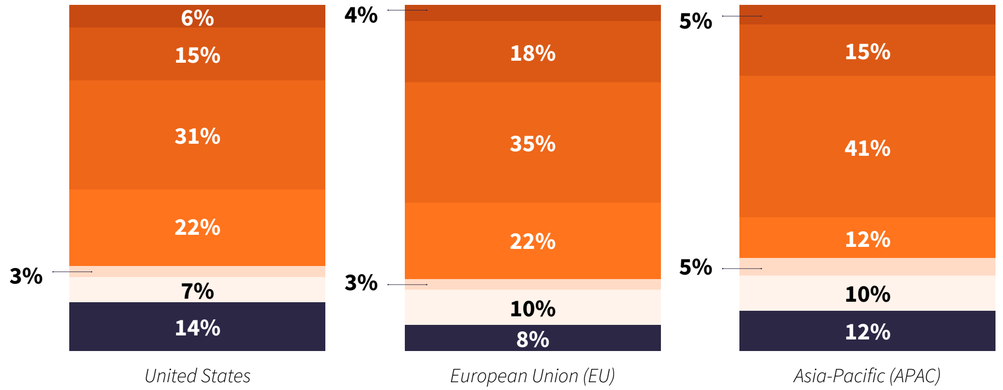


Q10: Approximately what return do you deliver for each $1 spent on data, analytics, and AI initiatives?

ROI on AI per $1 Spent — by Industry



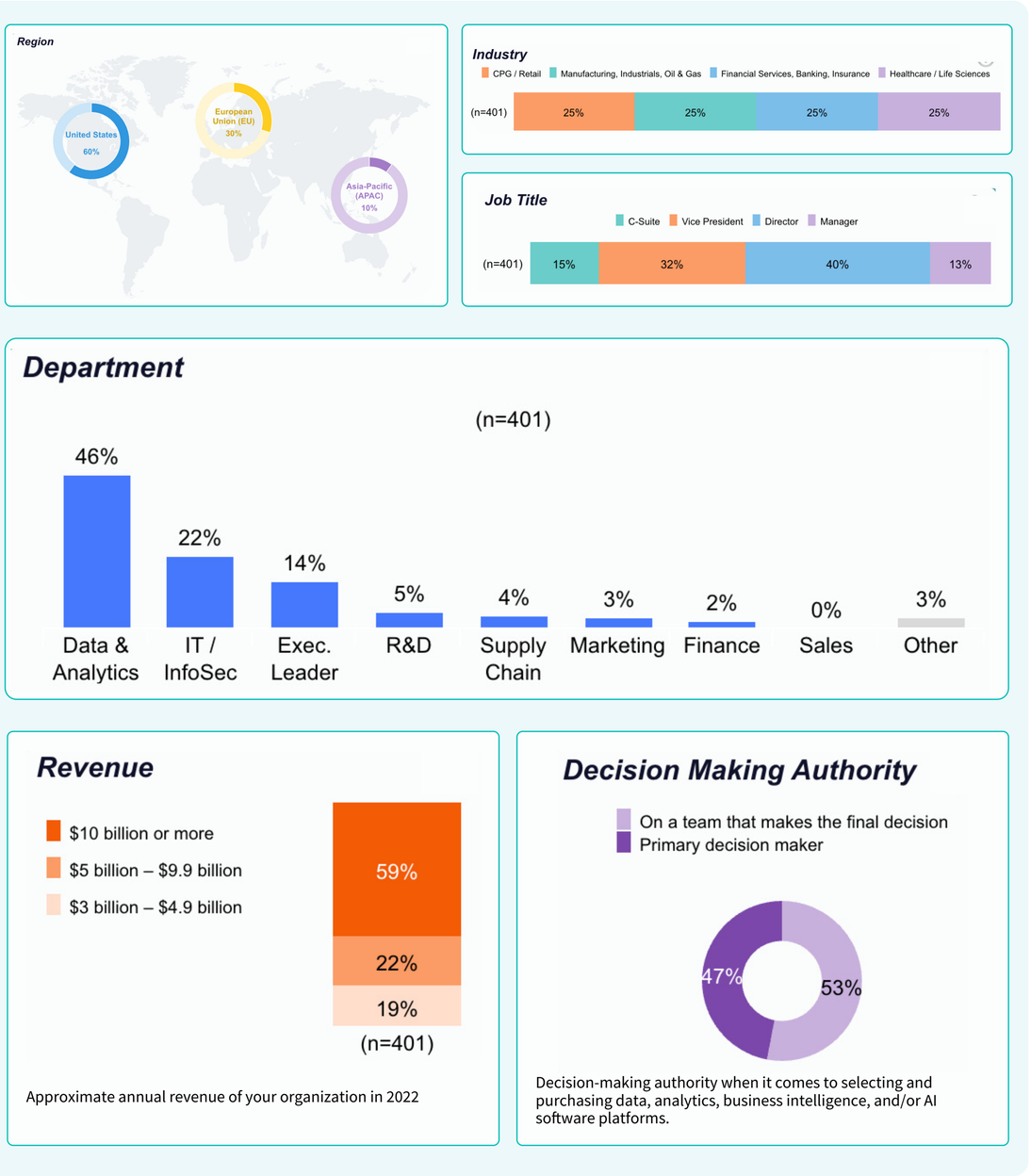
ROI on AI per $1 Spent — by Region



SAMPLE COMPOSITION

Length of Interview (LOI): ~8 minutes
Fielding: May 30, 2023 - June 8, 2023

- Data has been reported among total respondents unless otherwise noted.
- Small base sizes (n<30) are noted with an asterisk (*).



Dataiku Is the Platform for Everyday AI.

With Dataiku, data experts and domain experts design, develop, and deploy new AI capabilities at all scales and in all industries. Hundreds of companies worldwide use Dataiku to drive value from data and AI with diverse use cases, from predictive maintenance and supply chain optimization to quality control in precision engineering, marketing optimization, Generative AI use cases, and everything in between. To learn more, *follow Dataiku on LinkedIn.*

Databricks Is the Data and AI Company.

More than 10,000 organizations worldwide — including Comcast, Condé Nast, and over 50% of the Fortune 500 — rely on the Databricks Lakehouse Platform to unify their data, analytics and AI. Databricks is headquartered in San Francisco, with offices around the globe. Founded by the original creators of Apache Spark™, Delta Lake and MLflow, Databricks is on a mission to help data teams solve the world's toughest problems. To learn more, follow Databricks on *Twitter*, *LinkedIn* and *Facebook*.