

FORRESTER®

The Total Economic Impact™ Of Dataiku

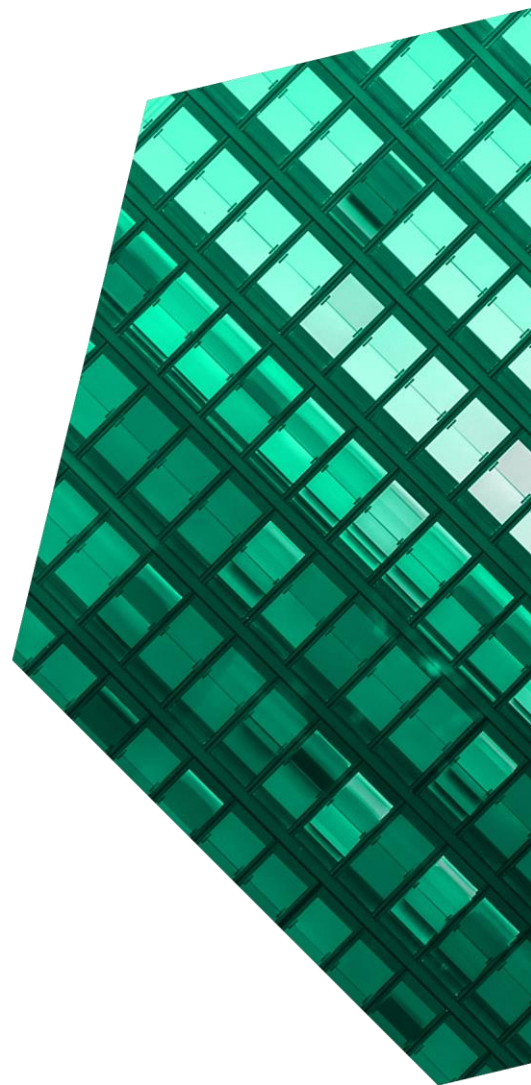
Cost Savings And Business Benefits
Enabled By Dataiku

JULY 2021

Table Of Contents

Consulting Team: Asha Dinesh

Executive Summary	1
The Dataiku Customer Journey	5
Key Challenges	5
Investment Objectives And Solution Requirements	5
Composite Organization	6
Analysis Of Benefits	7
Increased User Productivity: Data Engineering And Data Science	7
Reduced Costs: Data Analytics Tools And Consultants	9
Increased Operational Efficiency: Reporting	10
Increased Profit: Customer Targeting And Acquisition	12
Unquantified Benefits	13
Flexibility	13
Analysis Of Costs	15
Platform And License Fees	15
Implementation And Maintenance Costs	16
User Training Costs	17
Financial Summary	19
Appendix A: Total Economic Impact	20
Appendix B: Supplemental Material	21
Appendix C: Endnotes	21



ABOUT FORRESTER CONSULTING

Forrester Consulting provides independent and objective research-based consulting to help leaders succeed in their organizations. For more information, visit forrester.com/consulting.

© Forrester Research, Inc. All rights reserved. Unauthorized reproduction is strictly prohibited. Information is based on the best available resources. Opinions reflect judgment at the time and are subject to change. Forrester®, Technographics®, Forrester Wave, RoleView, TechRadar, and Total Economic Impact are trademarks of Forrester Research, Inc. All other trademarks are the property of their respective companies.

Executive Summary

The demand for enterprise artificial intelligence (AI) continues to skyrocket as companies seek to deploy AI solutions to become industry leaders. However, many struggle to deploy, monitor, and govern AI models in production applications.¹ To overcome barriers, data science and extended AI teams need vendors who can provide more automation to increase productivity, model operations for smooth deployment, and a product roadmap that makes breakneck machine learning innovations accessible.²

Dataiku is an enterprise AI platform that offers an end-to-end interface for data processing and preparation, data visualization, machine learning, and numerous other forms of advanced analytics. It is designed to support a wide range of use cases and serve multiple user personas, including data scientists, data engineers, business analysts, machine learning engineers, and business stakeholders.

Dataiku commissioned Forrester Consulting to conduct a Total Economic Impact™ (TEI) study and examine the potential return on investment (ROI) enterprises may realize by deploying [Dataiku](#).³ The purpose of this study is to provide readers with a framework to evaluate the potential financial impact of Dataiku on their organizations.

To better understand the benefits, costs, and risks associated with this investment, Forrester interviewed four customers with experience using Dataiku. For the purposes of this study, Forrester aggregated the experiences of the interviewed customers and combined the results into a single [composite organization](#).

Prior to investing in Dataiku, the interviewees' organizations typically had data silos across the organization and a fragmented set of applications including data extract, transform, and load (ETL) tools and data science platforms. This resulted in poor accessibility of datasets and an inability to collaborate on data and analytics projects.

KEY STATISTICS: THREE-YEAR RESULTS



Return on investment (ROI)

423%



Net present value (NPV)

\$10.43M

Additionally, too much manual work reduced the productivity of data engineers, data scientists, and analysts.

After the investment in Dataiku, the customers unified multiple approaches and tools for analytics and data science. Key results from the investment include improved productivity, reduced costs, and increased efficiency and revenue return of analytics and data science practices.

KEY FINDINGS

Quantified benefits. Risk-adjusted present value (PV) quantified benefits over three years include:

- **Increased data engineers and scientists' productivity, amounting to over \$8.4 million.** This was the biggest benefit organizations identified. In the case of the composite organization, data engineers and data scientists save 75% of the time they previously spent on

projects by using Dataiku. This results in a risk-adjusted, three-year present value of \$8,441,721.

- **Reduced costs for analytics tools and consultants, resulting in savings of over \$3.2 million.** Interviewees' organizations highlighted cost savings from reducing legacy tools used for analytics and data science. For the composite organization, the decommissioned tools and saved developer time amounts to a risk-adjusted, three-year present value of \$3,216,941.
- **Increased efficiency in reporting, amounting to over \$735,000.** The automation of manual, repeated tasks enabled operational efficiencies for interviewees' organizations. A key area of impact was in reporting. In the case of the composite organization, Dataiku reduces 90% of the manual, repeated tasks involved in regular reporting. This is equivalent to a risk-adjusted, three-year present value of \$736,168.
- **Increased profit from improved customer targeting, amounting to over \$500,000.** Interviewees' organizations reported that Dataiku AI models improved customer segmentation and targeting, bringing additional profit from recommendations and promotions. For the composite organization, this results in a risk-adjusted, three-year present value of \$503,080.

Unquantified benefits. Benefits that are not quantified for this study include:

- **Improved ability to work on data and analytics projects remotely.** Interviewees highlighted that, because Dataiku is cloud-based, simple browser access allows data engineers, scientists, and analysts to access all the data they need as well as their past work. This provides the flexibility and mobility needed during the extended period of remote working caused by the pandemic.
- **Improved work satisfaction for data engineers, data scientists, and analysts.** Some interviewees noted increased satisfaction

and a boost to morale as a result of Dataiku. For example, one interviewee noted: "People are excited to use it. There's this positive morale-boosting aspect that was unintended and I have a quote from a managing director saying, 'Work is fun again.'"

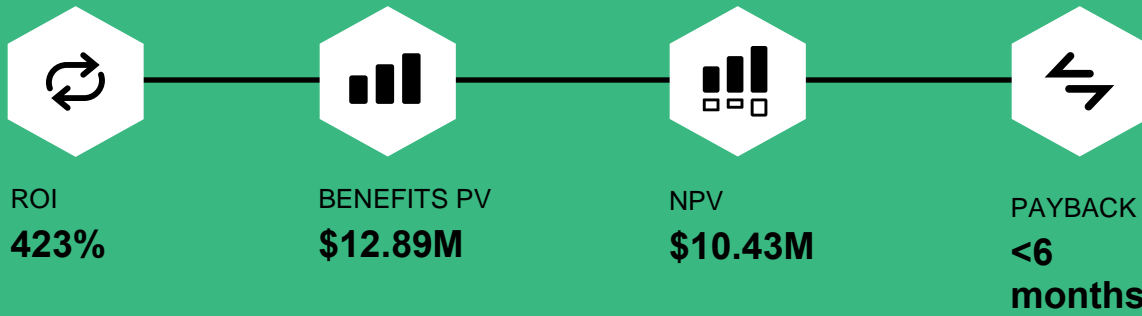
Costs. Risk-adjusted PV costs include:

- **Platform and license fees.** This was the biggest cost organizations incurred. For the composite organization, the risk-adjusted three-year present value of Dataiku platform and license fees comes to just under \$1,866,491.
- **Implementation and maintenance costs.** For the composite organization, the initial implementation and ongoing maintenance costs come to just under \$436,502. This includes the efforts involved for initial implementation and ongoing platform administration, maintenance, and user support resources.
- **User training costs.** Depending on their user persona, Dataiku users in the composite organization spend either 1 hour, 8 hours, or 16 hours on average on training and onboarding. This amounted to a cost of just over \$164,311.

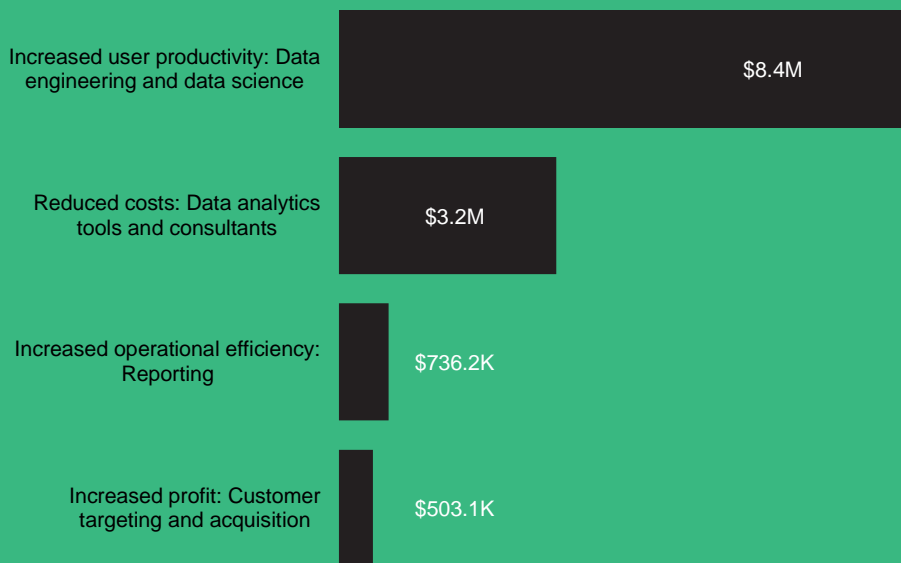
"By having these reusable data pipelines and data products, [we have] streamlined our operational side of development. We're talking about savings in the range of \$4 million plus."

Team Lead, Analytics Innovation, pharmaceutical company

The customer interviews and financial analysis found that a composite organization experiences benefits of \$12.89M over three years versus costs of \$2.47M, adding up to a net present value (NPV) of \$10.43M and an ROI of 423%.



Benefits (Three-Year)



The biggest benefit was the increased user productivity for data engineers and data scientists.

Organizations also experienced reduced costs, increased operational efficiency, and increased profit.

TEI FRAMEWORK AND METHODOLOGY

From the information provided in the interviews, Forrester constructed a Total Economic Impact™ framework for those organizations considering an investment in Dataiku.

The objective of the framework is to identify the cost, benefit, flexibility, and risk factors that affect the investment decision. Forrester took a multistep approach to evaluate the impact that Dataiku can have on an organization.

DISCLOSURES

Readers should be aware of the following:

This study is commissioned by Dataiku and delivered by Forrester Consulting. It is not meant to be used as a competitive analysis.

Forrester makes no assumptions as to the potential ROI that other organizations will receive. Forrester strongly advises that readers use their own estimates within the framework provided in the study to determine the appropriateness of an investment in Dataiku.

Dataiku reviewed and provided feedback to Forrester, but Forrester maintains editorial control over the study and its findings and does not accept changes to the study that contradict Forrester's findings or obscure the meaning of the study.

Dataiku provided the customer names for the interviews but did not participate in the interviews.



DUE DILIGENCE

Interviewed Dataiku stakeholders and Forrester analysts to gather data relative to the Dataiku.



CUSTOMER INTERVIEWS

Interviewed four decision-makers at organizations using the Dataiku to obtain data with respect to costs, benefits, and risks.



COMPOSITE ORGANIZATION

Designed a composite organization based on characteristics of the interviewed organizations.



FINANCIAL MODEL FRAMEWORK

Constructed a financial model representative of the interviews using the TEI methodology and risk-adjusted the financial model based on issues and concerns of the interviewed organizations.



CASE STUDY

Employed four fundamental elements of TEI in modeling the investment impact: benefits, costs, flexibility, and risks. Given the increasing sophistication of ROI analyses related to IT investments, Forrester's TEI methodology provides a complete picture of the total economic impact of purchase decisions. Please see Appendix A for additional information on the TEI methodology.

The Dataiku Customer Journey

■ Drivers leading to the Dataiku investment

Interviewed Organizations			
Industry	Region	Size	Interviewee
Pharmaceutical	USA	80,000 employees	Director — Team Lead, Analytics Innovation
Financial services	USA	48,400 employees	Head of Technology and Innovation
Aerospace	USA	48,000 employees	Product Owner
Retail	USA	15,000 employees	Data, Analytics, and AI Lead, Europe

KEY CHALLENGES

Prior to investing in Dataiku, the interviewees' organizations typically had data silos across the organization and a fragmented set of applications, including ETL tools and data science platforms. One customer mentioned that they "had every tool out in the market with different business functions investing in different tools." Some of the interviewees' companies had a centralized data science function.

The interviewees' organizations struggled with common challenges, including:

- **Poor accessibility of datasets and the inability to collaborate.** Interviewees described a number of pain points for data scientists and business analysts. Along with the fragmentation of tools, they faced challenges with finding, connecting, and accessing datasets and collaborating on analytics projects and initiatives.
- **Too much manual work involved in data extraction and transformation.** With the data silos and fragmented toolset, the process of extracting and transforming data required a lot of manual work. One interviewee described it as "almost like empires, [where] you had to know somebody to get data. There was a lot of emailing of spreadsheets and data not dynamically transmitted."

INVESTMENT OBJECTIVES AND SOLUTION REQUIREMENTS

Interviewees noted the overall need to transform their companies into data-led organizations, and harness the power of analytics, AI, and machine learning. The primary catalysts driving the investment in Dataiku were the need to rationalize the data science and analytics tool set, enable collaboration, and reduce pain points for data scientists and analysts. The interviewees' organizations searched for a solution that allowed them to:

- **Consolidate different applications and databases.** One interviewee described the requirement for a "data operating system to help orchestrate the mess of different systems and address compatibility."
- **Democratize data science and analytics.** One of the key requirements of interviewees' organizations was a collaborative solution that different types of users could access and utilize, including data scientists, analysts, and business stakeholders.

"We really wanted a data operating system to orchestrate the mess of different systems."

*Head of Technology & Innovation,
Financial services*

COMPOSITE ORGANIZATION

Based on the interviews, Forrester constructed a TEI framework, a composite company, and a ROI analysis that illustrates the areas financially affected. The composite organization is representative of the four companies that Forrester interviewed and is used to present the aggregate financial analysis in the next section. The composite organization has the following characteristics:

Description of composite. The global, multibillion-dollar business-to-consumer organization operates in the manufacturing sector. The organization has a strong brand and global operations with 45,000 employees and annual revenue of about \$20 billion. Prior to the investment in Dataiku, the organization had a highly fragmented tool set for data science and analytics with datasets held in silos in different parts of the organization, such as production, supply chain/logistics, marketing/sales, and finance.

Deployment characteristics. The organization has a centralized data engineering and data scientist team with 30 individuals initially, along with 20 analysts across different lines of business who carry out analytics tasks like reporting, dashboard building, and modeling. There are a further 20 business users that need to access outputs like dashboards and data visualizations. Each of these types of users roughly double every year for the organization.

KEY ASSUMPTIONS

- **Manufacturing**
- **45,000 employees**
- **\$20 billion in annual revenue**
- **Centralized data science and analytics team**

Analysis Of Benefits

■ Quantified benefit data as applied to the composite

Total Benefits						
Ref.	Benefit	Year 1	Year 2	Year 3	Total	Present Value
Atr	Increased user productivity: Data engineering and data science	\$1,516,320	\$3,032,640	\$6,065,280	\$10,614,240	\$8,441,721
Btr	Reduced costs: Data analytics tools and consultants	\$318,370	\$996,739	\$2,800,109	\$4,115,218	\$3,216,941
Ctr	Increased operational efficiency: Reporting	\$154,094	\$308,189	\$454,378	\$916,661	\$736,168
Dtr	Increased profit: Customer targeting and acquisition	\$0	\$216,000	\$432,000	\$648,000	\$503,080
Total benefits (risk-adjusted)		\$1,988,784	\$4,553,568	\$9,751,766	\$16,294,118	\$12,897,910

INCREASED USER PRODUCTIVITY: DATA ENGINEERING AND DATA SCIENCE

Evidence and data. All interviewees highlighted significant time savings on their organizations' pipeline of data projects. The ability to set up reusable data products and models has increased the efficiency of data engineering and data science tasks.

- Director – Team Lead, Analytics Innovation at a pharmaceutical company: “By having these reusable data pipelines and data products, [we have] streamlined our operational side of development. We’re talking about savings in the range of \$4 million plus. There’s efficiencies on the dollar savings side and the time side because, typically, a project would take anywhere from three to six months, and now that we’ve standardized our data products, we can turn things around in four to six weeks.”
- Head of Technology and Innovation at a financial services company: “In the past, people would have to run a structured query language (SQL) query against the database, save that as a spreadsheet, do the same thing against another database, save it as a spreadsheet, and then use

V-lookups to bring it all together. That is not a scalable solution, and you are limited by your computer's resources, so if the spreadsheet is over 70 megabytes you can't even open it. It [now] takes minutes for that sort of thing.”

- “Everybody can say it's saves time. There are different ways that it can save time — if there is a manual task that's been automated, then that's one example. In other cases, it's just the speed of accessing data. Even if you save a couple of minutes here and there throughout the day, it adds up quickly.”

Modeling and assumptions. For the composite organization, Forrester assumes:

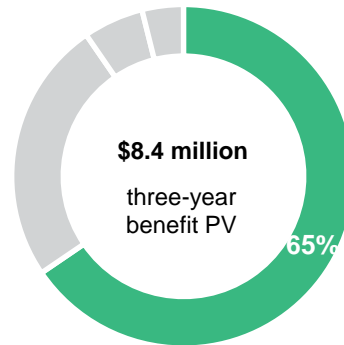
- A centralized team of data engineers and data scientists of 30 in Year 1, which doubles every year.
- Each data engineer or data scientists works on 10 projects on average per year. Each project takes 8 hours a week for six months on average, which totals 192 hours per year for each project.
- Data engineers and data scientists save 75% of time previously spent on these projects through

using Dataiku, and spend 50% of this saved time on other work tasks and projects.

Risks. Risks that could impact the realization of this benefit include:

- Establishing a range of use cases, which might take some time after implementation and can impact how quickly organizations realize this benefit in different scenarios.
- Experienced data engineers and data scientists may have strong preferences on ways of working and tools/platforms they like to use. This may delay the adoption of new platforms, and organizations will need effective change management to mitigate this.

Results. To account for these risks, Forrester adjusted this benefit downward by 10%, yielding a three-year, risk-adjusted total PV (discounted at 10%) of \$8,441,721.



Increased User Productivity: Data Engineering And Data Science

Ref.	Metric	Calculation	Year 1	Year 2	Year 3
A1	Number of data engineers and data scientists	Composite organization	30	60	120
A2	Hourly rate per person	Assumption	\$78	\$78	\$78
A3	Average number of projects per person (annual)	Composite organization	10	10	10
A4	Average hours spent per project before Dataiku	8 hours*24 weeks	192	192	192
A5	Time saved with Dataiku per project	Composite organization	75%	75%	75%
A6	Percent captured	Composite organization	50%	50%	50%
At	Increased user productivity: Data engineering and data science	$A1 \cdot A2 \cdot A3 \cdot A4 \cdot A5 \cdot A6$	\$1,684,800	\$3,369,600	\$6,739,200
	Risk adjustment	↓10%			
Atr	Increased user productivity: Data engineering and data science (risk-adjusted)		\$1,516,320	\$3,032,640	\$6,065,280
Three-year total: \$10,614,240			Three-year present value: \$8,441,721		

REDUCED COSTS: DATA ANALYTICS TOOLS AND CONSULTANTS

Evidence and data. With Dataiku catering to a wide range of use cases and user personas, interviewees' organizations have saved on the costs of previous tools used for ETL, statistics, and analytics. In addition to saving license fees, they also saved on developer and maintenance time for legacy tools. Other cost savings include a reduction in consultant fees for analytics projects that were previously outsourced. Interviewees shared that:

- Head of Technology and Innovation at a financial services company: "There have been a handful of examples where something was retired as a result of Dataiku. We were able to remove an application that was \$350,000 a year — and that was just the licensing cost. Because we removed a lot of dependencies on it, all it really did was one thing by the time it was retired. The time spent internally to develop it and maintain it has been removed as well."
- Director – Team Lead, Analytics Innovation at a pharmaceutical company: "There will be a lot of tool rationalization that will happen in the data wrangling and machine learning space. There will also be a consolidation in different environments that are being set up on cloud infrastructure."
- Data, Analytics, and AI lead at retail company: "In terms of reporting systems, from an analyst's point of view, things have been decommissioned."

Modeling and assumptions. For the composite organization, Forrester assumes:

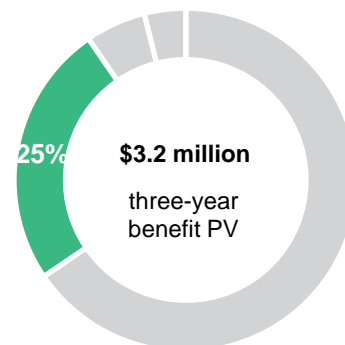
- The composite removes one tool in Year 1 and an additional four tools are removed over the next two years.
- Each tool has one developer associated, who spends on average 2 hours a week per tool, which totals 96 hours per tool.

- The composite organization previously outsourced ten analytics and data science projects on average per year to consultants, costing \$10,000 per project.

Risks. Risks that could impact the realization of this benefit include:

- Decommissioning legacy tools can be a long and complex process at large companies due to the contractual and legal requirements. This may delay the savings organizations realize from removing tools that have become redundant due to Dataiku implementation.

Results. To account for these risks, Forrester adjusted this benefit downward by 10%, yielding a three-year, risk-adjusted total PV of \$3,216,941.



Reduced Costs: Data Analytics Tools And Consultants

Ref.	Metric	Calculation	Year 1	Year 2	Year 3
B1	Number of assets	Interviews	1	2	3
B2	License fees per asset per user	Composite organization	\$5,000	\$5,000	\$5,000
B3	Number of users	Composite organization	50	100	200
B4	Percent captured	Assumption	100%	100%	100%
B5	Number of developers per asset	Composite organization	1	1	1
B6	Hourly rate for developers	Assumption	\$78	\$78	\$78
B7	Annual hours saved per developer per asset	2 hours*48 weeks	96	96	96
B8	Percent captured	Assumption	50%	50%	50%
B9	Number of outsourced projects	Composite organization	10	10	10
B10	Consultant fees per project	Interviews	\$10,000	\$10,000	\$10,000
B11	Percent captured	Assumption	100%	100%	100%
Bt	Reduced costs: Data analytics tools and consultants	$(B1*B2*B3*B4)+(B1*B5*B6*B7*B8)+(B9*B10*B11)$	\$353,744	\$1,107,488	\$3,111,232
	Risk adjustment	↓10%			
Btr	Reduced costs: Data analytics tools and consultants (risk-adjusted)		\$318,370	\$996,739	\$2,800,109
Three-year total: \$4,115,218			Three-year present value: \$3,216,941		

INCREASED OPERATIONAL EFFICIENCY: REPORTING

Evidence and data. The automation of manual, repeated tasks enabled operational efficiencies for interviewees’ organizations. A key area of impact was in reporting. Interviewees shared that:

- Head of Technology and Innovation at a financial services company: “We have one team that was using multiple databases: spreadsheets that were emailed to them, spreadsheets that were dropped out on network drives, and lots of lookup spreadsheets as well. They had a very complicated process with many steps where, eventually, they would come up with financial

data for revenue forecasting and internal reporting. And they have dropped that process down to 99% of what it was through this automation reproducibility.”

- Data, Analytics, and AI lead at a retail company: “Data analysts would [previously] spend entire days downloading data and this is terrible for data lineage, data privacy, and security. They were then storing these in [an information management tool] and building reports on top of these. Today, they have all the reports scheduled and they work on [other] projects. We’re talking about three people generating an incredible amount of work for all of Europe.”

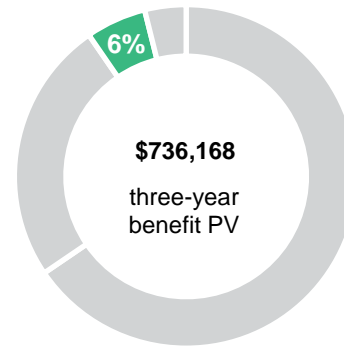
Modeling and assumptions. For the composite organization, Forrester assumes:

- Twenty analysts and business users who need to produce regular reporting are employed in Year 1 with this number doubling every year.
- Analysts and business users spend on average two days per month on manual reporting tasks, which totals 192 hours per year per user.
- Analysts and business users save 90% of time previously spent on manual reporting tasks through automated, recurring reports on Dataiku, and spend 50% of this saved time on other work tasks and projects.
- The organization reduces headcount for reporting tasks by 5 FTEs over three years.

Risks. Risks that could impact the realization of this benefit include:

- The time and expertise required to set up automated reports, which may not be immediately available and can delay the realization of this benefit in some scenarios or use cases.
- Experienced analysts and business users may have strong preferences and habits around reporting, which may slow down the adoption of new tools.

Results. To account for these risks, Forrester adjusted this benefit downward by 10%, yielding a three-year, risk-adjusted total PV of \$736,168.



Increased Operational Efficiency: Reporting					
Ref.	Metric	Calculation	Year 1	Year 2	Year 3
C1	Number of analysts and business users	Composite organization	20	40	80
C2	Hourly rate per person	Assumption	\$47	\$47	\$47
C3	Manual hours spent on reporting	16 hours*12 months	192	192	192
C4	Time saved on Dataiku	Assumption	90%	90%	90%
C5	Percent captured	Assumption	50%	50%	50%
C6	Reduction in headcount	Interviews	1	2	2
C7	Average salary (annual)	Assumption	\$90,000	\$90,000	\$90,000
C8	Percent captured	Assumption	100%	100%	100%
Ct	Increased operational efficiency: Reporting	$(C1 * C2 * C3 * C4 * C5) + (C6 * C7 * C8)$	\$171,216	\$342,432	\$504,864
	Risk adjustment	↓10%			
Ctr	Increased operational efficiency: Reporting (risk-adjusted)		\$154,094	\$308,189	\$454,378
Three-year total: \$916,661			Three-year present value: \$736,168		

INCREASED PROFIT: CUSTOMER TARGETING AND ACQUISITION

Evidence and data. Interviewees’ organizations reported that AI models in Dataiku improved customer segmentation and targeting, resulting in additional revenue and profit from recommendations and promotions:

- Director – Team Lead, Analytics Innovation at a pharmaceutical company: “[Dataiku led to us] creating customer 360 data products. So, [bringing] our syndicated purchased data that we have, along with some of our internal transaction data sets between sales and marketing systems, allowed us to start to connect information about a customer across these different sectors. Then, we can really publicize it and use a consistent and standardized data set.”
- Data, Analytics, and AI lead at a retail company: “We have promotions [where we have] recommendations of level of percentage discounts when running promotions and mark downs. Then we have pricing, so recommendations of full-season official pricing. Then, we have recommending systems for e-commerce and overall personalization of search and navigation.”
- “The promotions work stream was a result of that first lockdown work that was done very quickly. Between May and December 2020, we made \$3.6 million as an uplift from our work. And I don’t think we would have been live at the time had we not used Dataiku.”

Modeling and assumptions. For the composite organization, Forrester assumes:

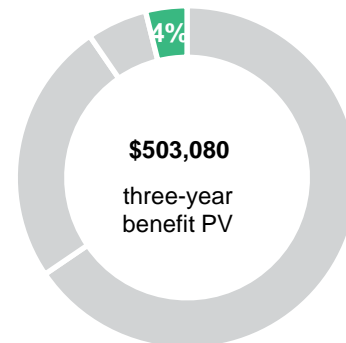
- The organization runs two promotions in Year 2, and four promotions in Year 3 using the AI-enabled targeting and personalization on Dataiku.

- An average gross profit margin of 3%, which is aligned to the manufacturing industry average.⁴

Risks. Risks that could impact the realization of this benefit include:

- Bringing together internal and external customer data, which might be complicated due to data governance and security processes in places. This may delay the ability of the organizations to deploy Dataiku to create new promotions or recommendation engines.
- The marketing team’s time and resources for projects related to trial campaigns and promotions may not be immediately available, which could delay the realization of this benefit.

Results. To account for these risks, Forrester adjusted this benefit downward by 10%, yielding a three-year, risk-adjusted total PV of \$503,080.



Increased Profit: Customer Targeting And Acquisition

Ref.	Metric	Calculation	Year 1	Year 2	Year 3
D1	Number of promotions	Composite organization	0	2	4
D2	Additional revenue per promotion	Interviews	\$0	\$4,000,000	\$4,000,000
D3	Average profit margin	Assumption	3%	3%	3%
Dt	Increased revenue: Customer targeting and acquisition	D1*D2*D3	\$0	\$240,000	\$480,000
	Risk adjustment	↓10%			
Dtr	Increased revenue: Customer targeting and acquisition (risk-adjusted)		\$0	\$216,000	\$432,000
Three-year total: \$648,000			Three-year present value: \$503,080		

UNQUANTIFIED BENEFITS

Additional benefits that customers experienced but were not able to quantify include:

- Improved ability to work on data and analytics projects remotely.** As one interviewee highlighted: “Now that we are all remote, teams have different ways of accessing their local computers. Sometimes they were remoting back into the office or a virtual environment where they could remote in and do some of their work. But if you are a coder and you rely on a local [tool], you were inconvenienced once you were sent home. But because Dataiku is cloud-based, all you need is the browser access, and you have access to all the data you need and all the work you completed in the past. So, a lot of people have been using this as way too move towards that flexibility and mobility that we need now, and untethering from the network at work.”
- Improved work satisfaction for data engineers, data scientists, and analysts.** Some interviewees noted increase satisfaction and a boost in morale as a result of using Dataiku. For example, one noted: “There’s this

positive morale-boosting aspect [with Dataiku] that was unintended and I have a quote from a managing director saying, ‘Work is fun again.’ We have a lot of people who brag about this and how much time it saved them.”

FLEXIBILITY

The value of flexibility is unique to each customer. There are multiple scenarios in which a customer might implement Dataiku and later realize additional uses and business opportunities, including:

- Embedding artificial intelligence and machine learning to gain business value.** One interviewee highlighted that harnessing the power of AI and machine learning was a key focus: “Dataiku makes it very easy to leverage some of the built-in, open-source AI and models, but also lets our analysts and data scientists customize and use packages they are familiar with. [We can apply] this to patient identification, patient diagnosis, treatment adherence, and types of use cases to help influence either sales and marketing strategies, or help the healthcare ecosystem identify patients for rare conditions.”

- **Predictive analytics for maintenance of supply chains and assets.** In industries with complex, multi-layered supply chains, predictive maintenance is critical to reducing operational risks and lowering costs associated with unplanned downtime. One interviewee summarized the potential impact of Dataiku in their supply chain: “We have hundreds and hundreds of suppliers out there that we need to source parts from. That is a very large logistical process. So, a data scientist would be building a project that would maybe look at how many suppliers or delivered parts on time versus not. And then taking a step further from the historical data predicting how late a supplier is going to be with the next part coming in. That is going to help a shop floor determine how many products can be built in the coming time.”

Flexibility would also be quantified when evaluated as part of a specific project (described in more detail in [Appendix A](#)).

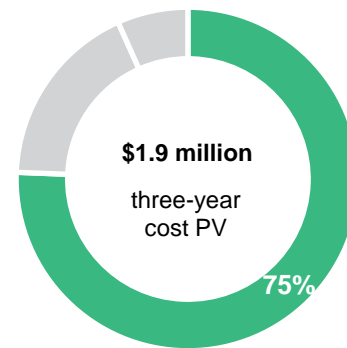
Analysis Of Costs

■ Quantified cost data as applied to the composite

Total Costs							
Ref.	Cost	Initial	Year 1	Year 2	Year 3	Total	Present Value
Etr	Platform and license fees	\$0	\$472,500	\$698,250	\$1,144,500	\$2,315,250	\$1,866,491
Ftr	Implementation and maintenance costs	\$24,960	\$143,770	\$143,770	\$215,654	\$528,154	\$436,502
Gtr	User training costs	\$0	\$50,490	\$51,007	\$101,497	\$202,994	\$164,311
	Total costs (risk-adjusted)	\$24,960	\$666,760	\$893,027	\$1,461,651	\$3,046,398	\$2,467,304

PLATFORM AND LICENSE FEES

In the case of the composite organization, the platform fee is \$240,000 per year. The license fee is \$4,000 per user per year for Designers and \$500 per user per year for Readers. These fees remain constant over the three-year period. There is a low risk level and the resulting three-year present value comes to just under \$1,900,000. This is equivalent to 75% of the total costs.



Platform And License Fees						
Ref.	Metric	Calculation	Initial	Year 1	Year 2	Year 3
E1	Platform fee	Interviews		\$240,000	\$240,000	\$240,000
E2	License fees per user (Designer)	Interviews		\$4,000	\$4,000	\$4,000
E3	Number of users (Designer)	Composite organization		50	100	200
E4	License fees per user (Reader)	Interviews		\$500	\$500	\$500
E5	Number of users (Reader)	Composite organization		20	50	100
Et	Platform and license fees	$E1+(E2 \times E3)+(E4 \times E5)$	\$0	\$450,000	\$665,000	\$1,090,000
	Risk adjustment	↑5%				
Etr	Platform and license fees (risk-adjusted)		\$0	\$472,500	\$698,250	\$1,144,500
Three-year total: \$2,315,250			Three-year present value: \$1,866,491			

IMPLEMENTATION AND MAINTENANCE COSTS

Evidence and data. The implementation journey varied by organization with resources ranging from teams of five to 15, and time ranging from three to six weeks. Implementation teams were internal resources in most cases. Ongoing maintenance required 1 to 2 FTEs according to interviewees' companies with some additional user support required in some cases. For example:

- Regarding initial implementation, one interviewee stated that their organization needed three people from the technology team and a project manager. Another noted that their organization had a team of five to 10 involved in the technology implementation, including people from engineering and DevOps. In one case, the interviewee highlighted that their organization had a group of 15 people involved in rolling Dataiku out to around 300 users.
- Regarding ongoing maintenance, one interviewee noted that 60% of a platform administrator's time is spent on maintenance. The Head of Technology and Innovation at the financial services company stated: "Eighty percent of my time is spent on something Dataiku-related and then the other 20% is spent on other projects."

Modeling and assumptions. For the composite organization, Forrester assumes:

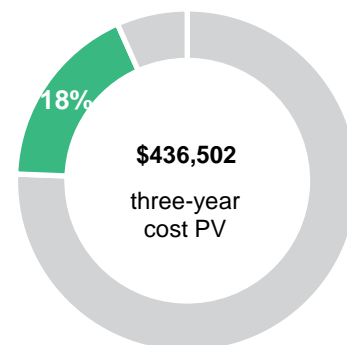
- An initial internal implementation team of five, consisting of a project manager, technology team members, and engineers.
- An implementation period of four weeks, during which the core team of five spent 50% of their time on implementation and technical configuration, which totaled 80 hours of effort.
- A team of two in Year 1, Year 2, and Year 3 responsible for ongoing maintenance, including platform administration and user support.

- The ongoing maintenance team spends on average 60% of their time on Dataiku maintenance tasks.

Risks. Risks that could affect the magnitude of these costs include:

- Approval processes may consist of many layers, involving different teams such as security committees, since in highly regulated industries, data is subject to several regulations, especially customer data. This could prolong the implementation process and result in a delay of the rollout of the platform to users.
- Establishing the connection and authentication of multiple data sources may be another source of delay to approvals and the implementation process. For example, different data sources might apply different security models, which need to be connected up ahead of implementation.
- Customers may also incur some professional services costs in the first year, which would increase the initial implementation costs.

Results. To account for these risks, Forrester adjusted this cost upward by 20%, yielding a three-year, risk-adjusted total PV of \$436,502. This is equivalent to 18% of the total costs.



Implementation And Maintenance Costs

Ref.	Metric	Calculation	Initial	Year 1	Year 2	Year 3
F1	Number of employees	Interviews	5	2	2	3
F2	Hourly rate per person	Assumption	\$52	\$52	\$52	\$52
F3	Hours for initial implementation	Initial: (20 days*8 hours)*0.5	80			
F4	Hours for ongoing management	Y1, Y2, and Y3: (48 weeks*5 days*8 hours)*0.6		1,152	1,152	1,152
Ft	Implementation and maintenance costs	$(F1*F2*F3)+(F1*F2*F4)$	\$20,800	\$119,808	\$119,808	\$179,712
	Risk adjustment	↑20%				
Ftr	Implementation and maintenance costs (risk-adjusted)		\$24,960	\$143,770	\$143,770	\$215,654
Three-year total: \$528,154			Three-year present value: \$436,502			

USER TRAINING COSTS

Evidence and data. User training on Dataiku varies according to the use profile. Along with initial training modules from Dataiku, users spend some of their own time gaining understanding of the platform. Interviewees shared that:

- “The official training module is two full days. For some types of users, [we] onboard them with the documentation and then have a two-to-three-hour personalized follow-up with Dataiku.”
- “If you’re a data scientist or power user of [similar] tools, you can be up and running in two to three days. For business scientists, the learning curve is more like three to four weeks.”
- “A data scientist who’s been writing Python, creating analytics on their desktop, or maybe in a different tool, it’s not going to take them much time to learn Dataiku — a few days, if that.”

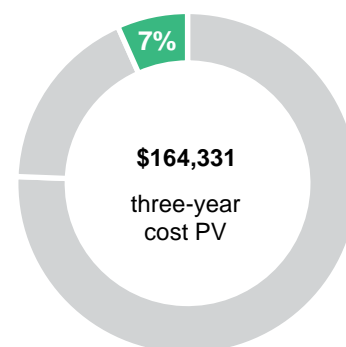
Modeling and assumptions. For the composite organization, Forrester assumes:

- 50 Designer users in Year 1, 100 in Year 2, and 200 in Year 3 with a 60-40 split between data engineers/scientists and analysts.
- 20 Reader users in Year 1, 50 in Year 2, and 100 in Year 3, comprising mostly of business users.

Risks. Risks that could affect the magnitude of these costs include:

- Onboarding and training could be delayed for some users due to a reluctance to move away from their existing tools, however this could be mitigated with effective change management.

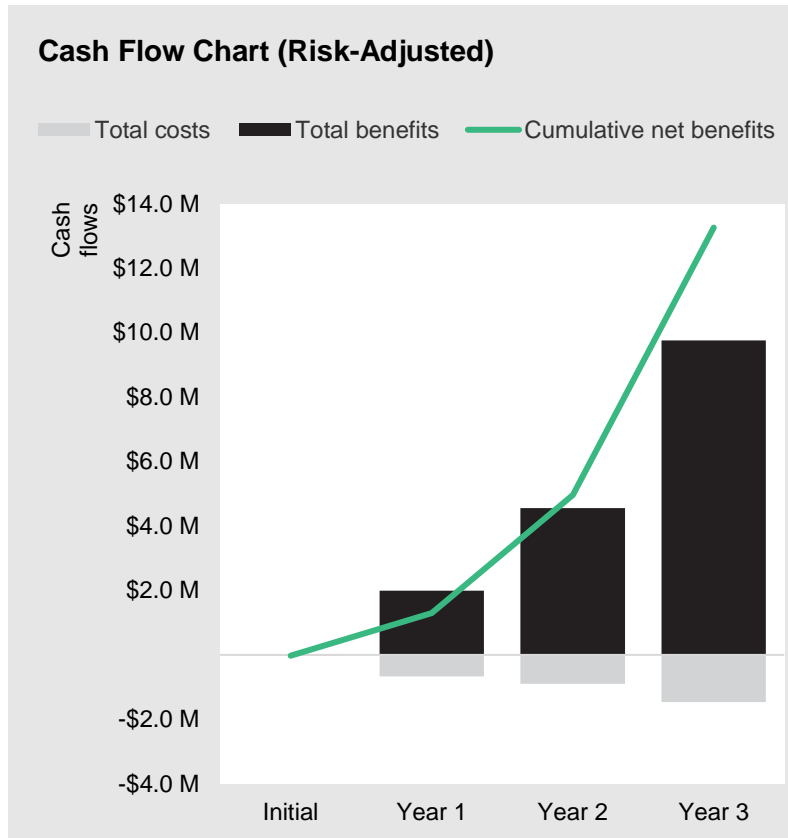
Results. To account for these risks, Forrester adjusted this cost upward by 10%, yielding a three-year, risk-adjusted total PV of \$164,331. This is equivalent to 7% of the total costs.



User Training Costs						
	Metric	Calculation	Initial	Year 1	Year 2	Year 3
G1	Number of Designer users	Composite organization		50	100	200
G2	Number of Designer data scientist/engineer users	Composite organization		30	60	120
G3	Net-new data scientist/engineer users	Composite organization		30	30	60
G4	Training hours per data scientist/engineer	Assumption		16	16	16
G5	Hourly rate per data scientist/engineer	Assumption		\$78	\$78	\$78
G6	Number of Designer analyst users	Composite organization		20	40	80
G7	Net-new analyst users	Composite organization		20	20	40
G8	Training hours per analysts	Assumption		8	8	8
G9	Number of Reader business users	Composite organization		20	50	100
G10	Net-new business users	Composite organization		20	30	50
G11	Training hours per business user	Assumption		1	1	1
G12	Hourly rate per analyst/business user	Assumption		\$47	\$47	\$47
Gt	User training costs	$(G3 \times G4 \times G5) + (G7 \times G8 \times G12) + (G10 \times G11 \times G12)$	\$0	\$45,900	\$46,370	\$92,270
	Risk adjustment	↑10%				
Gtr	User training costs (risk-adjusted)		\$0	\$50,490	\$51,007	\$101,497
Three-year total: \$202,994			Three-year present value: \$164,311			

Financial Summary

CONSOLIDATED THREE-YEAR RISK-ADJUSTED METRICS



The financial results calculated in the Benefits and Costs sections can be used to determine the ROI, NPV, and payback period for the composite organization's investment. Forrester assumes a yearly discount rate of 10% for this analysis.

These risk-adjusted ROI, NPV, and payback period values are determined by applying risk-adjustment factors to the unadjusted results in each Benefit and Cost section.

Cash Flow Analysis (Risk-Adjusted Estimates)

	Initial	Year 1	Year 2	Year 3	Total	Present Value
Total costs	(\$24,960)	(\$666,760)	(\$893,027)	(\$1,461,651)	(\$3,046,398)	(\$2,467,304)
Total benefits	\$0	\$1,988,784	\$4,553,568	\$9,751,766	\$16,294,118	\$12,897,910
Net benefits	(\$24,960)	\$1,322,024	\$3,660,541	\$8,290,115	\$13,247,721	\$10,430,606
ROI						423%
Payback period (months)						<6

Appendix A: Total Economic Impact

Total Economic Impact is a methodology developed by Forrester Research that enhances a company's technology decision-making processes and assists vendors in communicating the value proposition of their products and services to clients. The TEI methodology helps companies demonstrate, justify, and realize the tangible value of IT initiatives to both senior management and other key business stakeholders.

TOTAL ECONOMIC IMPACT APPROACH

Benefits represent the value delivered to the business by the product. The TEI methodology places equal weight on the measure of benefits and the measure of costs, allowing for a full examination of the effect of the technology on the entire organization.

Costs consider all expenses necessary to deliver the proposed value, or benefits, of the product. The cost category within TEI captures incremental costs over the existing environment for ongoing costs associated with the solution.

Flexibility represents the strategic value that can be obtained for some future additional investment building on top of the initial investment already made. Having the ability to capture that benefit has a PV that can be estimated.

Risks measure the uncertainty of benefit and cost estimates given: 1) the likelihood that estimates will meet original projections and 2) the likelihood that estimates will be tracked over time. TEI risk factors are based on "triangular distribution."

The initial investment column contains costs incurred at "time 0" or at the beginning of Year 1 that are not discounted. All other cash flows are discounted using the discount rate at the end of the year. PV calculations are calculated for each total cost and benefit estimate. NPV calculations in the summary tables are the sum of the initial investment and the discounted cash flows in each year. Sums and present value calculations of the Total Benefits, Total Costs, and Cash Flow tables may not exactly add up, as some rounding may occur.



PRESENT VALUE (PV)

The present or current value of (discounted) cost and benefit estimates given at an interest rate (the discount rate). The PV of costs and benefits feed into the total NPV of cash flows.



NET PRESENT VALUE (NPV)

The present or current value of (discounted) future net cash flows given an interest rate (the discount rate). A positive project NPV normally indicates that the investment should be made, unless other projects have higher NPVs.



RETURN ON INVESTMENT (ROI)

A project's expected return in percentage terms. ROI is calculated by dividing net benefits (benefits less costs) by costs.



DISCOUNT RATE

The interest rate used in cash flow analysis to take into account the time value of money. Organizations typically use discount rates between 8% and 16%.



PAYBACK PERIOD

The breakeven point for an investment. This is the point in time at which net benefits (benefits minus costs) equal initial investment or cost.

Appendix B: Supplemental Material

Related Forrester Research

“Introducing ModelOps To Operationalize AI,” Forrester Research, Inc., August 13, 2020.

“The Forrester Wave™: Multimodal Predictive Analytics And Machine Learning, Q3 2020,” Forrester Research, Inc., September 10, 2020.

Appendix C: Endnotes

¹ Source: “Introducing ModelOps To Operationalize AI,” Forrester Research, Inc., August 13, 2020.

² Source: “The Forrester Wave™: Multimodal Predictive Analytics And Machine Learning, Q3 2020,” Forrester Research, Inc., September 10, 2020.

³ Total Economic Impact is a methodology developed by Forrester Research that enhances a company’s technology decision-making processes and assists vendors in communicating the value proposition of their products and services to clients. The TEI methodology helps companies demonstrate, justify, and realize the tangible value of IT initiatives to both senior management and other key business stakeholders

⁴ “Gross margin — breakdown by industry,” Ready Ratios, 2020 (<https://www.readyratios.com/sec/ratio/gross-margin/>).

FORRESTER®