data iku

EBOOK

87 Requirements to Include in Your Predictive Analytics and Machine Learning RFP

Introduction

Whether you're a small team of data scientists or a large organization trying to make data analysis accessible even to those in non-technical roles, the tool you choose will help define how you work with data now and in the future. The balance you strike between addressing your immediate needs versus anticipating your future requirements will play a big role in creating demonstrable, sustainable value.

Immediate needs depend upon your team, your technology, your industry, and your specific organization. Teams have different skill levels and types of expertise, such as the coding languages they use and the techniques they are familiar with. Technologies range from data lakes to visualization solutions, and new technologies emerge all the time. Industries have regulatory requirements that evolve as different agencies weigh the role of policy in various sectors of the economy. And your specific organization has its own processes and use cases that will drive adoption of and value creation from data analytics.

Future requirements are much more difficult to define. Which skills will be required in three years? And how difficult will it be to recruit talent with those skills? Which new technologies will emerge, and will your existing technology be compatible with them? What will the regulatory and economic environment look like? And how will your own organization's culture and processes evolve?

We have created this data analytics RFP template to save you time and also to suggest some evaluation criteria on versatility and scalability. Feel free to add and subtract requirements in accordance with your needs and preferences. And don't hesitate to reach out to us directly if you have any questions about any of the lines included here.

Vendor Profile

The vendor profile provides the essential facts and context about the vendors you are evaluating. We emphasize the ability of the vendor to provide dedicated customer support, which is especially important in the still-nascent world of advanced data analytics.

Analytics RFP

This section contains the technical details of the RFP. Below, we will briefly describe the sections. For individual line items, please refer to the spreadsheet.

Data Access

Database Connectivity	Evaluate the breadth of connectivity of the solution, including traditional data sources, new data sources (NoSQL), and cloud.
Data Transformation and Preparation	Evaluate the ability of the solution to prepare and transform data. This includes parsing, normalising, and aggregating data. Evaluate how easy it is to enrich existing data with first-party or third-party sources.
Data Discovery	Evaluate the ability to search and discover data with the product

Coding and Extensibility

Programming Languages Support	Evaluate the product's language support both in terms of breadth of languages supported and quality of support for each
Development Environment	Evaluate the ability of the platform to support and maintain code. This includes every possible aspect of modern Integrated Development Environment: error detection, completion, versioning, etc.

Enterprise-Wide Adoption

The goal of this section is to evaluate the overall ability to deploy and achieve adoption of the product among a team.

Usability by Business User	Evaluate the overall ability of the product to be used by business analysts. The evaluation comprises evaluating the product documentation, user experience, and the overall learning curve to be expected with the product.
Project Management	Evaluate the ability of the product to create and maintain projects and collaborate within a project as a team. This includes the ability to document a project, reuse components, publish results, and organize overall collaboration and interactions among team members

Data Exploration and Visualisation

Simple and Advanced Data Exploration	Evaluate the ability to navigate within a dataset, as well as to perform statistical analysis of the dataset, in terms of data quality, data distribution, etc.
Charting and Reporting	Evaluate the ability to create charts and reports within a dataset. Evaluate ease of use of charting, overall quality, and completeness in terms of number and complexity of charts.
Interactive and Advanced Visualizations	Evaluate the ability to create more interactive and advanced visualization, including graph analytics, fractals charts, etc., that possibly leverage Javascript frameworks and customization capabilities.

Machine Learning

Algorithm Completeness	Evaluate the completeness of the platform in terms of algorithms and types of data supported for machine learning.
Feature Transformation	Evaluate the capabilities in terms of feature transformation suitable for machine learning algorithms. This includes possibly numerical transformation, text transformation, etc.
Model Evaluation and Interpretation	Evaluate the ability to evaluate a model and interpret its performance. This includes contributing variables analysis, technical performance analysis, and simulation
Model Optimization	Evaluate the ability of the solution to automatically optimize a model by means of selecting the best feature transformation, selecting the best parameters for machine learning algorithms, and selecting + combining the algorithms.

Natural Language Processing	Evaluate the capabilities of the solution for applying machine learning techniques to text. This includes text classification use cases as well as possibly more advanced NLP use cases (such as sentiment analysis).
Vision	Evaluate the capabilities of the solution for Image/Vision use cases.

Operationalization

Data Workflow Automation	Evaluate the ability to deploy, schedule, and monitor data creation workflows. This includes the ability to integrate with third party systems and check overall data consistency.
Model Lifecycle Management	Evaluate the ability of the solution to manage model evolution in terms of versioning and evolution of model performance.
Model Deployment	Evaluate the ability of the platform to deploy a model in various scenarios (batch or real-time). Evaluate the ability to embed a model in various applications, as well as the ability to securely deploy and monitor it.

Solution Architecture

Deployment Options	Evaluate the vendor's ability to support multiple deployment options, including laptop, on-premise server, and on-premise.
Scalability	Evaluate the scalability of the solution itself and the ability to sustain high growth in number of users or overall usage.
Cluster (Hadoop, Spark, Docker) Integration	Evaluate the ability of the solution to integrate with the Hadoop, Spark, or Docker ecosystem and to leverage existing clusters effectively for data scalability.

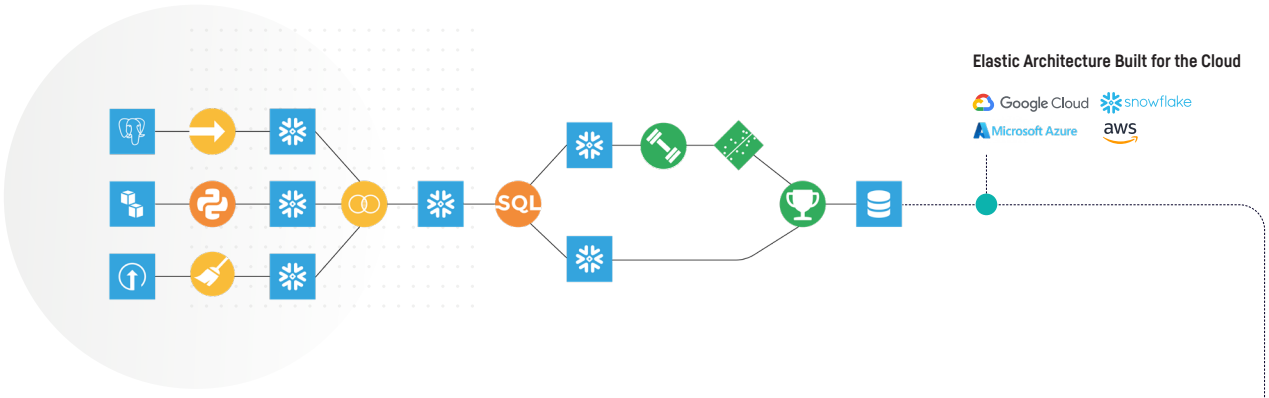
Extensibility / 3rd Party Integration

3rd party integration	Evaluate the capability of the platform to connect to third-party applications. This includes third-party applications such as data sources, data processing, and processing. This also includes evaluating the size of their currently supported third-party ecosystem.
Plugins and Extensibility	Evaluate the ability to easily extend the platform. This includes the ability to add new connectors, to analyze and audit the platform, and to pre-package processing

Security & Governance

Authentication, access control, network, governance	Evaluate the ability of the solution to provide fine-grain security and group-level security. This could include the ability to mirror and respect the security of the various data sources and systems accessed by the platform.
Auditability	Evaluate the ability of the platform to provide trace and audit logs for the actions performed by users.

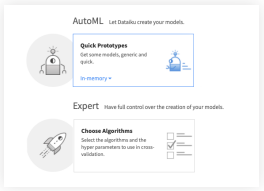
Everyday AI,
Extraordinary People



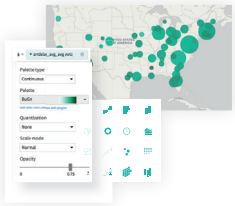
Elastic Architecture Built for the Cloud

Google Cloud snowflake
Microsoft Azure aws

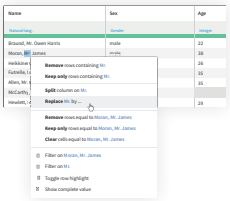
Machine Learning



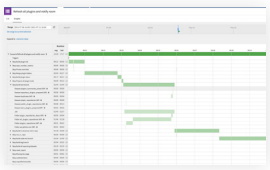
Visualization



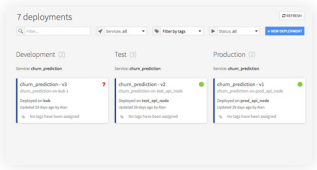
Data Preparation



DataOps



Governance & MLOps



Applications

45,000+
ACTIVE USERS

450+
CUSTOMERS

Dataiku is the platform for Everyday AI, systemizing the use of data for exceptional business results. Organizations that use Dataiku elevate their people (whether technical and working in code or on the business side and low- or no-code) to extraordinary, arming them with the ability to make better day-to-day decisions with data.

GUIDEBOOK

www.dataiku.com