



Product Profile:

# DataClarity

Eckerson Group is a research and consulting firm that specializes in data and analytics. We provide independent and objective research on emerging technologies, techniques, and trends in the field. Periodically, we profile products that we believe provide exceptional value to customers.

This publication may not be reproduced or distributed  
without Eckerson Group's prior permission.

## Executive Summary

DataClarity is a data and analytics integration vendor that packs a veteran punch. Its four-year-old product, which goes by the same name, provides analytics middleware that unifies disparate BI tools and data sources as well as robust data, analytics, and data science functionality. This versatility makes DataClarity suitable for large companies that need to integrate heterogeneous BI and data environments and software companies that leverage its containerized architecture to embed self-service analytics capabilities in their applications.

Prior to launching its product, DataClarity was a consulting firm that specialized in integrating BI and other products from IBM, Oracle, SAP, Tableau, Microsoft, Qlik, and others. Today, many companies use DataClarity to create a common semantic layer and analytics portal across multiple BI tools. Data analysts use its built-in data virtualization engine and data preparation environment to build custom data sets and pipelines from distributed data. And large software vendors use it as an embedded analytics tool with hundreds of native-built charts, dashboard capabilities, and data science functions.

DataClarity offers all this functionality at a fraction of the cost of most other data and analytics products. As a self-funded, employee-owned company, DataClarity has focused almost exclusively on product development, keeping a low profile. But this is changing as more user and vendor companies find a spot for DataClarity in their data and analytics portfolios.

## Company

DataClarity brings deep expertise and modern technology to solve enterprise data and analytics integration problems. Founded in 2002 after spinning out of a systems integration firm, DataClarity was a top IBM partner who helped brand name companies, such as Apple, Boeing, IKEA, and Nike, create dashboard and reporting solutions from a hodge-podge of IBM analytics products, including Cognos, SPSS, TM1, and others.

After a decade of analytics integration work, DataClarity's founder, Mark Mueller, teamed up with a BI and data-savvy software engineer, Dragos Georgescu, to develop a software product based on their cumulative knowledge and experience serving IBM customers. Launched in 2016, DataClarity is web-based BI and data middleware that unifies heterogeneous data and analytics environments across the enterprise. With its containerized, component-based, modular architecture, it also runs equally well in the cloud and on premises.

For a private, employee-owned company with a four-year old product, DataClarity offers a rich set of features at an affordable price. Its secret sauce? The company drives most research and development from offices in Eastern Europe where it recruits highly educated and experienced developers, software engineers, and architects and gives them stock incentives

and a comfortable, collaborative work environment. It also spends the vast majority of its revenue on research and development, not sales and marketing.

## Product

DataClarity offers whatever functionality a data or BI team needs to deliver a complete data analytics environment. Customers with multiple data sources and BI tools often select just the integration functionality, while those without an existing enterprise data or analytics footprint use DataClarity as an end-to-end BI and data science solution.

For organizations that have:

1. **Multiple BI tools** → DataClarity serves as a semantic hub to integrate multiple BI tools, creating a common content and data model that business users can consume and query no matter what BI tool they're using.
2. **Distributed data assets** → DataClarity serves as a data hub that delivers a common business view across those assets, federating queries across databases in real time.
3. **Lots of power users** → DataClarity provides a data preparation tool and a data catalog that enable data analysts and data scientists to find, combine, and transform data and create custom data sets and pipelines from data scattered throughout an organization.
4. **AI/ML requirements** → DataClarity provides data analysts code-less predictive and advanced analytic functions and data scientists the ability to write, run, and operationalize R and Python models.
5. **No embedded BI or analytics tool** → DataClarity offers a modular, microservices-based, containerized architecture that makes it easy for customers to embed its self-service BI and analytics environment within their applications.

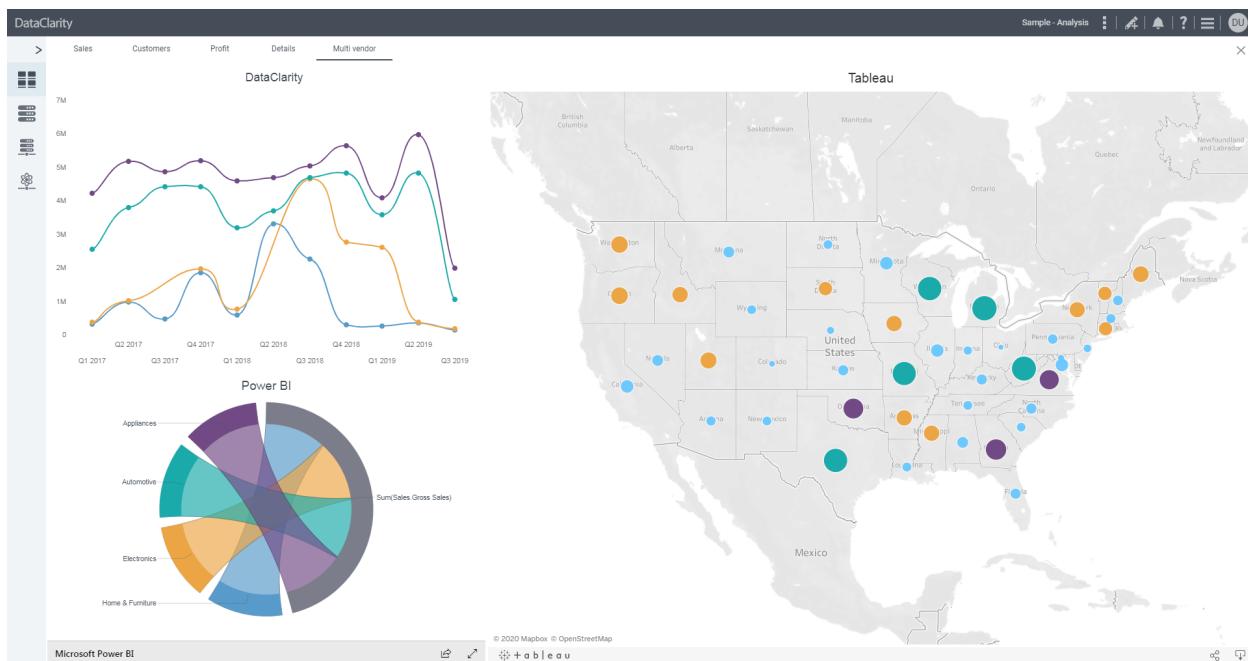
**Data virtualization.** DataClarity likes to call its product a data and analytics integration hub because it uses custom-built data virtualization technology to give users a common view of data stored in any BI tool or database, whether in the cloud or on-premises. Users can query and manipulate distributed data as if it resides locally in a single server. The product translates queries into the native language of the target database (i.e., pushdown queries) and joins data across remote data sets, if needed.

If a database does not support a specific function or operation, DataClarity executes the SQL in its own virtual database engine. This ensures that no matter what the data source, DataClarity always provides the same data manipulation capabilities. The data virtualization engine also caches queries using an in-memory grid that it OEMs from Hazelcast to ensure consistently fast query performance.

**BI integration.** DataClarity uses proprietary and standard (i.e., PostgreSQL application programming interface) database protocols to support bidirectional data connectivity with BI tools, including IBM Cognos, MicroStrategy, Power BI, Tableau, Qlik, and Business Objects. This allows users to query the semantic models of those tools as if they are just another data source. In addition, BI tools can query a DataClarity semantic model which might aggregate data and models from multiple BI tools and sources.

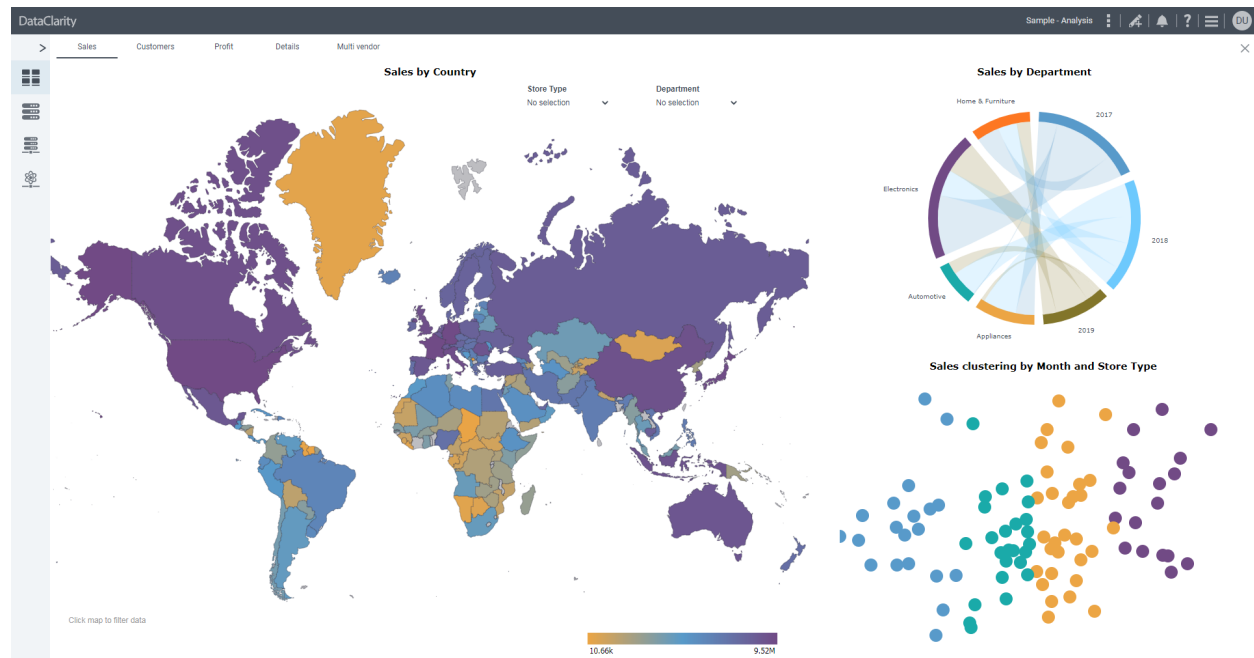
**BI and analytics.** DataClarity is more than just data and analytics middleware. It offers a full-fledged self-service analytics environment that makes it easy for business users to create, use, and modify interactive dashboards. These dashboards (which DataClarity calls Storyboards) can display analytic objects from multiple BI tools side by side. (See figure 1.)

**Figure 1. Side-by-Side Visualizations from Multiple BI Tools**



DataClarity also includes a custom-built geospatial engine for creating custom maps; a custom charting engine with 80+ visualizations; a data science engine that runs R and Python scripts; an alerting engine; and dozens of custom-built machine-learning algorithms. (See figure 2.) Finally, it provides a universal BI portal—or analytics catalog as some call it—that gives business users one place to go to consume all their reports, files, and web content, no matter what BI tool produced them.

Figure 2. DataClarity Geospatial and Data Visualization



**Data integration.** On the data side, DataClarity offers a data catalog that lets users search for data connections and data sets, explore and profile metadata, and view, pivot, and filter sample data. DataClarity supports connectors to 100+ data sources, including files, relational databases, big data sources, cloud applications, and BI semantic models.

DataClarity's built-in data preparation features enable data analysts to create custom data sets. The tool lets them combine and filter data, create custom groups, add calculations and fields, apply code-less machine-learning models, and write and execute their own custom scripts using R and Python. It also offers inline data-cleansing functions, such as the ability to split values into multiple fields, format text, and remove null values and punctuation marks. (See figure 3.) Authorized users can also certify shared data sets.

Figure 3. Inline Data Cleansing Functions

Department	Product Line	Product Type
abc		abc
Electronics		Televisions
Electronics		Players
Electronics	TV & Video	Theater
Electronics	Computers	
Electronics	Computers	N
Electronics	Computers	P
Electronics	Computers	L
Electronics	Computers	Software

## Customers

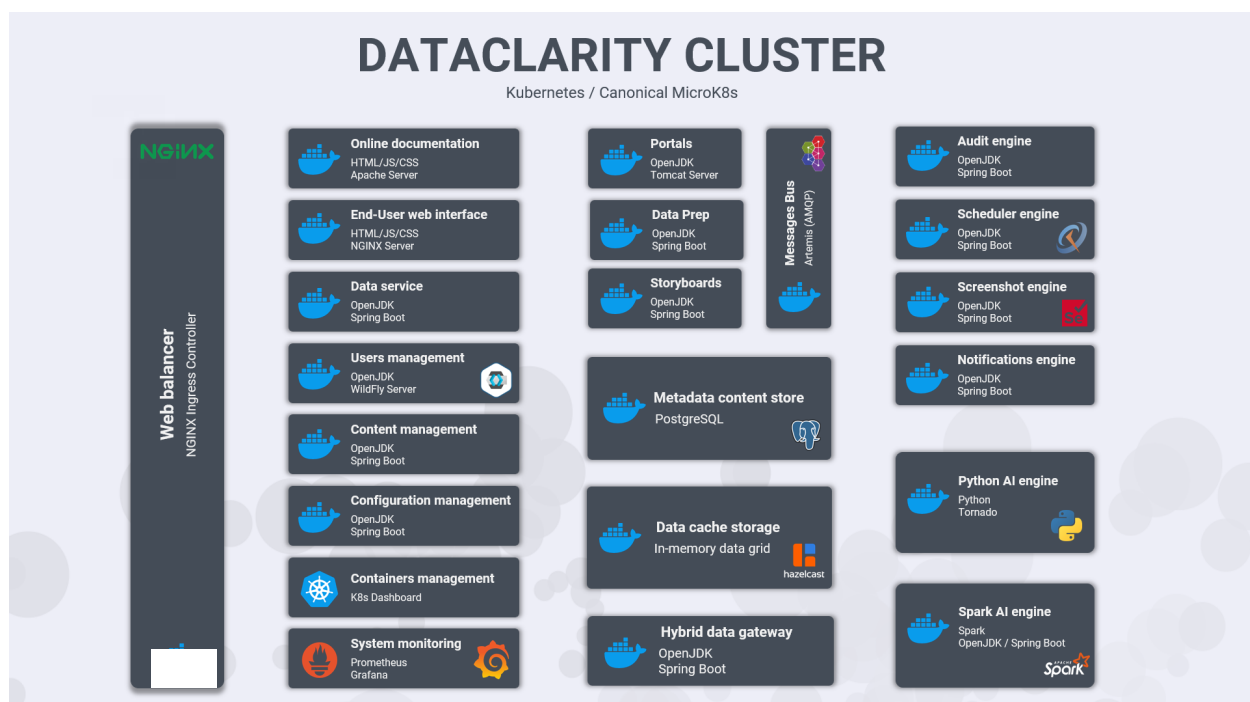
DataClarity customers span a range of sizes, from mid-market to large, and industries, including healthcare, banking, insurance, financial services, retail, distribution, energy, and technology. Customers fall into two camps:

1. Companies that need to integrate data and BI tools (i.e. data virtualization and semantic integration. These are larger organizations with multiple BI tools and distributed data who need data virtualization and/or semantic integration technology. Often, these customers want to connect Cognos or TM1 to Tableau, Power BI, or Qlik, or some other combination of tools.
2. Software vendors who need to embed some or all of DataClarity capabilities into their products. Vendors choose DataClarity because its modern containerized, microservices-based architecture is easy to embed, scale, and adapt to existing software environments compared to monolithic BI environments.

## Architecture

DataClarity is an entirely web-based product (i.e., no desktop) that runs on-premises or in the cloud. DataClarity offers a software-as-a-service version of its product or customers can host their own cloud instances on Amazon Web Services or Microsoft Azure. DataClarity uses a microservices architecture and runs in a Kubernetes or Docker cluster, making it easy to scale, change, or extend individual components. (See figure 4.)

Figure 4. DataClarity Architecture



## Pricing

DataClarity offers refreshingly simple and affordable pricing. Its Consumer license is \$11 user/month and applies to any user who consumes DataClarity content via a web browser, email, or another BI tool. Its Creator license costs \$31 user/month and applies to any user who creates DataClarity content. There are no charges for administrators, data connections, servers, or number of queries. Volume discounts can go as low as \$6 user/month and \$24 user/month respectively.

DataClarity offers two editions of its platform. The DataClarity Analytics and Data Science edition is geared to companies that want all DataClarity functionality. The DataClarity Data Platform is geared to companies that only need data integration capabilities.

## Differentiators

DataClarity has carved out a unique niche by serving as analytics middleware to unify heterogeneous BI and data environments. Its strategy is to land customer accounts with its integration offerings and then expand its footprint with its other features. Once customers recognize they can get comparable BI and data science functionality at a fraction of the cost, they often standardize on DataClarity.

DataClarity has many attractive features, some of which we've already touched on. Here is a summary of its most salient differentiators:

- 1. Semantic integration.** DataClarity is one of a handful of vendors that offers semantic integration with other BI tools. Customers can use DataClarity to create a universal semantic layer that connects to and builds upon the semantic layers of other BI tools. Essentially, DataClarity connects to BI tools as another data source, and BI tools can query DataClarity's semantic layer as if it were their own.
- 2. Data virtualization and cache.** Underlying the semantic integration hub is DataClarity's data virtualization engine that can connect to 100+ relational, big data, and cloud application sources as well as numerous BI tools, as mentioned above. The engine translates queries into database-specific SQL, which it pushes down to remote databases. Its cache runs on Hazelcast's scalable in-memory data-grid to ensure consistently fast performance for previously executed queries.
- 3. Custom engines.** Besides its data virtualization engine, DataClarity has built its own engines for charting, data science (Python, R, and Spark), geospatial analysis, geocoding, alerting, and report scheduling and distribution.
- 4. Universal BI portal.** Thanks to its BI connectors, DataClarity can serve as a company's analytics portal. It can catalog reports from multiple BI tools and personalize the display of those reports based on role, group, and user-selected

favorites. When users open a third-party report, DataClarity displays a live, interactive version of the report within the DataClarity portal page.

5. **Universal BI dashboards.** Similarly, DataClarity can display visualizations from multiple BI tools side by side within a single dashboard. Users can control the visualizations on a page from a single filter if the elements share common data.
6. **Granular security.** DataClarity provides administrators with granular control over every feature and function in the platform. Administrators can set permissions based on groups, roles, or on an individual basis. DataClarity also integrates with a customer's security framework and offers single sign-on to integrate content and models across multiple BI tools.
7. **Release cycles.** DataClarity releases new features and capabilities every six weeks, enabling it to deliver a comprehensive set of functionality worthy of a more mature player.
8. **Pricing.** DataClarity offers refreshingly simple and affordable pricing. Its Consumer license is \$11 user/month and its Creator license is \$31 user/month, with volume discounts as low as \$6 and \$24 user/month respectively.

## Conclusion

For organizations that need a complete data and analytics environment or the means to connect disparate BI tools and databases at an extremely affordable price, DataClarity is the perfect choice. DataClarity is a newer entry to the data and analytics space that offers a surprisingly mature set of capabilities built on a modern, modular architecture. Its development team designed the product using decades of in-the-trenches experience integrating diverse sets of data and analytics tools for Global 500 companies. As a 100% employee-owned company with no venture backers and an aggressive price point, DataClarity is playing the long game, making it a safe bet to bring in-house.



## About the Author

---



**Wayne W. Eckerson** has been a thought leader in the data and analytics field since the early 1990s. He is a sought-after consultant, noted speaker, and expert educator who thinks critically, writes clearly, and presents persuasively about complex topics. Eckerson has conducted many groundbreaking research studies, chaired numerous conferences, written two widely read books on performance dashboards and analytics, and consulted on BI, analytics, and data management topics for numerous organizations. Eckerson is the founder and principal consultant of Eckerson Group.

## About Eckerson Group

---

[Eckerson Group](#) provides research, consulting, and education services to help organizations get more value from their data. Our experts each have 25 years of experience in the field, specializing in business intelligence, data architecture, data governance, analytics, and data management. We provide organizations with expert guidance during every step of their data and analytics journey. Get more value from your data. Put an expert on your side. [Learn what Eckerson Group can do for you!](#)

# Eckerson Group

**RESEARCH****CONSULTING****ADVISORY**

GET • MORE • VALUE • FROM • YOUR • DATA