**Data Warehousing and Business Intelligence**

**INFO 7290**

**Fall 20'**

**HW02 - Gartner Magic Quadrant Data Integration Tools review:**

**By**

**Shrishti Khanna**

**Urjasvit Sinha**

**Smit Patil**

**Satya Sai Manikanta Borusu**

**Introduction:**

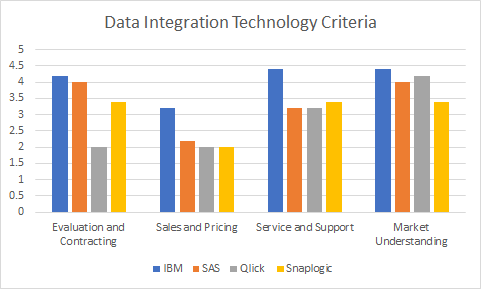
Data is the most vital resource for any organization. Few will argue against the idea that businesses operate better and accomplish their goals when they can use their data strategically. Businesses today generate huge amounts of data in their daily operations. Some of it is from the sales, marketing, and customer service arms of the business. Other parts may arise from the financial transactions, or perhaps its research, development, and production activities. Here is where Data Integration comes into picture. Data integration involves combining data from several disparate sources, which are stored using various technologies and provide a unified view of the data and this data can then be used by business applications to provide actionable insights. So, in today's competitive world choosing the right data integration toolkit provider who can meet the specific business needs is very important.

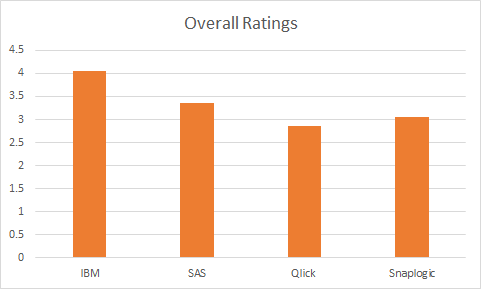
**Objectives:**

The aim of this review is to research and develop documentation of each of the Data Integration tools based on the customer reviews and industry standards and then find the most impressive vendor currently leading the market.

**Result:**

The pros and cons of each tool has been listed in detail in the Appendix, according to the ratings on Gartner.com the following charts have been deduced on a criterion by five reviewers. The final score for each tool was calculated using the average criteria of each tool. To visualize the ranking, the final scores of all the tools were plotted in a bar chart as shown below.





**Conclusion:**

IBM and SAS were the top two significant competitors in our selection. Even though SAS has a more extensive customer base than IBM, we ended up choosing IBM as our top pick due to its wide range of data integration tools and higher overall customer ratings. IBM's diverse data integration delivery styles and massive brand awareness set it up a notch better than our other picks. If SAS could improve some of its shortcomings in the pricing model, ease of use, and expand their vision from its current analytics only focus, SAS can become a more competitive vendor. We ranked Qlik and SnapLogic in the 3rd and 4th positions, respectively. Qlik has a more comprehensive data integration tool as compared to SnapLogic. It has also received higher customer ratings and is used in a greater number of organizations than SnapLogic. Even though Qlik has a higher overall ranking than SnapLogic, it lacks its pricing model and ease of use than SnapLogic. If SnapLogic improves its user interface and makes the code documentation more detailed, it can be placed as one of the magic quadrant's challengers.

Overall while looking for a data integration tool, the organization should focus on what they are trying to achieve from this project and choose a tool that fits their requirement. IBM is an excellent tool for larger corporations with diverse portfolios, whereas organizations can use Qlik with limited needs, especially in the data replication areas. Thus, choosing a tool that suits your needs should be the strategy of all the users.

**Appendix:**

1. **SAS**

Introduction:

Based in North Carolina since 1976, founded by James Goodnight as a project at North Carolina State University to create a statistical analysis system, primarily used by agricultural departments at the university. Vendor's customer base for this project is 14,000 organizations.

Pros:

* Open source metadata management projects.
* Allows exchange of metadata bidirectionally between SAS and Non-SAS data management technologies.
* Equipped with dealing with ML based self-healing capabilities which assists in distributing, optimizing, and recomposing integration processes.
* Timely and effective technical product support.
* SAS data preparation supports streaming data integration, MPP support , automated contact data tagging based on NLP analysis and cloud data exchange.
* Git integration available

Cons:

* High custom complex pricing models
* Limited contract flexibility during renewal process.
* Vendor lock-in and lack of contract transparency makes it difficult to negotiate.
* Data integration may not provide specific advantages above the other tools beyond use cases.
* Upgrade difficulties and deployment issues.

Conclusion:

SAS is a platform that offers several services such as Data management, data integration, federation server, data loader for Hadoop, data preparation, event streaming process etc. With changing times it has managed and tried to manage and inculcate all advancing technologies, for example incorporating ML for self-healing capabilities. The only drawback are more specifically sales and customer service related

1. **IBM**

Introduction:

IBM’s product set has a Vendor’s customer base of more than 11,000 organizations. It

operates globally and clients tend to be B2B and B2C organizations. Its data integration tool

suite comprises of IBM InfoSphere Information Server, IBM InfoSphere Classic Federation

Server, IBM InfoSphere Data Replication, IBM App Connect, IBM Streams and IBM Data

Refinery.

Pros:

* High Customer satisfaction rate among their Customer base has stemmed from a combination of product performance and Support for its integration tools in the form of accessibility of experts and responsiveness of the sales team
* IBM provides an extensive range of options with plenty of prebuilt functions and connectors which brings a holistic nature to its Data integration Suite
* They are praised for their multiple data integration delivery styles like traditional and modern and the tool portfolio can deliver on hybrid (on premises and cloud) integration scenarios.

Cons:

* The ease of upgradation and deployment is an issue and there are requests to focus more on container-based upgrade options
* It received lowest reference survey scores for self-service and data integration cloud services capabilities
* Their Complex Pricing structure is one of the main reasons for many surveyed customers to not choose IBM data integration tools

Conclusion:

IBM is a leading player in the current market with a lot of upsides to it and negligible

downsides. if it could resolve a few of its issues like pricing and upgradation complexity it

could become a force to be reckoned with.

1. **Qlik (Attunity)**

Introduction:

Qlik acquired Attunity in May 2019, and now is a division of Qlik. Attunity provides various tools in the data integration field, mainly Attunity Replicate® (now Qlik Replicate™), its data replication tool. Qlik has a customer base of more than 2500 organizations globally.

Pros:

* The biggest strength of Attunity appears in the data replication scenarios. Customers choose their data replication tool in most competitive circumstances and have frequently appeared in the Gartner surveys.
* It's CDC(Change Data Capture) based capabilities have also received very positive feedback. Features like automated code generation and nonintrusive implementation make it very easy to use.
* Attunity has also made a strong push in building its partner network for its cloud data replication and migration platform. Partnership with notable cloud infrastructure providers such as AWS(Amazon Web Services) and Microsoft Azure has enabled Attunity to provide robust OEM support.
* It has partnered with vendors like IBM, Oracle, SAP, etc. for its data lakes, big data integration, and data warehouse automation services.
* It has also partnered with global Sls and resellers such as Accenture, Infosys, Hewlett Packard Enterprise, etc.

Cons:

* One of the significant limitations of Attunity is its limited incorporation with other data integration styles such as data warehouse automation and data lake population.
* It also lacks bulk/batch (ETL) and data virtualization use cases and can be a point of concern for interoperable data delivery styles.
* Customers have also complained about Attunity's licensing model, where most of the consumers felt it is expensive.
* Qlik's acquisitions of Attunity has brought in uncertainty regarding the independent development and continued excellence of Attunity.

Conclusion:

Attunity is a formidable player in the data integration environment with a very competitive product in the data replication field. With its ease of use and a solid OEM support, it has led it to be the top pick in most competitive situations. Attunity has also tried to resolve issues regarding licensing costs by introducing a subscription-based licensing model. If Qlik could grow Attnuity's portfolio beyond its core use cases, it can be positioned as a leader in data integration tools.

1. **SnapLogic**

Introduction:

SnapLogic is an iPaas tool (Integration Platform as a Service) that delivers multi-tenant cloud services using HTML-5 based platform. It helps in improving business activities, drives better results and makes fast paced decisions by connecting applications and data across enterprise. It’s Intelligent Integration Platform (IIP) uses artificial intelligence powered workflows to automate stages of integration projects and brings speed, quality and accurate predictability to data driven decision making. The platform is easy to use and the self service enables the integrators to manage all the integration on a single scalable platform.

Features of SnapLogic:

* Prebuild connectors
* Connector modification
* Realtime and batch integration
* Data Quality services
* Monitoring console

Pricing Overview:

SnapLogic pricing starts at 24,000$/year. There is also a trail version offered by SnapLogic.

Some of the other vendors considered by reviewers before purchasing SnapLogic (According to

Gartner.com):

1. 86% considered Informatica
2. 57% considered Boomi
3. 43% considered MuleSoft

Pros:

* Scheduling tasks and data automation are the best features of SnapLogic.
* Data driven predictability using IIP.
* Great to build complex logic into ETL pipelines.
* Easy to use, doesn’t require much time to learn Snap pipelining.
* Easy customizations available on SnapLogic.
* Cheaper than other established Data Integration products like Informatica, Talend.

Cons:

* Parallel processing of data thread is hard to implement.
* Not well integrated with AWS.
* Difficult to stitch all the snaps together once the pipelines become complex.

Conclusion:

SnapLogic as a PaaS platform is robust, flexible and enterprise level tool that caters to all levels of organization namely from Developers and Coders to Data Architects and Data Analyst. The Platform provides intelligent integrations to nearly everything.