Modern BI Evaluation
Criteria Checklist

Modern BI platforms are designed to help you answer a deeper range of questions about your business using more of data and a greater variety of data. The objective is to move from an IT-lead approach to an IT-enabled one where citizen data scientists can easily apply their knowledge of data, analytics and the business to answer critical new business questions.

To achieve this, the modern BI platform you choose should contain key functionality that will allow you to:

• Use more of our valuable data, regardless of size, location and format
• Allow BI and analyst teams to be more productive through an analyst-friendly interface and agile analysis and data discovery
• The ability to quickly tell stories about the data to ensure the results are matching the business needs
• Put your insights to work across the organization through easy operationalization of data, strong governance and industrial strength enterprise features

Here are some specific functional items for you to look for when evaluating modern BI platforms.
### Works with All and Any of Your Data
- Works with structured, semi-structured and unstructured data
- Offers a complete set of data loading features designed for end-users
- Has 70+ pre-built connectors for structured and unstructured data that span the many data sources a business will use
- Has built-in data links to existing Hadoop data
- A fully self-service data ingestion and integration process – no IT needed for ETL
- Contains no limits to the size data ingested
- Has exception reporting on data ingestion processes

### Agile Analysis and Data Discovery
- Covers the full end-to-end analytic workflow: data connectivity, integration, preparation, analysis, visualization and operationalization
- Fluid workflow with the ability to operate on different analytic steps in parallel and immediately see downstream effects
- Dynamic modeling using schema-on-read
- Offers an easy way to profile datasets, then clean and prepare the data for analysis
- Easy to use advanced analytics for time-series, graph, path, text and sentiment analysis
- Smart data discovery via built-in advanced algorithms for clustering, decision trees, data dependencies and recommendations
- Supports the collaboration and sharing of analytic datasets with analyst colleagues
- Works with existing enterprise data warehouses and BI tools to extend existing analytics by adding new data or new analytic operations to the data
- Can extend an enterprise data warehouse by pushing results to the EDW for centralized management of analytic result-sets

### Analyst-friendly
- Fully self-service interface that enables your citizen data scientists to perform all functions, eliminating dependencies on IT
- Offers a familiar Excel-like spreadsheet-style interface
- Allows the analyst to apply functions via easy drag-and-drop and point-and-click operations
- Supports a large number of pre-built analytic functions (270+)
**Storytelling**
- Easy to use, built-in linked visualizations
- Supports over 30 drag-and-drop visualization widgets
- Integrated support for geo-spatial data display
- Automatically changes chart attributes depending on the data
- Free distribution of storyboards to business end-users without requiring the user to have a license
- Contains the ability to add external elements to the storyboard (images, video, etc.)
- Offers drill-down from the infographic into the analytic application data
- Allows the analyst to create slideshow-style storyboards
- Full HTML5 support for widgets and infographics for responsive design and seamless cross platform support – desktop, tablet and mobile
- Easy export and integration with leading visualization tools such as Tableau and PowerBI

**Operationalization of Data**
- Can export data to external systems or formats external systems can easily import for fast easy operationalization of the insights
- Offers an easy to use REST API for rapid integration with downstream applications and processes to deliver analytic results
- Centralized management and control through a web-based interface
- Supports both ad-hoc and scheduled jobs for the import of data, the execution of analysis, and export to downstream applications and systems

**Governance**
- Fully supports the five pillars to data governance: quality and consistency, data policies and standards, security and privacy, regulatory compliance, and retention and archiving.
- Offers complete data profiling, data statistics and metadata management tools for data quality and consistency
- Supports data policies that expose a subset of data to specific users and can apply masking, anonymization, or aggregation to sensitive data fields
- Allows policies and role-based security at every stage of the pipeline, from ingest to export
- Offers role-based controls on downloading or exporting data and accessing administrative functions
- Supports cross-artifact lineage features that can be combined with all relevant user and system events to produce a full audit trail for regulatory analytics
☐ Allows retention policies such as keeping data permanently, purging records older than a specific time window, or ones based on the number of ingests or analytic executions

☐ Supports security rules for retired data to be either instantly removed, retained until a specified time, or manually removed after system administrator approval

**Enterprise-ready**

☐ Built to easily scale via native integration with Hadoop storage and compute (doesn’t just connect to Hadoop)

☐ The platform provides linear scalability on large Hadoop clusters of up to 6000 nodes

☐ Uses Hadoop parallelism and scaling in all steps of the analytic cycle: data ingestion, data integration, data preparation, analysis and visualization

☐ Has an intelligent execution framework that:
  ☐ Hides the complexity of the underlying technology from the user
  ☐ Automatically selects the best execution engines based on the characteristics of the analytic workload
  ☐ Can breakdown jobs into smaller task and execute the individual tasks on the best engine for that task

☐ Has a future-ready architecture that can support new execution engines and technology as they mature

☐ Can integrate with enterprise identity management systems like Active Directory and LDAP

☐ Offers an extensible system with open APIs supporting a variety of languages

☐ Contains a plug-in SDK to add custom analytic functions.

☐ Offers flexible deployment options

  ☐ On-premise, supporting a variety of different Hadoop distributions (Cloudera, Hortonworks, IBM BigInsights, MapR, Microsoft HD Insights, Pivotal)

  ☐ In the Cloud on leading cloud infrastructure such as Microsoft Azure

  ☐ As a fully managed Software-as-a-Service (SaaS) solution requiring no infrastructure, IT or software management