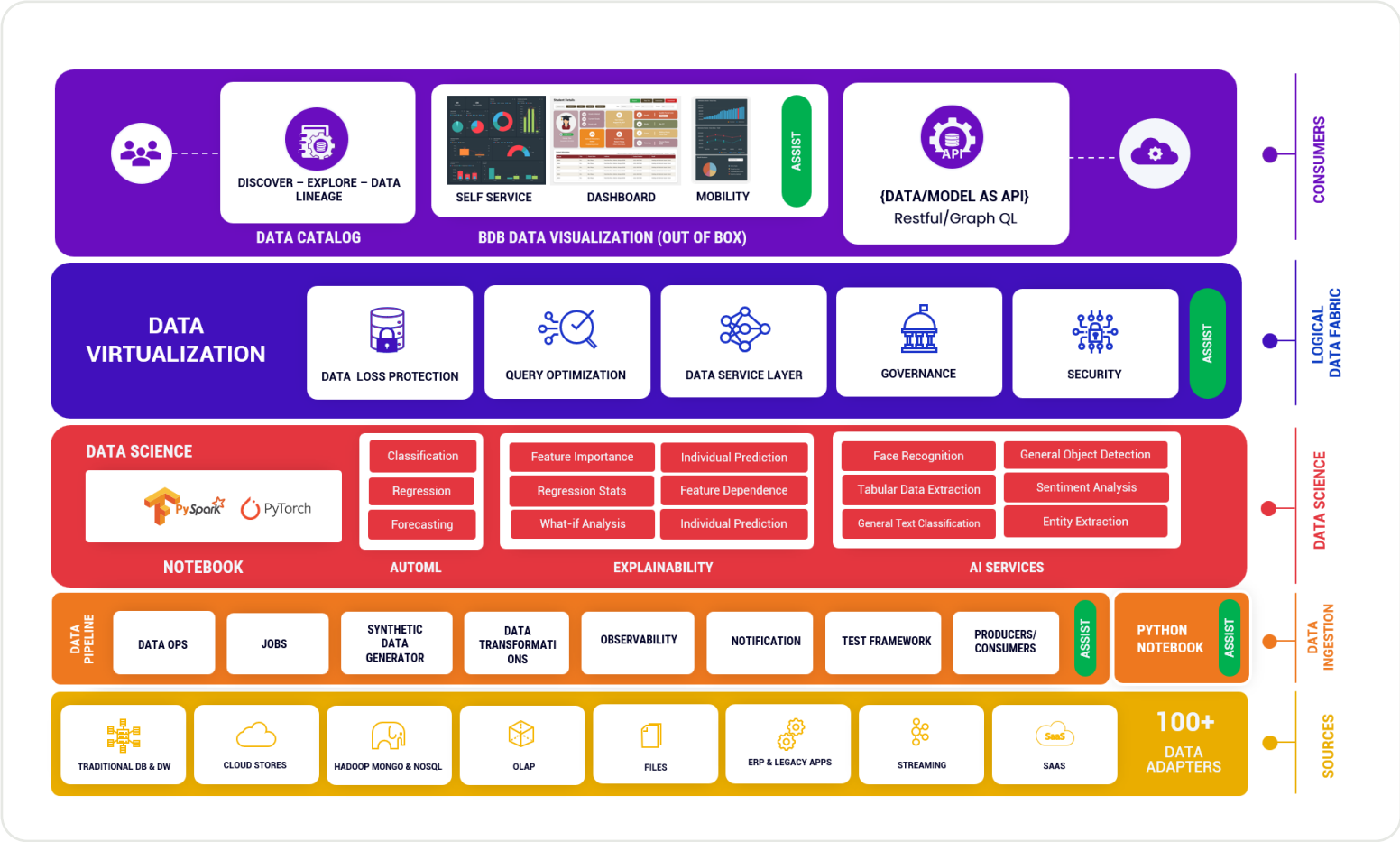




# DEEP DIVE + • BDB PLATFORM



# BDB IS MODERN, UNIFIED, INTEGRATED DATA ANALYTICS PLATFORM



**Cloud Agnostic**

Can get deployed in any Cloud (AWS, Azure, GCP), Hybrid or On-Premise environment(s)

**Data Storage / Lake Agnostic**

Select the data storage or data lake of your choice

**Data Ingestion**

Ability to ingest data from multiple sources, with real-time, batch (jobs) and micro-batch support

**Unified data dissemination**

Streamlined and cohesive distribution of data to ensure consistent and accessible information for users

**Data Democratization**

Empower individuals, regardless of their technical expertise, to access and use data for decision-making and insights through powerful Visualisations (Self Service/Governed Dashboards)

**Generative AI as Assist (LLMs/SLMs Integration)**

SLMs integrated as Assistant for Developers in Data Engr, Data Science and Data Viz modules. Systematic Search on Customer Data for Text based queries is possible.

*Fits in Budget | Scale with Usage | End-to-End | Flexible with White labelling | Easy to Integrate | Customization | Workflows Development | AI Services | Ease of Learning (Python/SQL)*



# BDB PLATFORM BENEFITS



## Plug-n-play with GCP, Azure and AWS

Seamless integration with existing Azure and AWS based platforms. Easy connect to major Azure and AWS components



## Non-disruptive & Non-invasive

Preserves enterprises investments in existing tools & technologies. No expectation of expensive phaseouts and migrations. Easy integration with hybrid environments



## No Lock-in

Built purely using open-source components. Platform is easy to resource for and support long term



## Future Proof

Our solution continuously adds new features and functionalities; thereby reducing technical debt



## End to End Low Code

Platform Products delivers sustained value for all the stakeholders and flexible to fuse connected and non-connected data together



## Secure and Compliant

Platform compliant with privacy guidelines, WPP and UN 155 standards, ISO 20022 (API), etc. to meet European and Global requirements



## Integrated within Customer

Platform will be white labeled and Customer stakeholders trained for organizational led innovation and product development



## Low-Risk

Easy to learn, easy to support. Mostly drag and drop interfaces. BDB committed to training and supporting enterprise teams, until steady state



## ROI Centric deployment approach

Easy to compare cost / effort / benefit of each use case being implemented. Scale incrementally once convinced



## Truly Affordable

Aggressively priced. Becomes dramatically attractive for mid and large-scale compute



## Leap Frog Innovation

Existing product portfolio can be used to deliver market leading connected products to markets quickly across different vertical segments



## Fast track Data Monetization

Out of the box product help to meet the fleet, government and third-party requirements in a short period



## Resource Optimization

Delivers increased value from existing resources and reorients resources in delivering business and customer value



## Scalable and Reliable

Platform is designed for high availability, regularly tested and deployed on Kubernetes for auto scaling

# WHY BDB FOR DATA ANALYTICS?

## GARTNER & OTHER CITATIONS ...

- Gartner Worldwide **Rank #30 amongst 1000+ Global Data Analytics** companies
- **FIRST company** to achieve **this rank in less than 4 years**
- BDB Product Secured **1<sup>st</sup> & 3<sup>rd</sup> rank in Data Analytics competition** organized by Analytics India Magazine
- View **our customer feedback** on **Gartner Peer Insights Portal** – [Click Here](#)

## SOLUTIONS & DEPLOYMENTS IN FOLLOWING VERTICALS

- Telecom
- Automobiles
- Healthcare, Lifesciences
- Education & EdTech
- Agro Industry
- Media & Entertainment
- Retail & e-Commerce
- Oil & Gas
- Hi Tech / CDPs

**We have global fortune 500 customers without Gartner references**

## UNIQUE DATA ANALYTICS PLATFORM (SELF RATING BY BDB EXPERTS)

**No Expert Agency Evaluates D&A Platform on all Following Points required in Modern Era**

| Parameters                  | Maturity Score    |
|-----------------------------|-------------------|
| Data Pipeline               | 7.5               |
| Data Preparation            | 7.5               |
| Data Science Lab            | 7.5               |
| Auto ML                     | 7                 |
| Data Virtualization         | 7.25              |
| Data Observability          | 7                 |
| Installation & Deployment   | 7.5               |
| Data Security               | 7.5               |
| Self-Service Reporting      | 7.5               |
| Dashboard                   | 7.5               |
| Migration & Version Control | 7.5               |
| Customization Capabilities  | 7.25              |
| Scalability                 | 7.5               |
| Performance                 | 7.5               |
| <b>Total Maturity Score</b> | <b>104.25/140</b> |



# BDB DEPLOYMENT STACK (ANY CLOUD, ON-PREMISE OR HYBRID DEPLOYMENT)



## Kubernetes Cluster

### BDB Platform

#### Platform Layer

Platform UI

Platform Gateway

Platform Core

Data as API

Dashboards

Reports

SSO

#### Orchestration Layer

Data Pipeline

Data Preparation

Internal Monitoring and Logging

Sandbox

Data Science Lab (Notebooks, Auto ML, Explainability etc.)

Generative AI (SLMs) (Optional – GPUs)

#### Tenant Execution Layer

Pipeline Components

DS Lab Components

Monitoring

API Ingestion

Web Socket

Data Services

Sandbox Storage

#### Namespace : Monitoring



Grafana



Fluentd



Prometheus, Zabbix etc.



Persistent Volumes



Horizontal Pod Auto scaler



Users



Http



Https



Ingress Controller



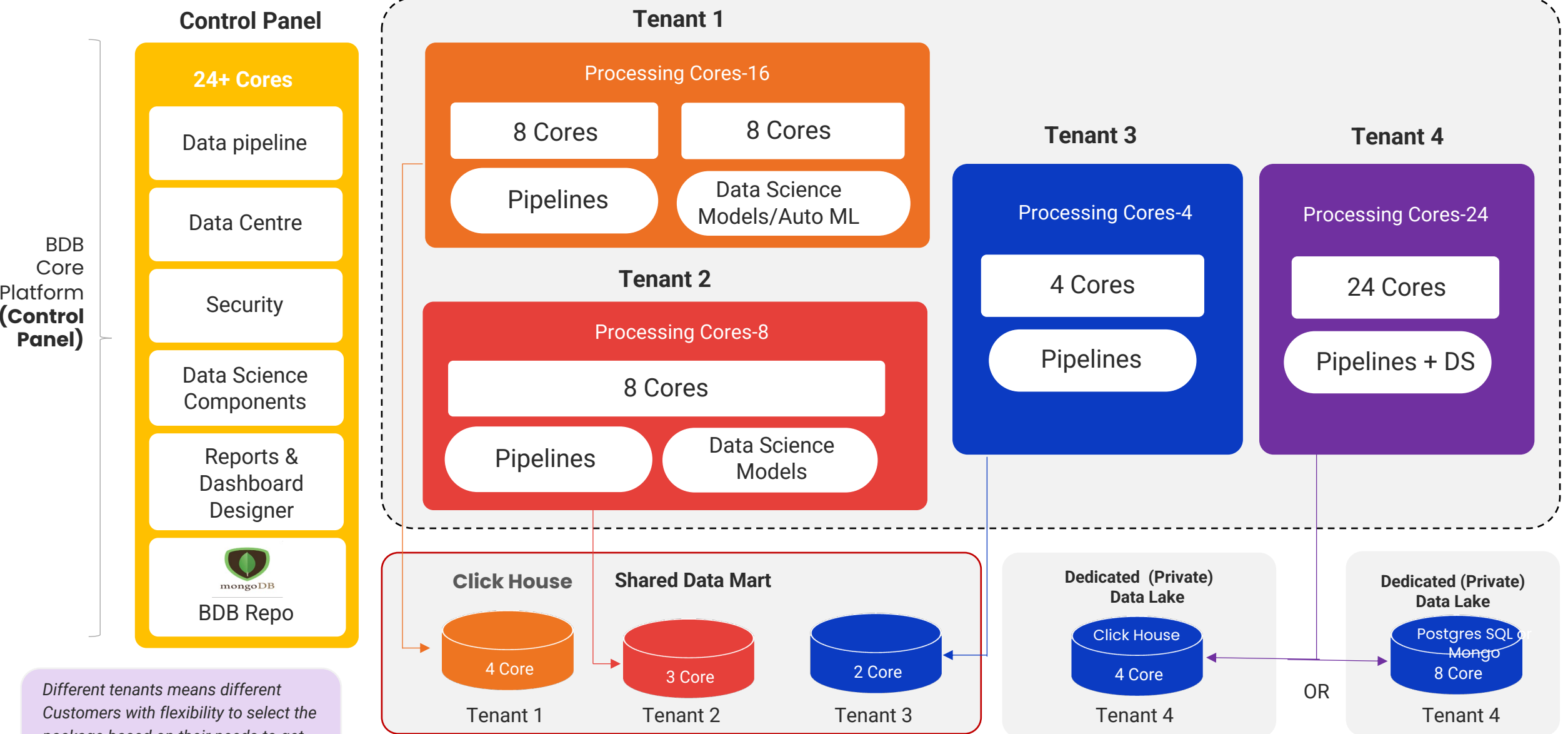
Firewall

Min 24 Cores for Core Platform (Control Panel)



# BDB SAAS DEPLOYMENT (MULTITENANT – MULTISITE)

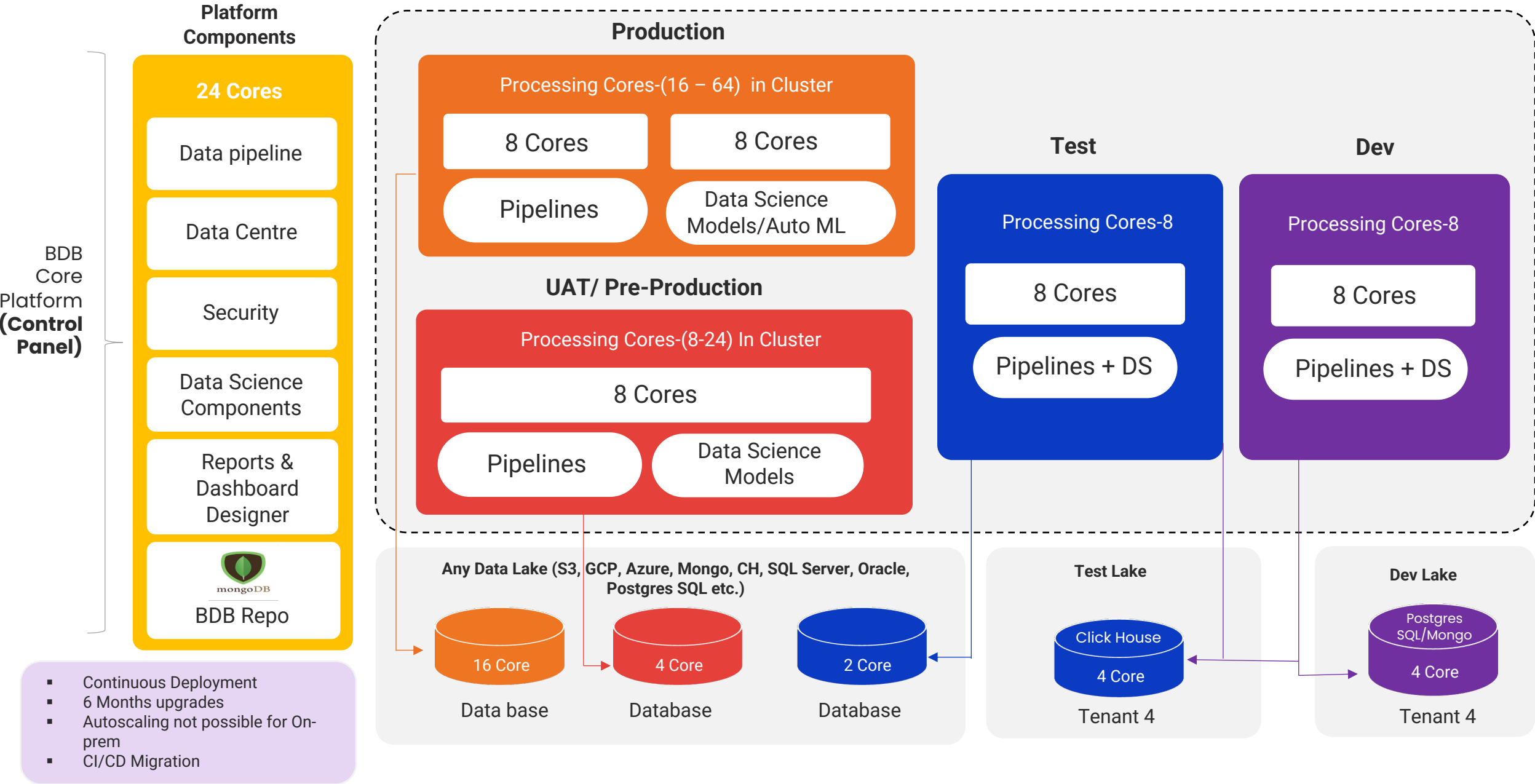
DIAGRAM SHOWING DEPLOYMENT FOR ONE SITE



Different tenants means different Customers with flexibility to select the package based on their needs to get started on BDB Cloud

Customer can select Data Mart based on their required cores- shared or private with flexibility to increase the processing core based on the data volume.

# BDB PRIVATE CLOUD OR ON-PREM DEPLOYMENT FOR ENTERPRISES



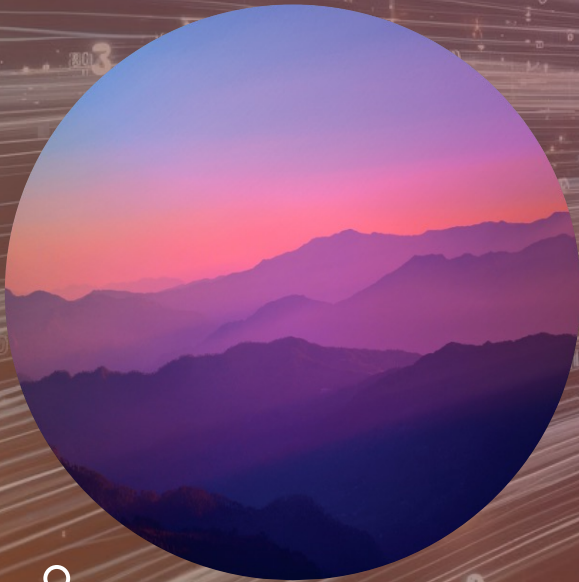
- Continuous Deployment
- 6 Months upgrades
- Autoscaling not possible for On-prem
- CI/CD Migration





## DATA SOURCES

“Data connectors in the BDB platform enable seamless and efficient communication between the Core Platform and various data sources. These connectors serve as bridges, facilitating the integration of diverse data types, including databases (SQL, NoSQL), APIs, and other sources.”



# DATA SOURCES



BDB Platform facilitates seamless data integration through native connectors.



## Connectivity at a Glance:

- Native connectors for diverse data sources: SQL, NoSQL, MySQL, APIs, and more.

## User-Friendly Interface:

- Intuitive connections with minimal user input based on data category.

## Effortless Customization:

- Create, edit, and delete connectors for adaptability to changing needs.

## Collaborative Efficiency:

- Share connectors for seamless teamwork on data projects.

## Continuity and Workspace Management:

- Reconnect to established connectors for workflow continuity.
- Efficiently delete unnecessary connectors for an organized workspace.

## Built-In Versatility:

- Access over 50 built-in connectors for databases, files, APIs, and more.

## User Control:

- Establish, modify, or remove connectors as needed.





## DATA PIPELINE

+ • “Data pipeline refers to a structured sequence of processes that transport, transform, and manage data from diverse sources to a designated destination”.

Modern era ETL, ELT to handle real time, batch and micro batch data





# BDB PIPELINE CHARACTERISTICS



DATA PIPELINE



## Speed

Experience swift data processing with our high-speed data pipeline, reducing latency and enhancing overall performance.



## Scalability

Grow your data operations effortlessly with our scalable solution that adapts to the evolving needs of your organization.



## Reliability

Trust in the robustness of our data pipeline tool, ensuring the reliability of your data workflows.



## Automation

Increase productivity through automated workflows, reducing manual intervention and streamlining data engineering tasks.



## Visibility:

Gain insights into your data pipelines with enhanced visibility, allowing for better monitoring and control.



## Security

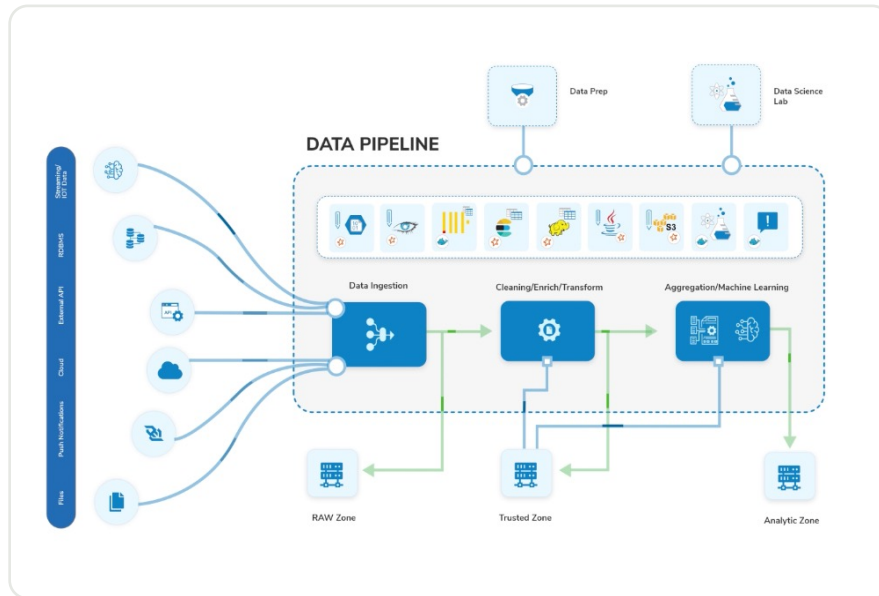
Safeguard your data with our advanced security measures, ensuring that your information is protected throughout the pipeline.



## Flexibility

Adapt to changing data requirements seamlessly, thanks to the flexible architecture of our platform.

# EMPOWERING DATA ENGINEERS FOR UNPRECEDENTED EFFICIENCY



## Event-Driven Orchestration

- Dynamic processes triggered by events for responsive data workflows.

## Intuitive Drag-and-Drop Interface

- User-friendly design simplifies pipeline creation and modification.

## Self-Service Low Code

- Empowers users with innovation through self-service and low-code development.

## Real-Time & Batch Orchestration

- Flexible processing for both real-time and batch data, adapting to diverse needs.

## MLOps and DataOps Integration

- Streamlined collaboration between data science and engineering for enhanced efficiency.

## Cloud Agnostic & Hybrid Deployment

- Flexibility with cloud-agnostic compatibility and support for hybrid deployments.

## Monitoring for Pipelines and Processes

- Comprehensive insights for proactive management and optimization.

## Reliability, Scalability, Maintainability

- Robust, scalable, and easily maintainable foundation for consistent data operations.

## Fault Tolerant with Auto Recovery

- Uninterrupted data processing with automatic recovery, minimizing downtime.

## Custom Integration and Extensibility

- Tailor the pipeline with custom integration and extensibility for unique requirements.

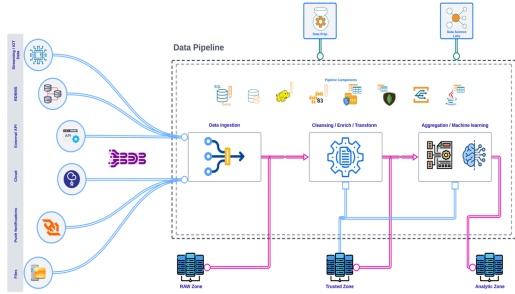
## Parallel and Distributed Processing

- Optimized performance through parallel and distributed processing.

# BDB DATA PIPELINE FEATURES



- Low Code Visual Authoring
- Real-time and Batch Orchestration
- Event based Process Orchestration
- ML and Data Ops



- Extensibility via Custom Scripting
- Version Control Management
- Pipeline Monitoring
- Pipeline TestFrame Work
- Pipeline Failure Analyse

## Simple Drag and Drop Interface

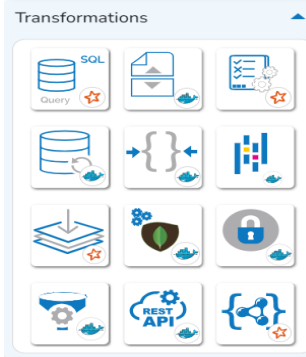
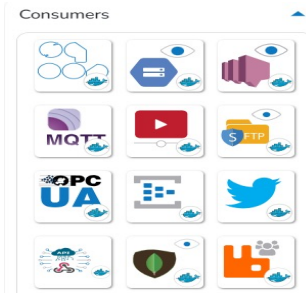
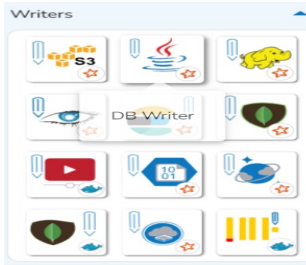
### Advanced Components

- DS Lab Runner
- Auto ML
- Synthetic Data Generator

## Spark Job

## Supported Ingestion Components

## 100+ Components



## Import Capabilities

- Data Prep Script
- Auto ML Models
- DS Lab Models

## Download Sample data

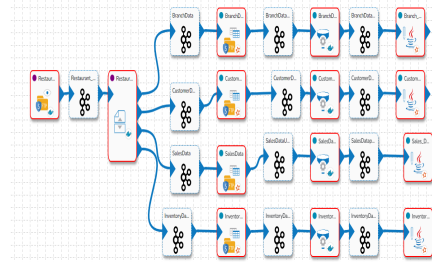
- CSV

## Live Data Ingestion Components

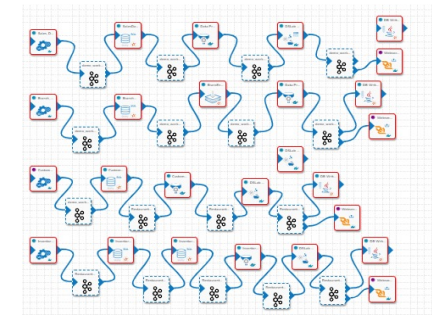
- API
- Data Sandbox
- Twitter
- SFTP Monitor
- Rabbit MQ
- OPC UA
- MQTT
- Event Hub Subscriber
- Mongo Change Stream

## Comprehensive Multi-Component Data Ingestion Pipeline

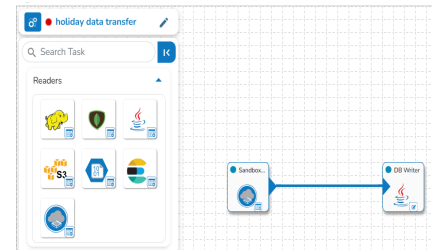
### Pipeline



### ML and Data Ops Pipeline



### Spark Job



## Effective Data Pipelines for Data Engineers

- Intuitive and attractive User Interface
- Complex Workflows can be created
- Gather data from diverse sources like databases, APIs, files, and streams.
- Preprocess and clean the data, ensuring it's in a suitable format for analysis and usage.
- Develop and apply machine learning models to gain predictive insights and optimize processes
- Monitor the pipeline's performance, detect anomalies, and trigger alerts for timely action.
- Implement mechanisms to handle errors gracefully and recover from failures without compromising the pipeline's functionality.
- Manage versions of data transformations, models, and other pipeline components for traceability and reproducibility.
- Scalability & High Performance
- User Level Access and Control
- Strong Migration Tool to handle Pipeline Content
- Affordable than any competitor
- Connects with any data – Cloud or on premise
- Constant Development – Managed by a disciplined, strong R&D Team
- Reliable & cost-effective Customer Support
- Real Time Data Analysis is possible



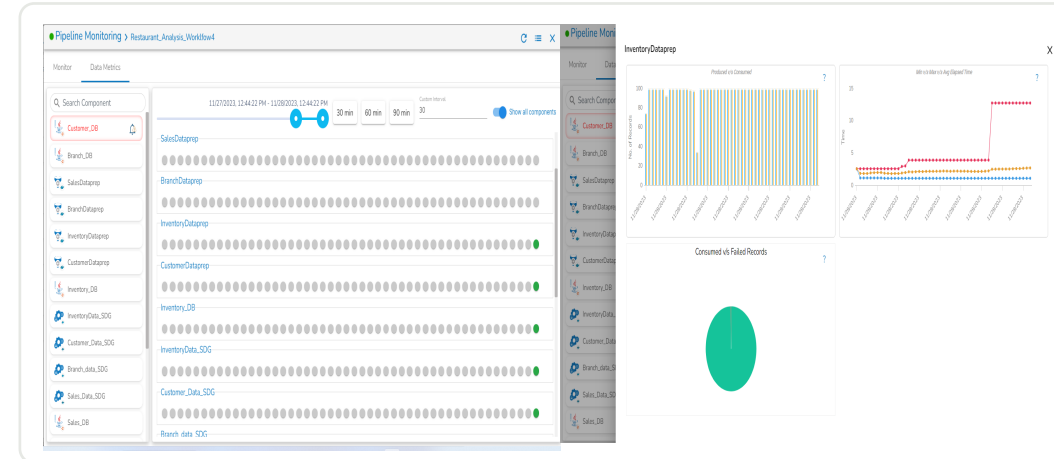
# BDB DATA PIPELINE MONITORING

- Users can track progress and performance metrics via the monitoring dashboard.
- Displays details including component status, types, activation and deactivation timestamps, total allocated and consumed CPU percentages, total allocated and consumed memory, number of records processed, and component logs.
- All information conveniently showcased on a single page for streamlined oversight and quick access to vital metrics.
- Provides users with a detailed overview of compute resource utilization and operational performance, enabling informed decision-making and proactive management of the pipeline process.
- All pipeline-generated metrics seamlessly integrate with Enterprise monitoring software, ensuring comprehensive oversight and alignment with broader organizational monitoring strategies.

Pipeline Monitoring > Restaurant\_Analysis\_Workflow4

Monitor Data Metrics

| Pipeline          | Restaurant_Analysis_Workflow4 | Last Activated   | Nov 28, 2023, 12:37:39 PM | Total Allocated CPU       | 13,200 Core         | Total Consumed CPU   | 16,885 Core             |                 |                    |
|-------------------|-------------------------------|------------------|---------------------------|---------------------------|---------------------|----------------------|-------------------------|-----------------|--------------------|
| Status            | true                          | Last Deactivated | Oct 30, 2023, 6:31:50 PM  | Total Allocated Mem       | 18634 MB            | Total Consumed Mem   | 13106.953 MB            |                 |                    |
| Component Name    | Running                       | Type             | Instances                 | Last Processed Time       | Last Processed Size | Last Processed Count | Total Number Of Records | CPU Utilization | Memory Utilization |
| BranchEnrichment  | OFF                           | batch            | 1                         | Oct 30, 2023, 4:38:24 PM  | 18.47 MB            | 11                   | 11                      | 2.208 / 1.500   | 965.93 / 2048      |
| SalesDataprep     | OFF                           | batch            | 1                         | Oct 30, 2023, 12:09:10 PM | 920 B               | 100                  | 0                       | 0.092 / 0.100   | 159.836 / 250      |
| InventoryData     | OFF                           | batch            | 1                         | Oct 30, 2023, 12:07:21 PM | 45.05 MB            | 79                   | 0                       | 1.417 / 1.500   | 1406.768 / 2048    |
| Sales_Data_SDG    | OFF                           | realtime         | 0                         | Oct 30, 2023, 12:06:06 PM | 0.00 KB             | 0                    | 0                       | 0.428 / 0.200   | 959 / 500          |
| Inventory_DB      | OFF                           | batch            | 1                         | Oct 30, 2023, 12:11:34 PM | 40.89 MB            | 74                   | 0                       | 1.915 / 1.500   | 1338.236 / 2048    |
| CustomerDataprep  | OFF                           | batch            | 1                         | Oct 30, 2023, 5:36:03 PM  | 248 B               | 20                   | 15360                   | 0.100 / 0.100   | 161.712 / 250      |
| CustomerDataSql   | OFF                           | batch            | 1                         | Nov 28, 2023, 12:38:23 PM | 39.60 MB            | 98                   | 16740                   | 2.476 / 1.500   | 993.84 / 2048      |
| InventoryData_SDG | OFF                           | realtime         | 0                         | Oct 30, 2023, 12:06:05 PM | 0.00 KB             | 0                    | 0                       | 0.333 / 0.200   | 925 / 250          |
| Customer_Data_SDG | OFF                           | realtime         | 0                         | Nov 28, 2023, 12:37:51 PM | 0.00 KB             | 0                    | 0                       | 0.182 / 0.200   | 953 / 250          |







# DATA SCIENCE LAB

BDB Data Science Lab, a collaborative powerhouse where data scientists converge to conduct joint experiments, exchange critical elements like Notebooks and models, and validate these resources for seamless deployment into Production. This dynamic environment fosters teamwork, ensuring efficient collaboration and precision in deploying innovative, data-driven solutions.





# DATA SCIENCE LAB- NOTEBOOK FEATURES



A Notebook within the Data Science Lab module comprises cells housing either explanatory text in Markdown format or executable code alongside its corresponding output."

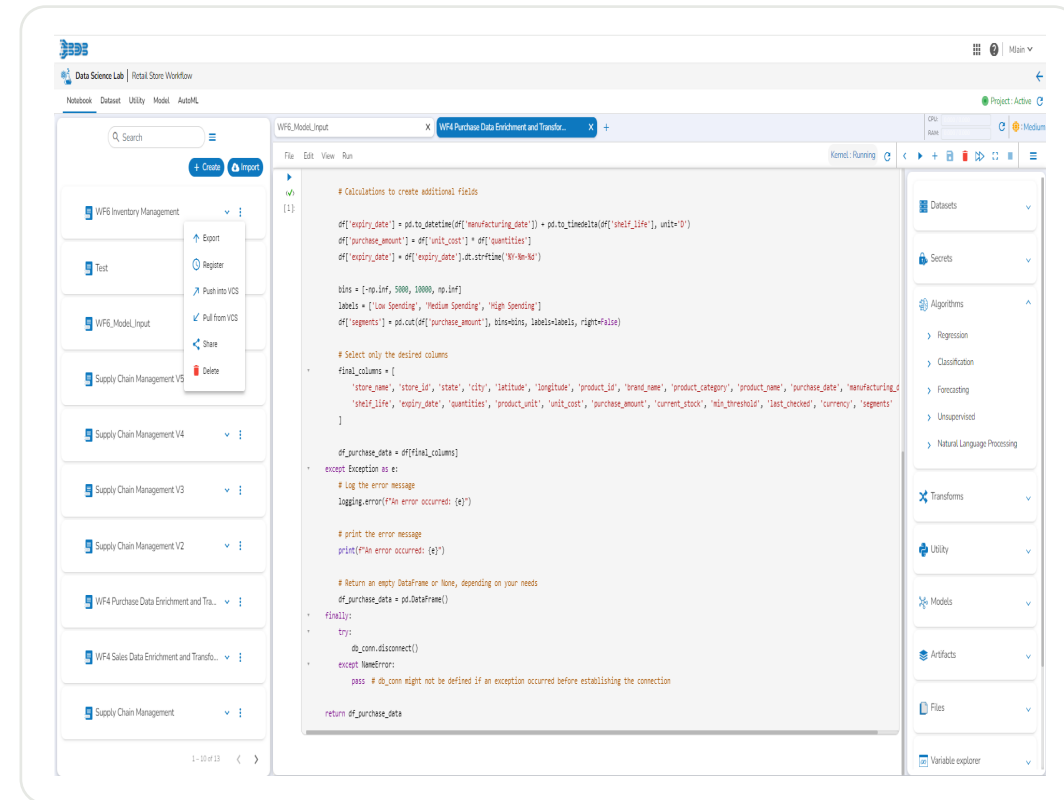
**Code Execution:** Facilitates direct execution of programming code, commonly in languages such as Python. The notebook environment processes this code and displays the output directly below the executed cell.

**Interactive Data Exploration:** Facilitates dynamic exploration of datasets through interactive widgets and libraries (Pandas, Pandas Profiling, etc.). Enables manipulation, filtering, and visualization of data to derive insights.

**Collaboration and Sharing:** Enables sharing and collaborative work through version control systems (GitHub, GitLab, etc.)

**Implemented a multi-tab interface:** supporting concurrent work on multiple notebooks, enhancing user productivity and flexibility in data analysis tasks. An Auto-Save option can be enabled or disabled for Notebooks.

**Markdown Support:** Allows the integration of explanatory text using Markdown formatting alongside code cells.



Compare this with Jupyter Notebook or Google Colab

# DATA SCIENCE LAB- GEN AI (SLMs) - BDB ASSIST



A BDB Assist is designed to be transparent and explainable AI assistant. Our notebook system guarantees that every AI recommendation transforms into transparent and replicable outcomes, enabling data teams to place unprecedented trust in AI.

## Generate Code Automatically:

Starting from scratch is no longer a hurdle. With BDB Assist code Generation capability, simply provide your prompts, questions, or instructions, and watch as an entire notebook— including code, SQL queries, and text — materializes before your eyes.

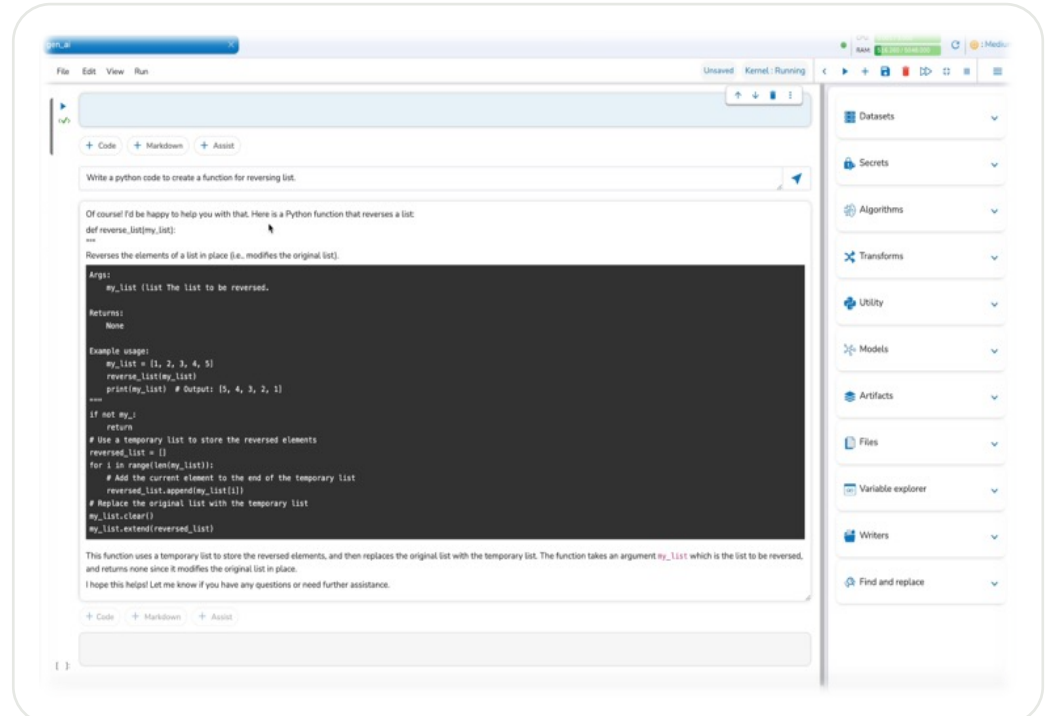
**Explain the code:** With BDB Assist don't let complex pieces of code baffle you anymore with concise, easy-to-understand explanation

## Debug & Edit the code:

BDB Assist help you to revise or refactor your code and pinpoints the issue and provides an immediate fix.

## Assist in all modules:

Like Data Preparation BDB Assist will be made available in all parts of the platform



BDB Assist (Generative AI) based search engine leverages both external and proprietary data sources to generate insightful responses and visualisations. It streamlines knowledge management in enterprises by processing user queries, thereby improving decision-making and enhancing productivity

# DATA SCIENCE LAB- ALGORITHMS



BDB DS Lab offers support for various machine learning algorithms across different domains:

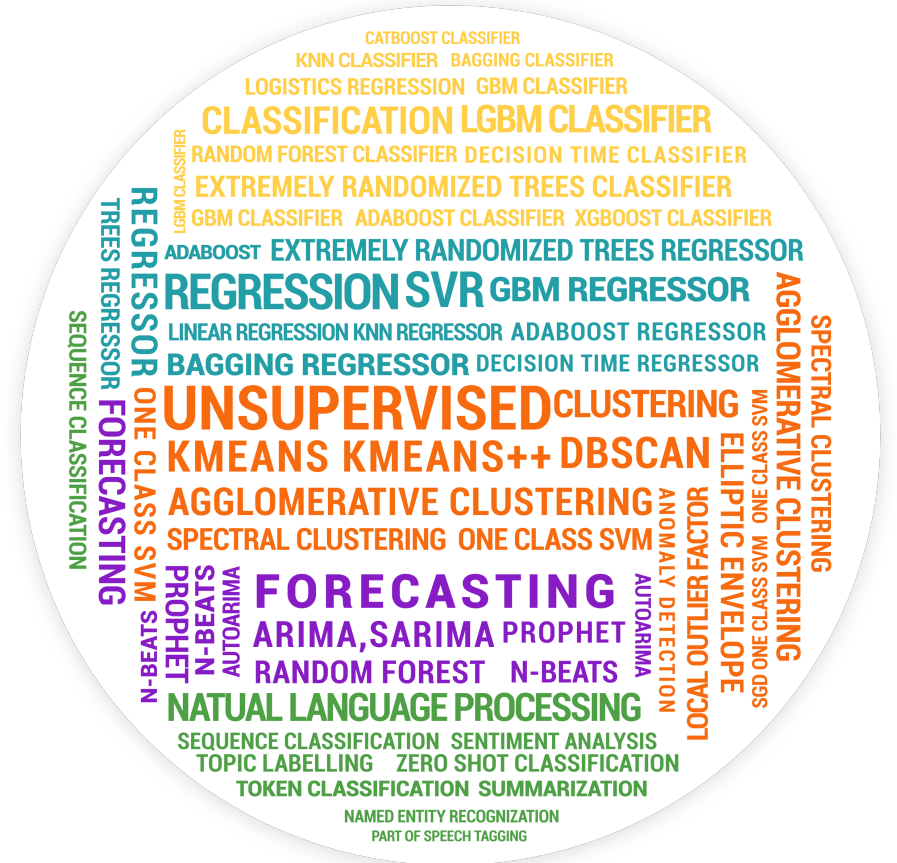
**Regression:** Used for predicting continuous numerical values, like prices or trends. It offers a suit of algorithms including Linear regression, SVR, KNN Regressor, Random forecast regressor and so on. Each algorithm operates uniquely, with approaches ranging from establishing linear relationships between variables to creating complex decision tree ensembles, offering varied ways to model and forecast numerical outcomes.

**Classification:** It offers a suite of algorithms including logistic regression, decision trees, random forests, SVM, and more

**Forecasting:** Predicting future values based on historical data, useful in finance, weather forecasting, etc.

**Unsupervised Learning:** Analysing patterns in data without labelled outputs, including techniques like clustering and anomaly detection.

**Natural Language Processing (NLP):** Handling text and language data, tasks like sentiment analysis, translation, and text generation.





# DATA SCIENCE LAB - AUTO ML FEATURES



The Auto ML empowers users to generate diverse experiments using their datasets and conveniently organizes a list showcasing all these created experiments for easy reference and comparison.

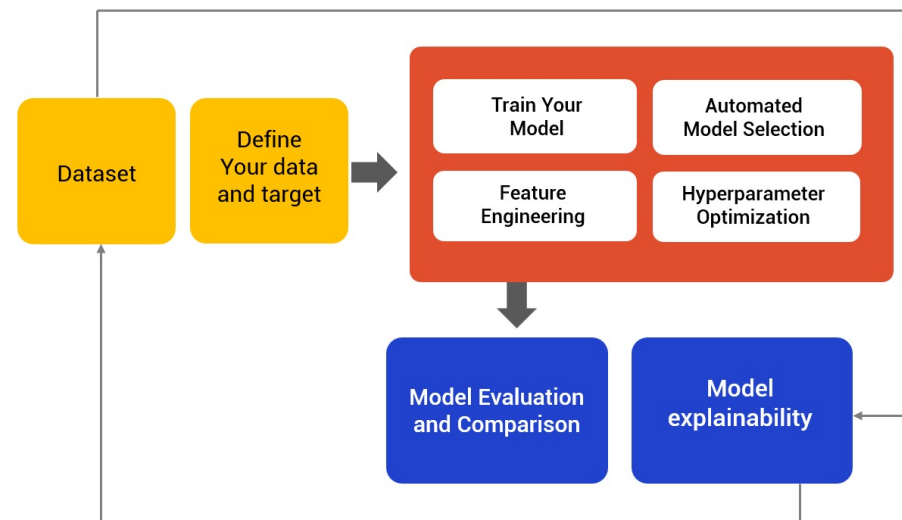
**Feature Engineering:** BDB's Auto ML automates feature selection, creation, and identification, crucial for improved model performance and insightful data interpretation

**Automated Model Selection:** Auto ML suggests ideal machine learning algorithms based on data characteristics and the problem type, optimizing performance.

**Hyperparameter Optimization:** BDB's Auto ML refines model settings by exploring various hyperparameter configurations, enhancing model performance for specific datasets.

**Model Evaluation and Comparison:** BDB's Auto ML streamlines model assessment, comparing performance across metrics to guide users in choosing the most suitable model for their needs.

**Model Explainability:** BDB's Auto ML platforms prioritize model explainability, offering insights into feature importance, dependencies, and "What if" analysis, fostering trust and informed decision-making.





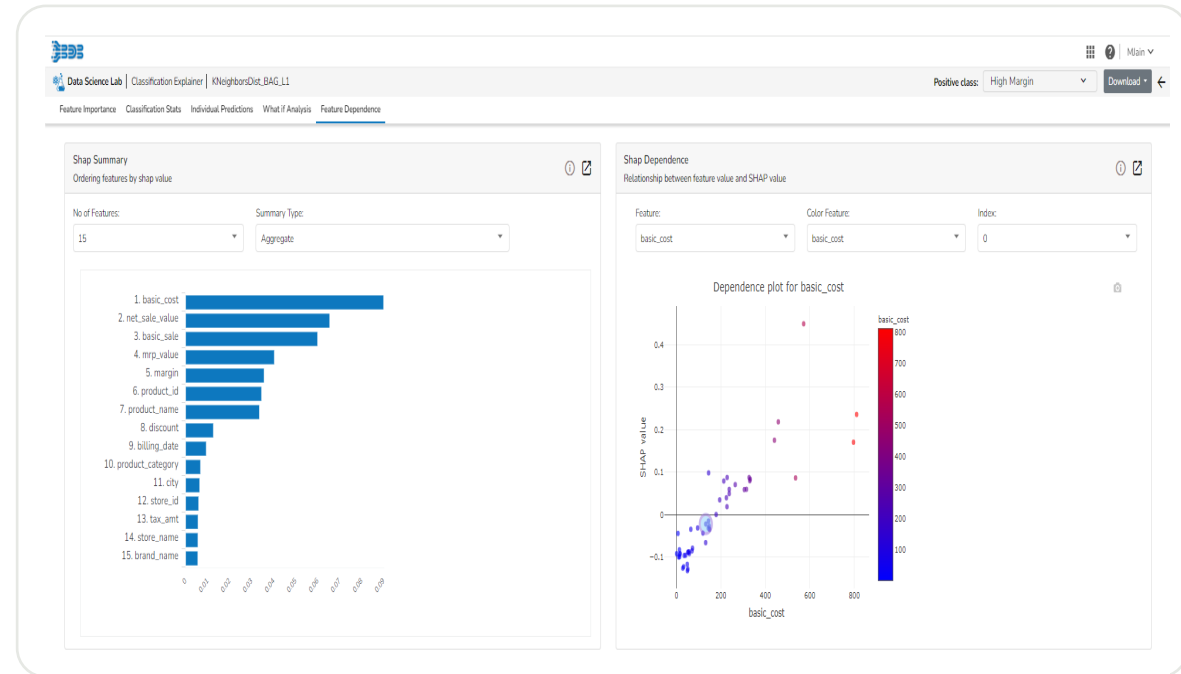
# AUTO ML- MODEL EXPLAINABILITY

**Feature Importance:** Feature importance analysis is a crucial aspect. This analysis helps determine which features had the biggest impact on the model's predictions. By examining feature importance, you gain insights into the key factors influencing the model's decision-making process.

**Regression stats:** The regression stats offer essential metrics for model evaluation. Predicted Vs Actual Plots visualize prediction accuracy, Residuals examine discrepancies across target values, and Plot Vs Features highlight model behaviour in different feature ranges, aiding in identifying limitations.

**Individual Prediction:** The individual prediction section displays the predicted probability for each target label. It provides insights into how the model assigns probabilities to different classes for a specific observation.

**What If Analysis:** The What If Analysis allows stakeholders to understand the potential consequences of different scenarios or decisions. In this analysis, you can change the values of selected variables to see how the outcome would change. It helps identify the sensitivity of the outcome to different inputs and which variables are most important.



**Feature Dependence:** The feature dependence analysis explores the relationship between feature values and their impact on predictions. This analysis allows you to investigate how the model uses features in line with intuitions or learn about the relationships the model has learned between input features and predicted outcomes.

BDB Auto ML supports such a diverse range of algorithms and techniques across forecasting, regression, and classification tasks. Let's break down what these models and techniques typically entail

**Regression Models** : Weighted Ensemble, Mighty GBM Large, and K Neighbors Dist—employ ensemble strategies, blending models or subsets for improved accuracy and robustness in regression tasks.

**Classification Models** : LightGBM, Weighted Ensemble, and K Neighbors—employ diverse strategies. LightGBM likely combines predictions from multiple LightGBM models, while Weighted Ensemble is tailored for classification. K Neighbors Dist uses Bagging to enhance accuracy by employing multiple KNN models on varied data segments.

**Forecasting Models** : Seasonal Naive, DeepAR, and PatchTST—cover varied approaches. Seasonal Naive relies on the last observed value within the same season for forecasts. DeepAR utilizes advanced deep learning for precise time series predictions, while PatchTST likely represents a specialized implementation of Time Series Transformers for specific forecasting tasks.

DEEPAR PATCHTST  
**FORECASTING** SEASONAL NAIVE  
WEIGHTED ENSEMBLE MIGHTY GBM LARGE  
**REGRESSION MODELS**  
K NEIGHBORS DIST  
**CLASSIFICATION MODELS**  
LIGHTGBM WEIGHTED ENSEMBLE  
K NEIGHBORS

# BDB DATA SCIENCE LAB FEATURES IN NUTSHELL



## Project setup

- Supported environments
  - TensorFlow
  - PyTorch
  - PySpark
- Ability to configure memory and CPU, GPU as per experimentation requirement.
- Specify library dependency.
- Role Based Data Access

## Auto ML

- Classification
- Regression
- Forecasting
- Model Explainability

## Notebook

- Create/Upload Notebook
- Model Explainability
- Model As API
- Artifacts
- Built-in Algorithms
- Files (user can create folder/file & upload data), Variable explorer, secrets,
- Automatic code generation for basic task saving time for repeated task.

## Data Preparation and Integration:

- Import and connect to various data sources.
- Data profiling and exploration.
- Data cleaning and transformation using visual recipes.
- Data wrangling with code-based transformations.
- Data enrichment and feature engineering.
- Feature store

## Data Visualization and Exploration:

- Interactive data visualization tools.
- Customizable dashboards and reports.
- Data exploration and analysis.

## Machine Learning and Modelling:

- Built-in machine learning algorithms.
- Feature selection and model validation.
- Hyperparameter tuning
- Supervised and Unsupervised learning
- Model training, evaluation, and tuning.
- Advanced model deployment options.

## Collaboration and Project Management:

- Project workspace for collaborative work.
- Role-based access control.
- Version control for projects and components.
- Comments and discussions within the platform.
- Sharing and collaboration features.

## Model Deployment and Operations:

- Single Click Deployment of ML models to Inference Pipeline
- Model As API
- Data drift, Model drift
- Model monitoring and performance tracking
- ML observability

## Advanced Analytics:

- Support for advanced statistical analyses.
- Time series analysis and forecasting.
- Text analytics/NLP
- Video/Image Analytics

## Data Governance and Security:

- Data lineage and auditing.
- Data quality assessment and profiling.
- GDPR and data compliance features.

## APIs and Integration:

- Integration with external APIs and services

## Cloud and On-Premises Deployment:

- Deployment options for cloud environments and on-premises setup

## Extensions and Customization:

- Customizable plugins and extensions.
- Integration with third-party tools.

## Training and Learning Resources:

- Training materials and resources for learning DS Lab.



# DATA SCIENCE LAB- AI SERVICES(ADVANCED AI-POWERED SOLUTIONS FOR IMAGE, TEXT, AND DATA ANALYSIS)

**Face Recognition:** Our AI-powered face recognition technology can accurately detect faces in images and videos, making it ideal for applications like security, identity verification, and more.

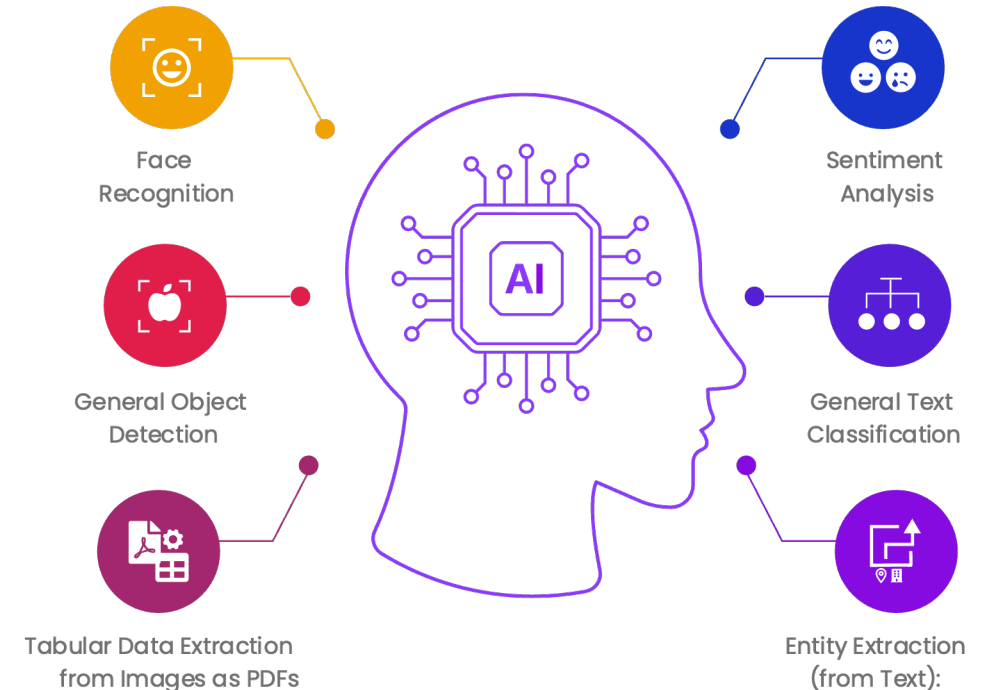
**General Object Detection:** Our advanced object detection technology can identify objects within images and videos, providing powerful capabilities for various use cases like image analysis, video surveillance, and more.

**Tabular Data Extraction from Images as PDFs:** Our solution uses state-of-the-art optical character recognition (OCR) techniques to extract tabular data from images, including scanned PDFs, making it ideal for data digitization and automation.

**Sentiment Analysis:** Our sentiment analysis technology Provides insights into emotions and attitudes expressed in text, useful for customer feedback analysis, social media monitoring, and more.

**General Text Classification:** Categorizes text data into predefined classes, beneficial for tasks like spam detection, news categorization, and similar applications.

**Entity Extraction (from Text):** Our entity extraction technology can Identifies and extracts crucial entities (names, locations, organizations, etc.) from unstructured text data, beneficial for various NLP-based applications.





# MODEL AS AN API



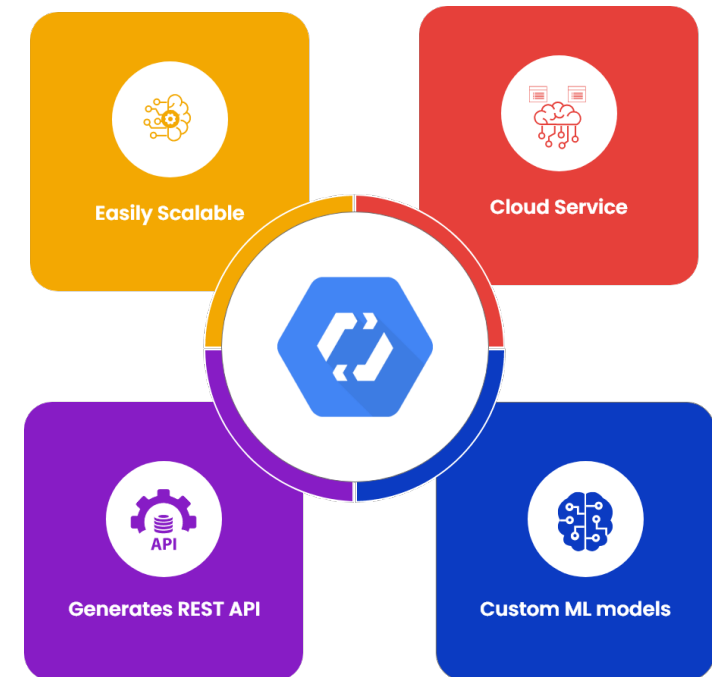
The Model as an API feature integrates machine learning models seamlessly into applications for real-time predictions and insights.

**Power of AI:** Users leverage AI capabilities for tasks like predictions and classifications through a simple and intuitive interface, requiring minimal coding or infrastructure setup.

**Versatile Applications:** From recommendation engines to fraud detection systems, this feature enables diverse uses of machine learning for intelligent decision-making.

**Unlocking Data Potential:** Businesses harness the full potential of their data, driving smarter decisions with AI-powered insights.

**Future Innovation:** Embrace the future of AI-driven applications, elevating business potential with the Model as an API feature.

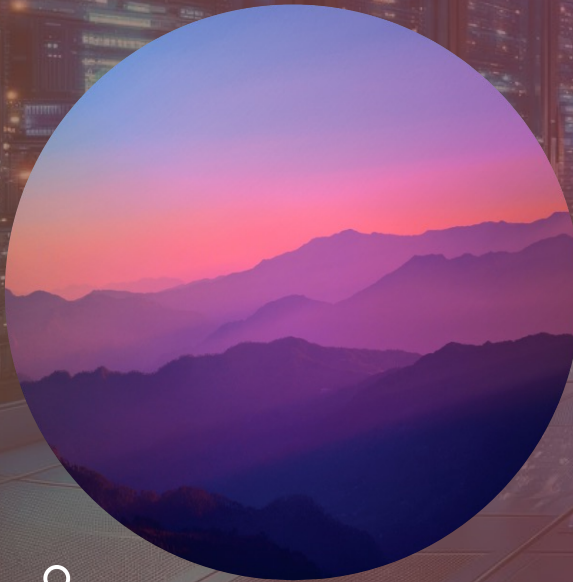






# DATA VIRTUALIZATION

Data Centre module functions as the core of the Platform, serving as a central hub for efficiently managing and organizing data collected, enriched from diverse sources"





# DATA CENTER KEY FEATURES



**Data Loss Protection:** In the dataset, users can leverage a range of data protection methods including redaction, masking, hashing, and data generalization, offering versatile ways to obscure or transform sensitive information while preserving data integrity and ensuring compliance with privacy standards.



**Query Optimizer:** Data centre incorporate a query optimizer to refine SQL query execution within datasets and data stores.



**Data as an API:** enabling user interaction with third-party and custom APIs via a Platform. Key features include creating/deploying custom APIs and seamless integration with external APIs for application connectivity.



**Governance:** BDB Data Center supports data transparency, access control, and some data quality monitoring aspects within data governance.



**Security:** The BDB Platform ensures data integrity by safeguarding against data loss or corruption throughout its lifecycle. It employs industry-standard encryption methods to secure client database connections, cached data, and other crucial security aspects.

# DATA CENTER KEY COMPONENTS

**Data Connector:** Provides optimized connectors for various data sources, easing data access and integration.

**Data Set:** Organizes structured data in a specific format for easy access, retrieval, and analysis.

**BDB Data Store:** Stores structured data with metadata, supporting governance, integration, and analysis.

**Data Sheet:** User-friendly spreadsheet for storing and sharing structured data, breaking down data silos.

**Data Sandbox:** Isolated environment for data science experiments, enabling secure work without impacting production systems.

**Data Catalog:** Centralized repository for managing metadata related to various data assets, including data sources, types, lineage, and quality.

**Data as an API:** Enables secure data sharing and integration with third-party applications.



The screenshot displays the Data Center interface. The top section, titled "Data Center", includes a navigation sidebar and a main content area. The main content area features an illustration of a person interacting with data storage units and a list of components: Data Connector, Data Set, BDB Data Store, Data Sheet, Data Sandbox, Data Catalog, Data as an API, and Data Preparation. Below this, the "Data Preparation" section is visible, featuring an illustration of a server rack and a list of capabilities.

**Data Center**

The central hub for managing and organizing data from various sources offers native connectors for data source access and data management. The Data Center Module includes:

- ✓ **Data Connector:** The BDB Platform offers access to a diverse array of native connectors optimized for databases, files, APIs, and various data sources, streamlining data access and integration.
- ✓ **Data Set:** Getting a structured collection of data organized and stored in a specific format for easy access, retrieval, and analysis.
- ✓ **BDB Data Store:** BDB Data Store stores data in a structured format, including databases, APIs, and flat files, with metadata supporting governance, integration, and analysis.
- ✓ **Data Sheet:** The Data Sheet stores data in a visually appealing & easy-to-use spreadsheet. It can act as a powerful database that allows to remove the data silos and store or share the structured data access. A user-friendly, visually appealing spreadsheet for storing and sharing structured data, breaking down data silos.

✓ **Data Sandbox:** A Data Sandbox is indeed an isolated environment used for data science experiments and analysis. It provides a controlled and secure space where data scientists and analysts can work with data without affecting the production systems. Seamlessly uploads CSV/Excel files in the Data Sandbox, empowering self-service reporting and unleashing data science experiments.

✓ **Data Catalog:** A Data Catalog is a centralized and structured repository designed to index, organize, and manage metadata related to various data assets. This metadata includes essential information about data sources, data types, data lineage, and data quality. The Data Catalog is the centralized repository to manage tagging, stewardship, and status of various Data Asset Types.

✓ **Data as an API:** Securely share your data with external entities or integrate their data into third-party applications.

✓ **Data Preparation:** Detect anomalous records and purge the unwanted data set records in a few clicks using Machine-learning-based smart techniques and sampling. Apply changes to data sets and export the analysis-ready data. Detect data anomalies and ensure data quality in a few clicks using the auto-prep feature and various other built-in transformers.

[Read Less](#)

**Data Preparation**

Data preparation is a critical step in the data analysis and machine learning pipeline. It ensures that the data is clean, reliable, and suitable for analysis, allowing organizations to derive meaningful insights and make informed decisions based on their data.

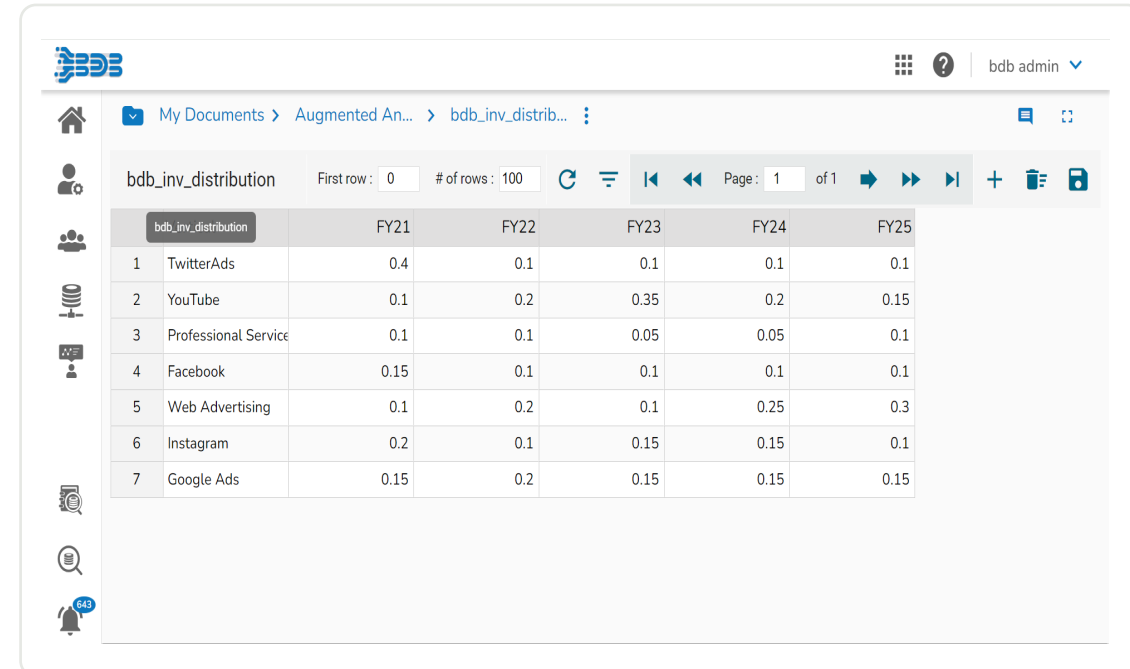
- ✓ The module employs machine learning algorithms such as clustering, outlier detection, and statistical methods to technically identify and flag anomalous records within the dataset, often with predefined thresholds or models.
- ✓ Including SQL queries or data filtering, are used to selectively remove unwanted data sets or records based on user-defined criteria, ensuring data cleanliness.
- ✓ Smart sampling methods like random, stratified, or systematic sampling are technically implemented to efficiently handle large datasets, allowing users to work with manageable subsets while preserving data representativeness.
- ✓ A robust data transformation engine is technically developed, offering a library of over 100 transformers. These transformers include data cleaning, normalization, aggregation, and the creation of new features to prepare data for analysis.
- ✓ Implement an auto-save mechanism within the Data Preparation module to periodically save the user's progress and changes made during data preparation. This may include saving the state of data transformations, configurations, and the applied scripts that can be exported to the Data Pipeline and Data Science Lab modules.
- ✓ The module is precisely integrated into the Data Center unit, providing users with a centralized point of access for data preparation tasks, streamlining workflows, and improving user experience.
- ✓ The Data Grid, Profile tabs, and in-built charts are technically designed to visualize and profile data. Data is carefully rendered within the user interface, providing insights into data patterns, distinct values, and occurrences through statistical summaries and smart visualization.

**Data Preparation:** Utilizes machine learning-based techniques to detect anomalies, remove unwanted data, and prepare analysis-ready datasets efficiently. This tool is integrated at Data Centre, Notebook, Pipeline and Reports module in a seamless manner.

# BDB DATA SHEET

Data Sheet acts as a visually appealing spreadsheet for storing structured data, fostering sharing and removal of data silos.

- Allows users to name the Data Sheet and input column names aligned with data sources.
- Options for various Input Types (text, numeric, dropdown, etc.) and Look Up for dropdown values from other datasheets.
- Data Restriction and Filter functionalities for defining custom fields and enabling filters.
- Add/Update function for adding new or modifying existing columns.
- Deletion option for removing added datasheet information.
- New columns or updates reflect in the preview space, enabling a visual understanding of changes.
- Save option becomes active after modifications, confirming the addition or updates to the Data Sheet.
- Users can leverage sandbox files to create datasheets, streamlining the process and enhancing usability.



The screenshot displays the BDB Data Sheet interface. The top navigation bar includes the BDB logo, a user profile icon for 'bdb admin', and a breadcrumb trail: 'My Documents > Augmented An... > bdb\_inv\_distrib...'. Below the navigation, there's a table titled 'bdb\_inv\_distribution'. The table has columns for 'FY21', 'FY22', 'FY23', 'FY24', and 'FY25'. The rows list advertising categories: TwitterAds, YouTube, Professional Service, Facebook, Web Advertising, Instagram, and Google Ads. The interface also shows a sidebar with various icons and a notification bell with '643'.

|   |                      | FY21 | FY22 | FY23 | FY24 | FY25 |
|---|----------------------|------|------|------|------|------|
| 1 | TwitterAds           | 0.4  | 0.1  | 0.1  | 0.1  | 0.1  |
| 2 | YouTube              | 0.1  | 0.2  | 0.35 | 0.2  | 0.15 |
| 3 | Professional Service | 0.1  | 0.1  | 0.05 | 0.05 | 0.1  |
| 4 | Facebook             | 0.15 | 0.1  | 0.1  | 0.1  | 0.1  |
| 5 | Web Advertising      | 0.1  | 0.2  | 0.1  | 0.25 | 0.3  |
| 6 | Instagram            | 0.2  | 0.1  | 0.15 | 0.15 | 0.1  |
| 7 | Google Ads           | 0.15 | 0.2  | 0.15 | 0.15 | 0.15 |



# DATA CENTER – DATA CATALOG FEATURE

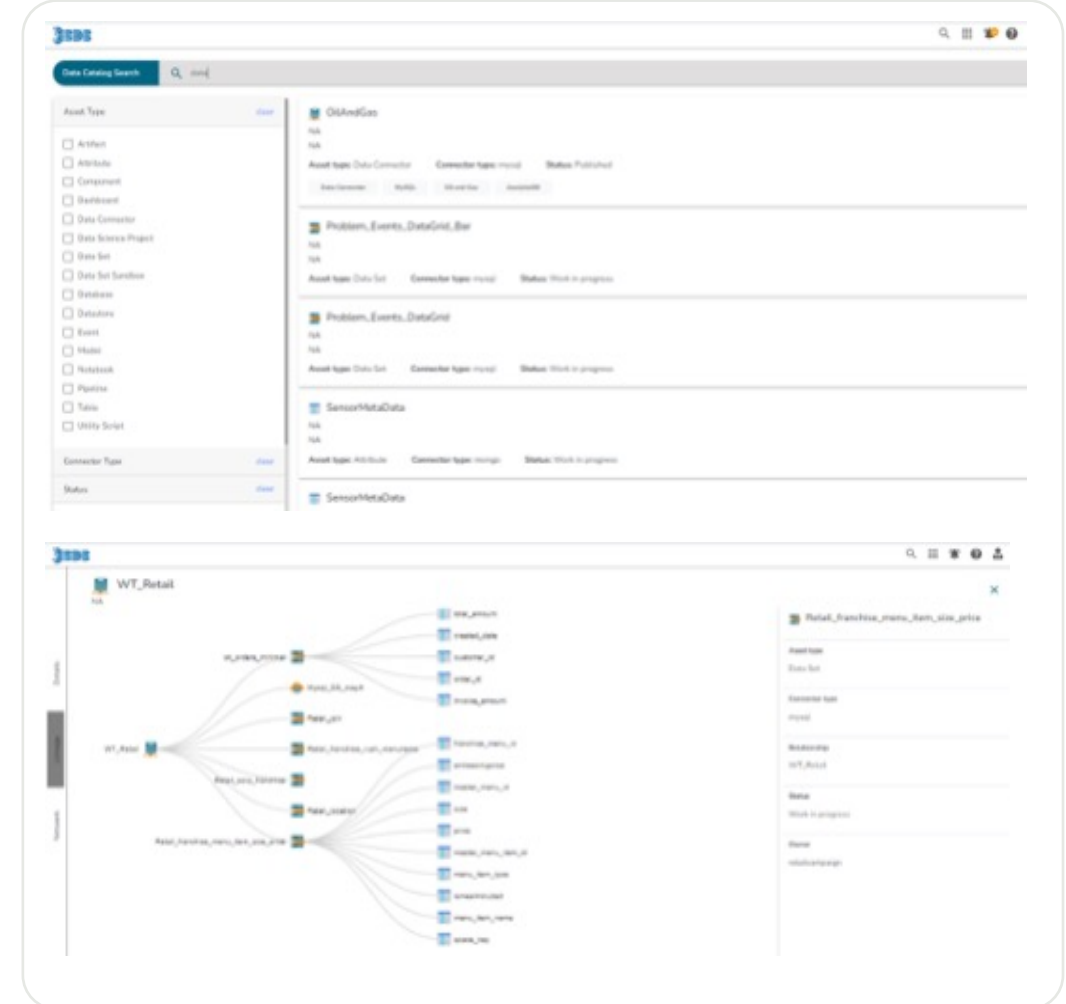
The **Data Catalog** feature in the Data Center module of the BDB Platform allows users to effectively manage and organize their data assets through the use of tags and data stewardship.

**Data Catalog** empowers the data user to have a holistic view of data and quickly discover data lineage that can drive towards insightful decisions optimizing their business worth.

BDB supports **Data lineage**, which allows users to track the origin and history of data as it moves through the platform, making it easier to manage and understand big data.

It Provides a robust metadata management capability. **Data Catalog Search** provides a centralized place to:

- Gain an organized view to decrease the time and effort of searching for the right data.
- Get enhanced data to speed up data-driven decision-making.
- Increase operational efficiency and productivity.
- Improve data management.
- Get ownership over the concerned data.
- Locate insightful dashboards and charts to save extra effort.
- Access an interconnected view of all your organization's data & assets.



# BDB DATA CENTER & DATA PREPARATION CAPABILITIES



- Micro Services based Architecture
- Rest API based Communication
- 70+ Data connectors

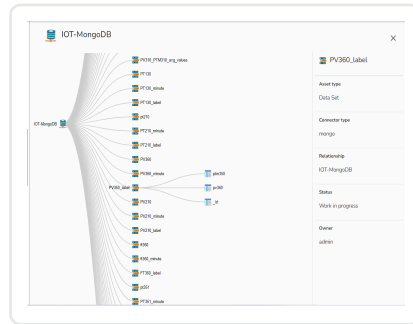


- Secure, Multitenant
- Hybrid Integration Capabilities
- Mobility and Diversity
- User Management & Manageability

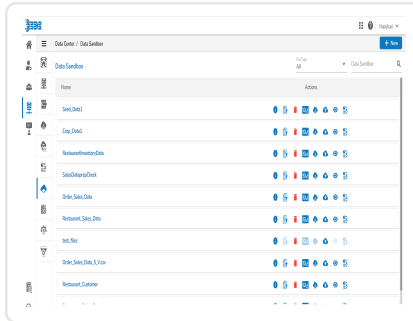
## Data Center Components

- Data Connector
- Data Set
- Data Store
- Data Store metadata
- Data Sheet
- Data Sandbox
- Data Catalog
- Data Preparation
- Data as an API

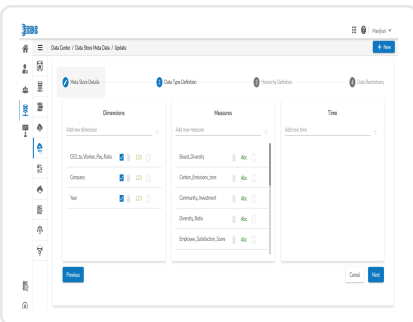
## Support Data Catalog



## Sandbox Environment



## Data Store metadata



## Data Preparation built-in Transformations

### Transformation Related Strings

- Add Prefix
- Add Suffix
- Change to lower case
- Change to title case
- Change to upper case
- Convert value to Column
- Extract substring at position
- Extract substring before delimiter
- Get Domain
- Get Host
- Get Subdomain Name
- Insert Character
- Is even
- Is Odd
- Negate Boolean Value
- Remove Accents
- Remove consecutive characters
- Remove part of text
- Remove Trailing and Leading Characters
- Search and Replace
- Split String

### Transformation related Dates

- Add Duration
- Add Interval To Date
- Convert from UTC
- Convert TimeZone
- Convert to UTC
- Extract Date Part
- Extract Time Unit
- Find Date Difference
- Format Date
- From Unix Time
- Sub Interval To Date
- Subtract Duration

### Transformation Related Numbers

- Max...
- Mean...
- Negate..
- Number Names...
- Remove Functional Part...
- Round Multiple Column using floor mode
- Round Value using ceil mode
- Round value using down mode
- Round value using floor mode
- Round value using half up mode

### Anonymization

- Data Hashing
- Data Masking
- Data Variance

### Advance Transformations

- Cluster & Edit
- Expression Editor
- Find Anomaly
- SQL Transform

### Data Cleansing

- Clear cells on matching value
- Delete rows on matching value
- Delete rows with empty cell
- Delete rows with invalid cell
- Delete rows with negative values
- Fill cells with value
- Fill empty cells with Text
- Flag duplicates in Columns
- Flag Duplicates in Table
- Remove Duplicates from Column
- Remove Duplicates from Table
- Remove letters
- Remove Numbers
- Remove Special Characters

### Transformation Related Functions

- Base64Decode
- Base64Encode
- Ceiling Columns
- Copy Column
- Count Distinct
- Degree to Radians
- Exact
- Extract Number
- Extract Number with regular expression
- Get URL Parameters
- If Missing
- If Null
- If Then Else
- IN
- Is Mismatched
- Is Missing
- Is Null
- Is Valid
- Log of Column
- Nest Column to Object/Array
- PI
- Power of Column
- Radians to Degree
- Round values of Columns
- Sign of Columns
- Split Invalid
- Sqrt of Column
- Trigonometric Functions

### ML

- Binarizer

## Data Centre and Data Preparation for Tech Staff

- Intuitive and attractive User Interface
- Complex Workflows can be created
- BDB Platform offers optimized connectors for databases, files, APIs, streamlining data access and integration.
- Publishing, Sharing and Collaboration Features
- User Level Access and Control
- Strong Migration Tool to handle data center Content
- Isolated environment for secure data science experiments, enabling self-service reporting and ana.
- Securely share and integrate data with external entities or third-party applications through Data as an API feature.
- Detect and handle anomalies, purging unwanted records, and ensuring data quality using machine-learning-based techniques and automated transformations.
- A visually appealing, spreadsheet-style interface for storing and sharing structured data, promoting collaboration and breaking data silos.
- More Affordable than any competitor
- Connects with any data – Cloud or on premise
- Constant Development – Managed by a disciplined strong R&D Team
- Reliable & cost-effective Customer Support
- Real Time Data Analysis is possible

# DATA CENTER – DATA AS AN API



REST APIs is the next big wave for the SAAS platform. In essence, API as a Service allows authorized users to interact with third-party APIs or manage their custom APIs via a software Platform.

The BDB Platform that provides this facility to use Data as an API Service has two key elements:

The ability to build, test, and deploy one's API services, and 2. The ability to connect an application with third-party APIs.

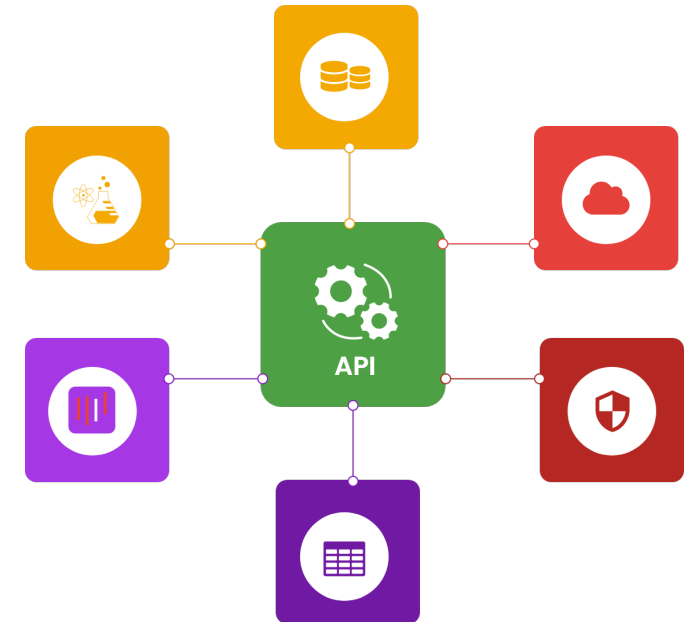
## Advantage of using Data in form of API Service

**Agility** - users can move quickly, due to the simplicity of data access, and not needing extensive knowledge of the underlying data. Data structures and location-specific requirements can be modified to meet user needs.

**Cost-effectiveness** - providers can build the base with the data experts and outsource the presentation layer, which makes for very cost-effective user interfaces and makes change requests at the presentation layer much more feasible. So, Customers /Partners can create Custom UI Layers or Use Other Visualization Tools to create their dashboards with these REST API calls.

**Data Quality** - data access is controlled through data services, which tends to improve data quality, as there is a single point for updates.

**Data Monetization** - This will help companies doing Data Analytics Implementation on BDB to use Rest API calls as a data monetization method to serve external Applications or Public Portals accessing that kind of data.







## DATA VISUALIZATION

BDB offers two powerful visualization plugins. The Report module enables non-technical users to create real-time visualizations and reports effortlessly, integrated with NLP-based search for insightful discoveries. With simple data selection and drag-and-drop features, users swiftly generate tailored views, saving them for personalized stories. The Designer module complements this, ensuring controlled dashboard creation and refined analytics design through a web-based interface”.





# BDB DASHBOARD VISUALIZATION CAPABILITIES



Micro Services based Architecture  
Rest API based Communication  
Horizontally & Vertically Scalable

BDBC CONNECTORS



Secure, Multitenant  
Hybrid Integration Capabilities  
Mobility and Diversity  
User Management & Manageability

**Simple Drag and Drop Interface**

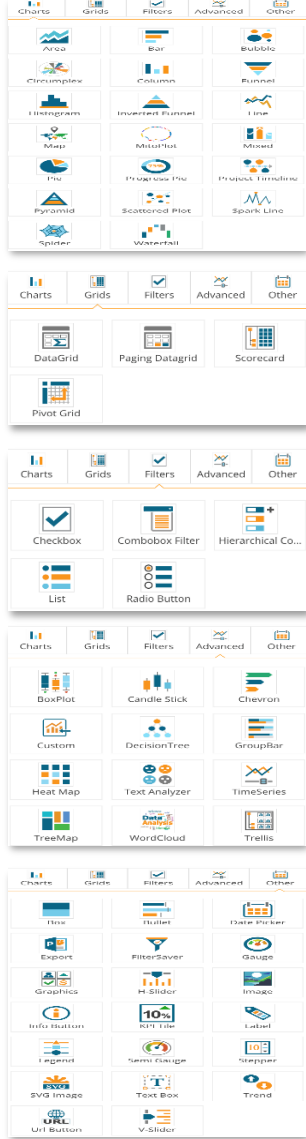
**Advanced Visualizations**

- Leaflets (for Geospatial)
- Mito Plot & Many More
- SDK based Java Scripting

**Filter Savers & Filter Panel**

**Highly Interactive and Drill Down**

**80+ Visualization Components**



**Import Custom Component Capabilities**

- Fusion Charts
- D3 Charts
- High Charts
- Data Table
- Am Charts etc.

**Strong Export Features**

- Excel
- PDF
- PPT
- CSV
- PNG

**Live Data Connectors**

- Excel & CSV
- Data Services
- Predictive Connectors
- Data Store (Elastic Based)
- Web Socket (real time data)
- Merged Data (at Client Level)

**Web, Mobile and Tablet**

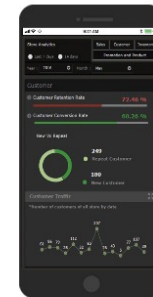
Governed Dashboard



- CSV
- Web Service
- Predictive
- Data Store
- Java Script

**Suitable for CXOs**

Governed Dashboard

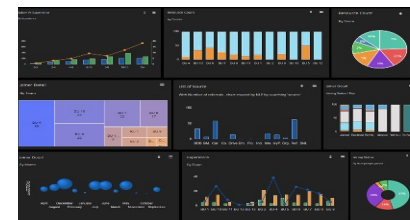


Self Service BI Dashboard



**Mobility iOS & Android**

Self Service BI Dashboard



**Suitable for Business User**

**Governed Dashboard Designer for Tech Staff**

- Intuitive and attractive User Interface
- Complex Workflows can be created
- Strong Visual Appeal with Animation
- Ease of Use, 100% mobile with Mobile App.
- Publishing, Sharing and Collaboration Features
- NLQ – NLP
- Embed the Analytics Content in other places
- Analytics Dashboards – No 1 in World
- Interactive Visual Exploration
- Scalability and High Performance
- User Level Access and Control
- Strong Migration Tool to handle BI Content
- Best Dashboarding tool for Enterprises
- Strong Predictive and Prescriptive Capabilities
- More Affordable than any competitor
- Connects with any data – Cloud or on premise
- Constant Development – Managed by a disciplined strong R&D Team
- Reliable & cost-effective Customer Support
- Real Time Data Analysis is possible

**Self Service for Non-Technical Staff**

- Simple Drag & Drop + Search based Charts
- Formula Editor for Calculations
- 100% Responsive, Mobile Authoring
- Ability to handle large Data sets in a Flash
- Security User Level access & Control
- Strong Copy and Share Features
- Ability to Deploy and Scale to large Audience with-in Days
- NLQ – Voice based search in Android and iOS

# BDB DASHBOARD KEY FEATURES

**Effortless Creation:** Drag-and-drop interface, rich property panel for customization.

**Visual Variety:** Over 60+ interactive visualization components from basic to advanced.

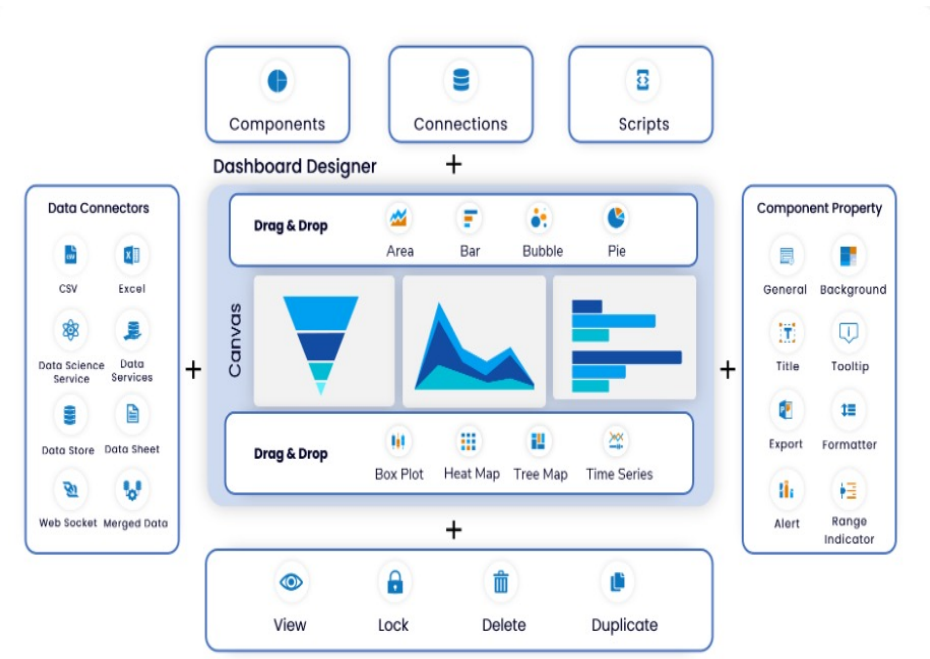
**Extensibility:** Incorporate custom charting libraries for unique visualizations.

**Robust Export:** Excel, PDF, PPT, PNG, CSV for seamless sharing.

**Data Connectivity:** Wide range for batch and real-time data from various sources.

**Real-Time Insights:** Monitor live data streams, track KPIs for immediate insights.

**Security:** User-level access control for safeguarding sensitive information.



**Migration Ease:** Smooth transition from development to QA and production environments.

**Device Flexibility:** Responsive design for optimal viewing across devices. (Mobile, Tabs etc.)

**Customization Power:** SDKs to extend and tailor the dashboard to specific needs.



# BDB DASHBOARD THEMES AND TEMPLATES

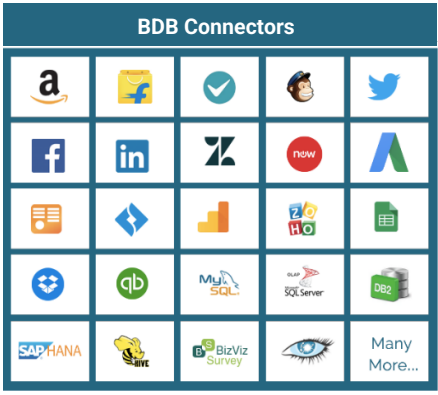
- Dashboard Theme:** Dashboard Theme enables users to effortlessly change their dashboard's appearance by selecting themes from an intuitive drop-down menu, instantly transforming its visual display to suit preferences and enhance aesthetics.
  - Dashboard Templates:** Enables users to select from various templates, facilitating the creation of new dashboards using pre-defined layouts or designs.
  - This streamlines the process by offering ready-made structures that users can customize, saving time and effort in building dashboards from scratch.
  - Streamline dashboard creation, ensure consistency, and leverage best practices.
- Users can select from various templates, saving time in building dashboards from scratch.



# BDB REPORT VISUALIZATION CAPABILITIES



- Micro Services based Architecture
- Rest API based Communication
- Horizontally & Vertically Scalable



- Secure, Multitenant
- Hybrid Integration Capabilities
- Mobility and Diversity
- User Management & Manageability

## Simple and Intuitive Interface

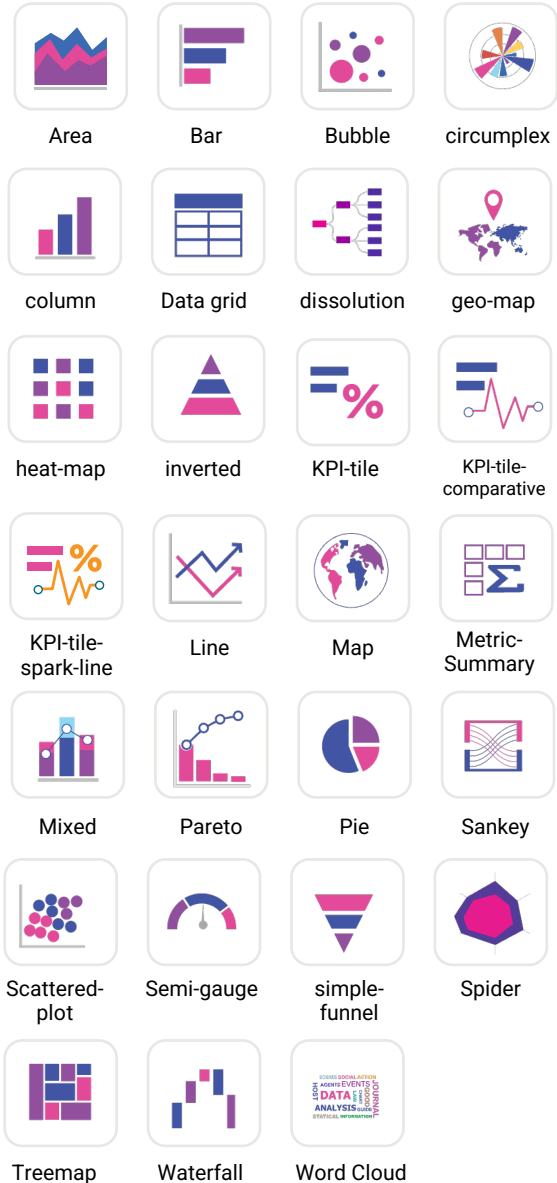
## Strong Export Features

- Excel
- PPT
- PDF
- CSV

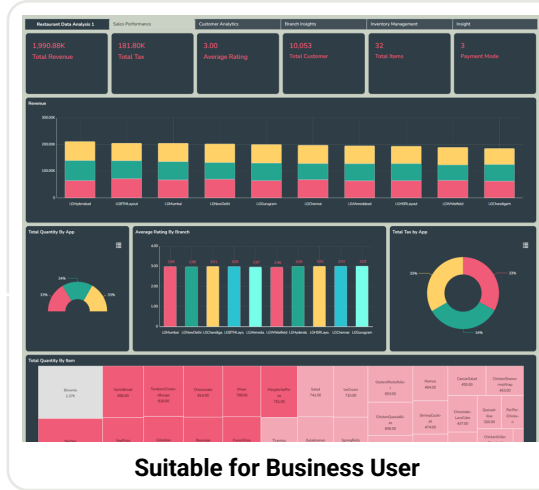
## Advanced Features:

- Interactions
- Global Filter
- Time-line Play
- Highly Drill down and drill through

## 25 + Visualization Components



## Modern day Reporting



Suitable for Business User

Web, Mobile and Tablet

## BDB Reports



Mobility iOS & Android

## Self Service Reports for Business User

- A simple and user-friendly interface.
- Easily generate a story directly from CSV files.
- Find and access data easily through a search interface.
- Streamlined data cleaning processes
- Create KPI's using multiple datasets within a single story
- Support for over 20 charting libraries.
- Easily customize chart properties.
- Customize headers for your reports.
- Receive suitable chart recommendations.
- Efficient data filtering options.
- Incorporate machine learning for advanced analytics.
- Perform calculations using a formula editor.
- 100% responsiveness for mobile authoring.
- Ability to work with extensive datasets swiftly.
- Control access with user-level security.
- Strong copy and sharing capabilities.
- Rapid deployment and scaling for large audiences.
- Natural language voice search on Android and iOS.
- Advanced search capabilities powered by AI.

# BDB REPORT KEY FEATURES

**Unified Data Story:** Compile a holistic narrative across multiple tabs, eliminating the need for numerous reports.

**Simple Interface:** Easily create charts by dragging dimensions/measures or using quick search.

**Data Connectivity:** Seamlessly connect using custom queries, data sandbox, or CSV/Excel uploads.

**Effortless Data Cleaning:** Over 100 data transformations for thorough data preparation.

**ML In Browser :** Swift predictions, anomaly detection, and segmentation for faster decision-making. (unsupervised Learning)

**Attractive Themes:** Generate captivating visual themes for effective storytelling.

**Versatile Export Options:** Tailor exports for easy accessibility and management.

**Global Filters:** Focus on specific datasets or parameters for clear analysis and data integrity.



**Drill-Down/Drill-Through:** Uncover patterns and outliers for deeper data exploration.

**Interactive Elements:** Enable targeted filtering and interactions for richer insights.

**AI-Powered Search:** Instant visual chart generation using natural language processing.



# BDB SURVEY

BDB Survey module streamlines survey creation, customization, and distribution for targeted data collection and analysis.

**Efficiency:** BDB Survey expedites survey creation, distribution, and analysis, streamlining the entire process for users.

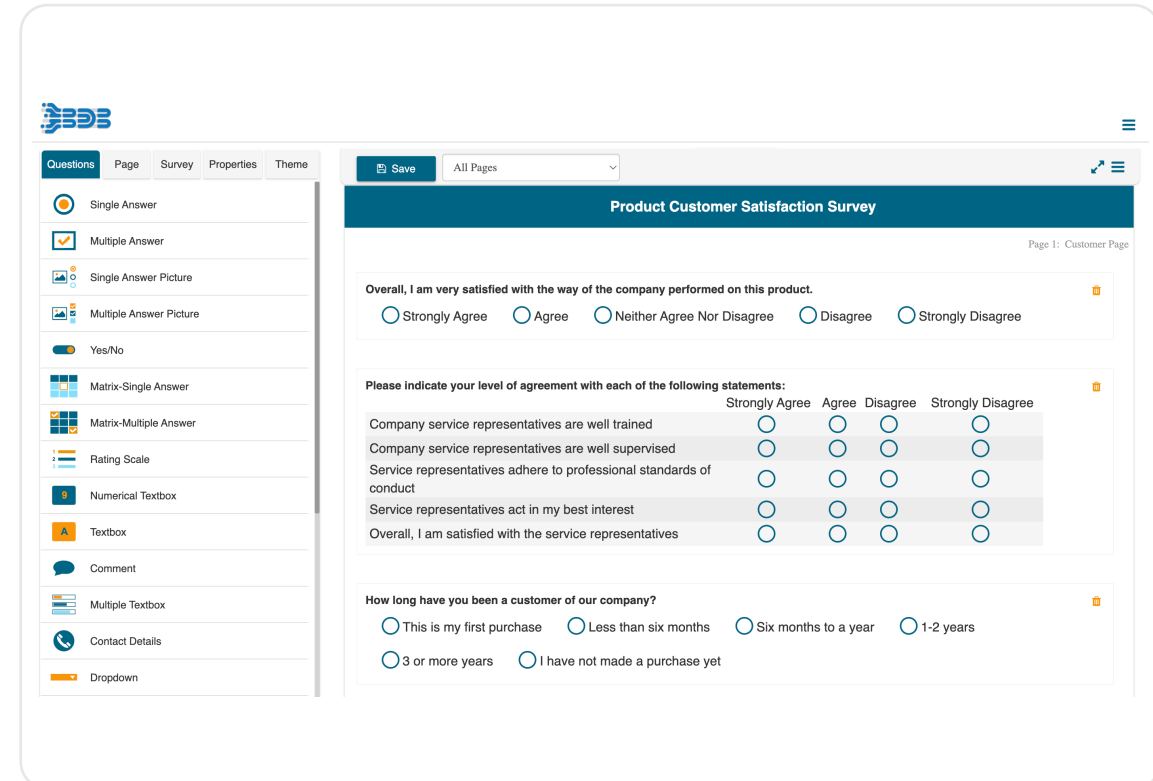
**Accessibility:** Respondents can access and complete surveys using various devices, such as computers, tablets, or smartphones, ensuring ease of participation.

**Versatile Question Types:** The module accommodates various question formats, including multiple-choice, rating scales, open-ended, and multimedia-based questions, enhancing survey flexibility.

**Multimedia Integration:** Users can include images, videos, and audio in surveys, enabling a richer and more engaging respondent experience.

**Targeted Distribution:** Surveys can be precisely directed to specific groups, allowing for more targeted data collection.

**Comprehensive Analysis:** The collected responses are amenable to thorough analysis, providing valuable insights and actionable data.



The screenshot displays the BDB Survey interface. On the left, a sidebar lists various question types: Single Answer, Multiple Answer, Single Answer Picture, Multiple Answer Picture, Yes/No, Matrix-Single Answer, Matrix-Multiple Answer, Rating Scale, Numerical Textbox, Textbox, Comment, Multiple Textbox, Contact Details, and Dropdown. The main area shows a survey titled "Product Customer Satisfaction Survey" with the following content:

**Overall, I am very satisfied with the way of the company performed on this product.**

Strongly Agree  Agree  Neither Agree Nor Disagree  Disagree  Strongly Disagree

**Please indicate your level of agreement with each of the following statements:**

|   | Strongly Agree        | Agree                 | Disagree              | Strongly Disagree     |
|---|-----------------------|-----------------------|-----------------------|-----------------------|
| Company service representatives are well trained                    | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Company service representatives are well supervised                 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Service representatives adhere to professional standards of conduct | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Service representatives act in my best interest                     | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Overall, I am satisfied with the service representatives            | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

**How long have you been a customer of our company?**

This is my first purchase  Less than six months  Six months to a year  1-2 years  3 or more years  I have not made a purchase yet

# BDB FORMS

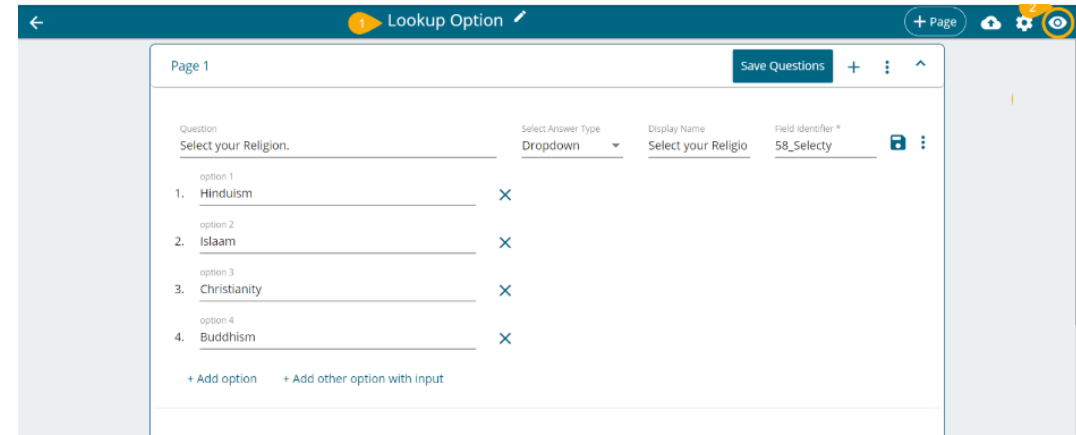
A "form" in the BDB Platform is a digital interface enabling users to create structured data collection systems, customize their appearance, settings, and oversee collected user responses.

**Creating a New Form:** The platform provides a user-friendly interface that allows users to design, customize, and manage forms for various purposes such as data collection, surveys, feedback gathering, or any structured information acquisition needed for business or organizational purposes.

**Tabular Response View:** Responses collected are presented in a tabular format, accessible to users. Filtering options include Page, Date, or Users, enhancing data organization and retrieval.

**Response Accessibility:** Ordinary users view only their responses, while admin/owners access all responses. Enables verification of data through the Form Response page.

**User-specific Instructions:** Users can add specific instructions, visible only to selected users during the preview.



The screenshot shows the BDB Forms editor interface for a 'Lookup Option' form. The form is titled 'Page 1' and has a 'Save Questions' button. The question is 'Select your Religion.' and the answer type is 'Dropdown'. The display name is 'Select your Religio' and the field identifier is '58\_Selecty'. There are four options listed: 1. Hinduism, 2. Islaam, 3. Christianity, and 4. Buddhism. Each option has a delete icon (X). There are also buttons for '+ Add option' and '+ Add other option with input'.

**Theme Customization:** Users can choose themes for Form Header, Page Header, and Form Body for aesthetic appeal.

**Publishing Forms:** Users can publish created forms on the BDB Platform's homepage, accessible under My Documents.

# BDB TRAINING AND BUILDING ECOSYSTEM

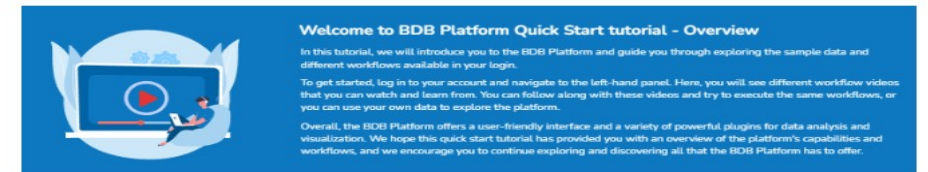
Here are workflows covering various plugins within the BDB platform:

**Data Pipeline:** The BDB platform's Data Pipeline efficiently manages data by identifying sources, establishing connections, and guiding data through extraction, transformation, and loading stages. It streamlines transformation with filtering and aggregation options and boosts capabilities by integrating DS Lab scripts. Finally, refined data gets loaded into chosen storage for further analysis.

**DS Lab:** Create/import analysis notebooks, develop scripts for data tasks, integrate datasets, implement ML models, and export finalized scripts/models for data pipeline processing.

**Dashboard Designer:** Access Dashboard Designer, select data sources, pick visualization elements, integrate processed data, configure connections, design layout and KPIs, customize visuals, and share/publish within BDB Platform.

**Report :** Access structured data sources, create data store/meta data, define report structure/content, integrate data for analysis, add visualizations, and include findings, trends, and recommendations. Export reports in PDF, Excel formats, and manage reports within the BDB Platform.



The first step is to download sample data in order to execute the workflows (1 to 5) listed below : [Click Here](#)

After downloading the data, upload it to the data sandbox with the name "HiringData", and follow the steps given in each workflow. To execute Workflow 6, begin by uploading the 'Training Data' to the data sandbox and creating an AutoML Experiment. Next, utilize 'dataset 1' and 'dataset 2' for future processing, and follow the steps provided in Workflow 6.

Here are the download links for the sample data files:

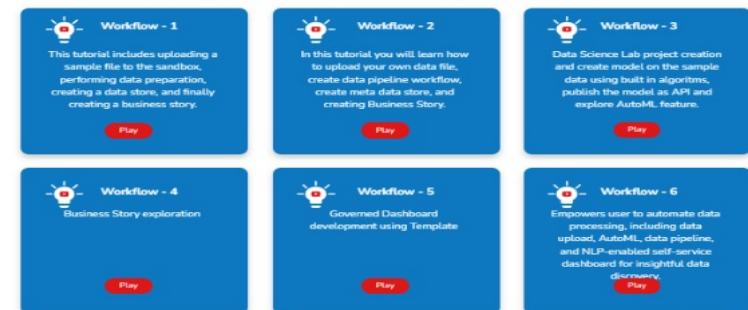
Training data - upload this data to sandbox, and use it for the Automl model creation : [Click Here](#)

Data Set 1 : [Click Here](#)

Data Set 2 : [Click Here](#)

You can download the guide by clicking on the link : [Click Here](#)

**Note:** To optimize resource consumption, it's essential to deactivate & unregister any active projects, pipelines, or models when you're finished with them. This step releases resources for other tasks and helps prevent potential issues.





# BDB TRAINING AND BUILDING ECOSYSTEM

BDB Platform operates across various domains catering to distinct roles:

**Business Intelligence:** Providing Dashboard and Report Plugins and functionalities for data visualization, reporting, and analytics to derive insights for business decision-making.

**Restaurant:** Analyzing sales data, customer preferences for menu optimization.

**Retail Store:** Utilizing sales data, inventory levels for stock management, and sales forecasting.

**E-Commerce:** Analyzing user behavior for personalized recommendations and marketing strategies.

**Hospital:** Tracking patient data, resource utilization for better operational efficiency.

**Manufacturing:** Monitoring production data, optimizing processes for increased efficiency.

**Education:** Analyzing student performance data, planning and improving educational programs.

**Banking/Insurance:** Analyzing risk, customer behavior for tailored financial products, and risk assessment.

**Media/Digital Marketing:** Analyzing audience data for targeted marketing campaigns and content personalization.

**Sales Analytics:** Utilizing CRM data for customer insights, sales forecasting, and pipeline analysis.

**Data Engineer:** Facilitating data handling, management, transformation, and pipeline creation to ensure efficient and streamlined data processing for various purposes.

**Retail Store/E-Commerce:** Managing databases for transactional data, ensuring data pipelines for inventory, and sales data.

**Hospital/Healthcare:** Designing systems to handle patient data securely and efficiently.

**Manufacturing/Agriculture:** Creating data infrastructure for sensor data, supply chain information.

**Telecom/Media:** Managing large volumes of customer interaction data and optimizing data flow.

**Security/Banking:** Building robust data systems for secure storage and access control.

**Data Scientist:** Offering resources and DS LAB capabilities for data exploration, model building, machine learning, and advanced analytics to uncover patterns and predictive insights from data.

**E-Commerce/Retail:** Building recommendation engines, predicting customer behavior.

**Healthcare/Life Science:** Analyzing clinical trial data, genomic research for disease prediction.

**Banking/Insurance:** Building risk models, fraud detection algorithms.

**Telecom/Digital Marketing:** Predictive analytics for customer churn, campaign optimization.

**Manufacturing/Agriculture:** Predictive maintenance models, yield optimization.

# BDB TRAINING AND BUILDING ECOSYSTEM

BDB Platform operates across various domains catering to distinct roles:

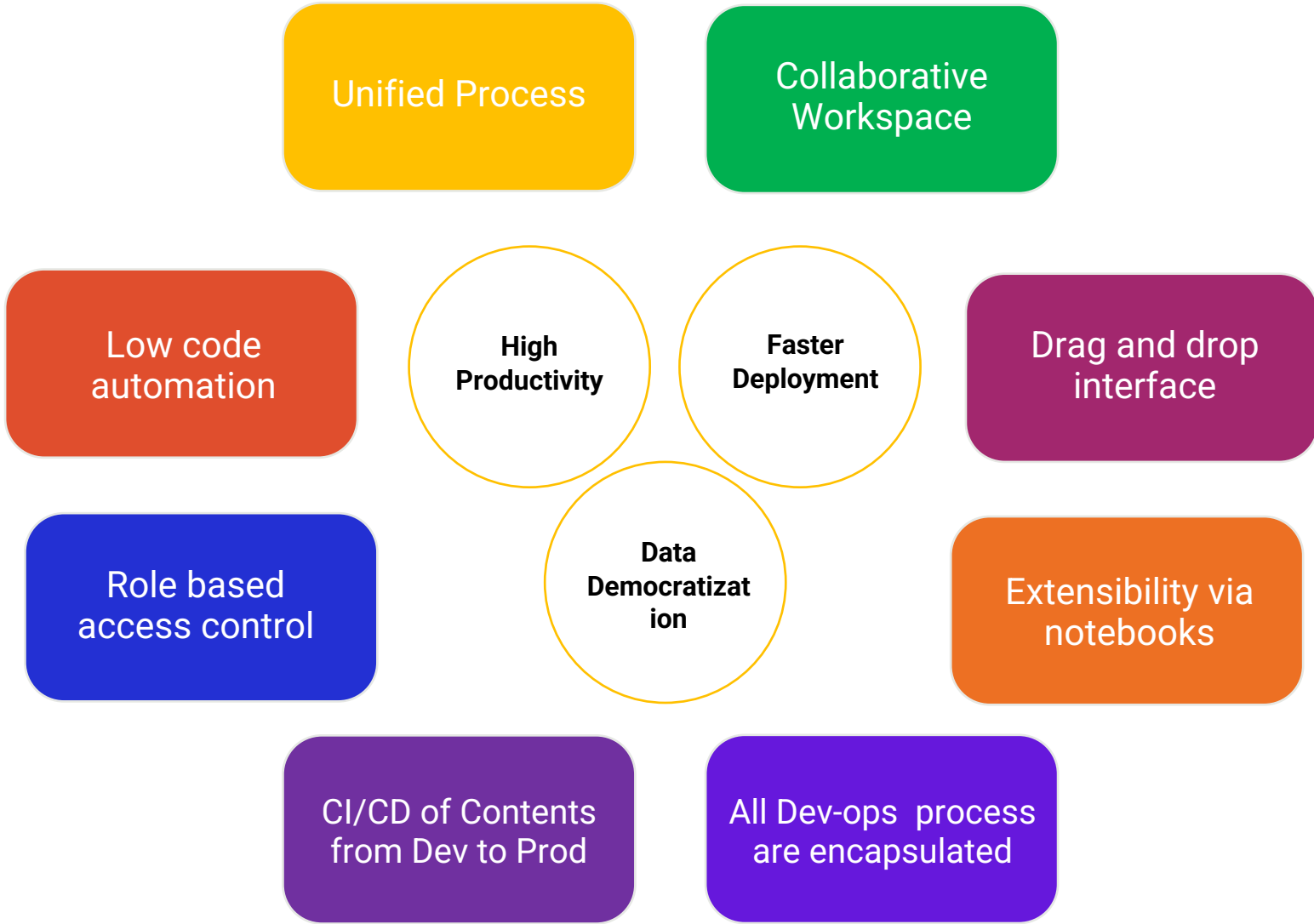
1. **Restaurant**
2. **Retail Store**
3. **E-Commerce**
4. **Hospital**
5. **Manufacturing**
6. **Life Science**
7. **Education**
8. **Automobile**
9. **Telecom**
10. **Agriculture**
11. **Insurance**
12. **Banking**
13. **Media**
14. **Security**
15. **Digital marketing**
16. **Sales Analytics –  
CRM, Customer 360, etc.**

**Business Intelligence (BI)** offers tailored Dashboard and Report Plugins across diverse industries, enabling efficient data visualization, reporting, and analytics. It provides insights into sales trends (restaurant, retail), customer behavior (e-commerce, digital marketing), patient data analysis (hospital, life sciences), production metrics (manufacturing, agriculture), and strategic decision-making (insurance, banking, media, security, telecom, sales analytics). Visualization Plugins serve as a crucial resource for informed decision-making and operational efficiency across these sectors.

**Data Engineers** orchestrate efficient data handling, transformation, and pipeline creation across various industries. They ensure smooth data flow for inventory management (restaurant, retail), customer insights (e-commerce, digital marketing), patient records (hospital), process optimization (manufacturing, agriculture), risk assessment (insurance, banking), audience analytics (media, telecom), and sales optimization (sales analytics). Their role spans tailored data management to support specific industry needs and streamlined operations.

**Data Scientists** leverage DS LAB capabilities to extract insights and build predictive models across diverse industries. In sectors like retail, e-commerce, healthcare, manufacturing, and finance, they uncover patterns and conduct predictive analytics tailored to industry-specific needs. Their role spans from customer behavior analysis and personalized recommendations to predictive maintenance and risk assessment, utilizing advanced analytics to drive informed decision-making.

# Unified Data Platform (BDB) advantages





+

○

●

# THANK YOU

For more information, contact:  
[Sales@bdb.ai](mailto:Sales@bdb.ai)