Project Design Phase-II
Technology Stack (Architecture & Stack)

Date	03 October 2022
Team ID	PNT2022TMID48417
Project Name	Project – Global Sales Data Analytics
Maximum Marks	4 Marks

Technical Architecture:

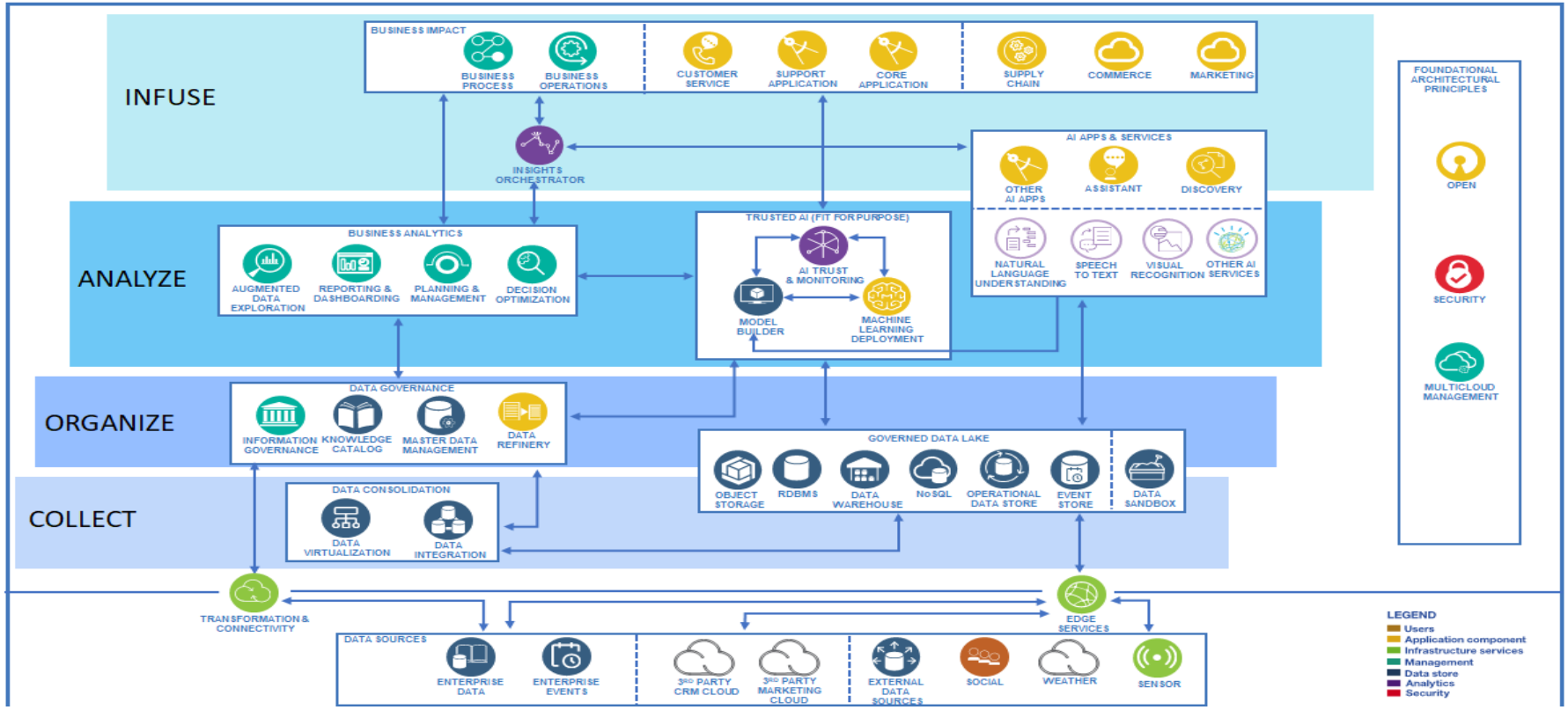

Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	How user interacts with application e.g. Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript / Angular Js / React Js etc.
2.	Knowledge catalog	The knowledge catalog helps you find, understand, and use needed data. It helps users to discover, curate, categorize, and share data assets, data sets, analytical models, and their relationships with other members of an organization. The catalog serves as a single source of truth.	**Products** IBM Watson® Knowledge Catalog
3.	• Master data management	Master data management is a method that is used to define and manage the critical data of an organization to provide, with data integration, a single point of reference. The data that is mastered can include reference data: the set of permissible values and the analytical data that supports decision-making.	**Products** Master Data Management
4.	Data refinery	Data refinery is a data-preparation capability in support of self-service analytics. It can be used for the quick transformation of large amounts of raw data into consumable, quality information that is ready for analytics.	**Products** Data Refinery, available by using IBM Watson® Studio and Watson Knowledge Catalog (Pro) IBM InfoSphere® Advanced Data Preparation
5.	Data RDBMS	A relational database management system (RDBMS) is a database that stores and processes data in structured, tabular format as a	**Products** IBM Db2® Family IBM Db2® on Cloud

		collection of tables that consist of columns and rows, with relational operators to query data through Structured Query Language (SQL).	IBM Db2® Hosted IBM Db2® for z/OS
6.	Data warehouse	A data warehouse is a consolidated repository of integrated, conformed, and aggregated data from multiple and disparate data sources in support of business analytics and reporting. Data warehouses typically process structured data in tabular or relational form, often with history, on scalable relational database technology platforms that support large numbers of concurrent users and complex queries across large data sets.	**Products** IBM Integrated Analytics System IBM Db2® Warehouse on Cloud IBM Db2® Warehouse IBM Cloud Pak® for Data System with IBM Performance Server for PostgresSQL
7.	NoSQL	NoSQL is a database technology that provides the storage and retrieval of data that is modeled in a means other than the tabular relations of a relational database management system (RDBMS). NoSQL databases can support SQL-like query languages ("Not only SQL") and often support data in structured, semi-structured, and unstructured form. Most NoSQL databases implement a model of eventual consistency compared to the transactional consistency that is offered by relational database technologies.	**Products** IBM Cloudant® IBM Cloud® Databases for MongoDB
8.	Data virtualizaton	Data virtualization is technology that connects all these data sources into a single self-balancing constellation. No longer are analytics queries performed on data copied and stored to a centralized location. The analytics application submits a query that is processed on the device where the data source is persisted. The results of the query are consolidated within the constellation and returned to the origin application. No data is copied. It remains persisted only at the source.	**Products** Data virtualization

9.	Data integration	Data integration copies and correlates information from disparate sources. AI technology can be used to semi-automate the process.	**Products** Data integration
10.	Edge services	Edge services provide network capability to deliver content through the internet (DNS, CDN, firewall, load balancer). They handle the request and get it to the right destination. When the web application server completes its tasks, it delivers the resulting content back through the firewall, which passes the content to the user's browser	**Products** IBM Cloud® Internet Services .
11. Security	Security	Security enables identity and access management, and data and application protection. It provides actionable security intelligence across cloud and enterprise environments.	**Products** Enterprise-specific

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Numpy,pandas,matplaotlip,seaborn.	**NumPy is an open source project** aiming to enable numerical computing with Python. pandas is a fast, powerful, flexible and easy to use **open source** data analysis and manipulation tool, built on top of the Python programming language.
2.	Security Implementations	Reliable and innovative security	e.g. SHA-256, Encryptions, IAM Controls, OWASP etc. IBM Cloud® offers superior cloud security with end-to-end capabilities and customizable solutions to help manage your data, all backed by expert support.
3.	Scalable Architecture	The **analytics** nodes collect and store **data** and provide this **data** through various REST API queries. **Scalability** is provided for the control nodes, ...	Horizondal scaling,Vertical scaling.
4.	Availability	Available 24/7/365	Achieve geographic redundancy. ...

			Implement strategic redundancy. ... Leverage failover solutions. ... Implement network load balancing.
5.	Performance	Maintain upto date tset infrastructure Use your live site Track client and server side traffic Pay attention to time consumption Distribution	Apache JMeter Load ninja Web load Load complete.