In the current data-driven environment, SBI must evolve its data strategy to remain competitive in the marketplace. In this regard we understand that the Bank is planning to modernize its technology infrastructure to support better decision making and customer experiences. Today, SBI is well positioned to unlock value from its data from various internal and external sources that will continue to grow and explode – this value should translate into achieving business outcomes of acquiring New to bank (NTB) Customers, strengthen and expand relationship with Existing to Bank (ETB) Customers, improve its risk and credit models to reduce Cost to bank and meet most of the regulatory requirements. Key constraints in achieving these outcomes lie in limitations of current platform in terms of scalability, concurrency, performance, agility and innovative use of AI / Generative AI Capabilities.

As trusted partner to SBI and sharing the experience of Data Warehouse and Data Lake program, we bring experience and learnings from dozens of large-scale data platform implementations to meet SBI's aspirations to become a data-driven organization. We present the architecture evolution in current data platform that will not only address current requirements but also enable Bank to be future ready by adopting the open lake house architecture.

While architecture plays an important role in Data Platform, more important aspects of Data Platform lie in its operating model that enables Self Service consumption and production of data as data products. Data security and governance are also areas which are important to prevent data breaches and ensure compliance with regulations.

IBM is pleased to present an approach that is pragmatic, drawn from our experience with our global customers across various industry verticals and our understanding of SBI objectives. Holistic security aspects should be taken into consideration on the proposed platform in terms of encryption and masking, Secure channel and comply to Bank's cyber security guidelines.

Based on learnings from current data platform the below requirements are planned to be addressed through the proposed architecture:

- Decoupling of Storage from Compute
- Decoupling of User-zone from Conformed zone
- Automated workload management and governance
- Data Governance with automated data quality checks
- Need for "Data as Product" approach to avoid duplication
- Oracle compatibility support for CDC
- Empower Business to perform data value exploration as self service
- Automate code review and deployment for continuous and faster code release
- Support open data, commodity hardware, no proprietary lock-in
- DR to account for two-way replication and 100% capacity

The principles for proposed architecture (outlined in Section 2.1) aim to address the above requirements. Designing the right data strategy and architecture is crucial for an organization to succeed in this age of data-driven decision making. SBI should move / evolve to an architecture which is future-fit to achieve its strategic goals and thrive in the marketplace. A Lakehouse architecture with cost-effective object storage, open data and table formats, metadata governance, and powerful query engines, is best suited as a cost-effective solution for evolving data growth and AI proliferation. IBM being leader in Data and Analytics Solutions and Services is well positioned to be the partner to evolve to the new proposed architecture.