*** KGiSL INSTITUTE OF TECHNOLOGY***

***NAAN MUDHALVAN***

***PROJECT TITLE:***

*Data Warehousing with IBM Cloud Db2 Warehouse.*

***TEAM MEMBERS:***

1. *Keerthana. J*
2. *Saruthi. T*
3. *Kishore. K*
4. *Jeevanandham. S*

***PROJECT DESCRIPTION:***

*Building a Scalable Data Warehouse with IBM Cloud Db2 Warehouse.*

***OBJECTIVE:***

*The primary objective of this project is to design, implement, and leverage IBM Cloud Db2 Warehouse for the purpose of building a scalable and efficient data warehousing solution that enables data-driven decision-making and advanced analytics.*

***PROJECT PHASES:***

1. **Project Initiation:**
   * Define project objectives, scope, and success criteria.
   * Establish a project team with roles and responsibilities.
   * Create a project plan and timeline.
2. **Requirements Gathering:**
   * Collaborate with business stakeholders to identify data sources and analytics requirements.
   * Determine the volume, velocity, and variety of data to be stored and analyzed.
   * Define key performance indicators (KPIs) and reporting requirements.
3. **IBM Cloud Setup:**
   * Sign up for an IBM Cloud account if not already done.
   * Provision the IBM Cloud Db2 Warehouse service in the desired region.
   * Configure security settings, including authentication and access controls.
4. **Data Ingestion and Integration:**
   * Identify and collect data from various sources, such as databases, files, and APIs.
   * Implement ETL (Extract, Transform, Load) processes to clean, transform, and load data into Db2 Warehouse.
   * Ensure data quality and consistency.
5. **Data Modeling:**
   * Design a dimensional data model (e.g., star or snowflake schema) that supports the analytics and reporting requirements.
   * Define data structures, tables, and relationships in Db2 Warehouse.
   * Optimize data structures for query performance.
6. **Analytics and Reporting:**
   * Develop SQL queries and analytics scripts to extract insights from the data.
   * Use reporting and visualization tools (e.g., IBM Cognos, Tableau) to create dashboards and reports.
   * Implement advanced analytics and machine learning models if required.
7. **Performance Optimization:**
   * Monitor query performance and optimize SQL queries for efficiency.
   * Adjust resource allocation (CPU, memory, storage) as needed to handle increasing workloads.
   * Implement data compression and indexing strategies for improved performance.
8. **Data Security and Compliance:**
   * Implement security measures to protect sensitive data.
   * Ensure compliance with data privacy regulations (e.g., GDPR, HIPAA).
   * Conduct regular security audits and assessments.
9. **Backup and Disaster Recovery:**
   * Set up automated backup and restore procedures to safeguard against data loss.
   * Create a disaster recovery plan to minimize downtime in case of system failures.
10. **Documentation and Training:**
    * Document the architecture, data model, and procedures.
    * Provide training to the project team and end-users on using the data warehouse.
11. **Monitoring and Maintenance:**
    * Continuously monitor system performance and resource utilization.
    * Perform routine maintenance tasks, including software updates and patches.
    * Address any issues or scalability challenges that arise.
12. **Project Closure and Evaluation:**
    * Review project objectives and success criteria.
    * Conduct a post-implementation evaluation to assess the impact on business operations and analytics capabilities.
    * Document lessons learned and recommendations for future enhancements.