PROBLEM DEFINITION AND DESIGN THINKING-PHASE 1

Problem Definition:

The project involves delving into big data analysis using IBM Cloud Databases. The objective is to extract valuable insights from extensive datasets, ranging from climate trends to social patterns. The project includes designing the analysis process, setting up IBM Cloud Databases, performing data analysis, and visualizing the results for business intelligence.

Proposed Solution:

To successfully address this project, we recommend the following steps and strategies:

1. Data Selection:

- ✓ Collaborate with domain experts to identify and select relevant datasets based on project objectives.
- ✓ Ensure that data sources are accessible and adhere to data privacy and security regulations.

2. Database Setup:

- ✓ Set up IBM Cloud Databases, selecting appropriate database types (e.g., relational, NoSQL) based on data characteristics.
- ✓ Define data schemas and implement data ingestion pipelines for regular updates and maintenance.

3. Data Exploration:

- ✓ Develop data exploration scripts and queries to gain a deep understanding of the data.
- ✓ Identify and address data quality issues, including missing values and outliers through data preprocessing.

4. Analysis Techniques:

- ✓ Choose appropriate analysis techniques, such as statistical analysis, machine learning, or a combination of both, based on the dataset's nature and project goals.
- ✓ Implement the chosen techniques using relevant tools and libraries.

5. Visualization:

- ✓ Utilize data visualization tools such as Tableau, Power BI, or custom Python libraries (e.g., Matplotlib, Seaborn) to create impactful visualizations.
- ✓ Design interactive dashboards for real-time monitoring and analysis.

6. Business Insights:

- ✓ Collaborate closely with stakeholders to comprehend the business context and project objectives.
- ✓ Interpret analysis results to extract actionable insights.
- ✓ Provide clear and concise recommendations aligned with the project's goals.

Timeline:

We anticipate completing this project in several phases:

- ➢ Phase 1 (2 weeks): Data selection, database setup, and initial data exploration.
- ➢ Phase 2 (4 weeks): In-depth data analysis and visualization.
- ➢ Phase 3 (2 weeks): Deriving actionable business insights, preparing comprehensive reports, and presenting findings.

Resource requirements:

- ➢ Subscription to IBM Cloud Databases.
- ➢ A team of data scientists, analysts, and domain experts.
- ➢ Access to data visualization tools and software.
- ➢ Collaboration with relevant stakeholders for domain knowledge and context.

Conclusion:

This project aims to leverage IBM Cloud Databases to extract valuable insights from extensive datasets, providing the organization with actionable business intelligence. The proposed solution includes a structured approach involving data selection, database setup, data exploration, analysis techniques, visualization, and the derivation of business insights. This approach is designed to meet the project's objectives systematically.