ASSIGNMENT 3: TAILORING THE ANALYTICS LIFECYCLE TO DOMAIN-SPECIFIC REALITIES

Chapter 3: Big Data Adoption and Planning Considerations

Instructor: Houssein Dhayne (houssein.dhayne@net.usj.edu.lb)

ASSIGNMENT OVERVIEW:

This assignment seeks to apply the principles discussed in Chapter 3 to practical scenarios within specific domains. Students are tasked with adapting the Big Data analytics lifecycle to the unique challenges and opportunities presented by their chosen area of study.

Task 1: Identify Domain-Specific Challenges:

Investigate and outline unique challenges and considerations related to Big Data adoption within your chosen domain. Consider aspects such as data sources, privacy concerns, security requirements, and specific performance challenges.

Task 2: Develop a Case Study:

Construct a hypothetical case study illustrating the application of the adapted Big Data analytics lifecycle in a real-world scenario within your chosen domain. Provide details about the nature of the data, challenges faced, and the impact of the analysis on decision-making.

Task 3: Utilization of Analysis Results:

Explore potential use cases for the analysis results within your domain. Discuss how insights gained from Big Data analytics can be leveraged to improve processes, optimize systems, or enhance decision-making.

The Big Data Analytics Lifecycle:

The Big Data analytics lifecycle, introduced in Chapter 3, consists of nine stages:

1.   Business Case Evaluation: Define the justification, motivation, and goals of the analysis, ensuring alignment with business requirements and Big Data characteristics.

2.   Data Identification: Identify internal and external datasets relevant to the analysis project, considering a wide variety of data sources.

3.   Data Acquisition & Filtering: Gather data from identified sources, automatically filtering out corrupt or irrelevant data.

4.   Data Extraction: Extract and transform disparate data into a format compatible with the underlying Big Data solution.

5.   Data Validation & Cleansing: Establish complex validation rules to ensure data accuracy, cleansing known invalid data.

6.   Data Aggregation & Representation: Integrate multiple datasets, considering differences in structure and semantics, and standardize data for reusability.

7.   Data Analysis: Conduct analysis tasks, exploring patterns and correlations iteratively through confirmatory or exploratory analysis approaches.

8.   Data Visualization: Use visualization techniques to communicate analysis results effectively, facilitating interpretation by business users.

9.   Utilization of Analysis Results: Determine how processed data can be further leveraged, considering input for enterprise systems, business process optimization, and alerts.

SUBMISSION GUIDELINES:

Prepare a report summarizing your findings for each task.

Include relevant examples, illustrations, or diagrams to support your explanations.

Provide proper citations for any external sources used.

Submit your assignment in a format suitable for academic presentation (e.g., a well-organized document).

GRADING CRITERIA:

Demonstration of understanding of Big Data concepts in the specific domain.

Depth and relevance of examples provided.

Clarity and organization of the report.

Integration of case study and real-world applications.

Thoughtful reflection on the interdisciplinary nature of Big Data analysis.