**MODULE 2**

**Student-Centric Activity:**

**Activity Plan**

Analytics starts with data. Identifying, accessing, obtaining, and processing of relevant data are the most essential tasks in any analytics study. You are tasked to find a large enough real-world data (either from your own organization, which is the most preferred, or from the Internet that can start with a simple search, or from the data links posted on KDnuggets.com), one that has tens of thousands of rows and more than 20 variables to go through and document a thorough data preprocessing project. In your processing of the data, identify anomalies and discrepancies using descriptive statistics methods and measures, and make the data analytics ready.

List and justify your preprocessing steps and decisions in a comprehensive report.

Explain about the dataset you have selected.

Perform STATISTICAL MODELING also.