**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 12 JULY 2024 |
| ID | 740069 |
| Project Title | Lymphography Classification using ML |
| Maximum Marks | 6 Marks |

**Data Exploration and Preprocessing Report**

|  |  |
| --- | --- |
| **Section** | **Description** |
| Data Overview | Dimension:  148 rows × 19 columns  Descriptive statistics: |
|  |  |

Dataset variables will be statistically analyzed to identify patterns and outliers, with Python employed for preprocessing tasks like normalization and feature engineering. Data cleaning will address missing values and outliers, ensuring quality for subsequent analysis and modeling, and forming a strong foundation for insights and predictions.

|  |  |
| --- | --- |
| Univariate Analysis |  |
| Bivariate Analysis |  |
| Multivariate Analysis |  |

|  |  |
| --- | --- |
| Outliers and Anomalies | - |
| **Data Preprocessing Code Screenshots** | |
| Loading Data |  |
| Handling Outliers |  |
| Training and Testing |  |
| Feature Engineering | Attached the codes in final submission. |
| Save Processed Data | - |