**Data Collection and Preprocessing Phase**

| Date | 15 March 2024 |
| --- | --- |
| Team ID | 740009 |
| Project Title | Student Adaptability Level of Online Education |
| Maximum Marks | 6 Marks |

**Data Exploration and Preprocessing Template**

Identifies data sources, assesses quality issues like missing values and duplicates, and implements resolution plans to ensure accurate and reliable analysis.

| **Section** | **Description** |
| --- | --- |
| Data Overview | Provide an overview of the data used to assess student adaptability, including sources like LMS logs, assessment scores, and survey data. Include basic statistics, dimensions, and structure of the data. |
| Univariate Analysis | Explore individual variables that indicate student adaptability, such as average time spent on the platform, quiz scores, participation in discussions, etc. Analyze their mean, median, mode, and distribution. |
| Bivariate Analysis | Investigate relationships between two variables, such as the correlation between time spent on the platform and quiz scores, or between participation in discussions and overall performance. Use correlation matrices, scatter plots, and other visualizations. |
| Multivariate Analysis | Identify patterns and relationships involving multiple variables to understand the factors contributing to student adaptability. Use techniques like PCA (Principal Component Analysis) or cluster analysis to uncover hidden trends. |
| Outliers and Anomalies | Identify and treat outliers or anomalies in the data that could skew the results. For example, extremely high or low engagement scores that are not representative of typical student behavior. |
| **Data Preprocessing Code Screenshots** | |
| Loading Data |  |
| Handling Missing Data |  |
| Data Transformation |  |
| Feature Engineering | . |
| Save Processed Data | . |