

Data Collection and Preprocessing Phase

| | |
|---------------|-----------------------------|
| Date | 15 March 2024 |
| Team ID | 739640 |
| Project Title | Work Force Retention System |
| 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 | Basic statistics, dimensions, and structure of the data. |
| Univariate Analysis | Exploration of individual variables (mean, median, mode, etc.). |
| Bivariate Analysis | Relationships between two variables (correlation, scatter plots). |
| Multivariate Analysis | Patterns and relationships involving multiple variables. |
| Outliers and Anomalies | Identification and treatment of outliers. |
| Data Preprocessing Code Screenshots | |
| Loading Data | Code to load the dataset into the preferred environment (e.g., Python, R). |

| | |
|-----------------------|-------------------------------------------------------------|
| Handling Missing Data | Code for identifying and handling missing values. |
| Data Transformation | Code for transforming variables (scaling, normalization). |
| Feature Engineering | Code for creating new features or modifying existing ones. |
| Save Processed Data | Code to save the cleaned and processed data for future use. |