

Data Collection and Preprocessing Phase

| | |
|---------------|--|
| Date | 15 th July 2024 |
| Team ID | 739823 |
| Project Title | Auto Foresight : A Predictive Model for Streamlining Car Loan Repayment Planning |
| Maximum Marks | 2 Marks |

Data Collection Plan & Raw Data Sources Identification Template

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Data Collection Plan Template

| Section | Description |
|-----------------------------|---|
| Project Overview | To develop an AI-powered system for predicting vehicle loan repayment outcomes to enable proactive measures in managing loan defaults and optimizing repayment plans. |
| Data Collection Plan | <ul style="list-style-type: none"> Search for datasets related to vehicle loan repayment, financial information, and applicant details. Prioritize datasets with diverse demographic information. |
| Raw Data Sources Identified | The raw data source for this project include dataset obtained from Kaggle, the popular platform for data science competitions and repositories. |

Raw Data Sources Template

| Source Name | Description | Location/URL | Format | Size | Access Permissions |
|----------------|---|---|--------|------|----------------------|
| Kaggle Dataset | Description of the data in this source. | https://www.kaggle.com/datasets/meastanmay/nbfi-vehicle-loan-repayment-dataset?select=Train_Dataset.csv | CSV | 15MB | Public |
| Dataset2 | Description of the data in this source. | LinkofDataset2 | Excel | 14MB | Private(with access) |
| ... | ... | ... | ... | ... | ... |