**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 4th July 2024 |
| Team ID | 739808ss |
| Project Title | Cost Prediction of Acquiring a Customer. |
| 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 | Businesses looking to maximize their marketing and sales tactics in a competitive market must be able to predict the cost of acquiring a client (CAC). In this process, the amount of money needed to turn a prospective lead into a paying client is estimated. This amount includes costs for sales, marketing, and advertising. With accurate CAC prediction, businesses may improve return on investment (ROI) by making well-informed decisions about client segmentation, pricing strategies, and budget allocation. |
| Data Collection Plan | * Search for datasets related to Customer Acquisition Cost. * Prioritize datasets with diverse demographic information. |
| Raw Data Sources Identified | The raw data sources for this project include datasets obtained from Kaggle & UCI, the popular platforms for data science competitions and repositories. The provided sample data represents a subset of the collected information. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Kaggle Dataset | This dataset consists of various parameters like store city, food category, food family etc which deteremine the cost. | https://www.kaggle.com/datasets/ramjasmaurya/medias-cost-prediction-in-foodmart?select=media+prediction+and+its+cost.csv | CSV | 16 MB | Public |