**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 9 July 2024 |
| Team ID | 740030 |
| Project Title | Leveraging Machine Learning For GDP Per Capita Prediction |
| 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 | Provide a brief overview of the "GDP Prediction" project, highlighting its goals and objectives related to leveraging machine learning for precise GDP Prediction. |
| Data Collection Plan | **Source**: Extract from database in the form of CSV files. |
| Raw Data Sources Identified | **Countries Dataset:** Focuses on various countries and their economic data.   **Goal:** Aims to predict future Gross Domestic Product (GDP) values.   **Features:** Likely includes historical GDP figures, along with potentially relevant economic indicators.   **Target Variable:** GDP (possibly for specific years) you want to predict.   **Use Case:** This data can be used for training machine learning models to forecast future economic performance of different countries. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Dataset 1 | Description of the data in this source. | Countries of world.csv: <https://drive.google.com/file/d/15qezCZPDSw6HMNbf7iisxXJe2Cl9XgY4/view?usp=drive_link> | CSV | 37 KB | Public |