

Data Collection and Preprocessing Phase

| | |
|---------------|--|
| Date | 2 july 2024 |
| Team ID | SWTID1720086522 |
| Project Title | Forecasting Economic Prosperity: 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 | In this project, we predict GDP per capita using historical data and relevant features. Our goal is to inform decisions for economic prosperity. |
| Data Collection Plan | <p>National Statistical Agencies: Official government data on GDP, population, and economic indicators.</p> <p>International Organizations: Reports from the World Bank, IMF, and UN.</p> <p>Research Papers: Academic studies related to economic development.</p> <p>Private Databases: Commercial databases or APIs.</p> |

| | |
|-----------------------------|---|
| Raw Data Sources Identified | <p>IMF World Economic Outlook: Provides macroeconomic data series from 1980 to the present.</p> <p>World Bank Data: Includes GDP per capita in current US dollars.</p> <p>Trading Economics: Offers economic indicators and historical data.</p> |
|-----------------------------|---|

Raw Data Sources Template

| Source Name | Description | Location/URL | Format | Size | Access Permissions |
|------------------------|--|---|--------|------|--------------------|
| Countries of the world | Information on population, region, area size, infant mortality and more. | https://www.kaggle.com/datasets/fernando/countries-of-the-world | CSV | 38kb | Public |