**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 15 March 2024 |
| Team ID | SWTID1720114403 |
| Project Title | Vitamin Vision: Unveiling the Spectrum of Nutrient Detection |
| 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 | The "Vitamin vision" project focuses on utilizing advanced analytical techniques and technologies to accurately identify and quantify the presence of various vitamins in samples such as food, supplements, and biological fluids. By leveraging methods such as chromatography, spectroscopy, and machine learning, the project aims to ensure the quality and safety of products, verify nutritional content, and support research in health and wellness. |
| Data Collection Plan | Dataset must consist of images of different food items with labels attached to it. Dataset can preferably be organized into sub-directories depending on the label (for the ease of data pre-processing). |
| Raw Data Sources Identified | The raw data sources for this project include datasets obtained from  Kaggle, which is a popular data-science platform conducting competitions and providing various repositories. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Kaggle Dataset | The dataset has been divided into two main folders-:  (a) For training and validation purposes (name) (the folder is further organized into sub-folders based on the labels of the dataset)  (b) For testing purposes (name) | https://www.kaggle.com/code/allulucky27/vitamin-detection | .jpg organized in the form of sub-directories | ≈ 92 MB | Public |