

## Data Collection and Preprocessing Phase

Date	3 July 2024
Team ID	740011
Project Title	Thyroid Classification using ML
Maximum Marks	2 Marks

### Data Collection Plan & Raw Data Sources Identification

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

Section	Description
Project Overview	The Thyroid gland is a vascular gland and one of the most important organs of a human body. This gland secretes two hormones which help in controlling the metabolism of the body. The two types of Thyroid disorders are Hyperthyroidism and Hypothyroidism. When this disorder occurs in the body, they release certain type of hormones into the body which imbalances the body's metabolism.
Data Collection Plan	Search for datasets related to Thyroid Classification.
Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle, the popular platforms for data

	science competitions and repositories. The provided sample data represents a subset of the collected information, encompassing variables.
--	---

### Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle	Description of the data in this source.	<a href="https://www.kaggle.com/code/dario-mandarino/thyroid-classification">https://www.kaggle.com/code/dario-mandarino/thyroid-classification</a>	CSV	373 KB	Public