

Data Collection and Preprocessing Phase

Date	8 July 2024
Team ID	739996
Project Title	Identification of Methodology Used In Real Estate Property Valuation
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	This project aims to leverage machine learning techniques to predict real estate property values. By analyzing historical data on property sales, characteristics, and market trends, we intend to build a robust predictive model that can offer insights into future property values and trends.
Data Collection Plan	Data will be collected from various sources, including historical property sales databases, market reports, and publicly available datasets on real estate transactions and trends. We will also use online APIs to gather real-time data on property sales and market conditions.
Raw Data Sources Identified	We have identified three key data sources: a comprehensive Property Sales Database containing historical sales data by location, property type, and transaction details; detailed Market Reports offering insights into real estate trends; and an Online Real Estate API providing real-time data on current property sales. These diverse and rich data sources will be integral in building an accurate and robust predictive model.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
UC Irvine Dataset	The dataset comprises details of age of house,distance from the nearest mrt station,latitude,longitude etc	Real Estate Valuation - UCI Machine Learning Repository	CSV	32.1kb	Public