**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 03 July 2024 |
| Team ID | 740709 |
| Project Title | House Rent Price Prediction Using Machine Learning. |
| Maximum Marks | 2 Marks |

**Data Collection Plan & Raw Data Sources Identification Template**

Creating a data collection plan &Raw Data Sources involves several key steps to ensure that the data gathered is reliable, relevant, and suitable for analysis.

**Data Collection Plan Template**

|  |  |
| --- | --- |
| **Section** | **Description** |
| Project Overview | This project aims to predict house rent prices using machine learning techniques. The objective is to build a model that can accurately estimate the rent based on various features such as location, size, number of bedrooms, and other relevant factors.. |
| Data Collection Plan | Data will be collected from various real estate websites, rental property listings, and publicly available datasets. Sources will include websites like Kaggle, Zillow, Craigslist, and government housing data portals. |
| Raw Data Sources Identified | The raw data sources will include descriptions, URLs, formats, sizes, and access permissions.  Raw data i.e., Dataset for House Rent Price Prediction is collected from Kaggle which consist city ,monthly rant, Baths ,BHKS… |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Kaggle  Dataset | A comprehensive dataset from Kaggle containing historical rental prices and property features. . | https://drive.google.com/file/d/1MeX69BpRYAMO0BvZPqRkP7R84qEuKm7u/view?usp=sharing | CSV | 3MB | Public |