**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 24 April 2024 |
| Team ID | team-739848 |
| Project Title | Identifying Airline Passenger Satisfaction Using Machine Learning |
| Maximum Marks | 2 Marks |

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

Creating a data collection plan 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 | The objective of this machine learning project is to predict and understand factors influencing airline passenger satisfaction. By leveraging historical data and applying machine learning algorithms, the aim is to develop a predictive model which will be useful to airlines. They can enhance their service offerings, improve customer retention, and ultimately drive business success through improved customer satisfaction metrics. |
| Data Collection Plan | * Collect dataset which is related Airline Passenger Satisfaction. * Collect the information regarding airlines which will be helpful for model development. |
| Raw Data Sources Identified | Raw data i.e., Dataset for Airline Passenger Satisfaction is collected from Kaggle which consist of Gender ,Type of Travel, Cleanliness,.. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Kaggle  Dataset | This dataset consist of information regarding different passengers and how they felt about airline services . | https://www.kaggle.com/datasets/teejmahal20/airline-passenger-satisfaction | CSV | 3MB | Public |