Data Collection and Preprocessing Phase

Date	05 June 2024
Team ID	739975
Project Title	To Predict Consumer Price Index
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 objective of this project is to predict the Consumer Price Index (CPI) for various categories including food items, pulses, vegetables, fruits, milk products, meat, fish, snacks, sweets, beverages, and housing. Accurate CPI predictions will help understand inflation trends and assist in economic decision-making and budget planning. The project will use historical price data, economic indicators, and other relevant data to build predictive models using machine learning techniques.
Data Collection Plan	Data will be collected from various sources: - Government Statistics Bureaus (e.g., Bureau of Labor Statistics) - Market Research Firms (e.g., Nielsen, Euromonitor International) - Online Marketplaces (e.g., Amazon, Walmart) - Economic Data Sources (e.g., World Bank, IMF) - Surveys and Reports (e.g., Consumer Expenditure Survey)
Raw Data Sources Identified	 - **Historical Retail Price Data** - **Source:** Bureau of Labor Statistics (BLS) - **Description:** Monthly prices of various consumer goods and services over

the past decade.

 - **Market Research Reports**
 - **Source:** Nielsen, Euromonitor International
 - **Description:** Detailed reports on market trends, price dynamics, and consumer behavior.

 - **Online Marketplace Data**
 - **Source:** Amazon, Walmart
 - **Description:** Daily or weekly price data scraped from online retailer websites to get the latest pricing information.

 - **Economic Indicators**
 - **Source:** World Bank, IMF
 - **Description:** Data on inflation rates, GDP growth, currency exchange rates, and other macroeconomic indicators.

 - **Consumer Surveys**
 - **Source:** Consumer Expenditure Survey
 - **Description:** Data on household spending patterns, income, and demographics.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	The dataset comprises of patient details (rural&urban)	https://www.kaggle.com/datasets/fiq423ubf/all-india-consumer-price-index-ruralurban/data	CSV	20kB	Public