MILESTONE 1

**Task 1: Nutrient Recommendation System for Personalized Diet Planning, Gather user preferences and dietary goals to customize recommendations.**

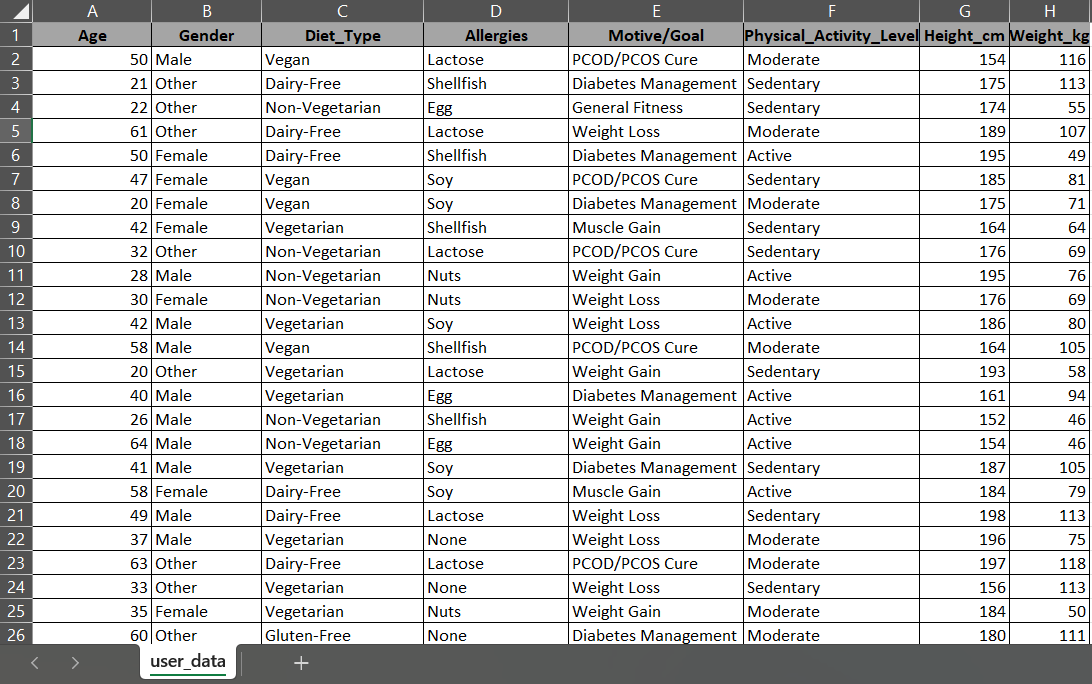
We are asked to collect user preferences and goals in-order to customize diet recommendations. I’ve searched Kaggle and GitHub and other sources on google like hummingbird for the data which has user data and couldn’t find.

So I decided to first write what all attributes I need and give them to ChatGPT to generate sample data.

The attributes I considered are:

* **Age**: Determines nutritional needs at different life stages.
* **Gender**: Influences calorie and nutrient requirements.
* **Diet Type**: Options include vegan, vegetarian, non-vegetarian, and specific dietary restrictions like gluten-free or dairy-free.
* **Allergies and Intolerances**: Identifies any foods to avoid (e.g., nuts, lactose).
* **Motive/Goal**: The user selects a specific health goal for the system to focus on, such as:
  + **Weight gain**
  + **Weight loss**
  + **Diabetes management**
  + **PCOD/PCOS cure**
  + **General fitness or muscle gain**
* **Physical Activity Level**: Sedentary, moderate, or active lifestyles impact daily calorie and nutrient needs.
* **Height and weight**: to calculate body mass index

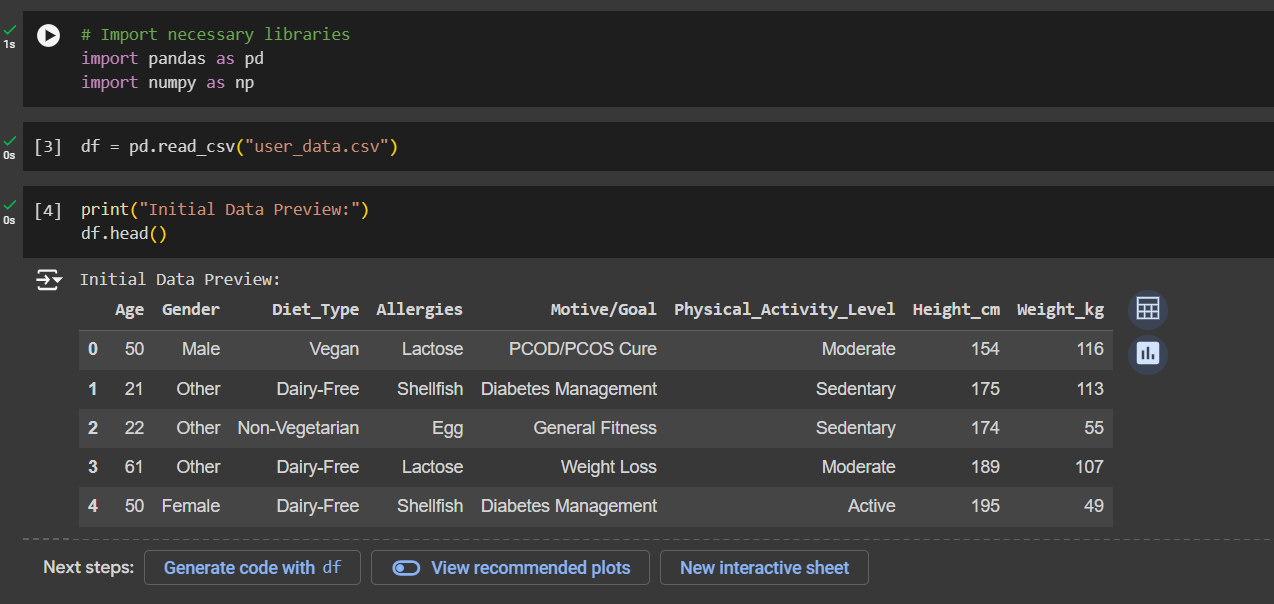
The screenshot of sample data is attached: (The data has a sample of 1000 users)

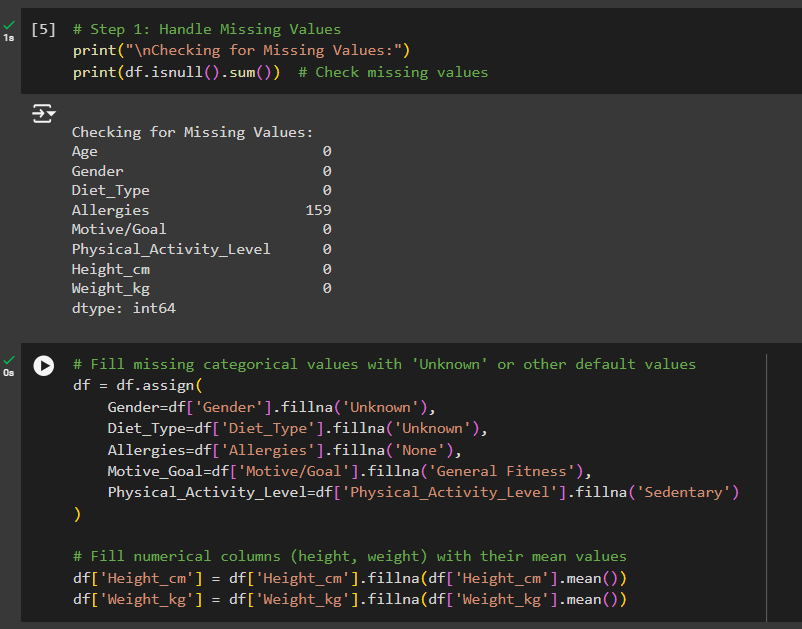


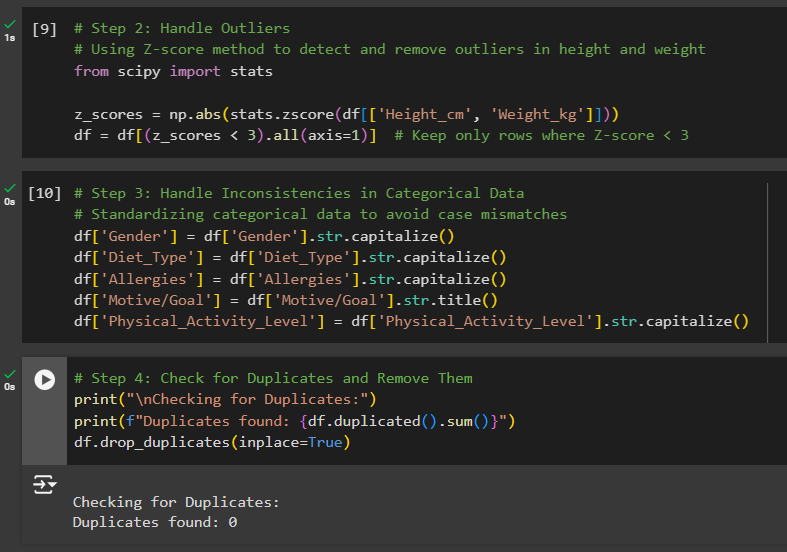
The drive link of the data:

<https://drive.google.com/file/d/1U_OjTKU-k4mDBLYMl_xdxo4xyDBpfp3m/view?usp=sharing>

**Task 2: Clean and preprocess the collected data by handling missing values, outliers, and inconsistencies.**

****

****

****

I’ve cleaned the user data and ensure there are no duplicates or empty values.

**Task 3: Gather Split the cleaned data into training and testing datasets, ensuring both sets are balanced and representative. Select a suitable machine learning or deep learning model architecture for nutritional analysis (e.g., neural network, decision tree, or regression model).**

*From here,*

*I've realized that we need additional nutrient information for various foods and details on which foods are suitable for different dietary categories. Integrating these datasets is essential for us to accurately test the system and provide meaningful recommendations. Right now, testing solely on user data might not yield the best insights. I'm currently working on merging both datasets to ensure our recommendations are accurate and relevant.*

*So my idea was to include an output column which has predefined diet plans (I considered 15 predefined diet plans for different categories and generated the output column for all 1000 entries). Hence I optimised the dataset and created new data set as follows:*



And I did all the tasks from beginning again

Below is the google collab link for the same.<https://colab.research.google.com/drive/18OU67saLNC3S-052dZOgIdLGJph-WrUX?usp=sharing>

