Ashoka Horizons

Week 1 Assignment

Data Collection Project: Kitchen Ecosystem

By Suhaani Singh

For this data science project, I chose to explore the data-rich environment of the kitchen. The primary objective was to familiarize myself with data storage practices by collecting, organizing, and structuring data related to meals prepared in our house over four days, including breakfast, lunch, and dinner.

By focusing on this field, I aimed to capture diverse data points ranging from ingredient usage and inventory levels to cooking duration, images of dishes, and appliance utilization patterns.

Data Collection Process

Planning and Structure

To ensure comprehensive data collection, I planned to record various aspects of each meal prepared in our house over four days. The data points included:

- Nutrient details: Calories, proteins, fats, carbohydrates.
- Categorical variables: Meal type (breakfast, lunch, dinner), meal taste (sweet, savory), cuisine (e.g., Italian, Chinese, Indian).

Tools and Environment Setup

1. Installing WSL:
   Windows Subsystem for Linux (WSL) was installed to provide a Linux-like environment on my Windows machine. This step was crucial for running various data processing tools and scripts.
   Command: `wsl --install`

2. Installing Miniconda:
   Miniconda, a minimal installer for conda, was installed to manage Python packages and environments efficiently.
   Steps:
   - Download Miniconda installer for Linux.
   - Run the installer: `bash Miniconda3-latest-Linux-x86_64.sh`

3. Creating a Conda Environment:
   A dedicated conda environment was created to ensure a clean workspace with all necessary dependencies.

Commands:
conda create --name data-science-fundamentals python==3.9
conda activate data-science-fundamental

Data Recording

Over the span of four days, I recorded details of each meal:

- Meal Details: The name of the dish, meal type, cuisine, and taste.
- Nutrient Information: Calories, proteins, fats, and carbohydrates were noted for each meal, either from food labels or reliable online databases.

Data Organization

The CSV file contained the following columns:

- Date
- Food Item (Name)
- Meal Type
- Cuisine
- Flavour
- Calories
- Proteins
- Fats
- Carbohydrates
- Fibre