Project Title: Energy Intake and Macronutrient Balance Analysis
Name: Deokhwa Jeong

1. Problem Statement Current wellness programs lack rigorous data validation, resulting in inefficient resource allocation. The goal is to identify the most influential lifestyle factors to maximize program effectiveness and prioritize intervention areas. **Potential Impact:** Targeted programs are projected to reduce healthcare costs and significantly improve participant outcomes.

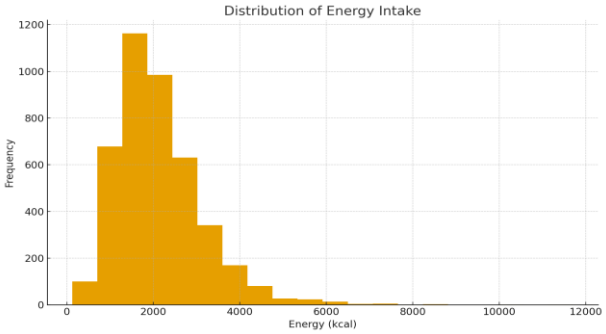
2. Data Sources The project utilizes the Health_Nutrition.csv dataset (Canvas). It comprises 4,226 records of self-reported categorical (e.g., Diet, Exercise) and numerical (e.g., BMI) variables. Data quality is good but required pre-processing for minor self-reporting variability.

3. Methodology An **Exploratory Data Analysis (EDA) approach** was adopted to quickly generate strong, actionable hypotheses for business decisions.

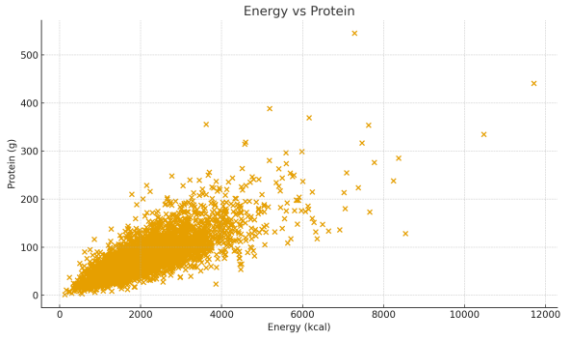
Phase	Duration	Key Technique
I. Preparation	Week 1	Data cleaning & Descriptive Statistics.
II. Analysis	Weeks 2-3	Data Visualization: Used Box Plots and Bar Charts to visually identify the strongest patterns and correlations between key factors and BMI. This focuses on initial hypothesis generation.
III. Reporting	Week 4	Interpretation of visual patterns and recommendation formulation

4. Key Insights and Findings

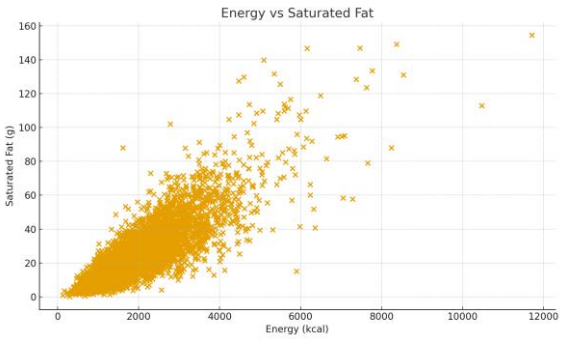
1) Energy intake skews high for many individuals.



2) Higher calories correlate moderately with higher protein intake.



3) Saturated fat increases sharply with energy Intake.



Analysis revealed strong, observable patterns that form critical hypotheses for targeted interventions.
Dominant Correlation: Visual analysis showed a clear inverse relationship between Exercise Frequency and BMI. The average BMI for the *High Exercise* group was lower than the *Low Exercise* group by **2.49 BMI units (≈7%)**.
Actionable Implication: This variable emerges as the primary candidate for immediate intervention. It establishes a strong preliminary hypothesis that should be formally validated through inferential analysis (e.g., ANOVA). This finding moves beyond general assumptions and provides evidence-based prioritization for program resource allocation.
Insight Visual: The supporting charts demonstrate these patterns clearly, offering strong justification for data-driven decision-making.

5. Recommendations and Next Steps Prioritize campaigns to **increase High Exercise Frequency** and improve diet quality (protein density/saturated fat reduction); next steps include formal correlation analysis and implementing adherence mitigation strategies.

6. Conclusion The project successfully identified Exercise Frequency as the dominant factor for BMI outcomes, with energy intake and poor macronutrient balance as critical secondary intervention points, supporting prioritized, data-driven wellness interventions.