Nutrition Analysis (Part-1 of Bachelor Thesis Project)

Authors: Gaurav Shresth(20BT10018), Kshamit Bombarde(20BT3AI07)

Advisor: Prof. Neeraj Kumar Goyal, Subir Chowdhury School of Quality and Reliability

Introduction: Food and nutrition are the way that we get fuel, providing energy for our bodies. We need to replace nutrients in our bodies with a new supply every day. Water is an important component of nutrition. Fats, proteins, and carbohydrates are all required.Nutrition is the science that interprets the nutrients and other substances in food in relation to maintenance, growth, reproduction, health and disease of an organism. It includes ingestion, absorption, assimilation, biosynthesis, catabolism and excretion.

Knowing and eating mindfully is not only essential for a healthy gut but also for peace of mind. Also,A diet filled with vegetables, fruits and whole grains could help prevent major conditions such as stroke, diabetes and heart disease.More often than not, we like to gorge on our favourite foods which are not exactly the best for our bodies.While it is okay for such binges to occur occasionally, such diets can be extremely harmful if the person does not strike a balance with healthy foods.

This notebook analyses the most common available foods and the nutritional facts in them.

Objective: Perform an in-depth data analytics study on a comprehensive food nutrition dataset to uncover patterns, relationships and insights into dietary landscapes. The goals are to categorize foods based on nutritional composition, utilize data visualizations to highlight findings, provide a nuanced understanding of nutrition to inform dietary choices, and identify foods with distinct nutritional qualities to support health goals.

Dataset Overview:
- 333 rows of food items, 10 feature columns
- Columns include name, calories, total fat, carbohydrates, protein, vitamins, cholesterol etc.
- Features are string, integer and float data types

Data Cleaning Process:
- Handling missing values through interpolation methods
- Identifying and removing duplicate entries
- Fixing inconsistencies by cross-validating with external food databases
- Converting data types into appropriate formats
- Documenting all data modifications for reproducibility

Exploratory Data Analysis Findings:
- Fats and oils have the highest calorie density per gram, reinforced through bar charts plotting calorie contributions across food groups. The visual dominance of fats & oils emphasizes the importance of monitoring fat intake.
- Comparative line charts showcase whole wheat having high proteins but also high caloric content compared to refined grains. This highlights the need for balance in nutritional prioritization.
- Treemaps effectively showcase sources of various macronutrients and micronutrients across food categories, facilitating an interactive overview.

Food-Specific Insights:
- Amongst high calorie foods, fortified milk and white bread top the list, revealed through sorted horizontal bar charts.
- For fats, lard has the maximum calories per gram.
- For protein-rich foods, beef slices and white rice emerge as optimal options through scatter plots.
- All visualizations are aimed at data-driven dietary decision making.

Future Work:
- Predictive machine learning models for personalized nutrition planning
- Integrating fuzzy logic to handle uncertainty and variability in nutrition data
- Expanding current dataset by assimilating external nutrition databases
- Developing a nutrition advisor application for tracking intake and guiding food choices

In summary, the thesis leverages data analytics methodologies to provide in-depth, visualized insights into the nutritional properties of foods to assist healthy dietary decision making.