CASE FILE: PROJECT UDAAN

Presented by: SPJIMR Analytics Club
Client: NutriLife India Pvt. Ltd.
Dataset: `Bharat_Consumption_Data.csv`

Context & Problem Statement

For five years, **NutriLife India Pvt. Ltd.** has led the premium breakfast market in India's six largest metro cities. Its strategy is based on one assumption: health-focused packaged food is consumed mainly by the urban elite. This assumption is now limiting growth.

Sales in metro markets have slowed. At the same time, internal supply chain data shows consistent orders for oats and muesli from Tier-2 towns and rural districts where NutriLife invests no marketing spend. The distinction between Urban India and Rural Bharat is weakening. A segment with rising purchasing power exists outside metros, but it remains under-identified because it does not fit standard "Urban, High-Income" definitions.

Current State & Strategic Need

NutriLife India Pvt. Ltd. maintains household-level data for **50,000 homes** (`Bharat_Consumption_Data.csv`), covering income, asset ownership, and expenditure. Current targeting relies primarily on reported income. This approach is unreliable in informal and semi-formal economies.

Consumption patterns and asset ownership often provide stronger indicators of purchasing power. NutriLife does not require another descriptive market report. It requires a **decision-oriented targeting framework** that identifies real demand beyond income labels.

The strategic need is to identify:

- Households with understated income but high consumption capacity.
- Families whose food choices indicate a shift from traditional staples to modern packaged alternatives.

Case Challenge

How can NutriLife India Pvt. Ltd. use household-level data to:

1. **Identify** high-potential non-metro customer segments.
2. **Estimate** their likelihood of purchasing oats and muesli.
3. **Recommend** an effective route-to-consumer strategy suited to their context.

Deliverables

Teams must submit:

1. **One presentation** (maximum 6 slides, including opening and closing slides).
2. **One code file** (Jupyter notebook export or Python/R script with comments, exported as PDF or HTML).

Submissions must address the following:

Target Persona

A data-backed profile of the priority customer segment. The persona should be derived from observable behavior such as consumption choices, asset ownership, and lifestyle signals. Generic demographic labels should be avoided.

1. **Wealth Identification Logic**: A transparent and explainable approach to inferring household purchasing power using available data. The approach should not rely solely on reported income.
2. **Route-to-Consumer Strategy**: A recommendation on how NutriLife should reach the identified segment. The rationale should be grounded in patterns observed in household behavior and access conditions reflected in the data.
3. **Feasibility and ROI**: Identification of the **top five districts** where this strategy can be piloted, with justification.

Evaluation Criteria

Submissions will be evaluated on:

Criteria	Weight	Description
Analytical Rigor	30%	Appropriate treatment of the income versus purchasing power problem.

Criteria	Weight	Description
Insight Clarity	25%	Findings and personas understandable to a non-technical business leader.
Business Viability	25%	Alignment between data insights and recommended strategy.
Interpretability	20%	Preference for clear, explainable logic over opaque or black-box models.

Data Dictionary

A separate file, `Data_Dictionary.csv`, is provided with the dataset. It contains definitions and descriptions of all variables in `Bharat_Consumption_Data.csv`.

Important Notes:

- Participants are expected to review this file before analysis.
- Missing values in expenditure columns should be treated as **zero spend**.
- **Values marked as `-99` represent blank/unavailable data points.**
- The target variable for oats/muesli consumption is `EXP_FLAKES_MUESLI_OATS`.

Event Terms & Conditions

- The decisions of the judges are final and binding.
- Any plagiarism of code or analytical output will result in disqualification.
- The organizing committee reserves the right to determine eligibility and authenticity of submissions.
- The dataset provided is synthetic and intended solely for competition use.

Organized by SPJIMR Analytics Club