Analyzing Fast Food Consumption: Trends, Health Impacts, and Machine Learning Solutions

Laurens Fontaine, Cooper Mastin, Amaia Modolo, Madeleine Page

Problem statement

- High fast-food consumption is known to increase the risk of a variety of metabolic diseases such as obesity, cardiovascular disease, diabetes, and high blood pressure.
- As fast food consumption across the world increases, there is a lack of information on what drives young people to consume fast food, and what effects it has on other aspects of their lifestyle.
- Improving identification and prediction of at-risk groups could help in the prevention of these conditions.
- Identifying relationships between fast food intake and other at risk behaviors is also significant as it can aid in understanding the connection between negative habits in different locations, age groups, and sexes.

Datasets

- Global School-based Student Health Survey
- Data pertaining to fast-food, cannabis, alcohol, and tobacco use and more.
- Contains data from countries across the world.
- Focuses on school-aged children in the age range of 13 to 16.

Proposed solution

Data extraction

- Data pertaining to fast-food consumption, cannabis, alcohol, tobacco, and cigarette use were extracted from our dataset.

Data organization and labelling

Dataset was labelled according to geographical region, age, and sex.

Dimensionality reduction

Umapp was employed to reduce dimensionality.

Clustering

Kmeans was used to create clusters from the reduced data.

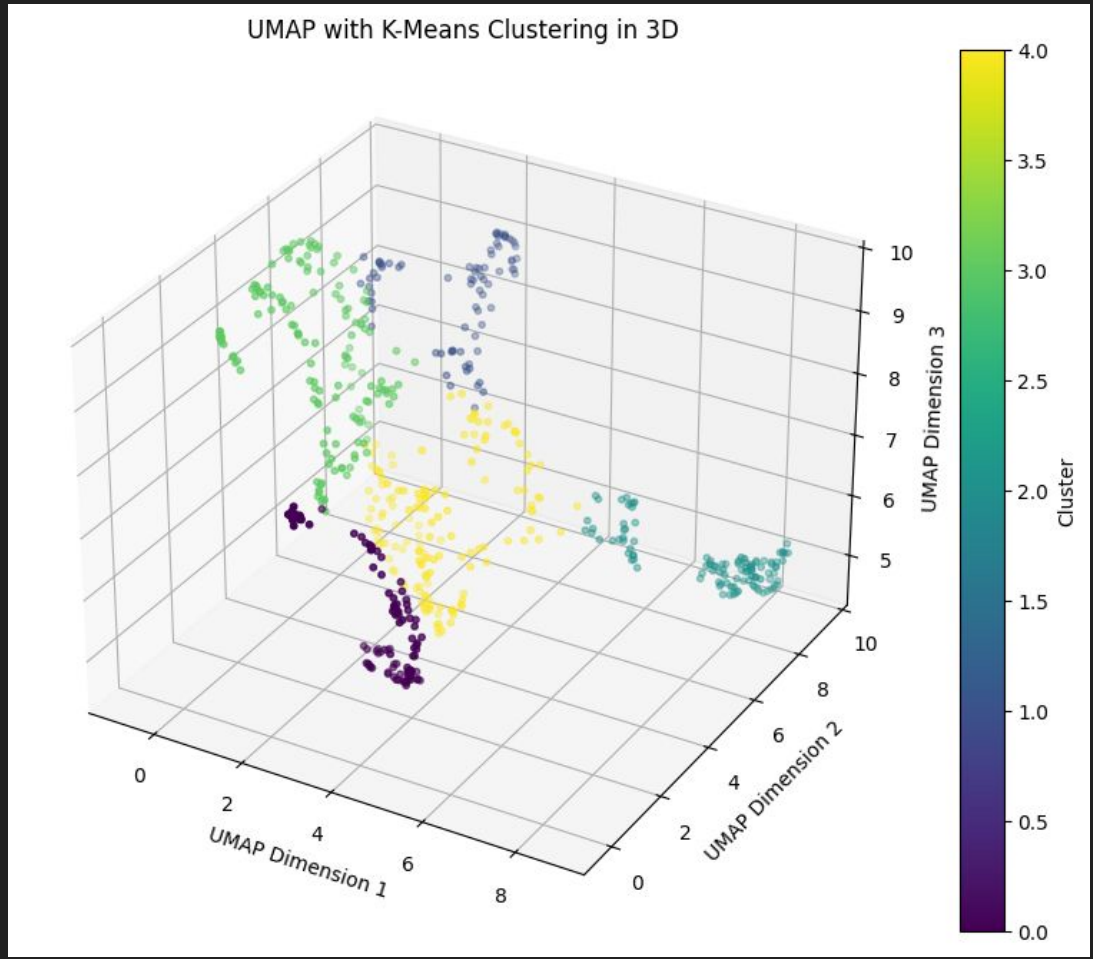
Graphing

Umapp was used for graphing and cluster analysis.

Clustering Results

- Selecting 5 clusters in the K Mean analysis lead to the best results as cluster sizes were similarly distributed.

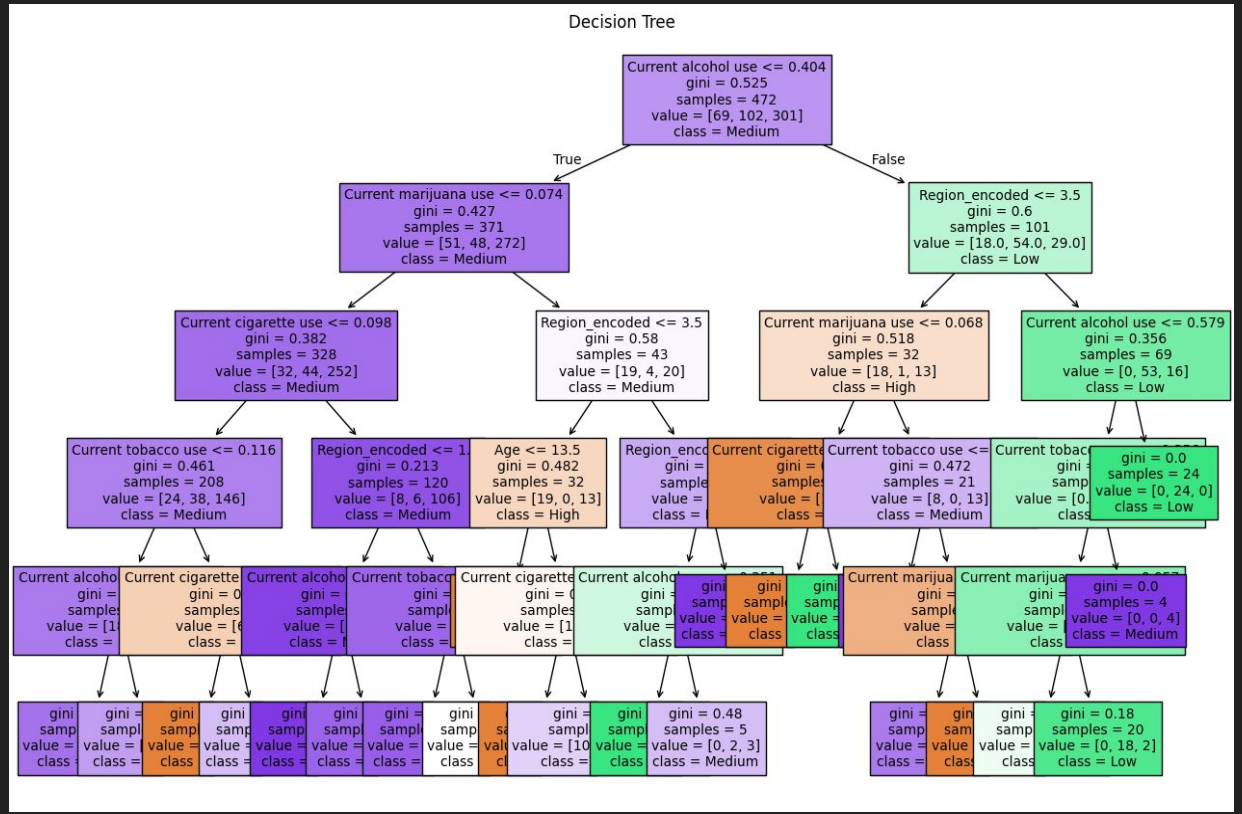
```
Cluster Sizes:
Cluster
0    104
1     80
2    107
3    146
4    154
```



01	**Largely East and West Asian.**	• Low cigarette, alcohol, and cannabis use • Moderate fast-food consumption • Predominantly female
02	**Largely East and West Asian with some South American and African representation.**	• Medium-high cigarette and tobacco use, low-moderate alcohol and cannabis use • Medium fast-food consumption • Mostly male
03	**Almost entirely South American.**	• High cigarette, alcohol, cannabis, and tobacco use • Low fast-food consumption
04	**Largely East and West Asian, some African representation.**	• Low cigarette, tobacco, cannabis, and alcohol use. • Low-moderate fast-food consumption • Lowest average age
05	**Predominantly North American and East Asian. Highest Oceanic representation.**	• Moderate alcohol, tobacco, cannabis, and cigarette use • High fast-food consumption • Highest age

Predictive Modeling using Sklearn

- Used to predict low, moderate, or high fast food consumption given other factors.
- Accuracy of around 70%.



Practical applications

- **Identifying High-Risk Populations:**
  - Enables identification of populations at higher risk based on clustering behaviors.
- **Intervention:**
  - Nuanced clustering provides insights that can tailor health interventions for specific demographic and behavioral groups.
- **Prediction:**
  - Predict future health risks by identifying individuals likely in high-risk clusters.
- **Prevention:**
  - Proactively direct health resources to populations at risk, potentially preventing adverse health outcomes.