NATURAL GAS CONSUMPTION IN UŞAK (2015–2018)

Dataset Description:

Column Name	Description
Year	Year of the data record (e.g., 2015, 2016...)
Longitude, latitude	Geographical coordinates of the user's gas meter
Jan usd/tr - Dec usd/tr	Monthly USD/TRY exchange rates
Jan - Dec	Monthly natural gas consumption (in cubic meters)
Total	Annual consumption total

Note: Each row represents a yearly or regional consumption profile.

Project Objective:

In this project, students are expected to apply **unsupervised learning** techniques, particularly **clustering**, to identify patterns, segment usage profiles, and detect anomalies in the dataset. You are expected to deliver insights in the following four areas:

1. Consumption Pattern-Based Clustering

- Use monthly consumption percentages to group different types of usage profiles.
- **Goal:** Identify distinct seasonal or consistent consumption behaviors.
- *Hint:* Use KMeans, Agglomerative Clustering, or dimensionality reduction with PCA or t-SNE for visualization.

2. Clustering Based on Exchange Rate Influence

- Combine monthly USD/TRY exchange rates with consumption data.
- **Goal:** Identify whether certain consumer profiles are sensitive to currency fluctuations.
- *Hints:* Normalize features before clustering. Explore correlations between exchange rate spikes and usage drops.

3. Anomaly Detection

- Detect rows (years or entries) that significantly deviate from the rest.
- **Goal:** Highlight unusual consumption behaviors or potential data anomalies.
- *Hints:* Apply algorithms like DBSCAN, Isolation Forest, or calculate Z-scores for outlier detection.

4. Geospatial Consumption Profiling

- Dataset includes location or regional metadata as longitude and latitude, use it to map consumer profiles geographically.
- **Goal:** Identify areas with similar consumption behaviors or outliers by location.
- *Hints:* Use folium, geopandas, or plot clustering results on a map using coordinates if available.