# Fine-Tuning Clustering Model to Determine Earthquake Locations Along the East Anatolian Fault

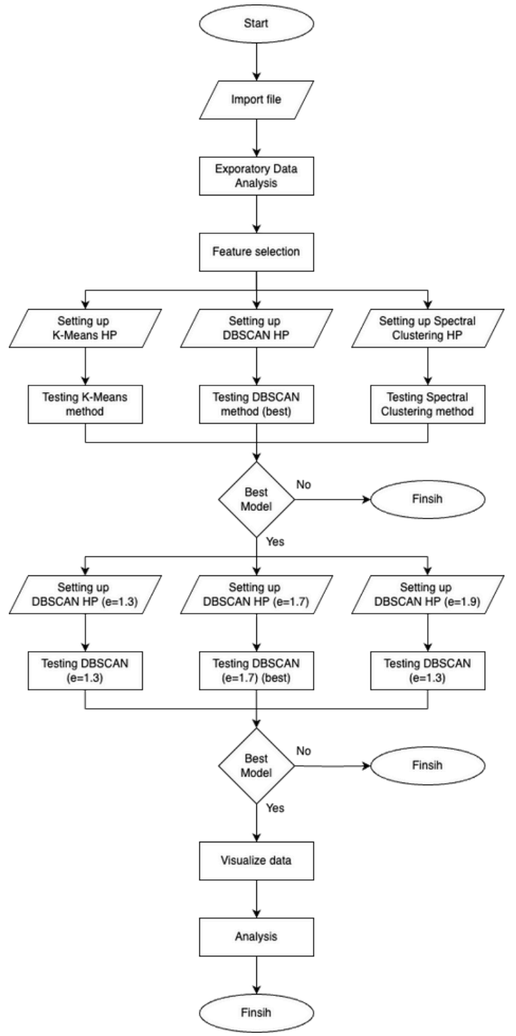
**Dziand Dafi Ginandjar**
**12321072**

# Project Overview

The devastating earthquake that struck Turkey on February 6, 2023, had a magnitude of 7.8 and was centered in Nurdağı, near Gaziantep. The tremors caused massive destruction in Turkey and Syria, claiming the lives of more than 50,000 people and injuring tens of thousands of others.

Many buildings were destroyed, leaving residents in emergency conditions amidst the winter season. Relief efforts included rescue operations, distribution of food, water, and psychosocial support, carried out by various local and international organizations to help survivors cope with the trauma and challenges of life after the disaster.

# Flowchart Diagram



# Dataset Being Used

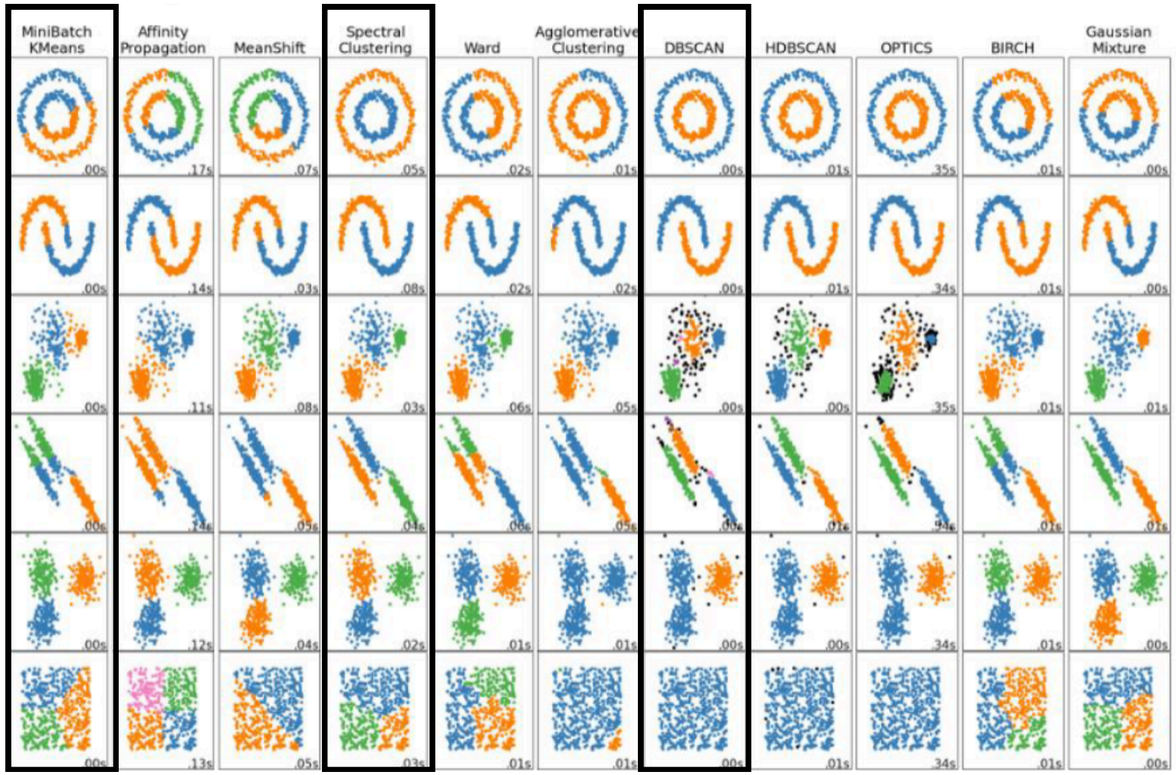

| Parameter | Value |
| --- | --- |
| Range of Time | 2020-01-01 -> 2024-01-01 |
| Minimum scale (magnitudo) | 5.0 |
| Location | Turkiye and its surroundings |

# Clustering Algorithm

K-MEANS

SPECTRAL CLUSTERING

DBSCAN



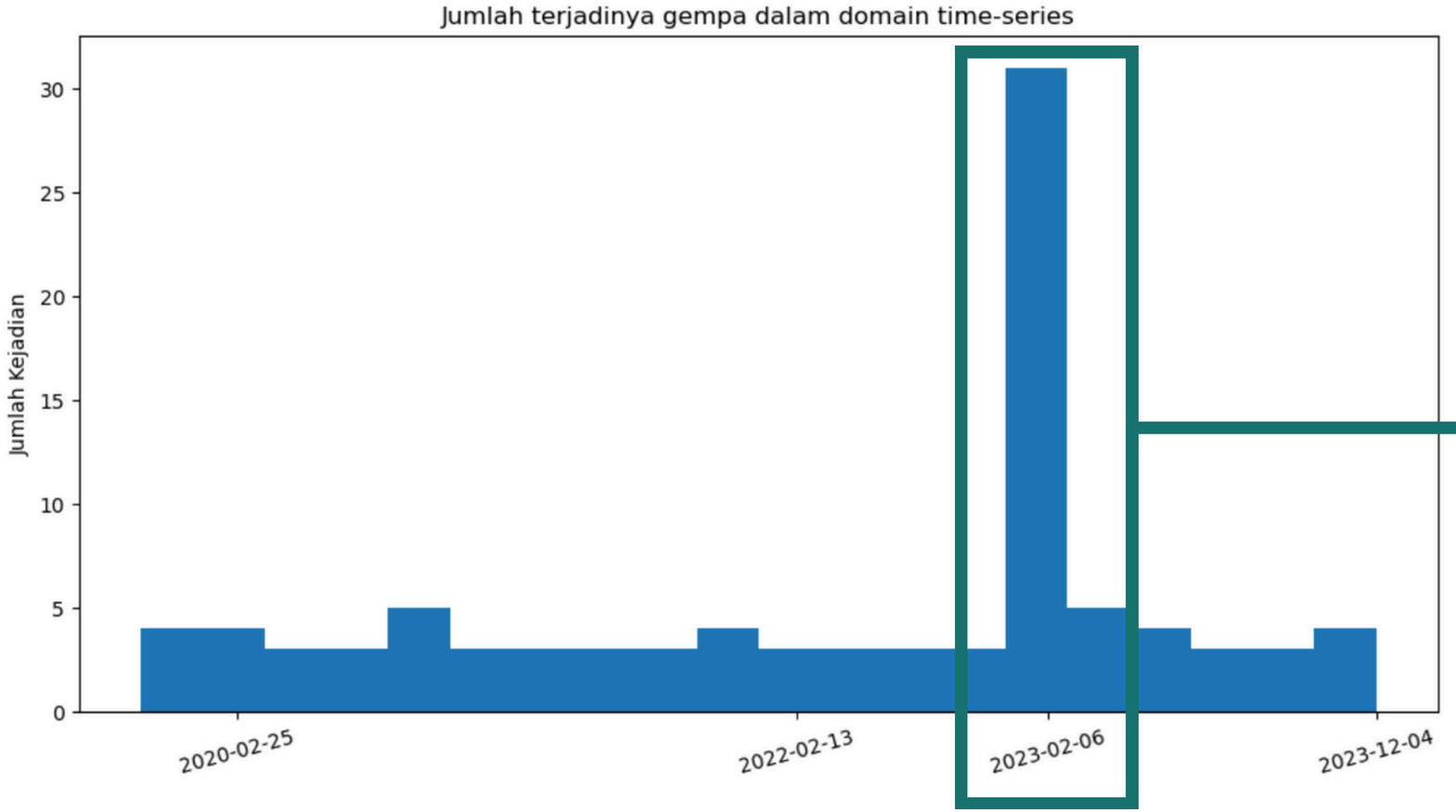

https://scikit-learn.org/stable/auto_examples/cluster/plot_cluster_comparison.html

# Feature Selection

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 97 entries, 0 to 96
Data columns (total 22 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   time             97 non-null     object
 1   latitude         97 non-null     float64
 2   longitude        97 non-null     float64
 3   depth            97 non-null     float64
 4   mag              97 non-null     float64
 5   magType          97 non-null     object
 6   nst              59 non-null     float64
 7   gap              97 non-null     int64
 8   dmin             97 non-null     float64
 9   rms              97 non-null     float64
 10  net              97 non-null     object
 11  id               97 non-null     object
 12  updated          97 non-null     object
 13  place            97 non-null     object
 14  type             97 non-null     object
 15  horizontalError  97 non-null     float64
 16  depthError       97 non-null     float64
 17  magError         97 non-null     float64
 18  magNst           97 non-null     int64
 19  status           97 non-null     object
 20  locationSource   97 non-null     object
 21  magSource        97 non-null     object
dtypes: float64(10), int64(2), object(10)
memory usage: 16.8+ KB
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 97 entries, 0 to 96
Data columns (total 5 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   time       97 non-null     object
 1   latitude   97 non-null     float64
 2   longitude  97 non-null     float64
 3   mag        97 non-null     float64
 4   depth      97 non-null     float64
dtypes: float64(4), object(1)
memory usage: 3.9+ KB
```

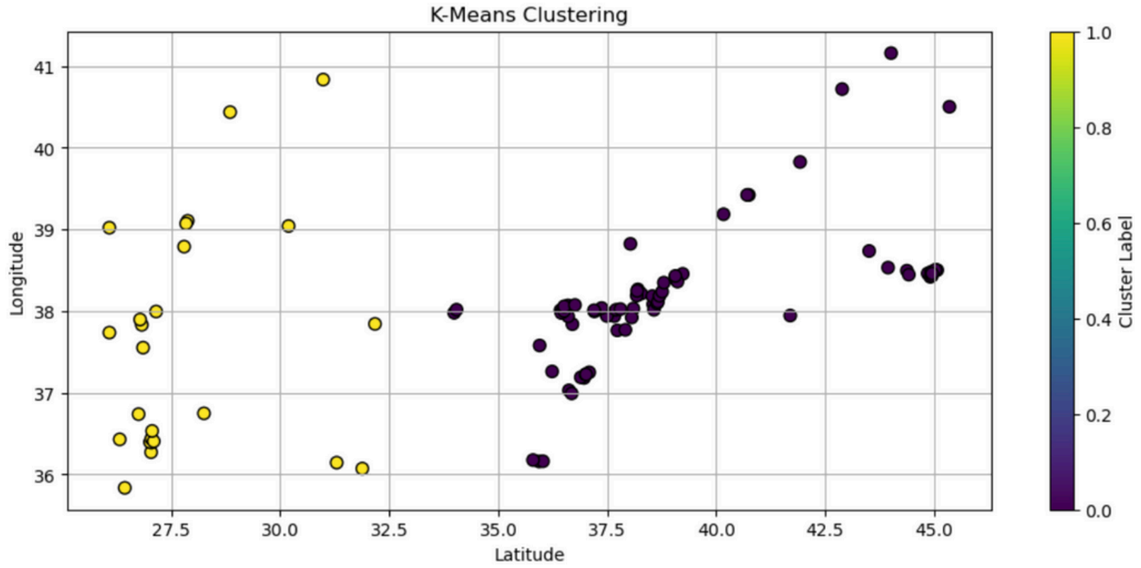
# Exploratory Data Analysis



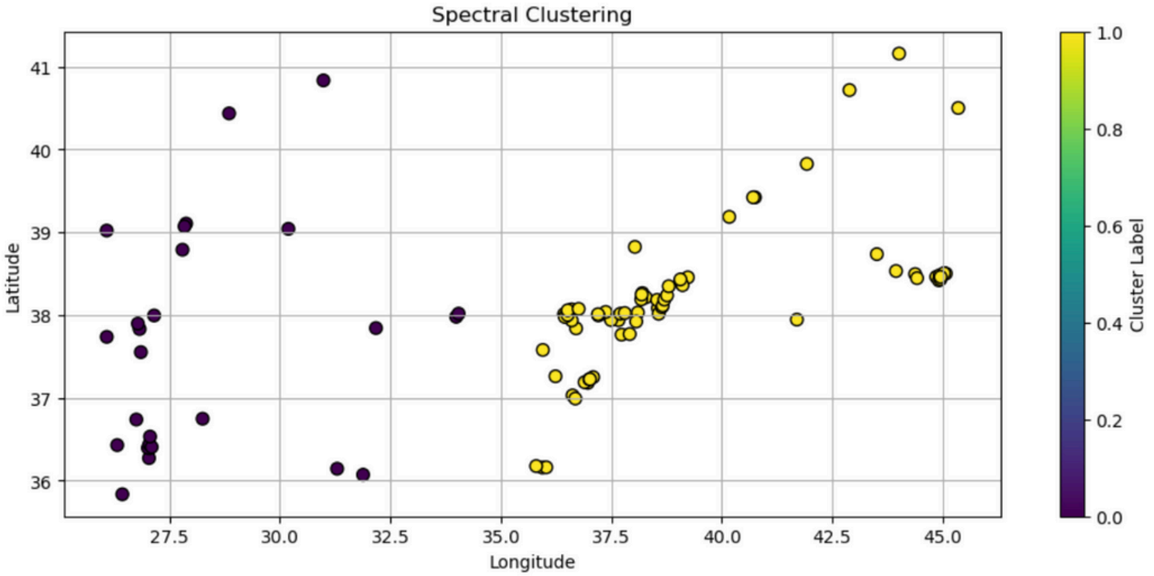

**THE 2023 TURKIYE EARTHQUAKE**

# Analysis

## K-MEANS




## SPECTRAL CLUSTERING

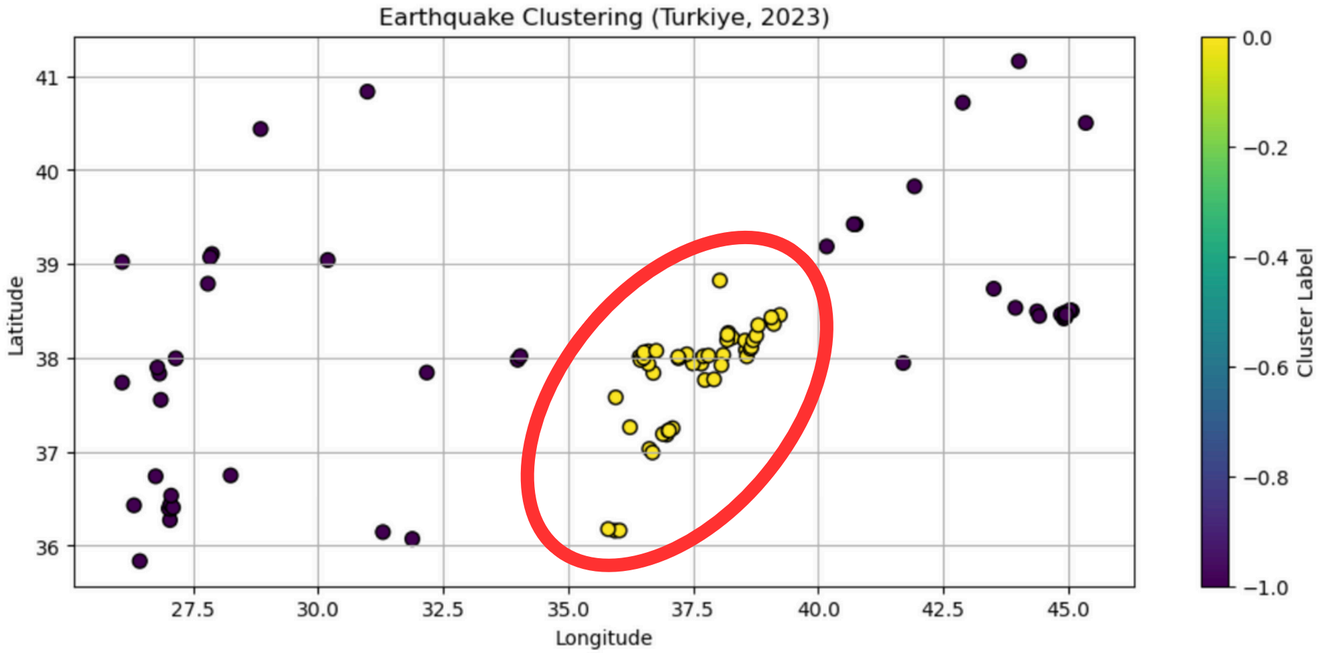



**ANALYSIS**

The K-means clustering method (k=2) and Spectral Clustering are unable to estimate target locations with high-density levels.

# Analysis

## DBSCAN




| Hyperparameter | |
|---|---|
| Epsilon (contoh) | 1.3 |
| Minimum Samples | 30 |

**ANALYSIS**

**The DBSCAN clustering method can estimate target locations with high-density levels.**
**Therefore, the most suitable clustering method is DBSCAN.**

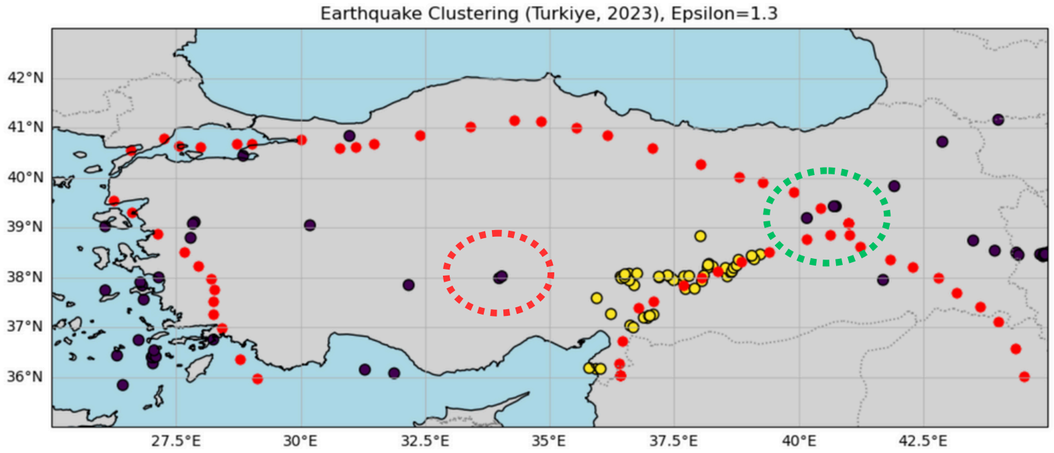
.

# Analysis

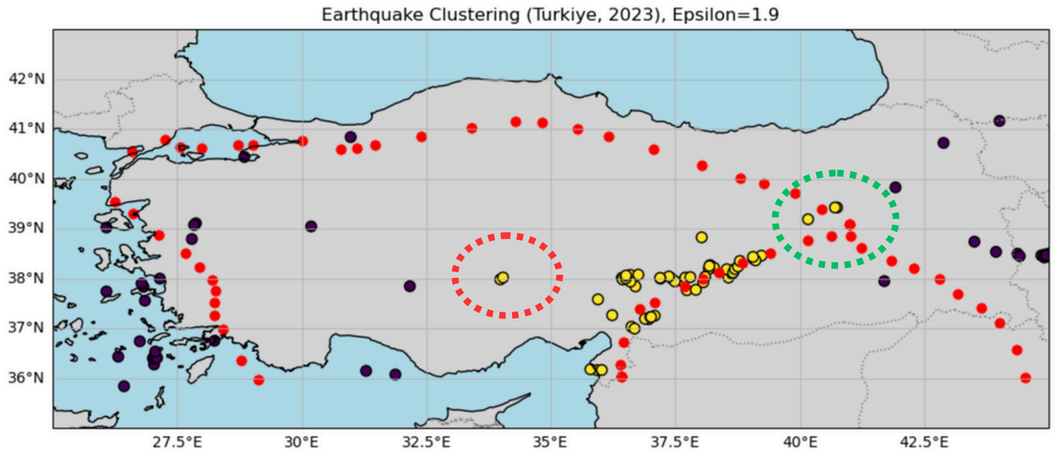
Anatolian Fault

EQ location (non-target)

EQ location (target)




| Hyperparameters | |
|---|---|
| Epsilon | 1.3 |
| Minimum Samples | 30 |



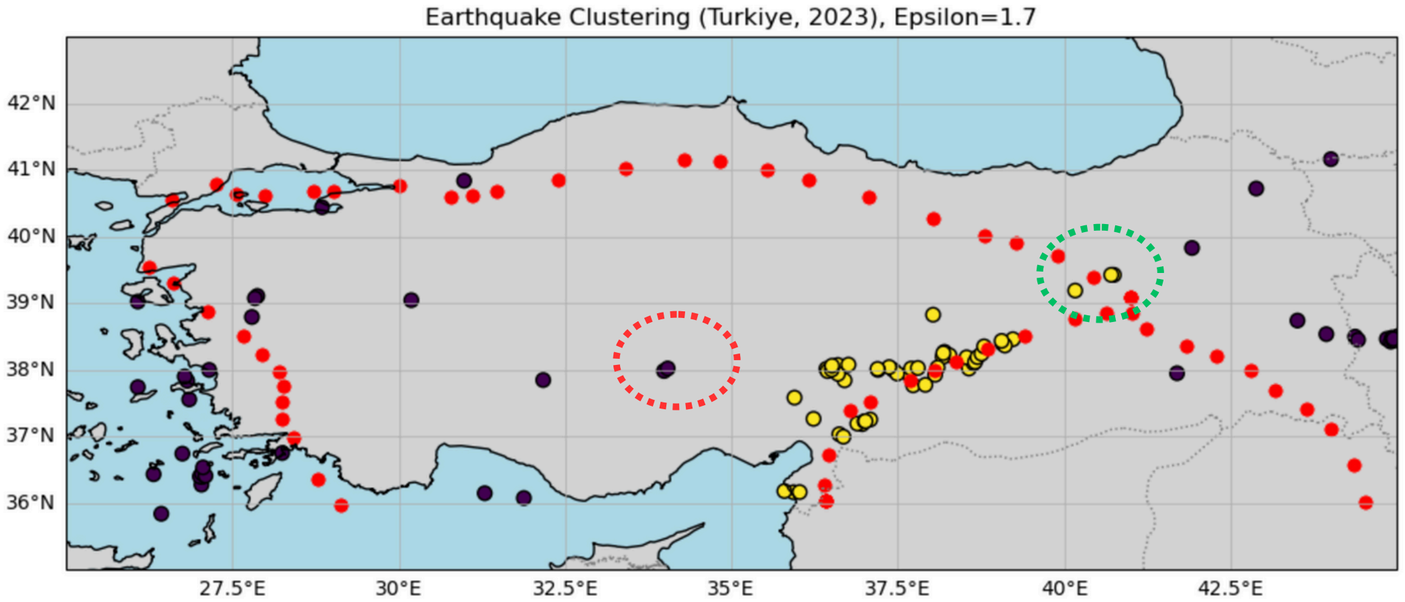

| Hyperparameters | |
|---|---|
| Epsilon | 1.9 |
| Minimum Samples | 30 |

Marker 1

Marker 2

# Analysis

Anatolian Fault  EQ location (non-target)  EQ location (target)




| Hyperparameter | |
|---|---|
| Epsilon | 1.7 |
| Minimum Samples | 30 |

Marker 1  Marker 2

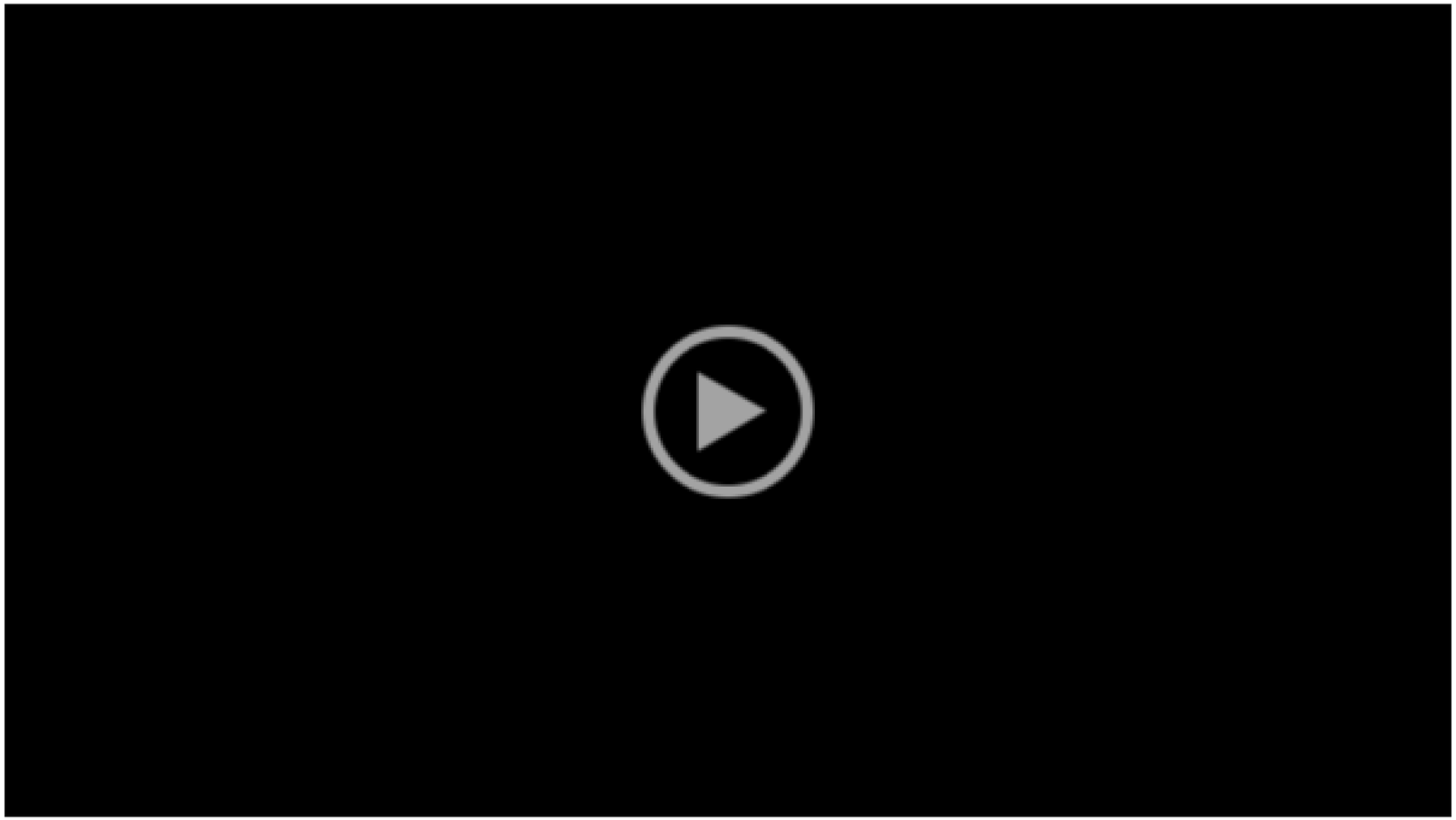
# Analysis

The DBSCAN model can cluster earthquakes in the Marka 1 area but does not include Marka 2 because Marka 1 is located along the fault line, whereas Marka 2 is far from the fault line and therefore should not be included in the cluster.
Would you like me to keep anything from this conversation for future reference?

| Model | **Marker 1** | **Marker 2** |
|---|---|---|
| DBSCAN (epsilon=1.3) | X | X |
| DBSCAN (epsilon=1.7) | V | X |
| DBSCAN (epsilon=1.9) | V | V |

The best DBSCAN model is the one with an epsilon value of 1.7 because it successfully includes earthquakes in Marker 1 within the cluster while excluding earthquakes in Marker 2 from the cluster.

# Demonstration



# Thank You