Project Name: Clustering project for the municipality of Breda

<Picture>

Context of project1:

The municipality of Breda examined the state of segregation in Breda and the extent to which Breda's policy is effective in combating segregation in the city. There are several indications that segregation and social inequality between neighborhoods have increased in the past 10 years. I will use the skills and knowledge I have gained over the course of the first year to build an example of how to cluster the neighborhoods for the City of Breda to help combat rising segregation in a completely data-driven manner.

**Data cleaning & EDA**

I downloaded raw data from various sources and conducted Exploratory Data Analysis (EDA) to gain insights into the dataset.I carried out data cleaning tasks, including handling missing values, removing duplicates, and handling outliers. After cleaning

**Data processing**

I applied some of Pre-processing techniques, I applied Data Scaling using StandardScaler from Sklearn library, to have zero mean and unit variance. which is crucial for the next step, Next step is applying PCA "Principal Component Analysis" its a typical process for dimensionality reduction, lastly I applied Neares Neighbors, The Nearest Neighbors algorithm used to find the nearest neighbors of each data point in the PCA-transformed space. This can help in clustering

Implementation

Phase1

Phase2

Phase3

Skills used