

Visual Analytics Tool for Air Quality Index

Visual Analytics project 2023/24

ELISA DE BELLIS 1858927

Contents

1 Introduction

2 Related works

3 Data

4 Visualizations

5 Analytics

6 Application & Utilities

Introduction

Air pollution emissions have declined in the last decade, resulting in better air quality. Despite this improvement, air pollution remains the largest environmental health risk in Europe.



Air Quality Index Visual-Analytics tool has the aim to analyze concentration of pollutants and explore insights to capture an overview of the last 10 years about the situation of Italy.

Related works

1. The TRAFAIR air quality dashboard

The TRAFAIR air quality dashboard is a web tool that monitors urban air quality using low-cost sensors. It tracks pollutants like CO, NO₂, NO, and O₃ in real time, updating every two minutes. The data is visualized through interactive graphics, helping decision-makers understand air quality trends and implement policies.

([BACHECHI et al., 2020](#))

2. AirVIS

AirVIS is a web-based system for analyzing air quality data using visual analytics. It offers three main views: a GIS View to show pollution distribution on a map, a Scatter Plot View to track temporal air quality patterns, and a Parallel Coordinates View to explore correlations between pollutants and their impact on the Air Quality Index (AQI). The system helps users understand pollutant trends, identify anomalies, and explore relationships between different pollutants.

([LIAO et al., 2014](#))

3. Air Quality Index (AQI)

The Air Quality Index (AQI) monitors air quality in real time, with values ranging from 'good' to 'dangerous' based on pollutants like PM₁₀, PM_{2.5}, NO₂, and O₃. Used by agencies like the EPA and EEA, it helps the public understand pollution levels. The EEA's portal provides data, analysis, and visualizations of air quality across Europe.

([EEA](#))

Dataset

The dataset is taken from ISPRA
(Istituto Superiore per la Protezione e
la Ricerca Ambientale).

It is divided into four files, one for
each pollutant: PM2.5, PM10, NO2,
O3

- station_eu_code
- id_regione
- id_provincia
- id_comune
- station_code
- Regione
- Provincia
- Comune
- nome_stazione
- tipo_zona
- tipo_stazione
- TIPO
- Lon
- Lat
- yy
- n
- sup25
- sup15
- media_yy
- minimo
- massimo

Dataset

To establish air quality on the basis of previous pollutants, reference is made to the European Environment Agency (EEA)

Pollutant	Index Level (based on concentration in $\mu\text{g}/\text{m}^3$)					
	Good	Fair	Moderate	Poor	Very Poor	Extremely poor
PM2.5	0-10	10-20	20-25	25-50	50-75	75-800
PM10	0-20	20-40	40-50	50-100	100-150	150-200
NO2	0-40	40-90	90-120	120-230	230-340	340-1000
O3	0-50	50-100	100-130	130-240	240-380	380-800

Pre-Processing

2016

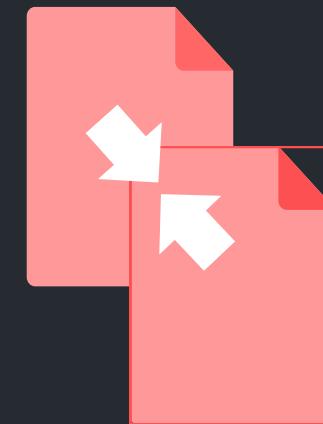
Divide data by year



Group data by
Regions



Select numerical value for
t-SNE analysis



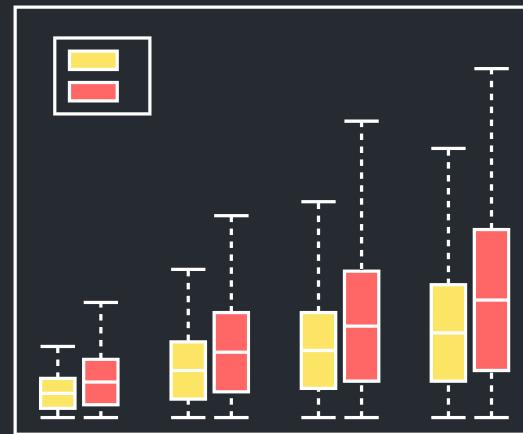
Create one file for all
pollutants divided by years.

Visualizations

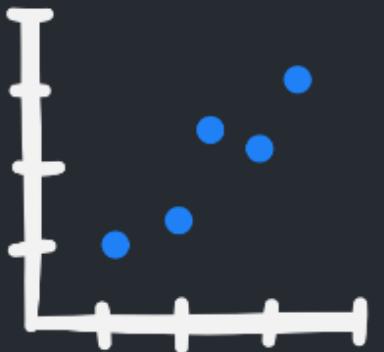
Map



Box-Plot



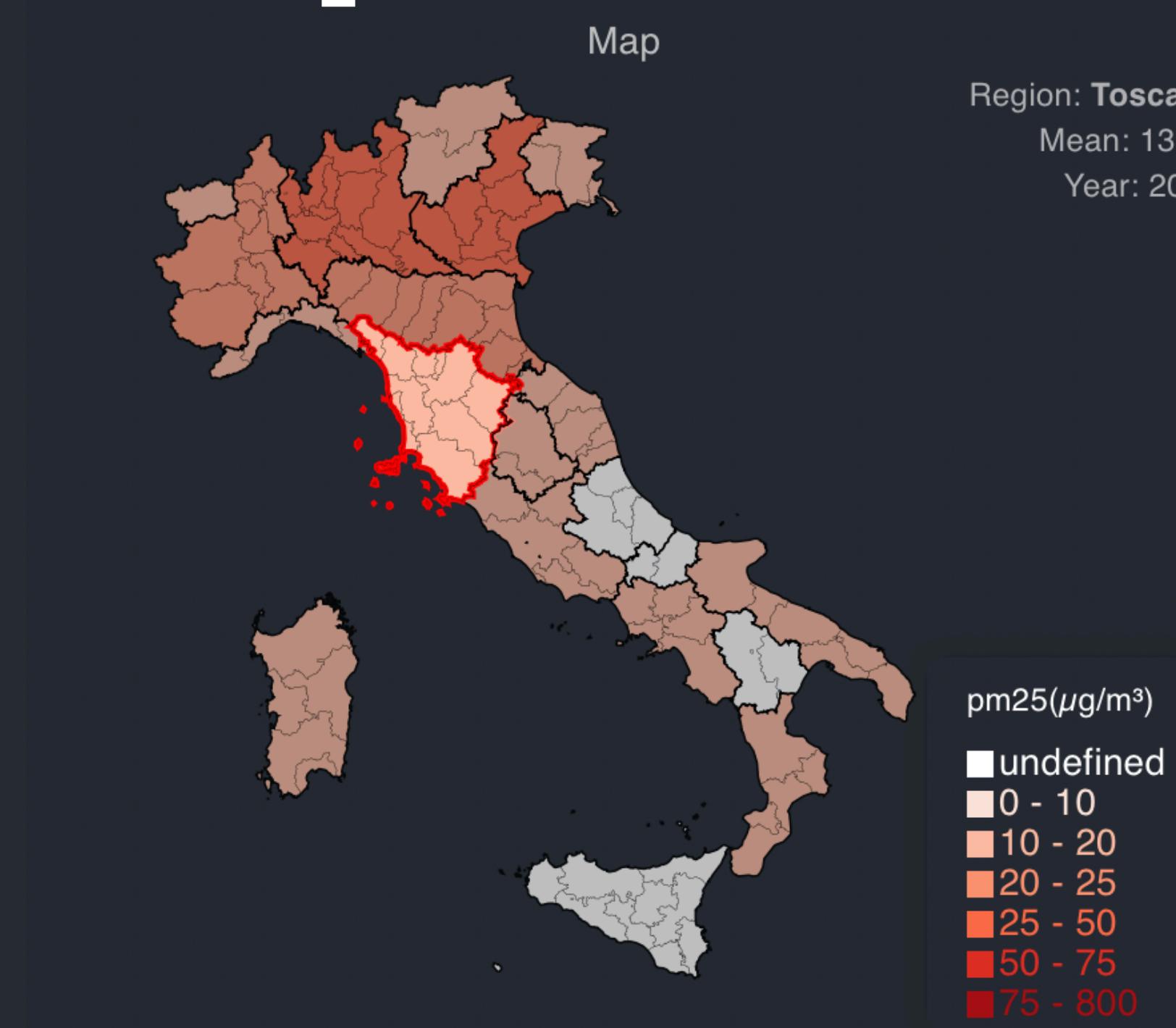
Scatter-plot



Time-Series



Map



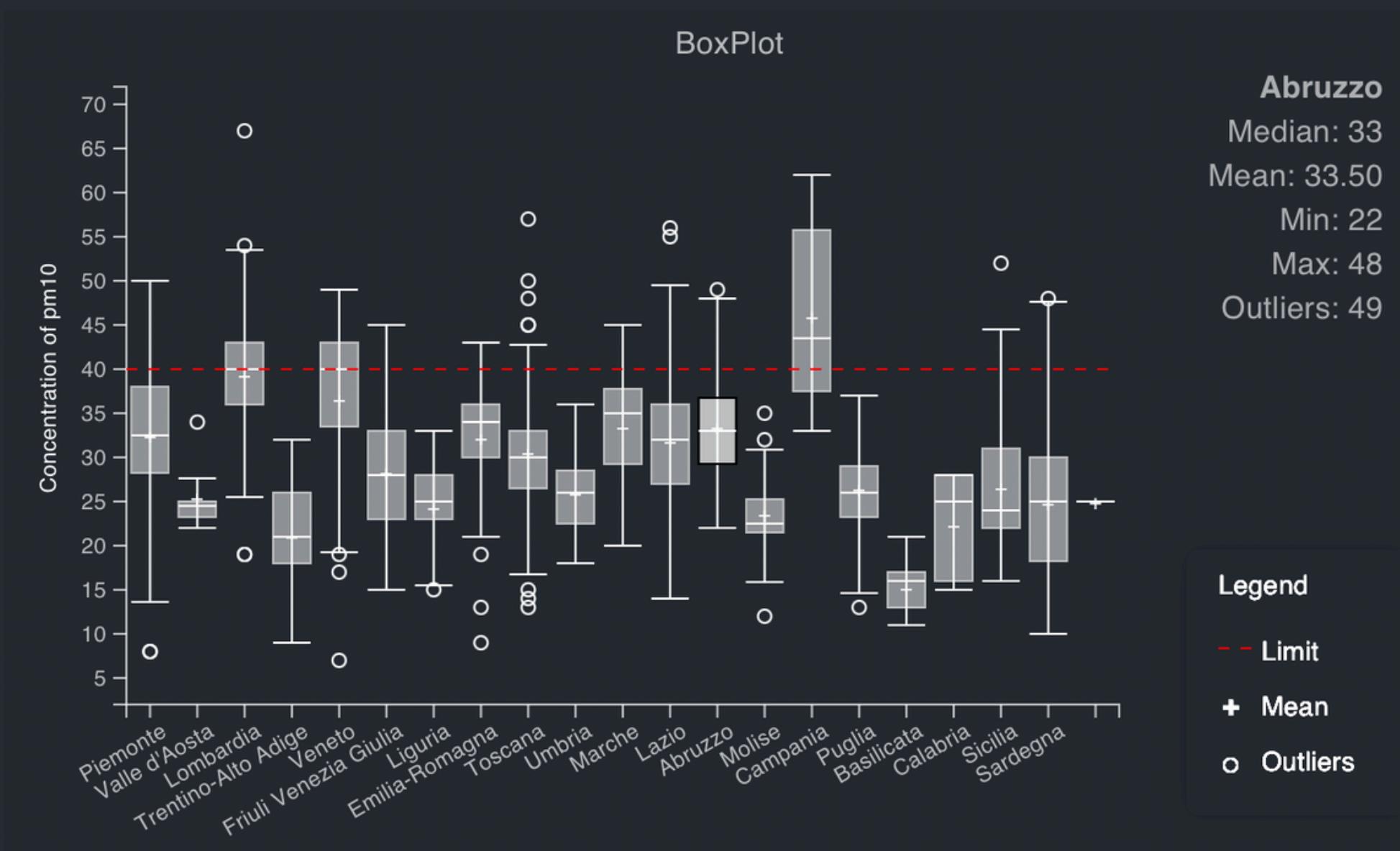
Visualization

The Map allow to show the mean concentration of the different pollutants in the Italy's regions. It consist of a map divided into 20 regions, each colour according to a shade of a red scale, according to the concentration of the selected pollutant.

Insight

- Critical regions and province
- Overview of the general situation

Box-Plot



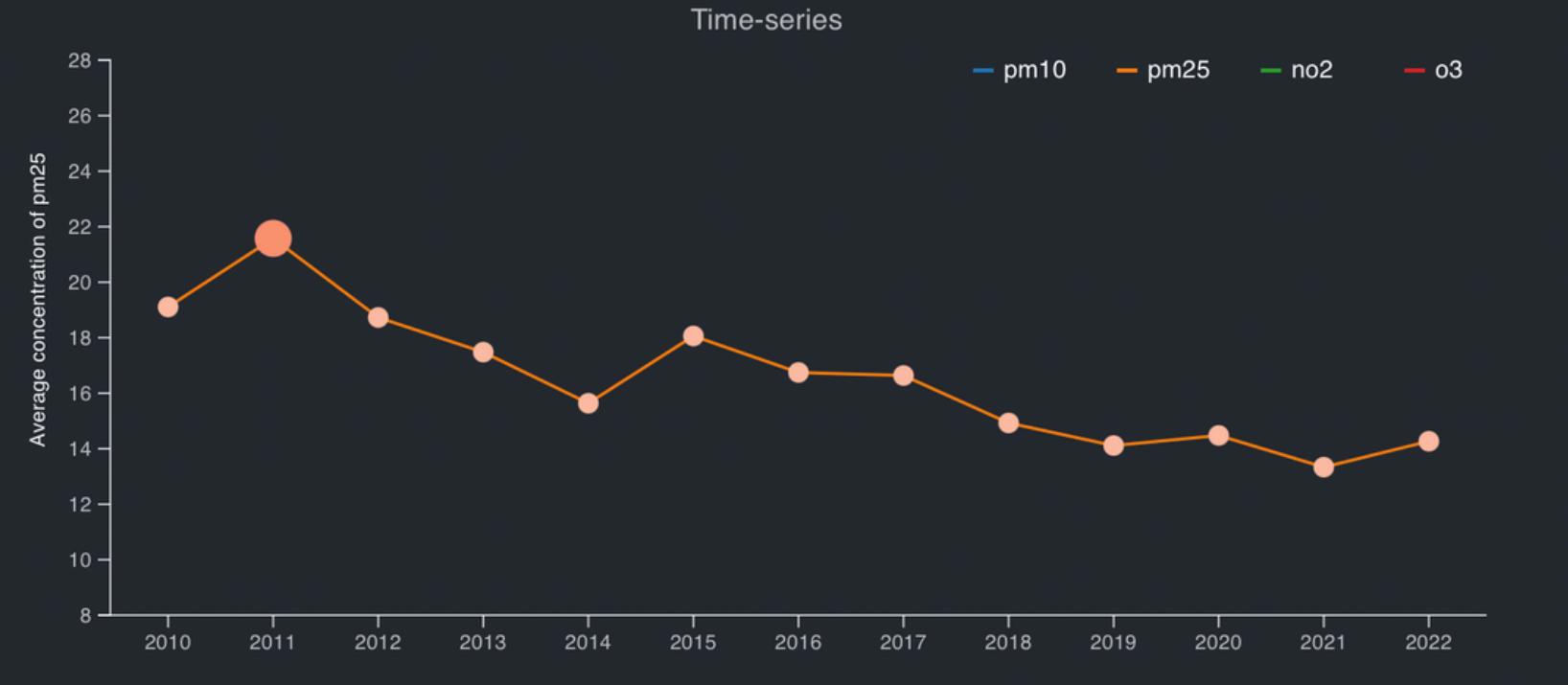
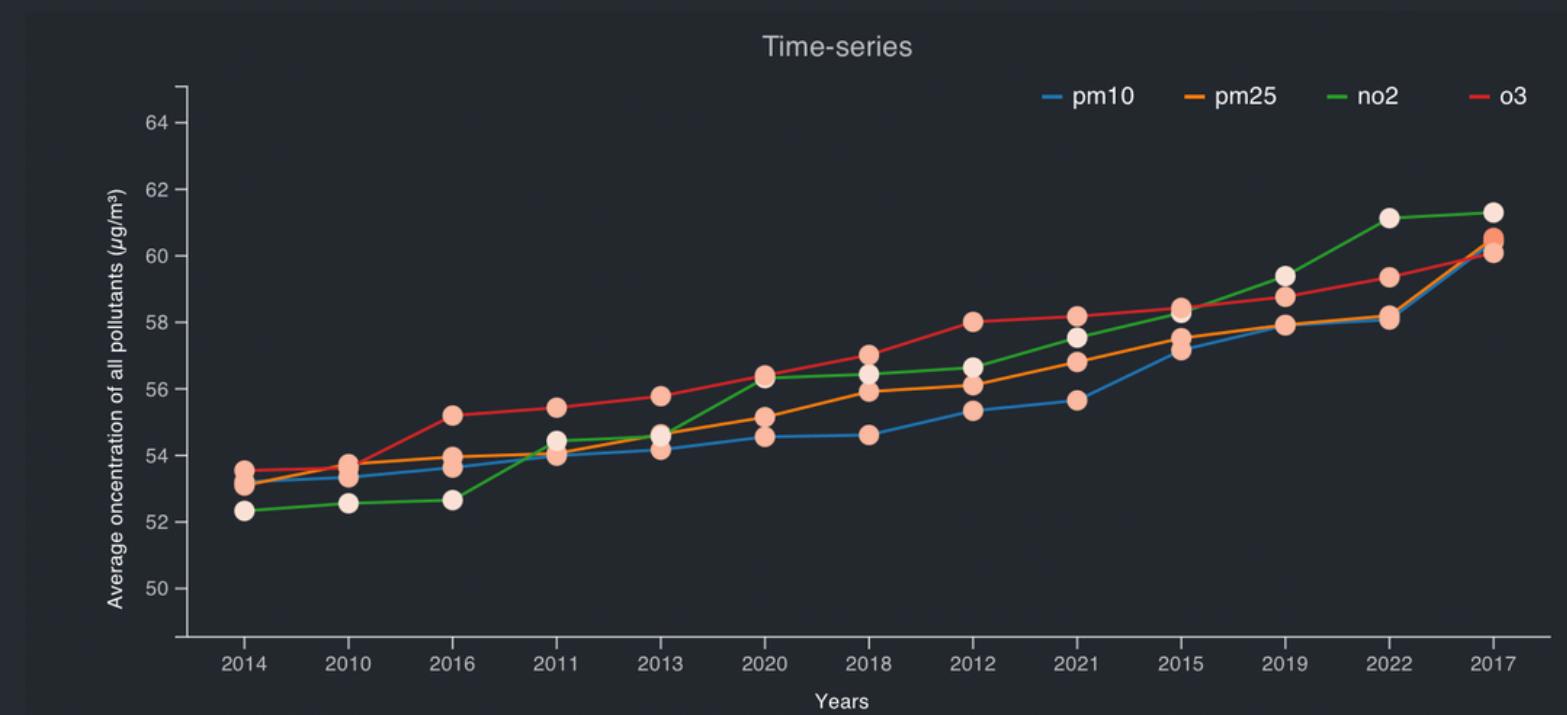
Visualization

The box-plot depicts the data for each region and allows comparison of data from all regions, showing the mean, median, maximum and minimum values and outliers. The graph also shows a threshold indicating the European limit allowed for that pollutant.

Insight

- Critical regions and province
- Comparison between different regions and provinces

Time-Series



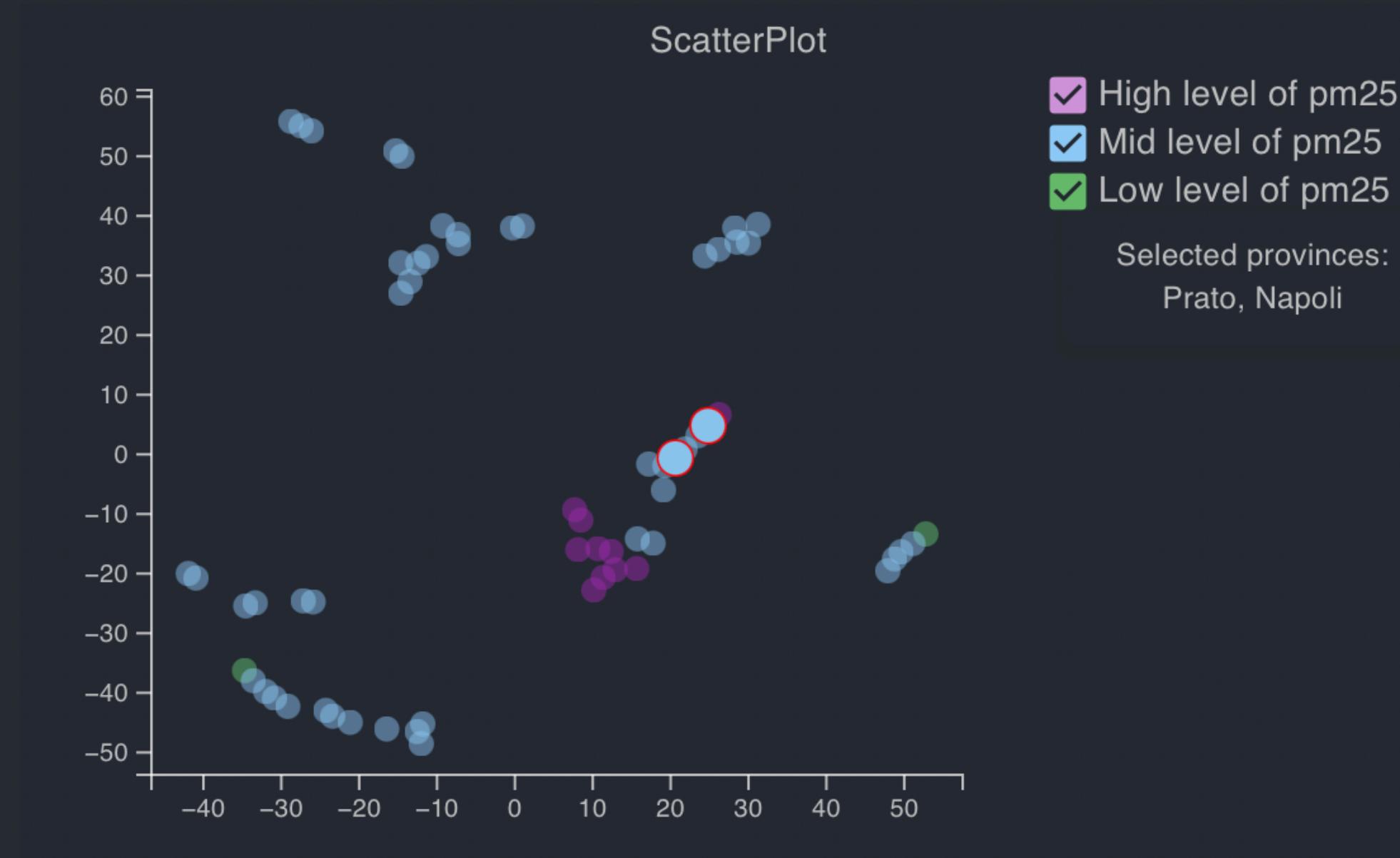
Visualization

The Time Series depicts how the average concentration in Italy in general has changed over the years. On the x-axis we have the different years taken into consideration, from 2010 to 2022, on the y-axis the average of all regions.

Insight

- Decrease in pollutants over years
- Impact of Lockdown

Scatter-Plot



Visualization

The Scatter-Plot is the result of dimensionality reduction using t-SNE. It helps to amplifies the separation between clusters of the data, that are arranged on the 2D space of the scatterplot.

Insight

- Cluster models based on agent concentration
- Analysis of specific provinces

Analytics

The analytical process of the project focuses on applying t-distributed stochastic neighbor embedding (t-SNE) as a technique for dimensionality reduction. The following steps were undertaken to achieve the objective:



PRE-PROCESSING



STANDARDIZATION



t-SNE

Since my aim was to find clusters of the data in order to capture some insight, this solution was the one that worked the best for my purposes wrt to PCA or MDS.

Application & Utilities

Intended Users

The Visual Analytics tool for the Air Quality Index is designed to be a valuable resource for environmental analysts, policymakers.



ANALYSTS



POLICYMAKERS

A black and white photograph of a dense forest. The trees are tall and thin, their branches reaching upwards. The sky is bright, creating a strong contrast with the dark trunks and branches. The overall effect is one of a dreamlike or surreal landscape.

DEMO