

Milestone I - Dataset

Computing Project 2016/17
14D003/14D004

Carlos Isaac Rodriguez Prado
Hans-Peter Höllwirth
Veronika Kyuchukova

The project will analyze annual pollution data in Europe, provided by the European Environment Agency (EEA). We intend to combine two independent datasets.

The first dataset [1] captures annual air pollutant concentrations at station, city, and national level from 36 states (EU28, BIH, ISL, LIE, MKD, MNE, SRB, SUI, TUR). The dataset assesses the population's exposure to 5 types of air pollutants: NO₂, O₃, Particulate matter < 2.5µm (PM_{2.5}) / < 10µm (PM₁₀), and Benzo(a)pyrene (BaP) as measured from roughly 1,000 official monitoring stations in European cities over several years (depending on station) until 2012. Data for 2013 is provided in a separate data file and shall be added. The stationary dataset consists of roughly 50,000 records. The data comes in an Excel format.

The second dataset [2] captures national annual emissions reported to the convention on long-range transboundary air pollution by 31 states (EU28, ISL, NOR, SUI). The dataset breaks down emissions in terms of pollutants (including NO₂, PM_{2.5}, PM₁₀) as well as sectors (such as road transport, residential, agriculture, natural emissions) for the years 1990 - 2014. The data consists of over 3 million records and is provided in a CSV file (> 400MB).

Intended analytics on the datasets are to observe annual trends in measured air pollutants, differences between countries and regions over time, and correlations between different air pollutant types. If helpful, we might consider to add geographical coordinates for each city. The stations and cities in [1] are associated with an affected population size and we will try to link the air pollutant concentrations to the size. We further hope to find some correlations between the reported national annual emissions [2] and the measured air pollutant concentrations [1].

Appendix - Links to the datasets:

- [1] <http://www.eea.europa.eu/data-and-maps/data/air-pollutant-concentrations-at-station#tab-statistics>
- [2] <http://www.eea.europa.eu/data-and-maps/data/national-emissions-reported-to-the-convention-on-long-range-transboundary-air-pollution-lrtap-convention-10#tab-european-data>