IA - 06 - Analyzing time series

Objectives of workshop 6 (Thursday, Oct 27)

- Quick Quiz 1 / discussion
- Debrief video capsules 15 (linear regression), 16 (Non linear Regression and Regularization);
- Student presentations and debates (*comparing models*)
- Introduce next project (see below).

Video lectures ("capsules")

Please view capsules 13 and 14 before the workshop.

Mini-project: (to be presented Thursday, November 4)

*To be completed, this project will need capsule **16** and all the previous ones.*

The RATP is the organization in charge of public transports in Paris (subway network, tramways and buses). It published a lot of data on a dedicated web portal: https://data.ratp.fr/explore

This project aims at analyzing the air quality in three subway stations.

Datasets

Air quality at subway stations Auber, Châtelet and Franklin-Roosevelt (2013-2020):

- https://data.ratp.fr/explore/?sort=modified
- Auber: https://data.ratp.fr/explore/dataset/qualite-de-lair-mesuree-dans-la-station-auber/information/
- Châtelet: https://data.ratp.fr/explore/dataset/qualite-de-lair-mesuree-dans-la-station-chatelet/information/
- Franklin Roosevelt https://data.ratp.fr/explore/dataset/qualite-de-lair-mesuree-dans-la-station-franklin-d-roosevelt/information/

Additional data sources: surface pollution measurement:

- https://data-airparif-asso.opendata.arcgis.com/
- http://www.airparif.fr/en/telechargement/telechargement-station

Airparif is a non-profit organization accredited by the Ministry of Environment to monitor the air quality in Paris and in the Île de France region (the capital city region). Airparif continuously monitors the air quality and contribute to the assessment of health risks and environmental impacts of air pollutants.

Questions

Here is a list of relevant questions that can be tackled by analyzing the published data. feel free to ask more questions or reformulate these ones.

1. Predict CO2 level in a selected station from time of day, past values, temperature and humidity.

2. Predict NO/NO2/PM10 levels (+PM25 for Auber) in a selected station from past values, temperature and humidity. Does it help to also consider current CO2 level?
3. Compare models for each stations.
4. Can current levels in two stations help predict levels in third station?
5. Can surface air quality data help predict NO/NO2/PM10/PM25 levels in subway stations?