



Funded by
the European Union

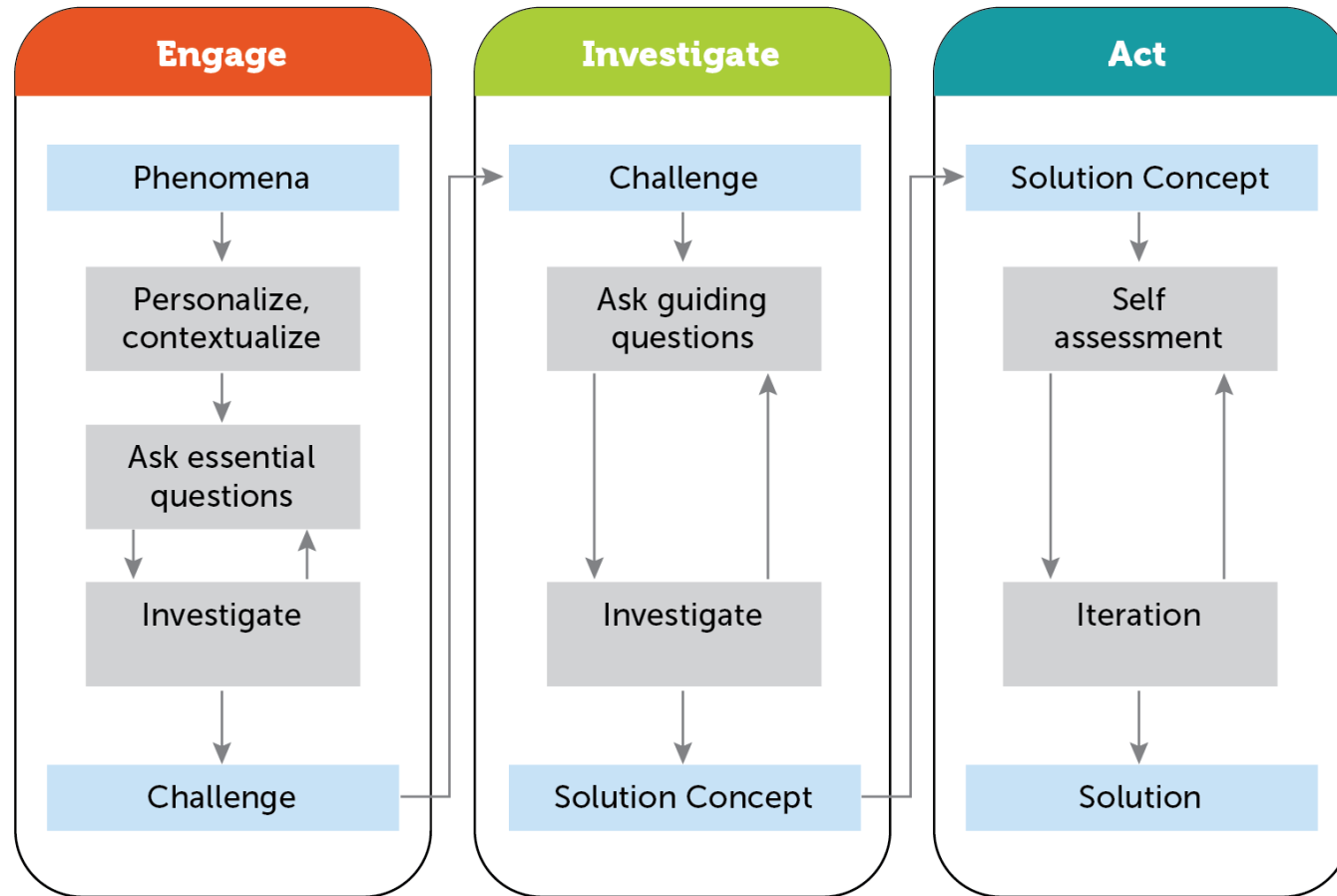


Evaldas Vaičiukynas (evaldas.vaiciukynas@ktu.lt)

<https://www.assistant-erasmus.eu>

Big Data Challenge: Urban Air Pollution

Challenge-based Learning (CBL): Work in Groups of 2-4 Students



Challenge Type: MINI



Challenge Topic: Analysis and Now Casting of Urban Air Pollution

World Health Organization has worked to ensure that health-relevant indicators of household and ambient pollution exposure and disease burden are included in the formal system of SDG indicators. SDG targets of relevance to ambient and household air pollution include:

- **SDG #3** – a substantial reduction in deaths and illnesses from air pollution;
- **SDG #11** – to reduce the environmental impact of cities by improving air quality.

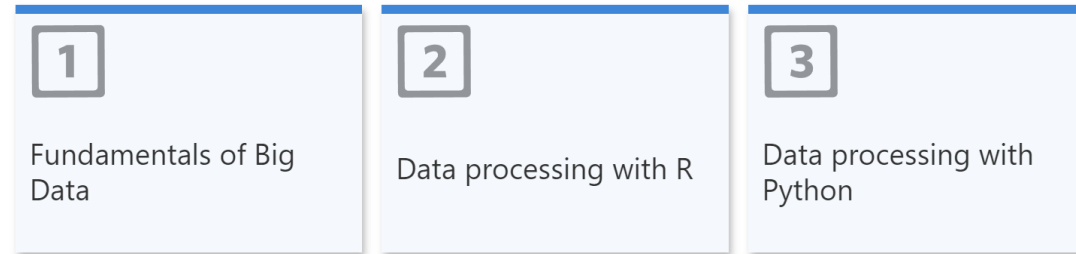
More granular data (for example, daily) would be useful for urban citizens for the purpose of monitoring situation as well as planning activities or travel trips.

Challenge Variants: Simple (1) or Complex (2)

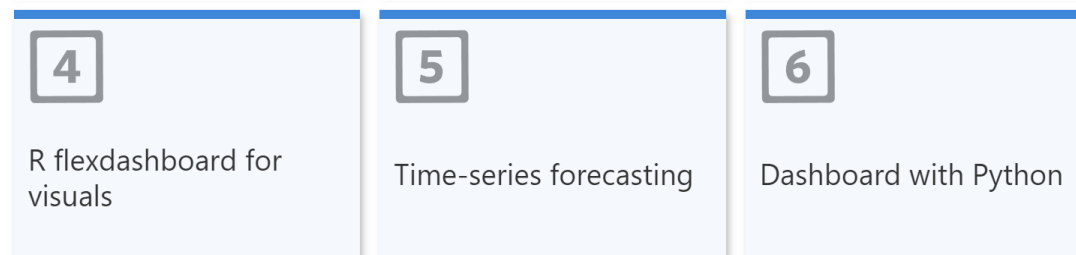
1. Finding dangerous pollution patterns. Analytics dashboard, visually summarizing historical tendencies of air pollution at a selected city. Solution should help to answer which months of the year and which days of the week have the largest levels of pollutants.
2. Analysis of historical patterns and forecasting pollution (for example, a week ahead). Analytics dashboard, allowing to make historical insights and forecast of air pollution at a selected city. Forecast should be obtained using multivariate model.

Challenge Timetable: self-learning mixed with team meetings (at least 2)

- Week #1: choose a team and plan a meeting, go through topics #1-3 and do TEST #1-3



- Week #2: the 1st team meeting, essential and guiding questions, topics #4-6, do TEST #4



- Week #3: the 2nd team meeting, programming solution, preparing group presentation

Final grade of Big Data Challenge

- Final grade (10-point scale) composition:
 - 40% = Moodle tests from learned topics with certainty-based marking (CBM);
 - 20% = inside peer (team members) and lecturer's evaluation of a teamwork;
 - 40% = outside peer (of other teams) evaluation of a mini-challenge solution.

Moodle Tests: Certainty-Based Marking (CBM)

- Certainty-Based Marking (CBM) aims to improve both the reliability of computer-based assessment and the effectiveness of student selftests as learning tools.
- When answering question, i.e. selecting A/B/C/D, student needs to select confidence: Low/Mid/High

Certainty Level	Mark if Correct	Penalty if wrong
C=1 (low)	1	0
C=2 (mid)	2	-2
C=3 (high)	3	-6
[No Reply]	0	0

- All correct answers can result in scores such as 10, 20 or even 30, but it's enough to exceed 10 😊

Internal Evaluation: lecturer's grade slightly adjusted by team self-reflection

- One member of the team uploads results. Recommended formats: .pptx, .pdf, .html, .zip
- All members can see uploaded results and must grade team colleagues: how helpful he/she was?
- Aspects that need to be evaluated individually:
 - **Participation:** attended all meetings/discussions.
 - **Fulfillment:** completed all assigned tasks on time
 - **Involvement:** cooperation and communication.

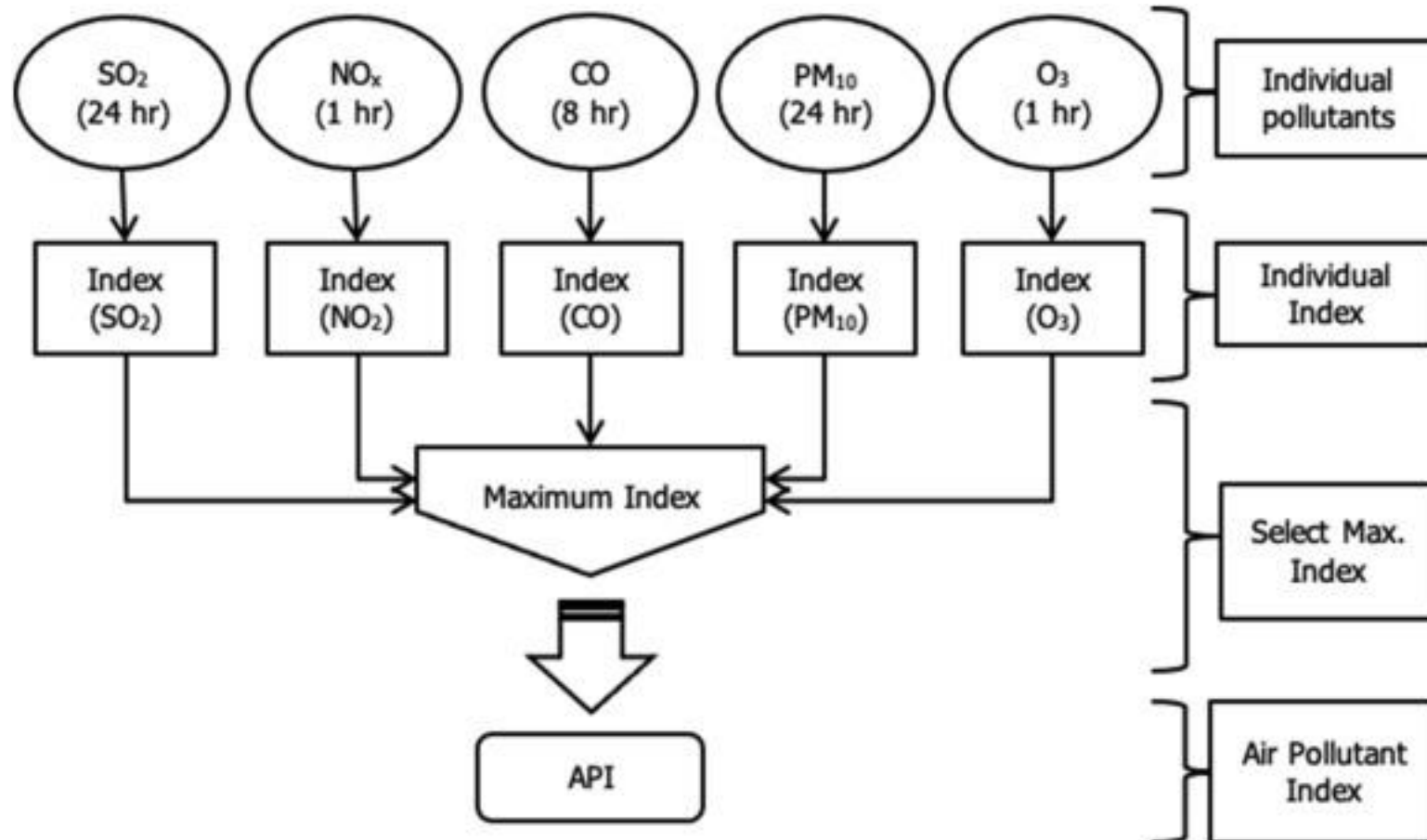
	0	1	2	3	4	5	6	7	8	9	10
Participation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Fulfillment	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Involvement	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

External Evaluation: Moodle workshop (hackaton) for uploaded solutions

- Each team member needs to upload result of his team. Recommended formats: .pptx, .pdf, .html, .zip
 - Submissions close: (date will be specified)
 - Assessments close: (date will be specified)
- Each student will get assigned 3 uploaded solutions from members of other teams (team can repeat).

CRITERION	LEVELS			
Architecture of solution:	nothing was found for this criterion	plots in slides, Power BI, .html from Colab or Markdown	.html from R flexdashboard, Python jupyter-flex or Panel	link to dashboard by Shiny, Dash, Panel, Streamlit, Gradio, etc.
Objects of analysis:	nothing was found for this criterion	a single city is analysed	2-3 cities are analysed	any city can be selected
Scope of analysis:	nothing was found for this criterion	insights made from analysis of historical data	historical insights, univariate forecasting, no backtesting	historical insights, multivariate forecasting, backtesting
Visualizations:	nothing was found for this criterion	only one type of plot was used	several types of plots were used	plots, drop-down lists, tabs, sidebar, etc.
Questions/conclusions:	nothing was found for this criterion	single question/conclusion	2-3 questions/conclusions	many questions/conclusions

Challenge Formula: $API = \text{MAX}(PM10, PM2.5, NO2, O3, SO2, CO)$



Thank you for attention

