



## Group 2 - Air pollution in Poland

### Current state of the project

	W1					W2					W3					W4		
	8	9	10	11	12	15	16	17	18	19	22	23	24	25	26	29	30	31
1. Define the problem				✓														
2. Acquire data							✓											
3. Explore and understand data								✓										
4. Data wrangling												✓						
5. Data analysis																		
6. Model and develop algorithms																		
7. Communicate and visualize results																		
8. Finishing product and portfolio																		

### Adjusted objectives:

1. Analyze the relative importance of different sources of air pollution in Poland in 2019.(almost done)
2. Investigate the relationship between traffic and air pollution using machine learning (ML). ()
3. Visualization of the relationship of air pollution and respiratory diseases
4. Visualization of the Polish resident perception of the air pollution



## Group 2- Air pollution in Poland

# Problems and solutions

Problems	Solutions
Data in different administrative level (county, province, national).	Aggregating data from lower to higher administrative units.
Data with mismatched time frames. For e.g. Facebook data for 2023, pollution data for 2019	Choosing timeframe that has the most available data.
Data inconsistency due to different measurement units (PM10 and/or PM2.5).	Changing the pollutant (CH <sub>4</sub> , NO <sub>x</sub> , NO <sub>2</sub> , etc.)

### Successes:

- Successfully collected data for most factors that we want to analyze.
- Split complex research objective to feasible ones.



Group 2- Air pollution in Poland

## Questions to advise on from the other groups (or tutors)

- Is "**distance**" from a pollutant source sufficient for modeling the relative contribution of air pollution, or is it also necessary to include the "**emission rate**" of the pollutant source?