

# project-02

October 18, 2020

## 1 EPA-1316 Introduction to *Urban* Data Science

### 1.1 Project 2: Identifying the Health Vulnerability in a City

**TU Delft Q1 2020 Instructor:** Trivik Verma **TAs:** Aarthi Meenakshi Sundaram, Jelle Egbers, Tess Kim, Lotte Lourens, Amir Ebrahimi Fard, Giulia Reggiani, Bramka Jafino, Talia Kaufmann  
**Computational Urban Science & Policy Lab**

---

COVID-19 pandemic has revealed very clearly how our health infrastructure is not resilient. For global health policy, the challenges for decision-makers are not only to understand which communities are vulnerable and where they are located but also assess how resilient the city is to various shocks (natural disasters, evacuations, health emergencies, etc). People who are exposed the most to certain events like extreme heat are also the most vulnerable to it because they are mostly low-income single households with kids and elderly with mobility issues [Ref]. The lack of greenery in their residential areas may even amplify such shocks. It is not an easy task to build infrastructure in a city that is going to be optimally located for all citizenry. Knowledge of spatial and temporal variability of infrastructure use can provide policy-makers with insight about potential interventions.

Project goals: You are tasked to identify trends in ambulance calls across a city and/or in time and measure its relationship with indicators of socio-economic status of households, neighbourhoods or regions.

#### DATA

1. We will provide you with an emergency **calls** dataset (ambulance, firefighter, police, and coastguards) of The Netherlands collected from January 2017 to September 2020. That also includes calls made to these services during the period of lockdown and release cycles in the pandemic. **Since the data is not open-source yet, please send an email to me requesting access to it. Kindly do not share it with anybody outside of your project group.**
2. A few selected OpenData resources [Den Haag Cijfers](#) Statistics Bureau of The Netherlands [CBS](#)

The calls dataset has the following indicators and was obtained from [112-Nederland](#),

pmeId - Related to the id on the website  
pmeTimeStamp - Time of call registered  
pmeProtocol1 - P2000 protocol related info

pmeProtocol2 - P2000 protocol related info  
 pmeTarget - P2000 protocol related info  
 pmeMessage - Original message  
 pmePrio - Priority of the message  
 pmePrioLevel - Conversion of priority to a normalized level  
 pmeDienst - A=Ambulance, B=Brandweer, P=Politie  
 pmeStrippedMessage - Cleaned Message of the call  
 pmeStraat - Street  
 pmeHouseNumber - Mid-point of the street closest to location of call (for data privacy)  
 pmeHouseNumberAdd - Addition to house number  
 pmeANRoad = A=highway, N=Provincial Way  
 pmeHectometerpaal = Location on A or N way  
 pmeZip - Zip code of the area  
 pmeRegionName - Name of region  
 pmeLatitude = wgs84  
 pmeLongitude = wgs84  
 pmeHash - Ignore internal indexing (for collection purposes only)  
 pmeGeoAccuracy = H=Housenumber, S=Street  
 pme\_strId = Street Id  
 pme\_wplId = Woonplaats Id  
 pme\_gemId = Gemeente Id  
 pme\_proId = Province Id  
 pme\_vrgId = Veiligheidsregio Id  
 pmeCapCodes = Capcodes  
 pmeLifeLiner - True/False (Extraction by air)  
 wplName = Woonplaats  
 gemName = Gemeente / Municipality

## High-level project goals

1. Explore, Investigate, and Visualize various factors of the emergency services call data: geographic differences and time trends. Example: compare time trends, city trends, or both. You may also choose one city and focus on that.
2. Describe and incorporate additional data sources that you will use to help you understand or use the calls data. Example: demographic data from municipalities in The Netherlands or CBS data as used during assignments.
3. Train and evaluate models using the calls data (either as predictors or as the primary response in some fashion) and your data sources.
4. Or, Cluster calls based on some socio-economic indicators of urban regions around The Netherlands.
5. Use and interpret your models to discuss the causes and correlations of or effects due to differences in calls across cities or within a city of your choice. (You may use the Pyrosm package to find locations of health infrastructure and other amenities in a city and use the Osmnx package to find road-distance to nearest locations from a certain region in a city).