**Assignment 1 – Water quality Amsterdam**

When considering data sources available on water quality and traffic in Amsterdam it's hard to find holistic and detailed data sets online. However, for this assignment we have considered a combination potential data sets available and other sources for information which might create a complete image or at least a starting point for more specific measurement in water quality later.

When consider the need to not interfere with water transport we haven’t found easily accessible raw data however we have found a pdf map which shows no go zones in pink (<https://sloepdelen.nl/downloads/information_amsterdam_en.pdf> ). Furthermore, to track main traffic routes throughout this website which track the pathways of ships in the world and also Amsterdam can create an idea of what areas are frequently frequented (<https://www.marinetraffic.com/en/ais/home/centerx:4.910/centery:52.375/zoom:13> ). Through a scraper tool one could also over time derive all of the data through time which would allow a more detail consideration for the event managers to consider. However, neither of these data sources are directly usable in python in their current format.

However, to consider feasibility of the event the main focus point is water quality. There are several different monitoring points. Primarily we found specific data about water quality as mentioned on websites thus hard to pull together as a holistic data set (<https://www.amsterdam.nl/toerisme-vrije-tijd/parken/sloterpark/waterkwaliteit-sloterplas/> ). Similarly, Waternet has a map which shows the oxygen and salt level in the waters surrounding it AGV. However again this map is not downloadable and thus the information would have to be manually entered. Thus, here some scraping tool might be useful.

However, for the sake of this exercise we have focused on data sets which are downloadable and while not perfect might give an image of areas in Amsterdam with clean water as well as the water quality in the ground water in certain other areas which might be used to infer overall water quality and available locations.

**Data on swimming locations**

Our primary starting point for water quality is the map of possible swimming locations in Amsterdam (<https://maps.amsterdam.nl/zwemwater/>). This data set can be downloaded in several formats but predominantly csv and GEOjson. Both are easily legible also without a program as they can be opened in excel and an online browser. However, because this information is geographic in nature, for later processing the geographic format GEOjson might make more sense with python. The file is point data but set up as a standard number file. The file formats are both quite straightforward to open and geopandas or pandas are some of the most straightforward libraries in python which can likely read the data. For analysis plotly might be the best option to graph but for geographic analysis one should remain with geopandas. To situate the points withing a largely reference using mapbox plug in with plotly, or omnx data from open street map might be useful extras,

Overall, the file contains, name of the place, category of what it is, an ID code and the coordinates of the location. Important to consider this data set contains both indoor and outdoor locations which are safe for swimming so in categories one should filter for outdoor locations and canal. Furthermore, when looking at the file it is semicolon delimited and has commas thus when using pandas an assuming a csv file one either needs to specify definition or resave it from excel or set semicolon as the dividing feature when reading it.

**Data on surface water quality**

The secondary data set is on groundwater quality in Amsterdam. This data set is again a csv file and can thus be read alone and predominantly opened through pandas. It has information for location but no geographic reference thus making it unsuitable for geopandas. Plotly or matplotlib might be nice libraries to create graphs around the data. It si temporal data which describes the water quality at one specific occasion in time.