

CASE STUDY

Drugs in municipal wastewater
in selected European cities

OPEN SOURCE DATA PROJECT by Viktoria Dyk



GENERAL OVERVIEW

SCENARIO

Wastewater analysis is a rapidly developing scientific discipline with the potential for monitoring illicit drug use. The European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) works closely with the Sewage analysis group (SCORE) which has been collecting data indicating the amount of drug residues in wastewater in major European cities since 2011. In this open source data project we explore geographical and temporal trends in illicit drug use covering cannabis, cocaine, amphetamine, methamphetamine and MDMA (ecstasy) and examine the relationship between drug usage and wealth (GDP per Capita).

DATASETS

- 2 datasets including information about measured values indicating the amount of drug residues between 2011-2021 and one about the sites (wastewater treatment centers) where the measurements have been conducted. Datasets accessed from the [European data portal](#)
- A customised dataset including GDP per Capita for the examined years and selected European countries, which have cities participating in the wastewater analysis study. Dataset accessed from the [World Bank Databank](#)

GENERAL OVERVIEW 2.

TOOLS USED

- Python/ Pandas for data cleaning and manipulation
- Python / Matplotlib, Seaborn, Folium for data visualisations
- Python / Scikit-learn for regression analysis and clustering
- Python / Statsmodels for time series analysis
- IDE: MS Visual Studio Code
- Tableau for final presentation

LINKS

- Tableau [Storyboard](#) (designed for full screen)
- [Github repository](#) with scripts, datasets and additional documents

CHALLENGES

- Following data limitations made it difficult to conduct accurate or wider spectrum analysis:
 - not all European countries provided data
 - some missing data (some cities started providing data for the study in a later year)
- Finding proper variables to explore relationships:
 - GDP per Capita data was merged to our measurements dataset but it required significant data manipulation so that it has a format that can be merged
- Due to special characters (e.g ä, ü, ö) in some city names, manual work was required to get Tableau recognise them as geographic items

TASKS PERFORMED

- 1
- 2
- 3
- 4

Performed data cleaning, consistency checks and merged the 3 datasets in a Jupyter Notebook

Explored relationships among variables, Performed regression analysis, geospatial analysis, time series analysis and clustering

Created visualisations and interactive dashboards in Tableau to answer key questions

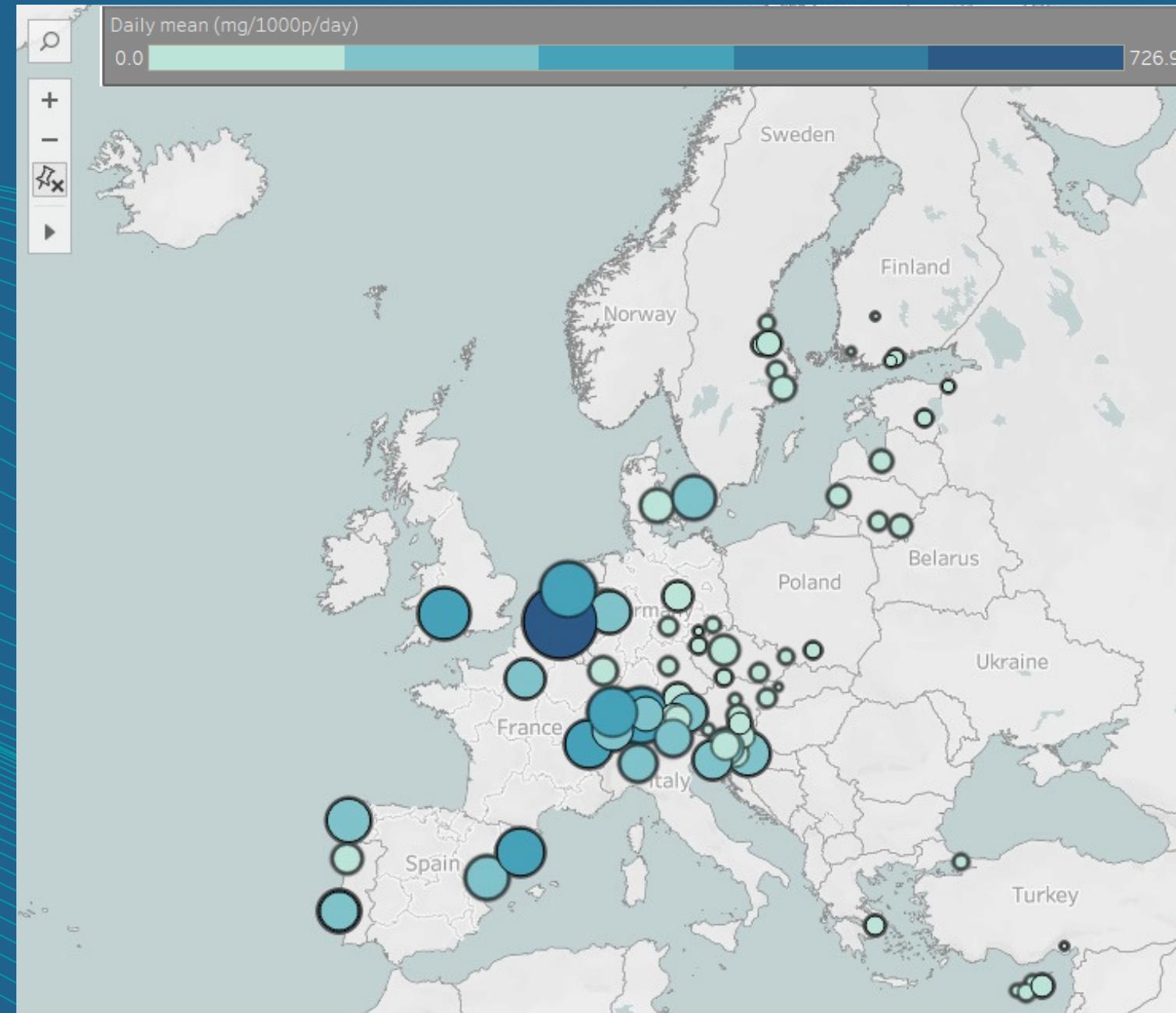
Storytelling in Tableau: presented results to the stakeholders

KEY QUESTIONS:

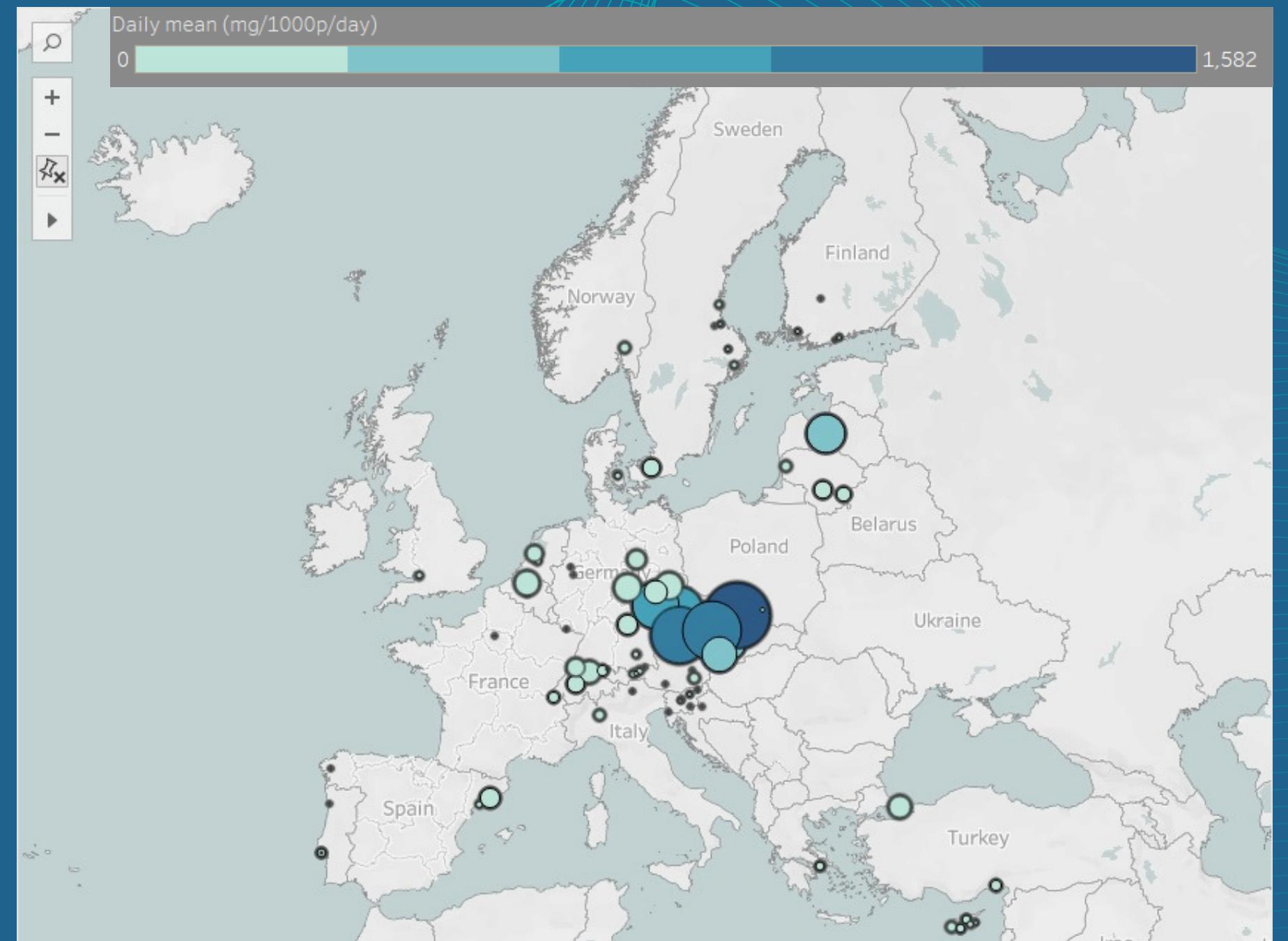
- Which illicit drugs are the most used in which area of Europe?
- What temporal trends can be observed?
- How did the pandemic affect the drug use in 2020-2021 when social events and interactions were strongly limited?
- Is there a connection between illicit drug use and the country's wealth (GDP per capita)?

RESULTS/ANSWERS

Extracts of the interactive spatial dashboard



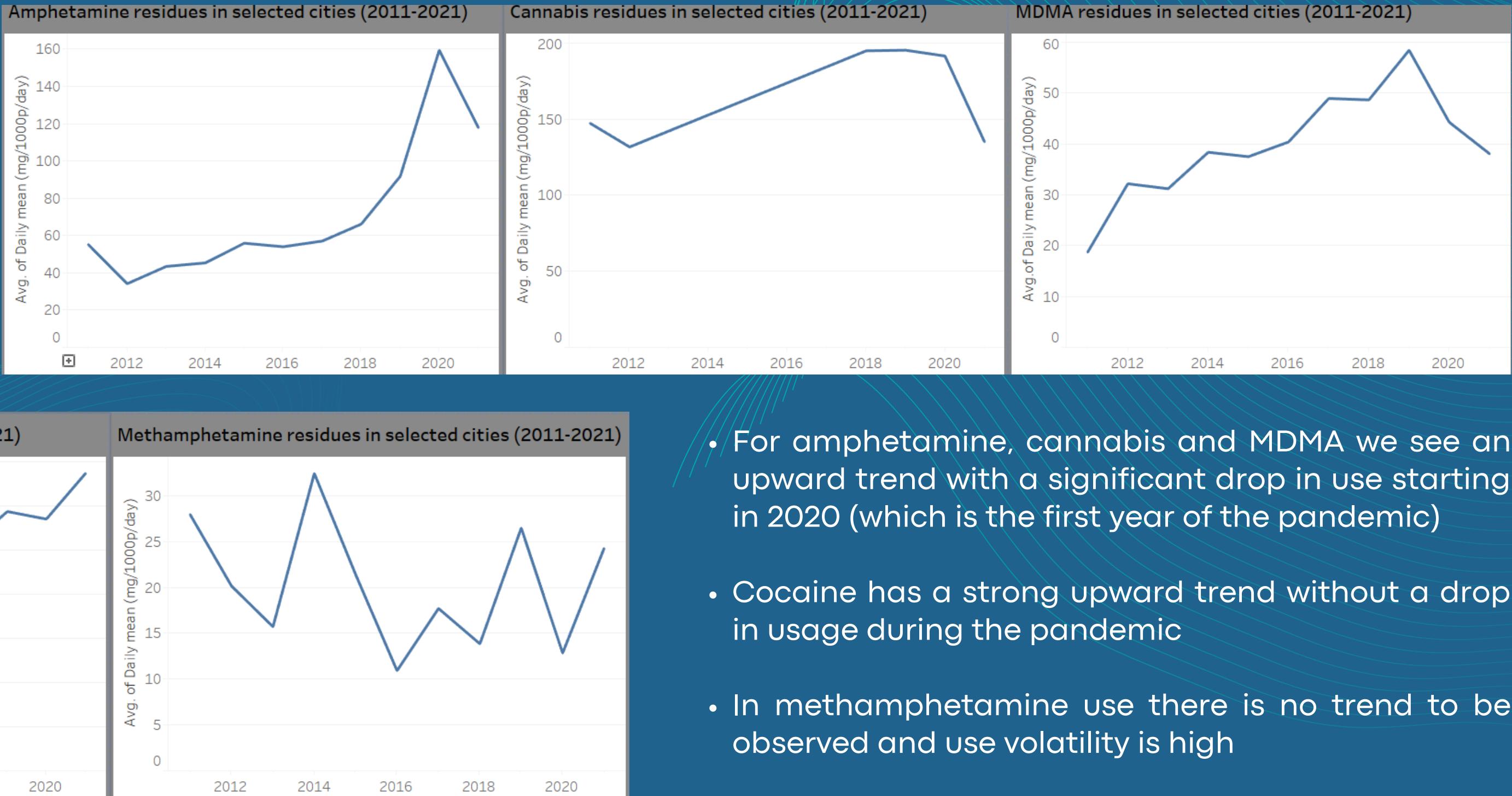
Cocaine use in 2021 - highest in western
and southern European cities



Methamphetamine use in 2021 - historically
highest in Czech Republic and Slovakia

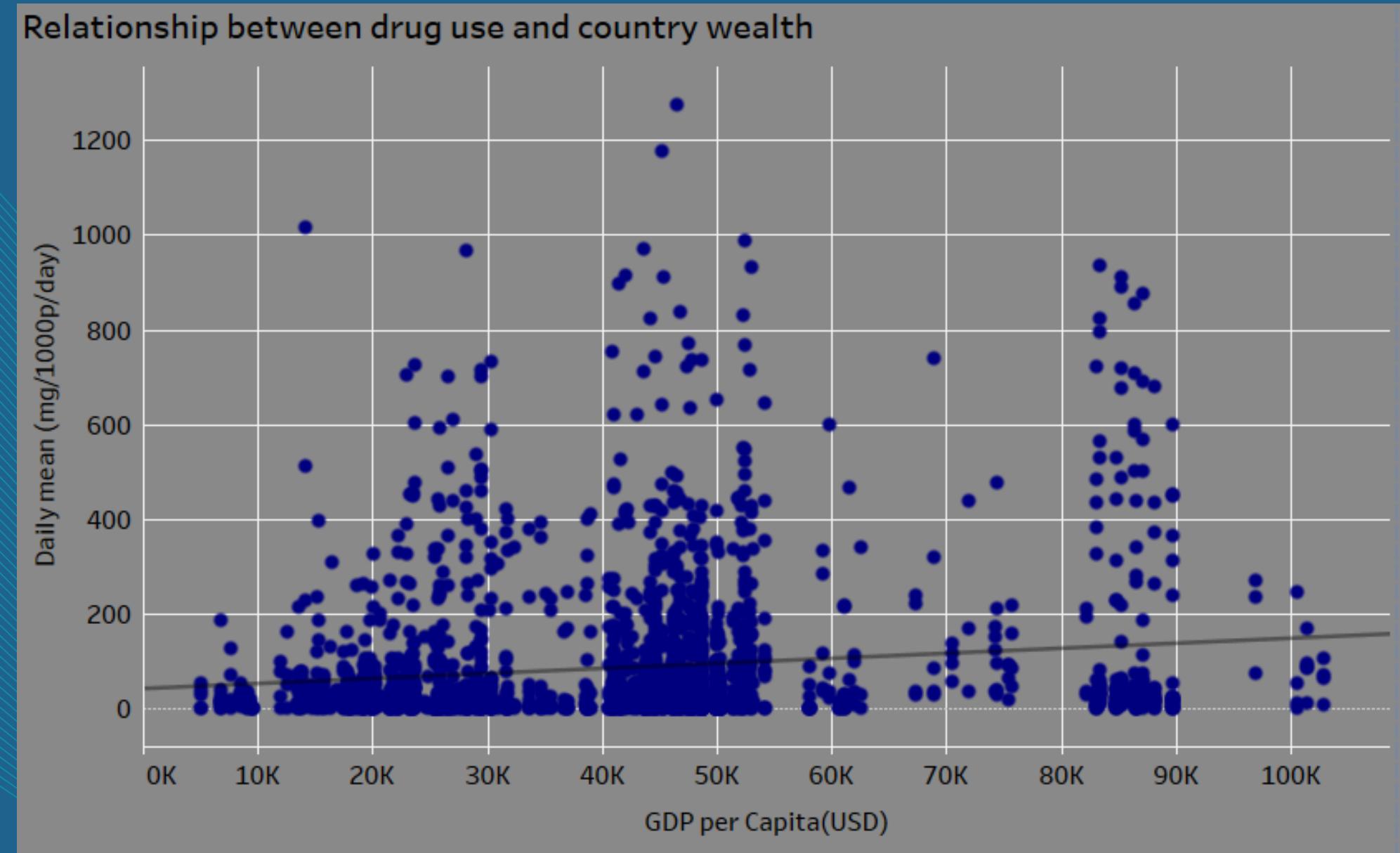
RESULTS/ANSWERS

Temporal trends



RESULTS/ANSWERS

Drug use vs. GDP



The trend line (in grey) in the above scatterplot indicates a weak positive relationship between drug use and the country's wealth. Wealthier countries tend to have higher use but this correlation is weak, and there's a great degree of variability.

KEY FINDINGS

- Geographic and temporal trends vary strongly across different metabolites
- During the pandemic we did not find a drop in use for all the metabolites, only for 3 of them
- There is a weak relationship between the drug use in a city and the wealth of that country but the relationship is not linear and other factors must also influence the drug use