Soil geochemistry take-home exercise

The objective of this exercise is to infer and plot a continuous spatial distribution of trace metal concentrations for an entire area from a set of soil samples taken in the field.

This exercise should not take more than 1-2 hours to complete.

Task

Geochemical soil surveys are commonly used in mineral exploration. Workers walk or drive around a (potentially large) area and collect soil samples at specified locations. These samples are analyzed to determine the concentration of certain trace metals. The metal content in soils depends both on the composition of the bedrock directly beneath the soils and by surficial processes like fluid flows.

Your task is to create a continuous map of your best guess for the soil metal concentrations across the area, not just at the sampled locations. This is useful to visualize and can be used as an input to other data science workflows.

Data

We have provided a set of data that is representative of a typical geochemical soil survey. It is a CSV file with one row per soil sample from some area of exploration interest. Each sample has a location (x,y) and a measured concentration of copper (Cu) and/or cobalt (Co) given in units of "ppm" (parts-per-million) or "pct" (percent, or parts-per-hundred).

What to submit to us

1. **Two image files**, one showing a map of Cu concentration, the other Co concentration. Both should show estimates of soil concentrations throughout the region.

A static map or screenshot is fine. Please include plot titles, axis labels, and scale bars where appropriate.

2. All **code** written to create the maps with enough information/documentation for us to be able to run it ourselves. Please use Python for this task as it is the dominant language used for data science at KoBold. You are welcome to submit a script or an interactive notebook. You are welcome but not required to use the packages numpy, matplotlib, pandas, scipy.interpolate, and/or pytest.

3. A short **description** (~2-3 paragraphs) discussing your solution. Explain the assumptions you made to produce these maps, any limitations of the method, and resulting uncertainty in the results.

What we're looking for

- **Correct answer**: the maps should clearly show your best estimate of soil metal concentrations across the region and should accurately reflect the underlying measurements.
- **Clear communication/documentation**: Please include a writeup describing how you approached this problem and the steps you took to get to the answer, including whatever data exploration you did to formulate your answer.
- **Coding best practices**: Please write at least one function with a docstring and unit-test for functionality. Code should be clear to understand, run, and modify, and should demonstrate professional coding standards.