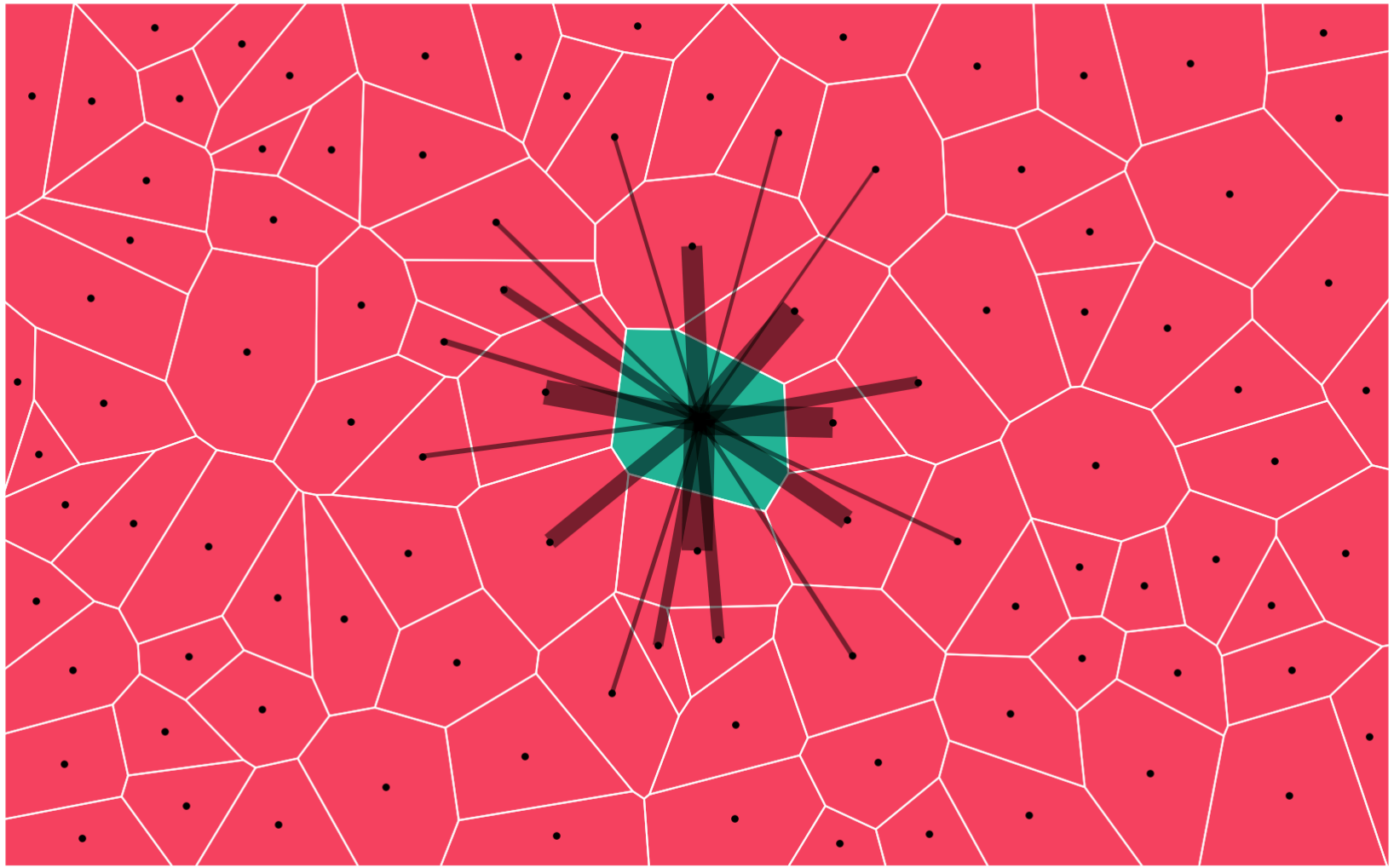
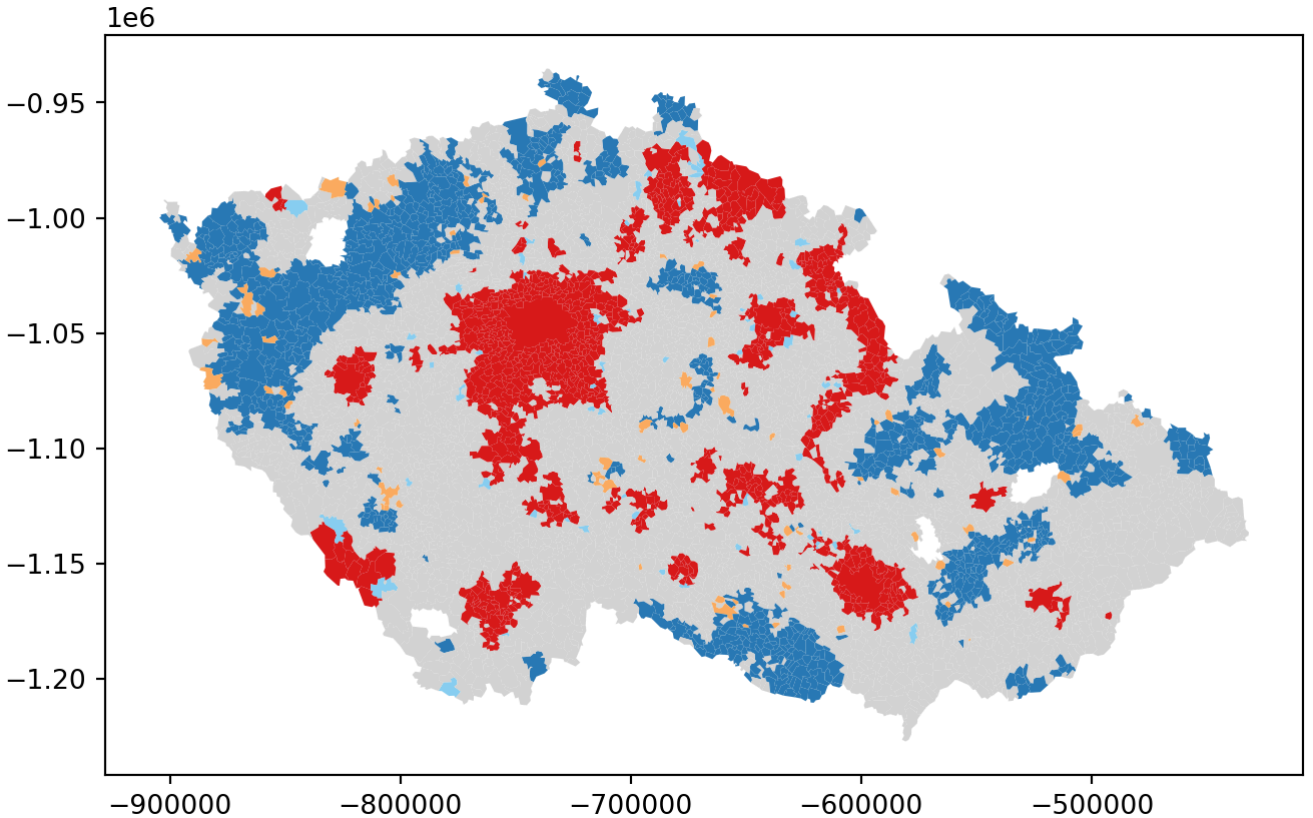
# Spatial relationships and how to use them

> **Everything is related to everything else, but near things are more related than distant things.**

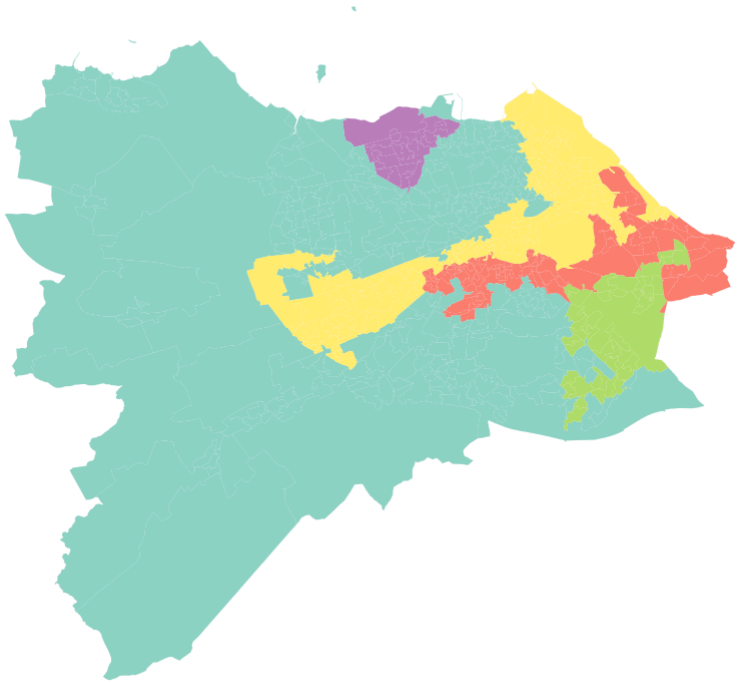
**Waldo Tobler (1970)**



Spatial calculations.




Spatial autocorrelation.



Spatially constrained clustering.

# How to include space in statistics?

We need its mathematical representation

**geometries are not**

graphs are

# What is a graph?

**a data structure that consists of a set of objects called nodes and a set of connections between them called edges**

# Matrix

$$\begin{pmatrix} 0 & w_{12} & \dots & w_{1N} \\ w_{21} & \ddots & w_{ij} & \vdots \\ \vdots & w_{ji} & 0 & \vdots \\ w_{N1} & \dots & \dots & 0 \end{pmatrix}$$

# Matrix

$$\begin{matrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{matrix}$$

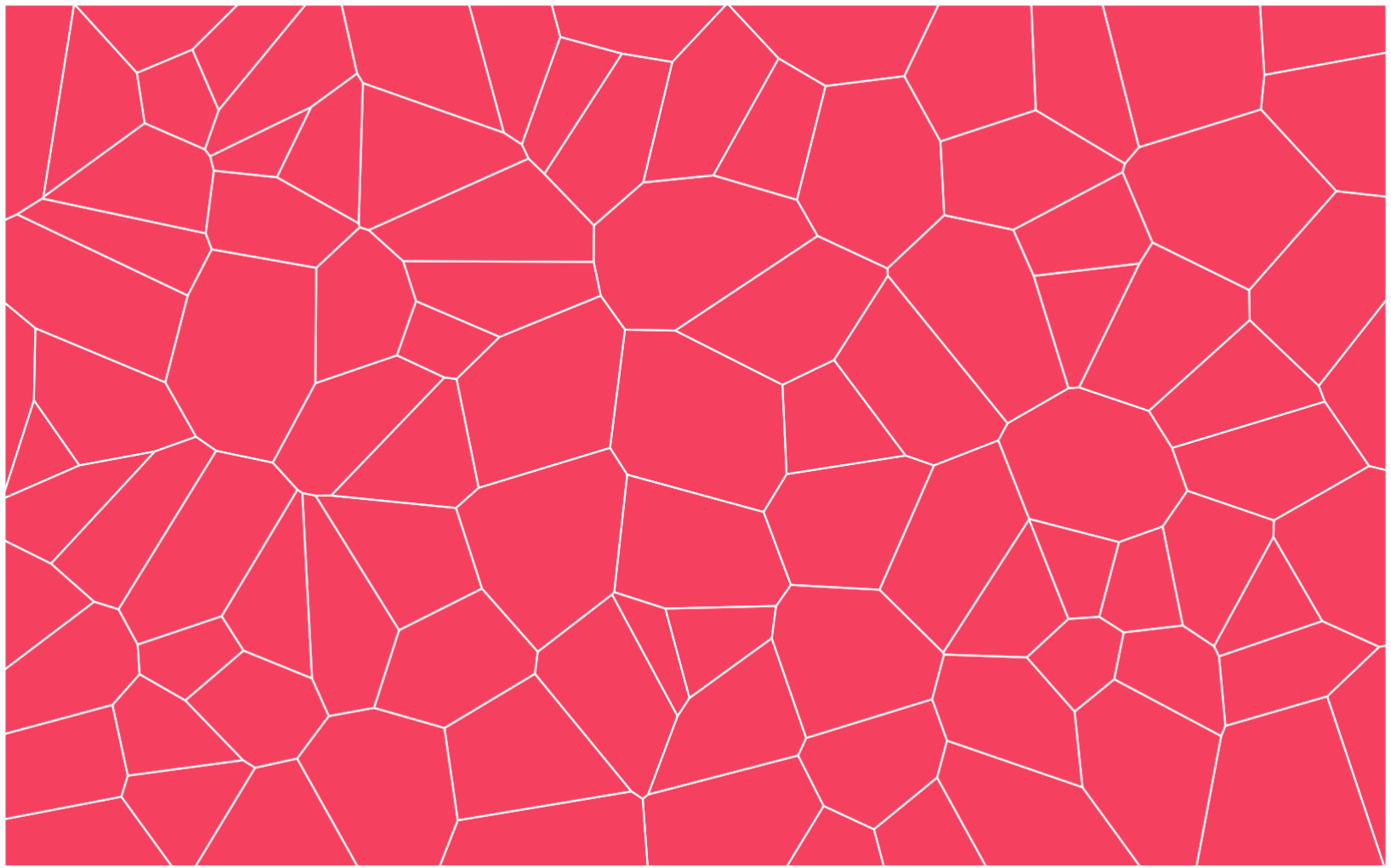
mathematical structure capturing spatial relationships

# What is a neighbour?

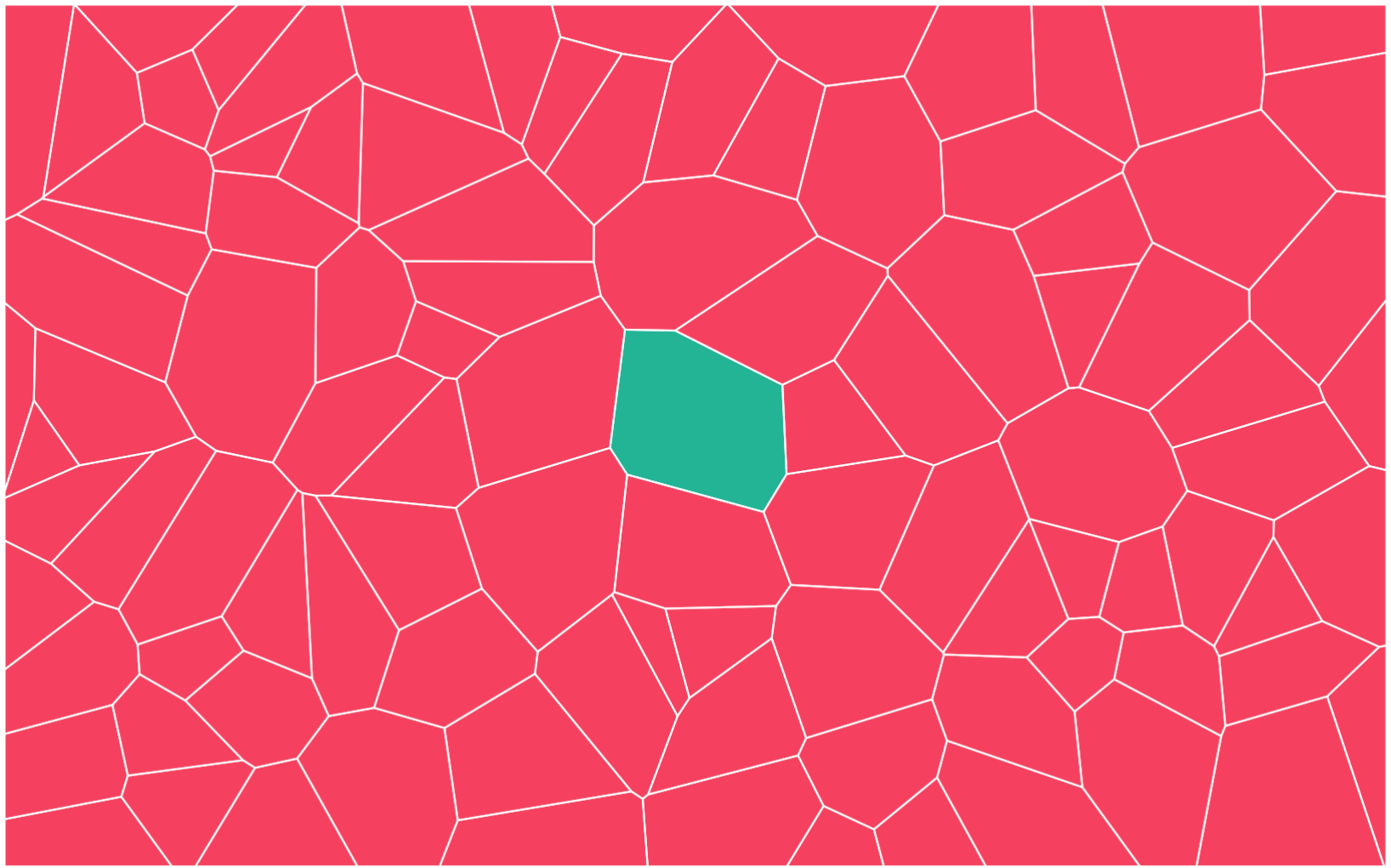
it depends

# Contiguity

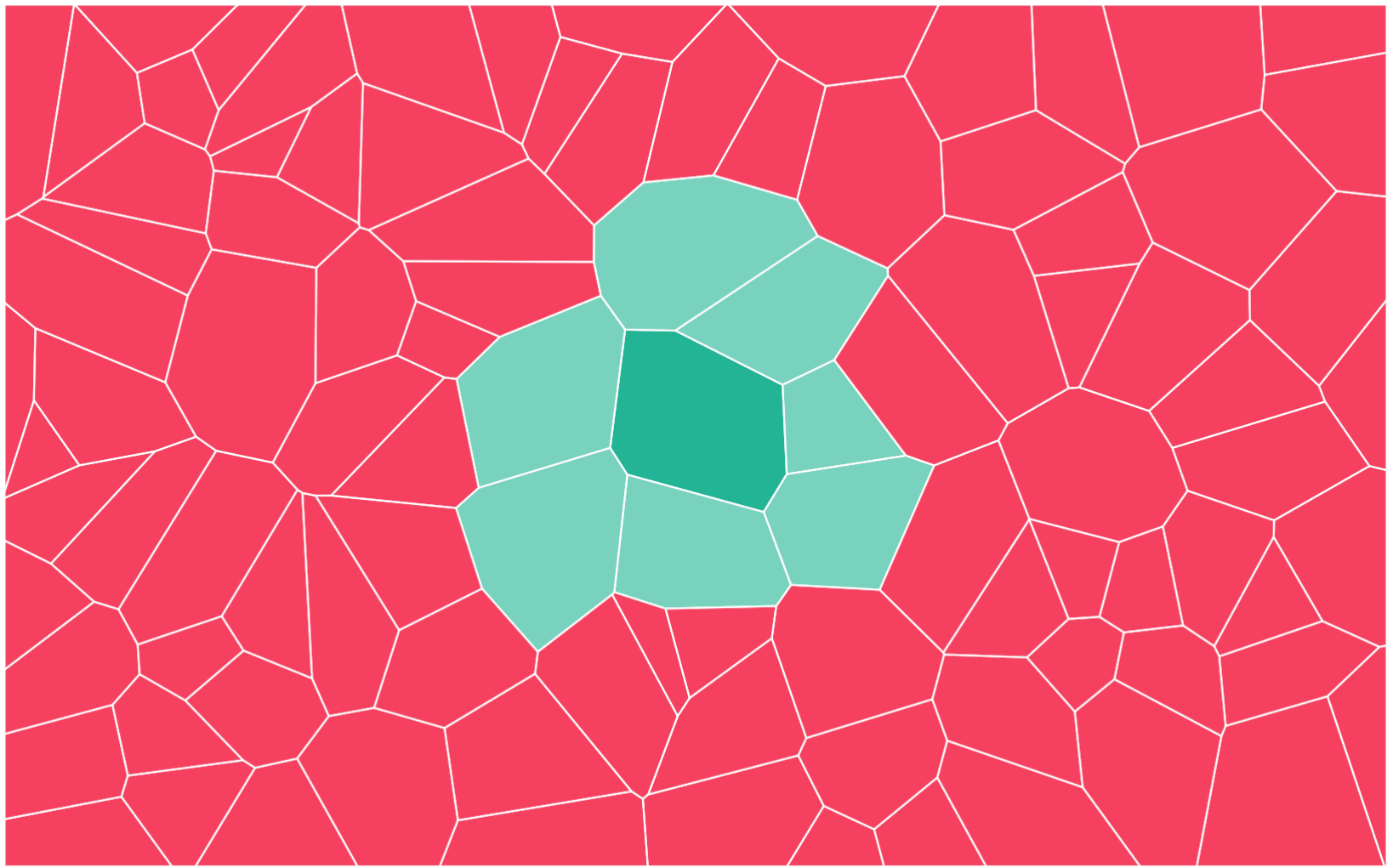
**Two geometries are considered neighbours if they share at least one vertex (*queen contiguity*) or one edge (*rook contiguity*).**



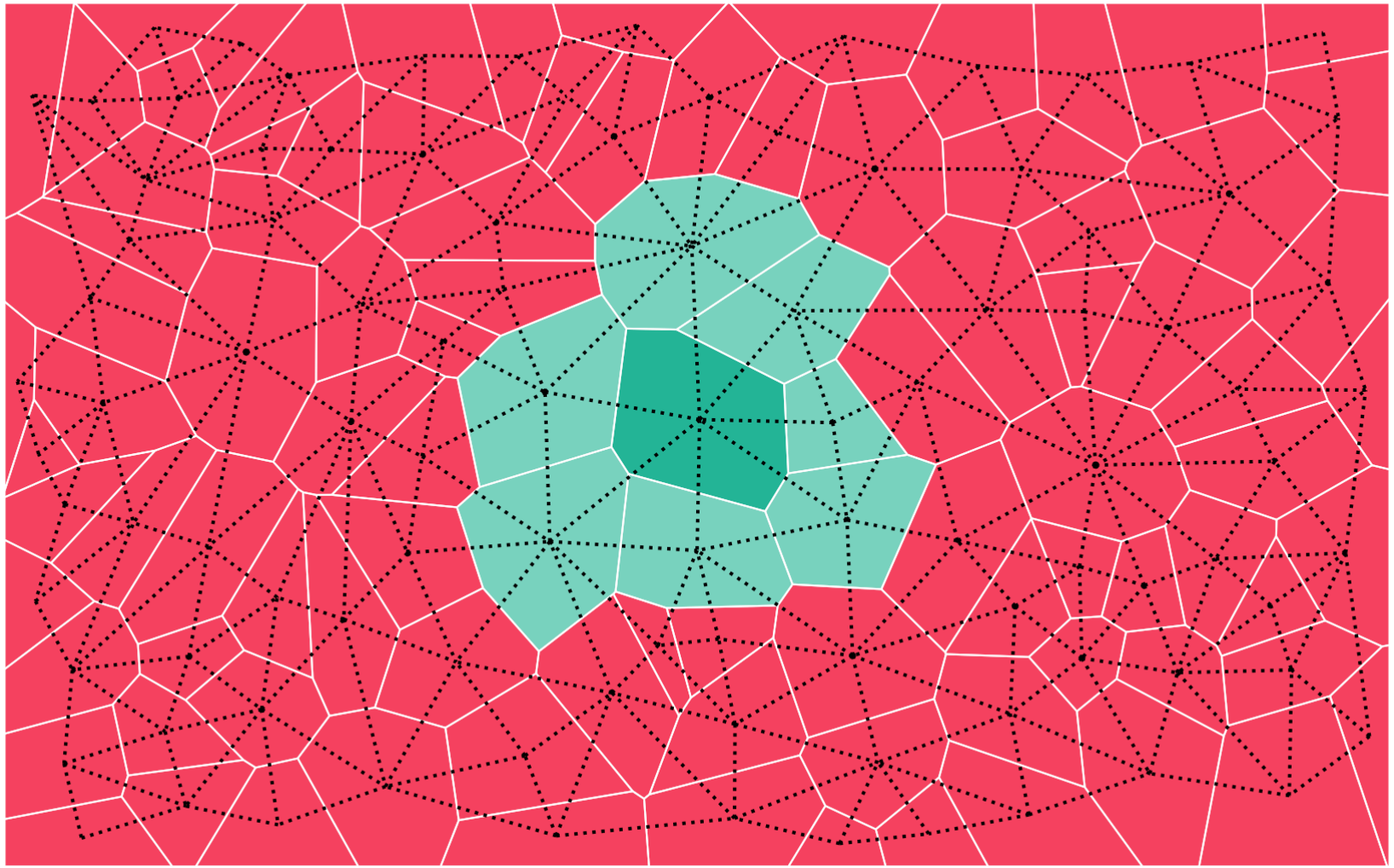
Starting with some polygons…



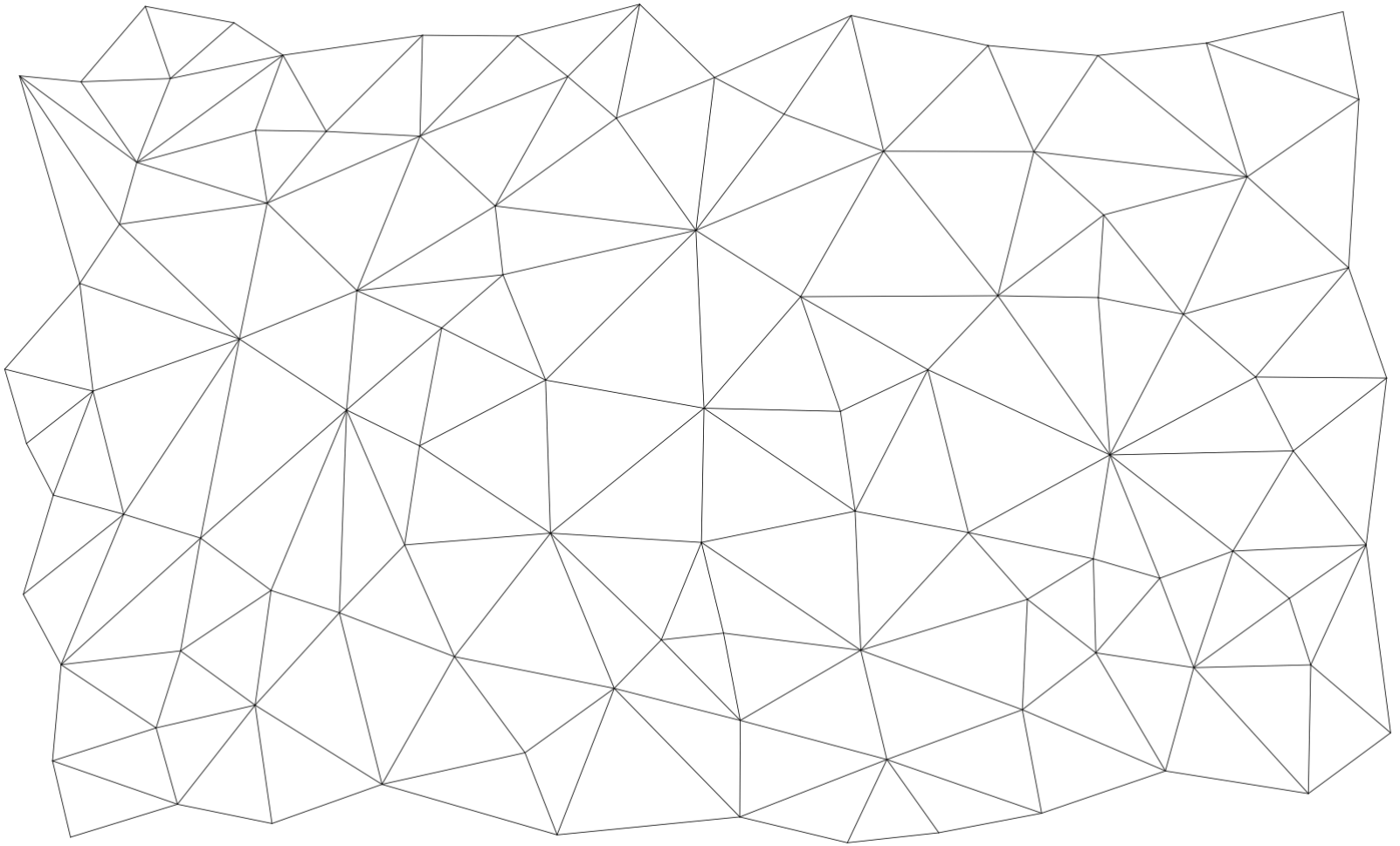
Focus on this one. Which are its neighbours?



Based on (rook) contiguity, these.



Here is a complete graph.

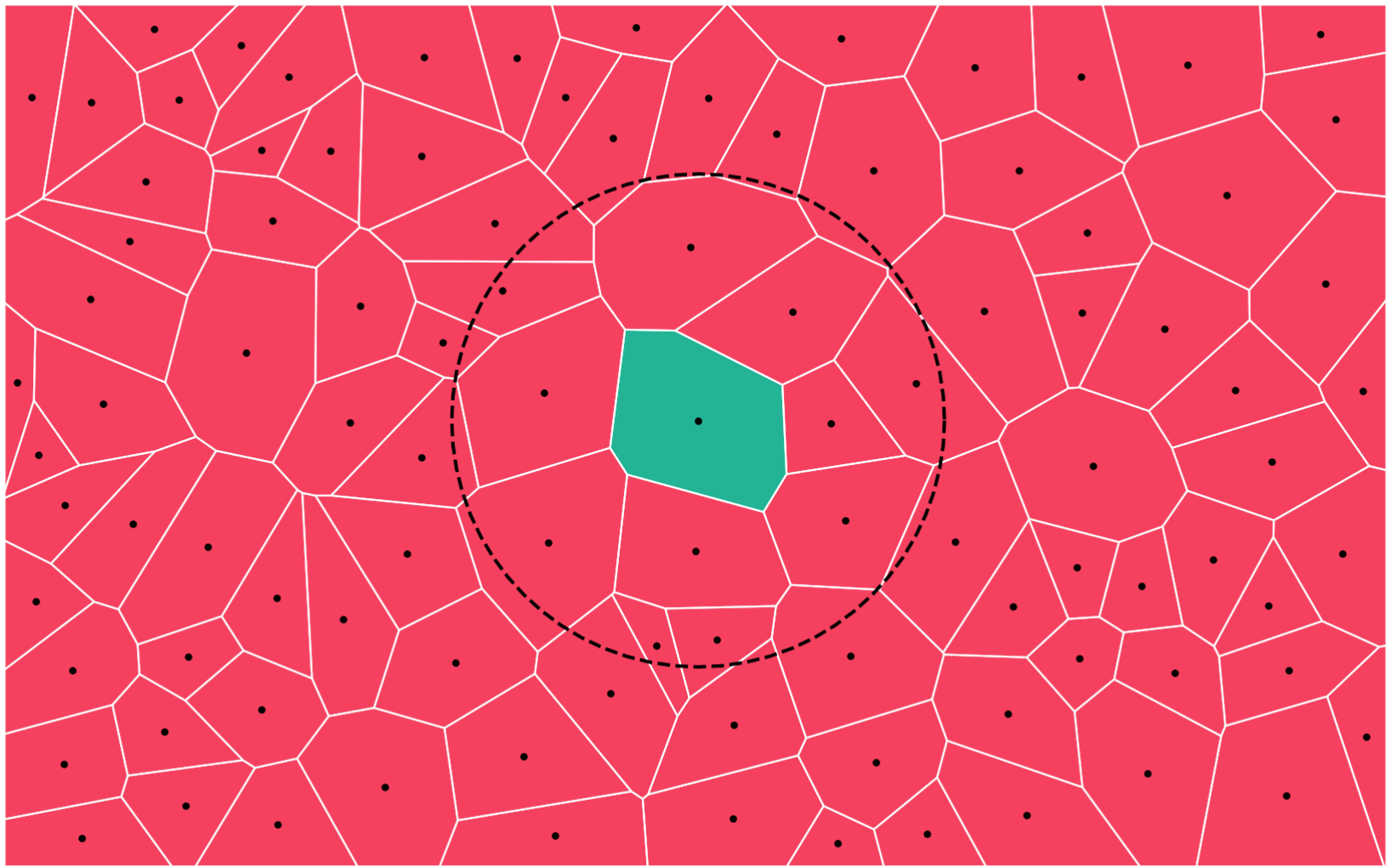


Here is a complete graph.

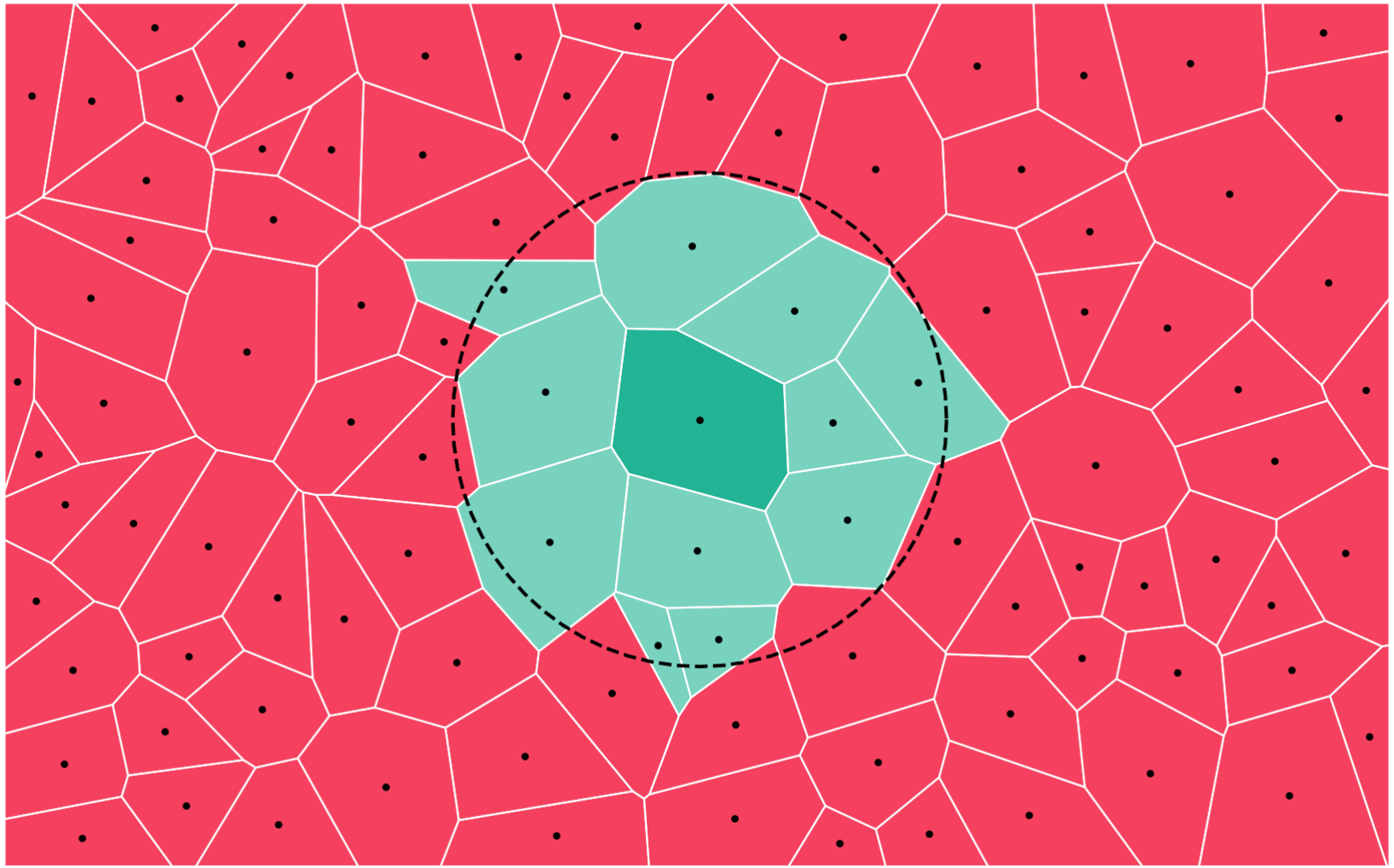
# Distance

## Two geometries are considered neighbours if they lie within a set threshold from each other

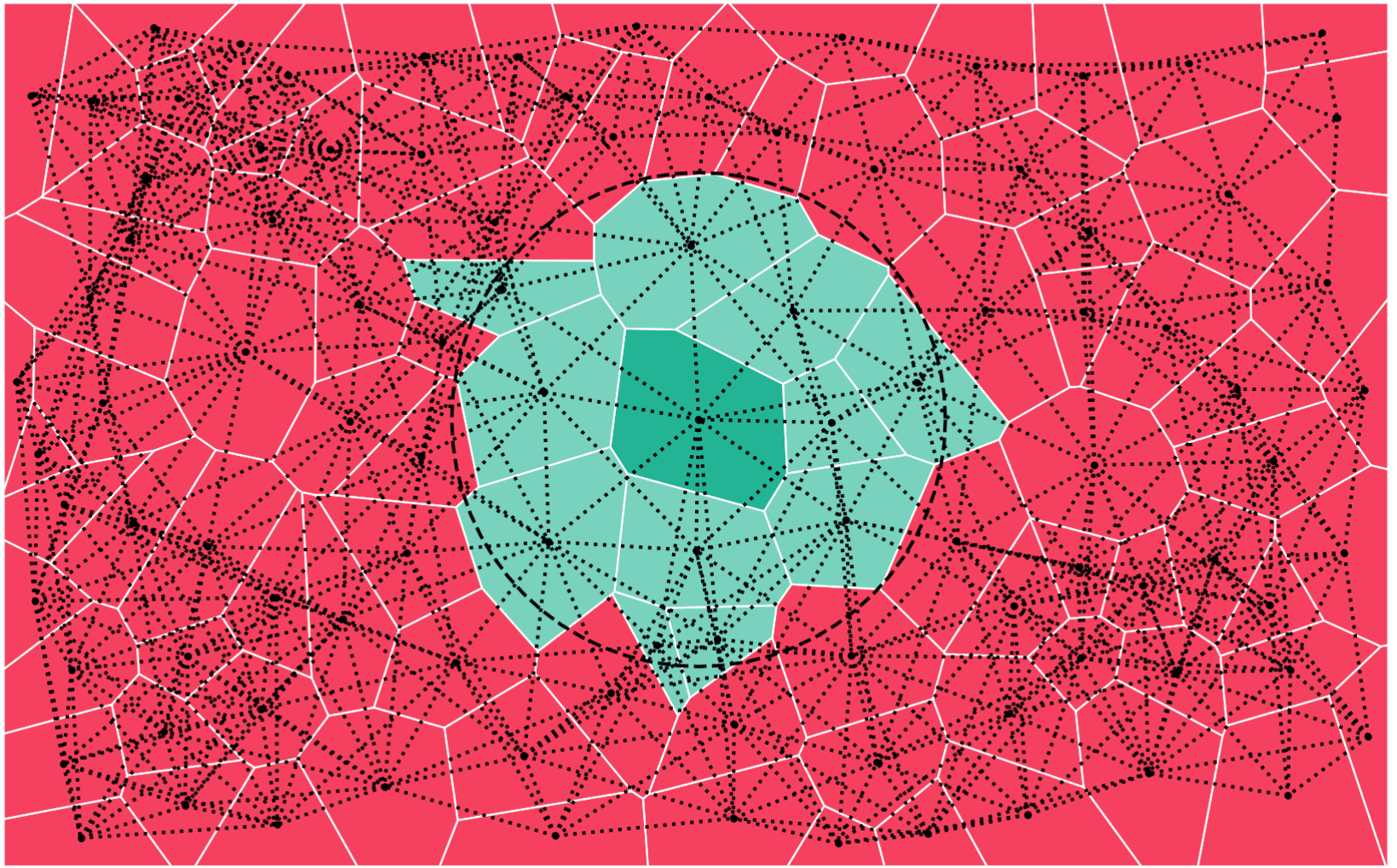
### Distance band



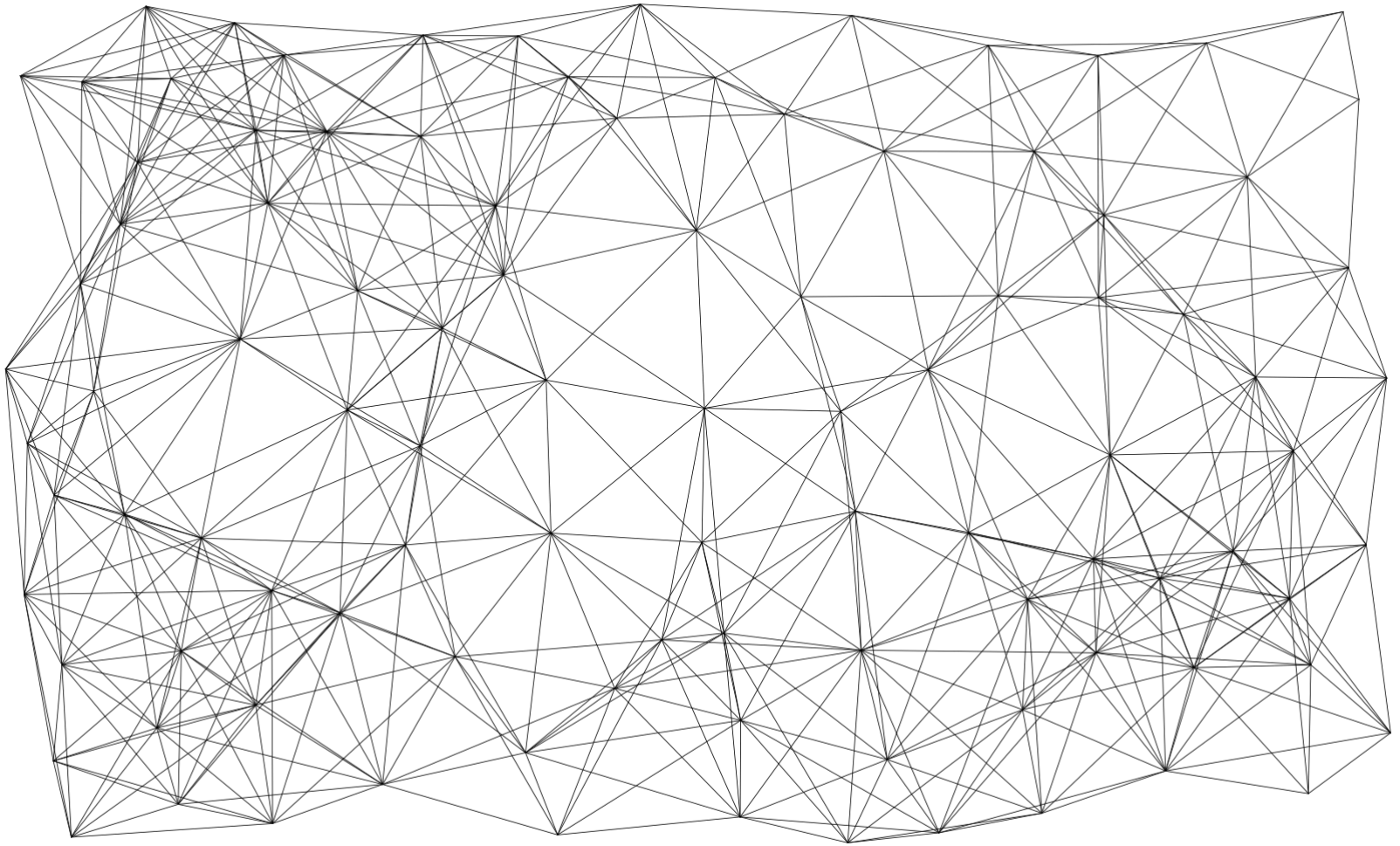
Everything within a distance is a neighbour.



Just like this.



Here is a complete graph. A bit denser.

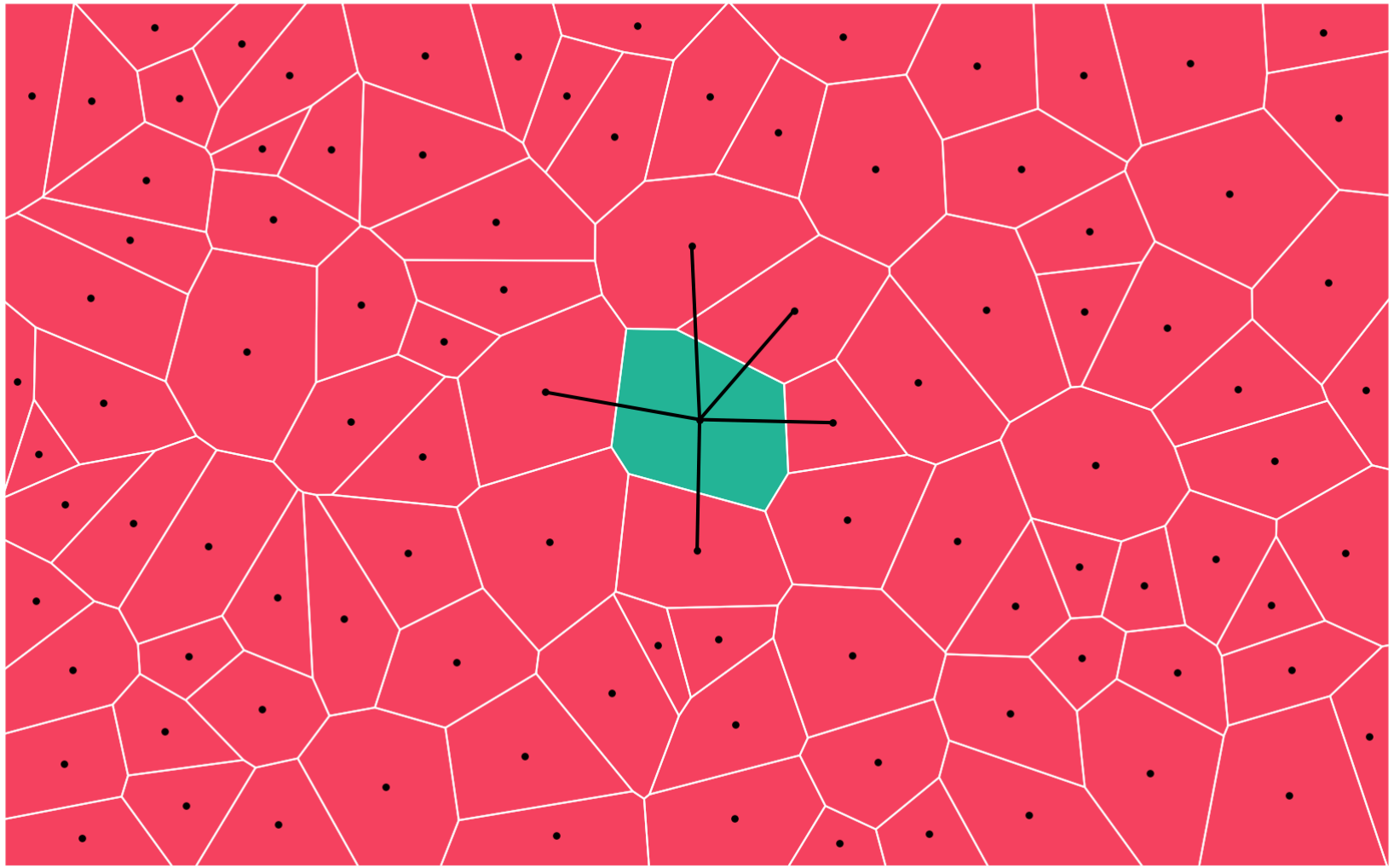


Here is a complete graph. A bit denser.

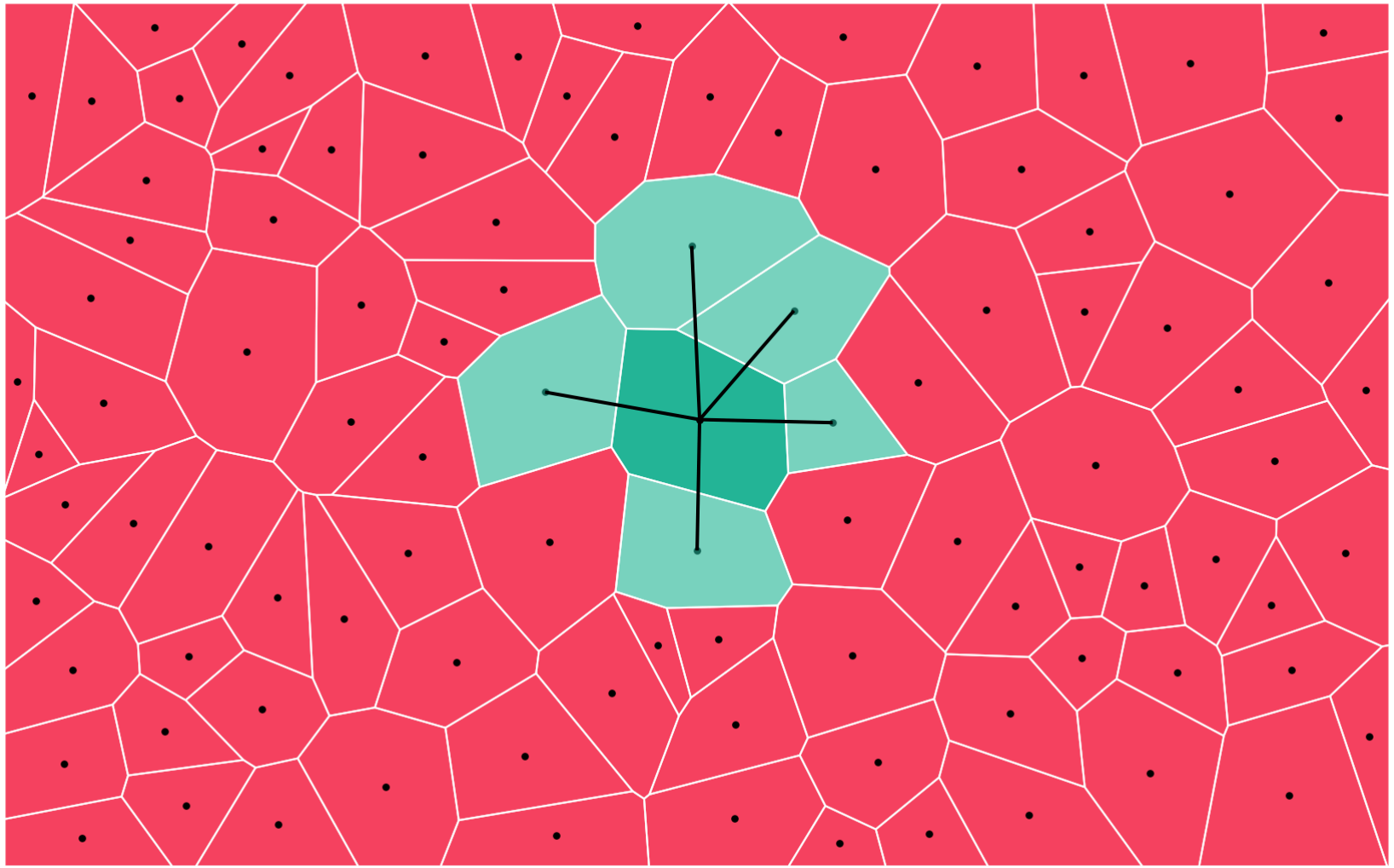
# Distance

## A geometry is considered a neighbour if it is no further than K-1 others

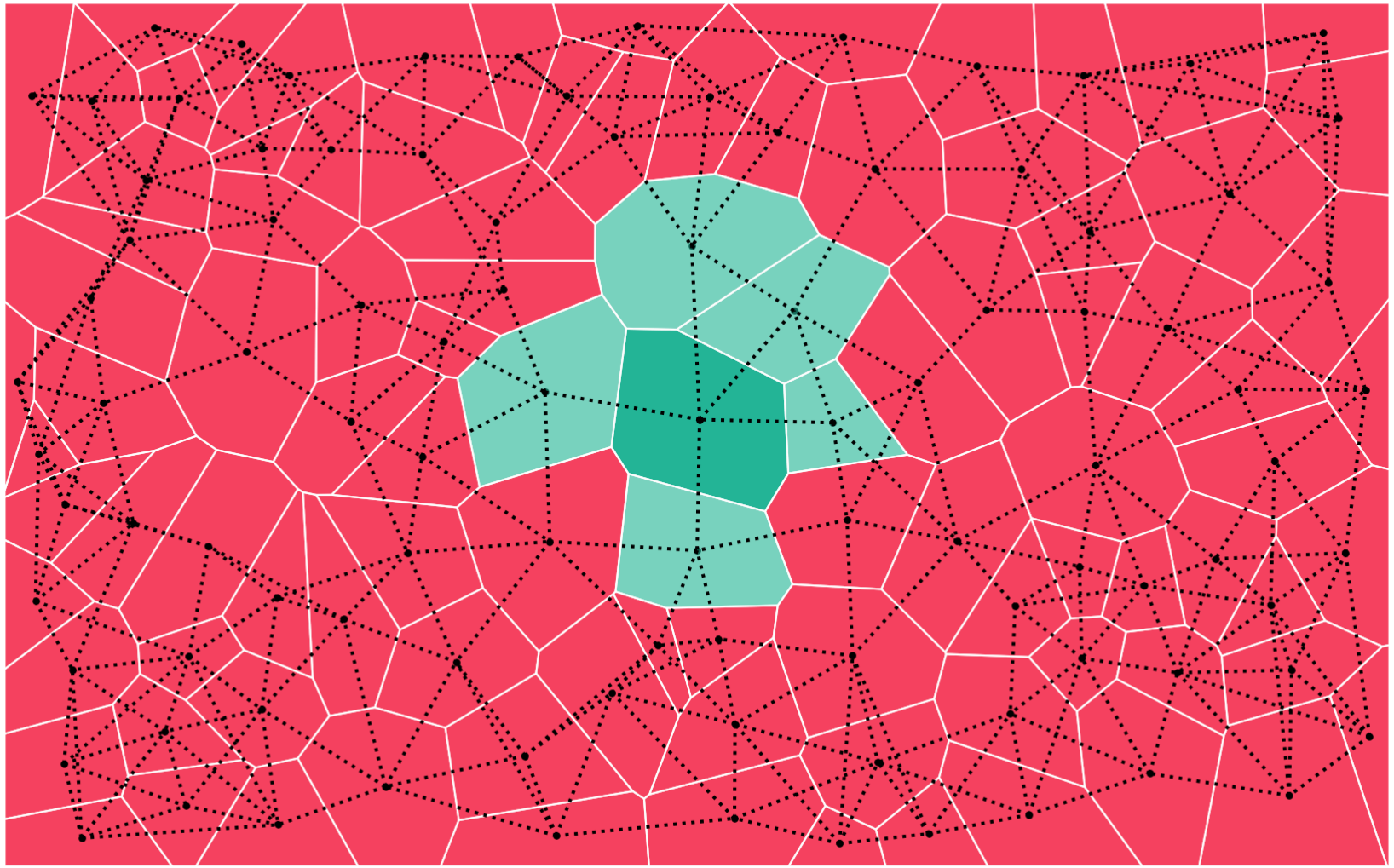
### K-nearest neighbour



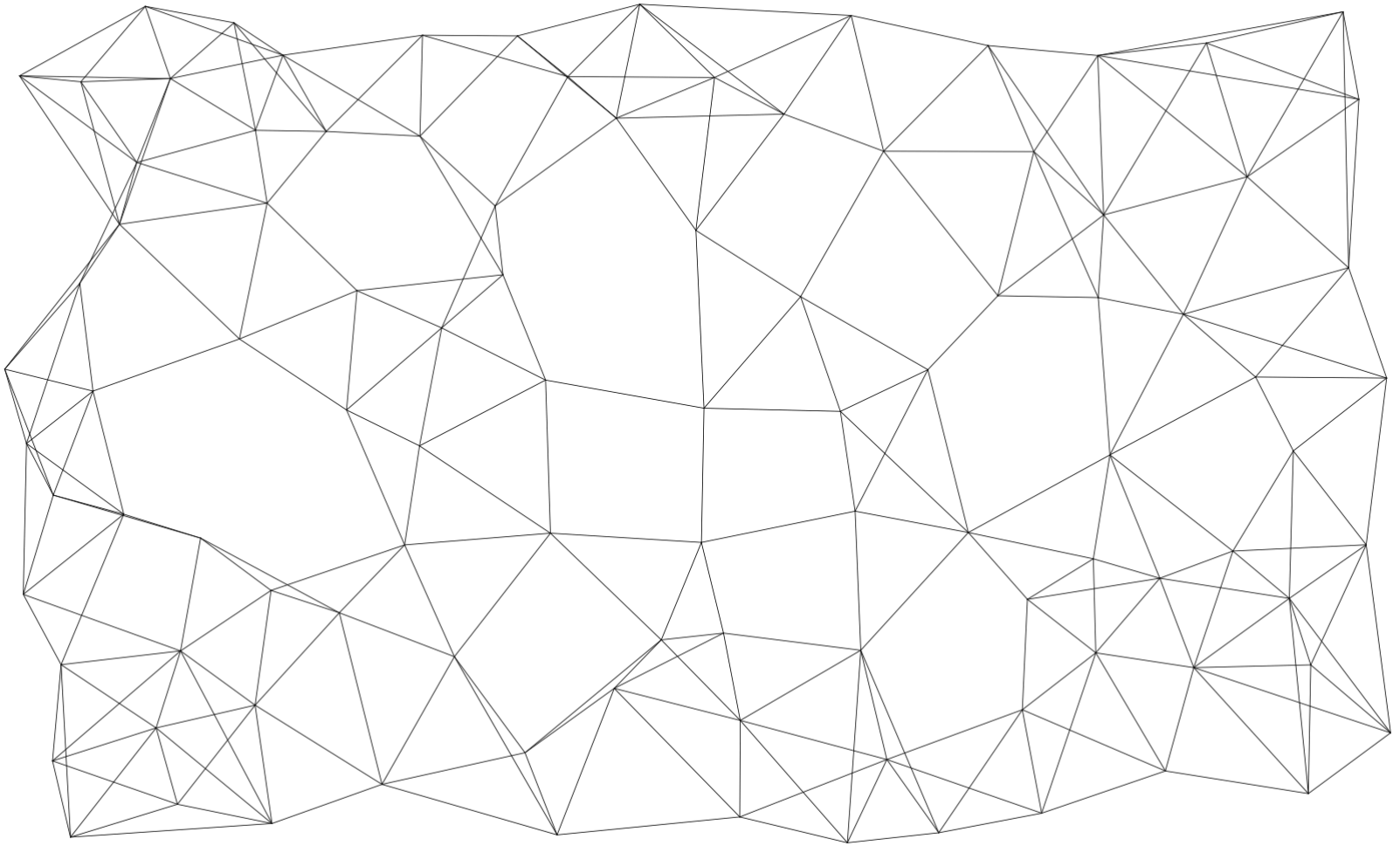
Only a set number of nearest geometries are neighbours.



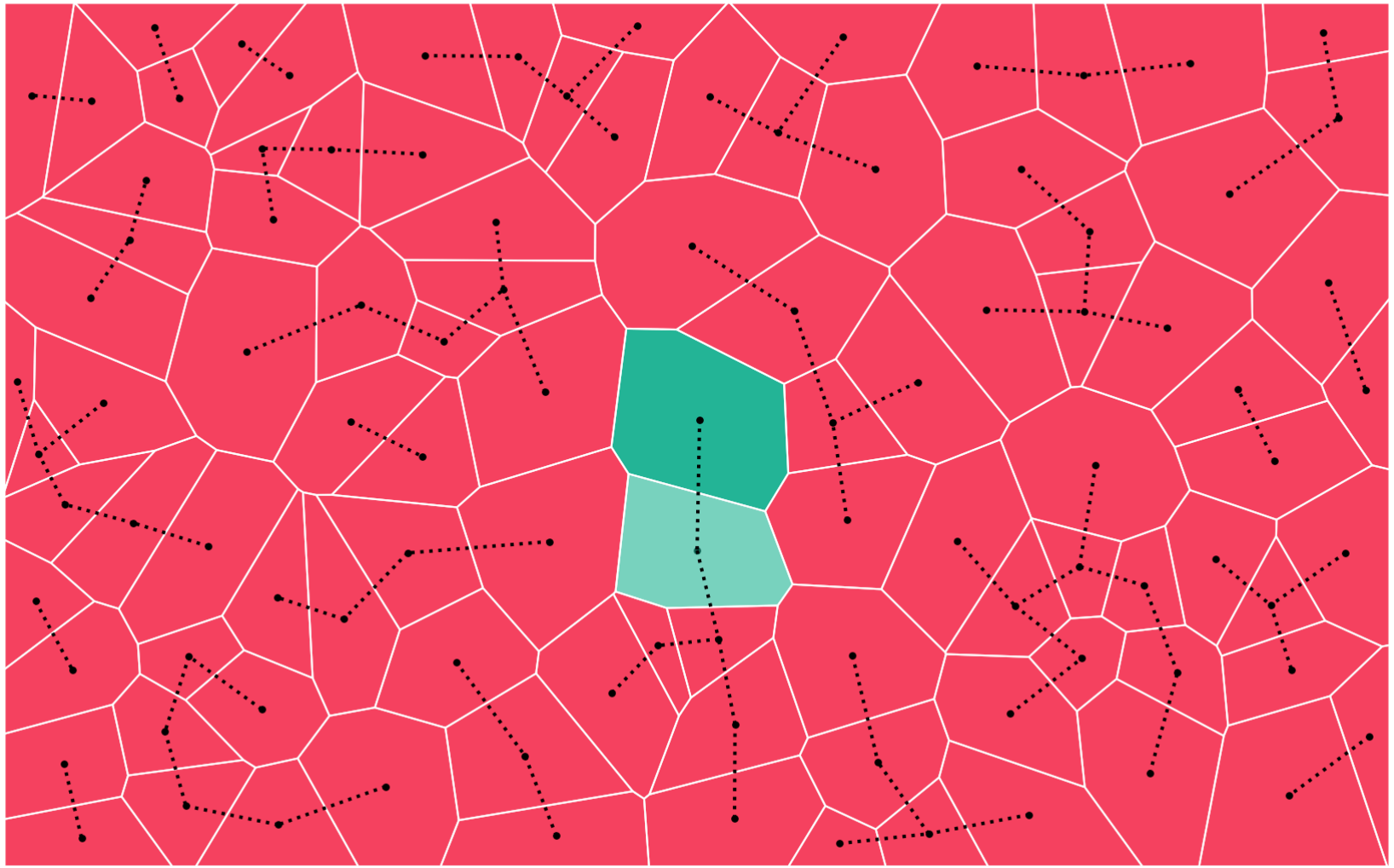
In this case $k = 5$.



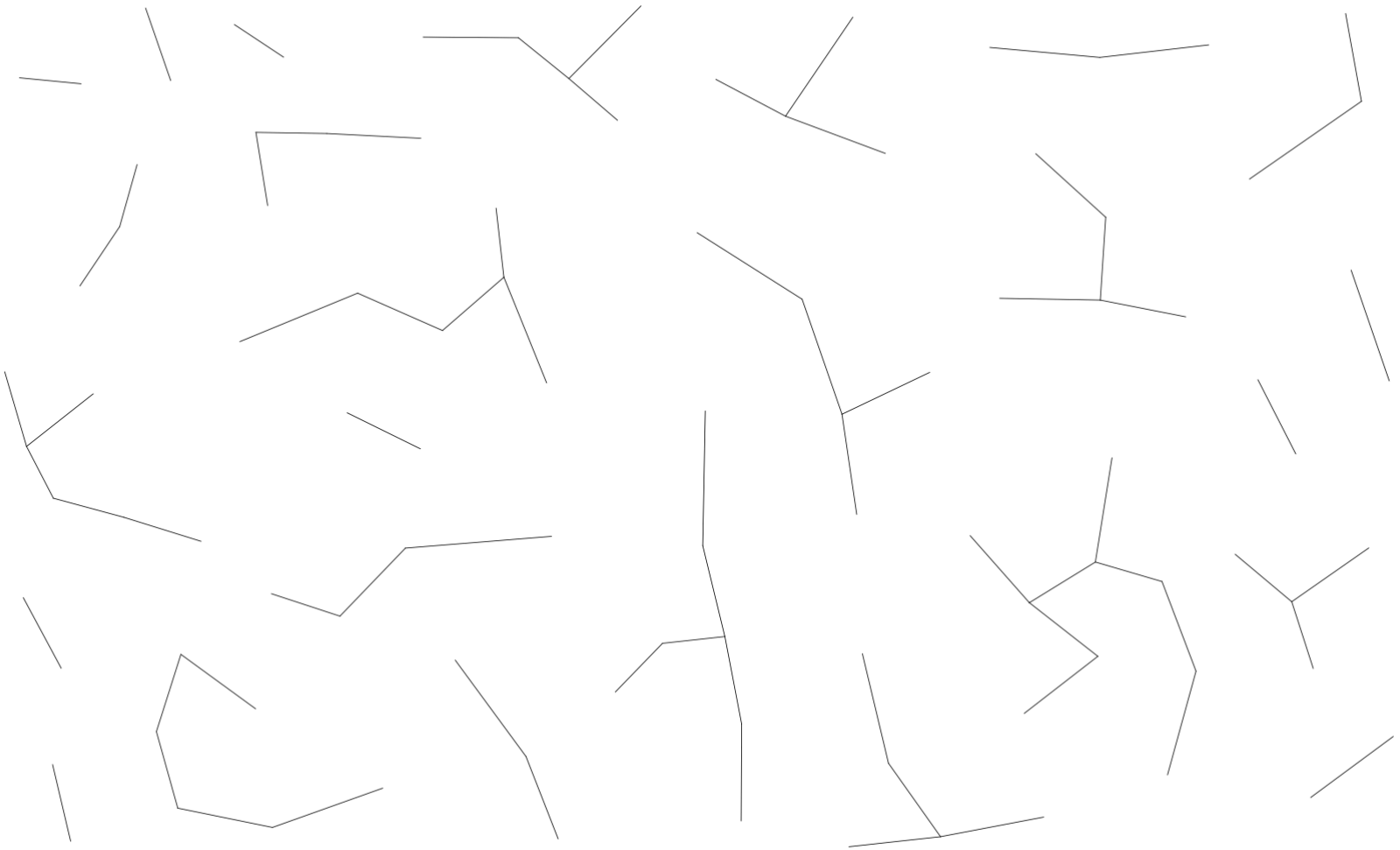
And a graph.



And a graph.



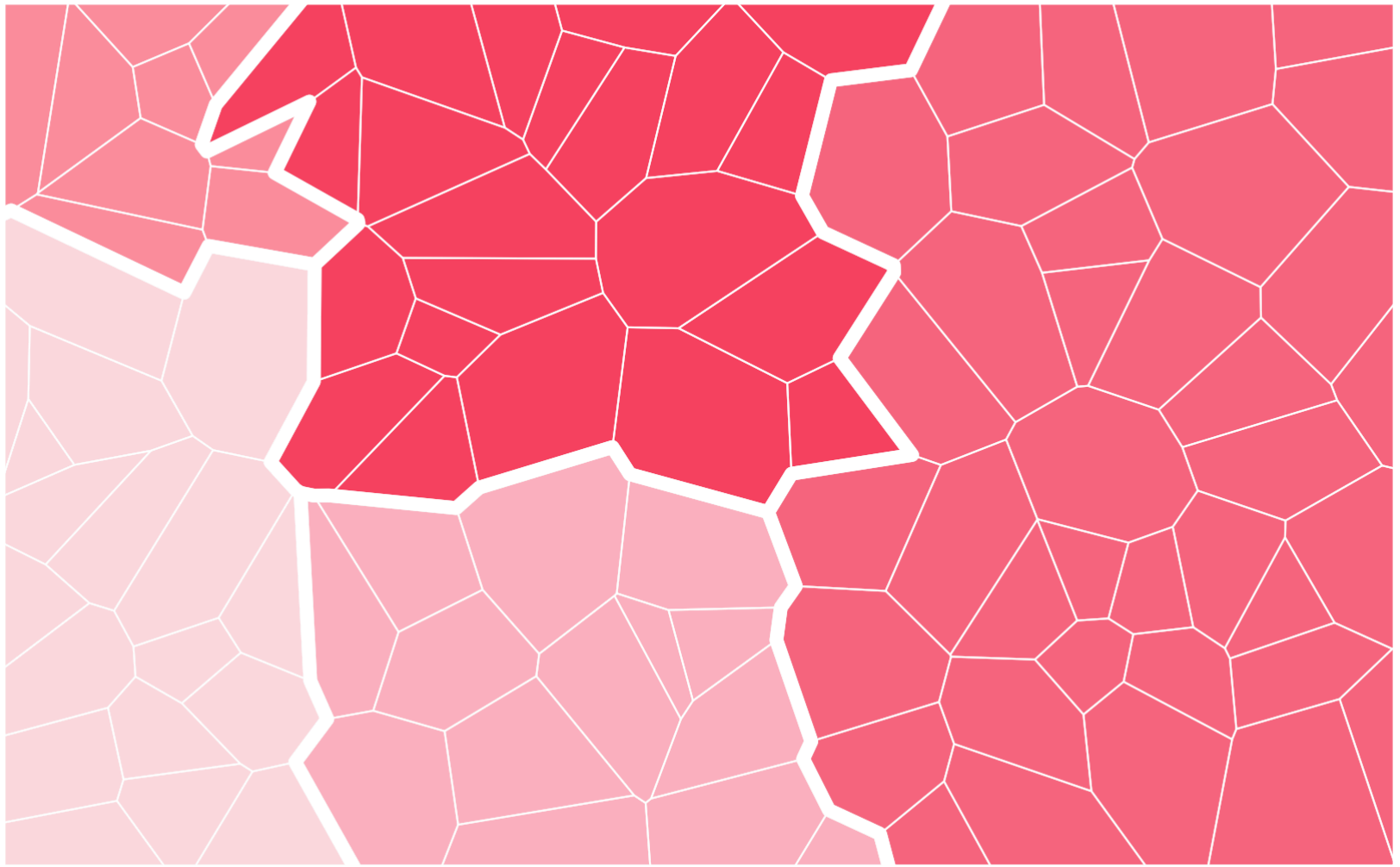
This would be the same with $k = 1$.



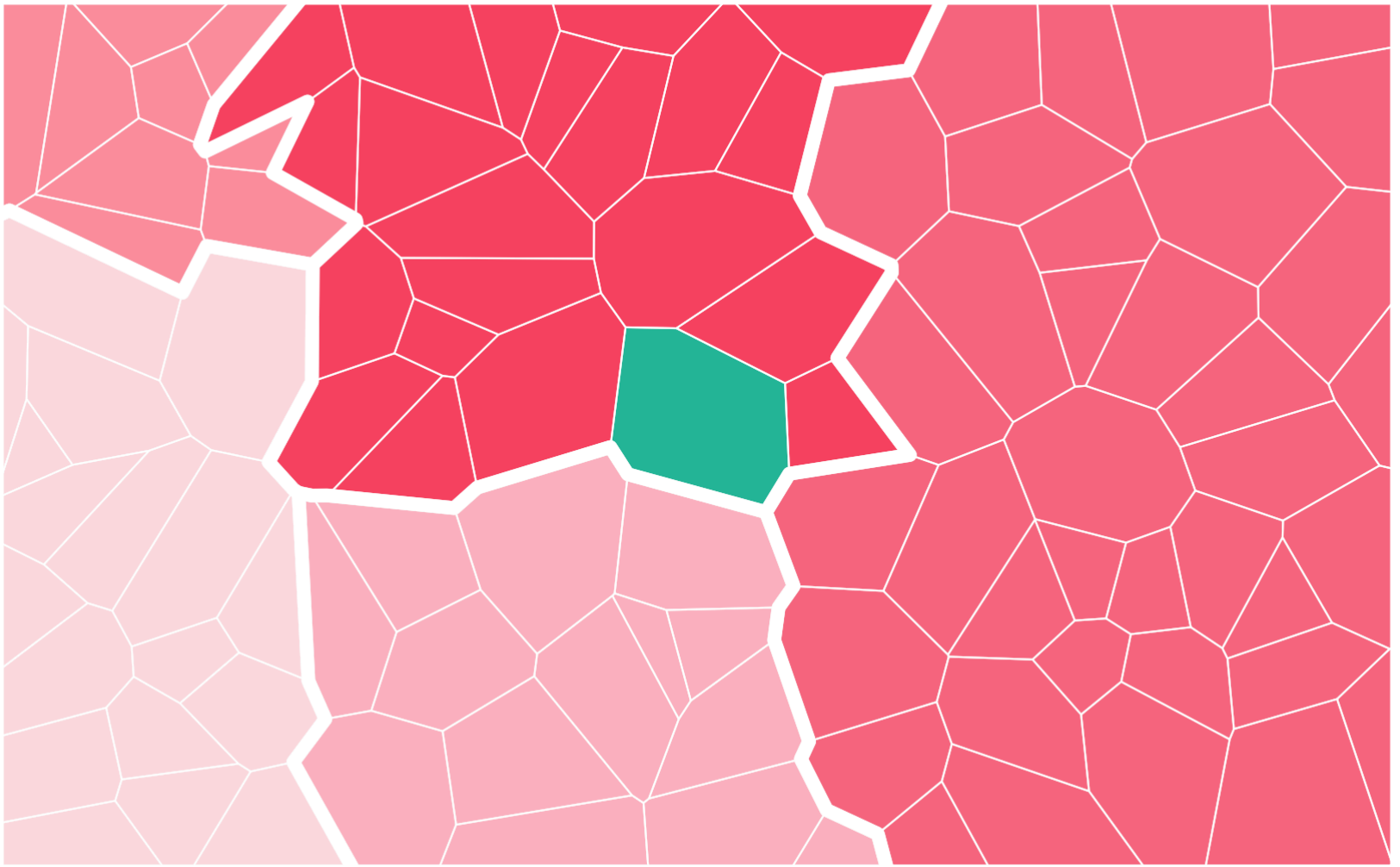
This would be the same with $k = 1$.

# Block contiguity/regions

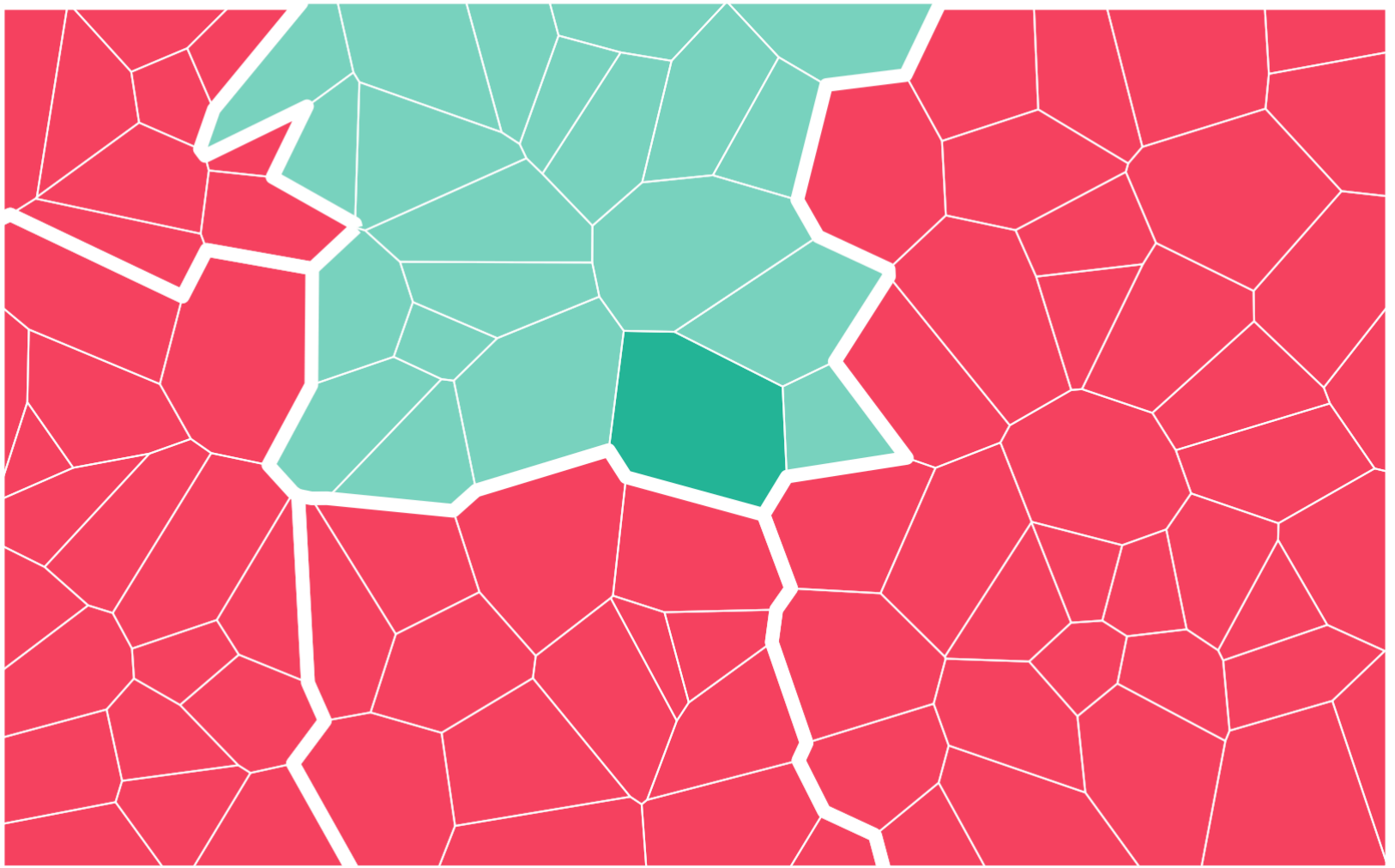
**Two geometries are considered neighbours if they belong to the same group**



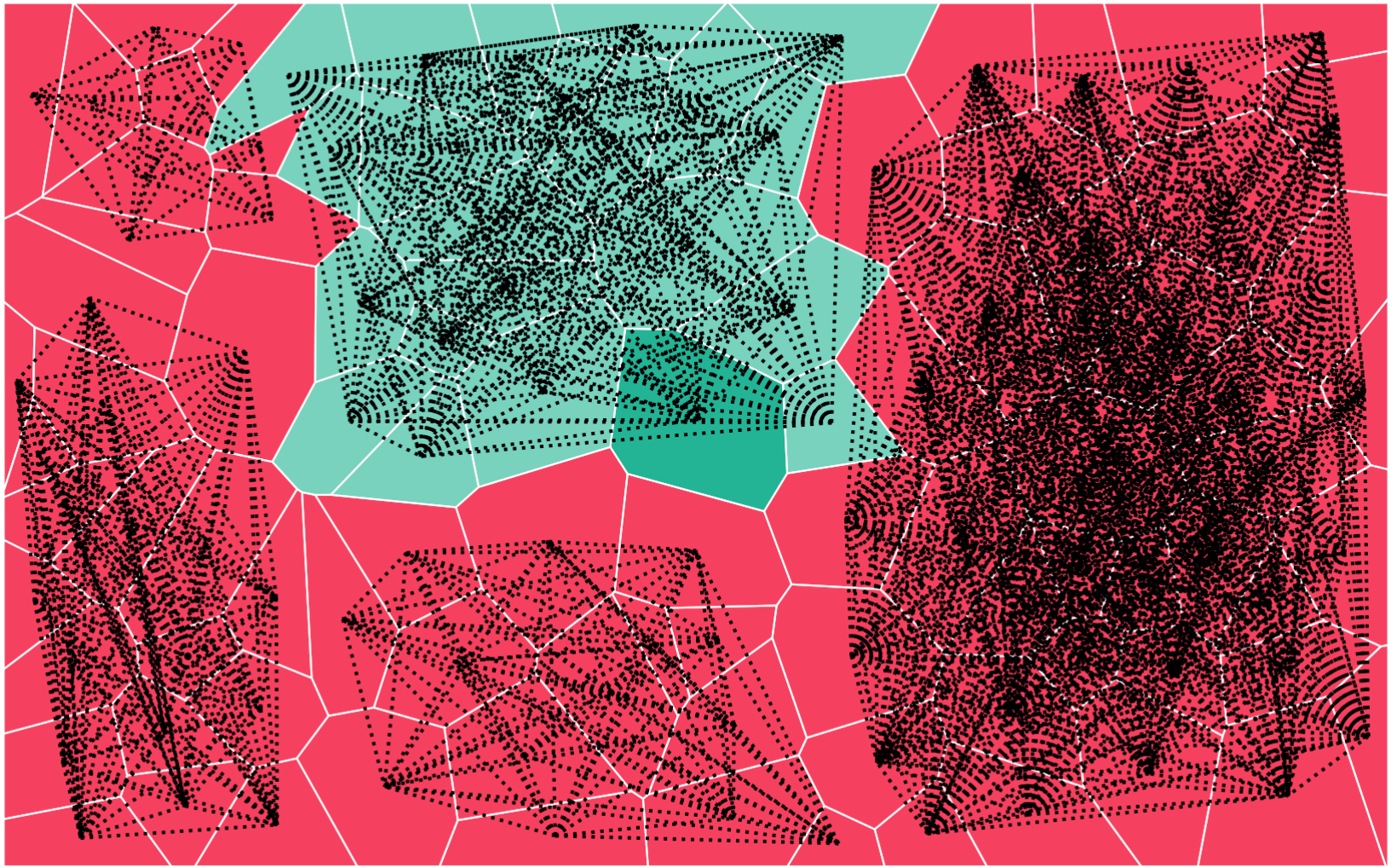
Let’s assume five regions.



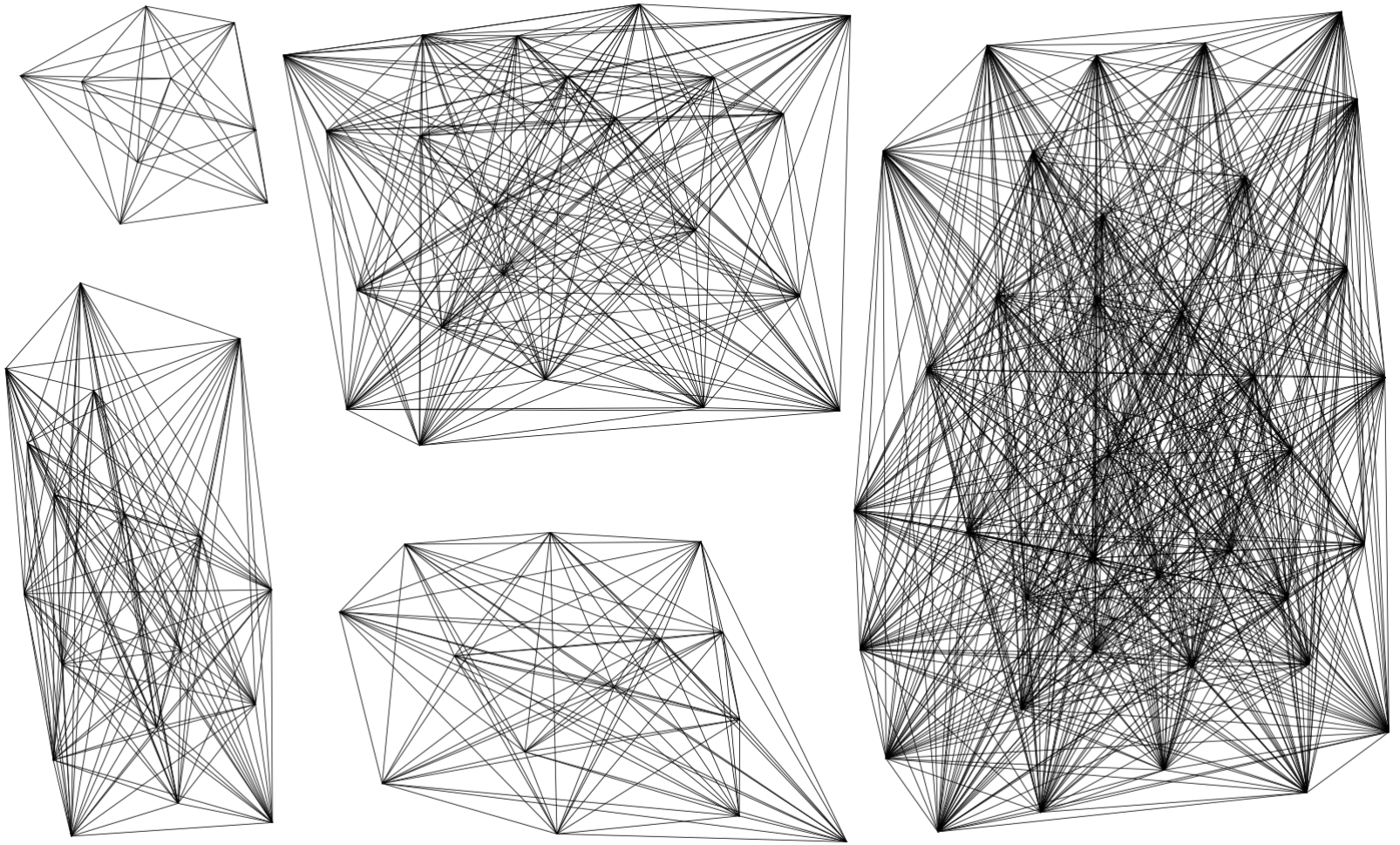
Let’s assume four regions. Each polygon belongs to one.



All geometries within a region are neighbours.



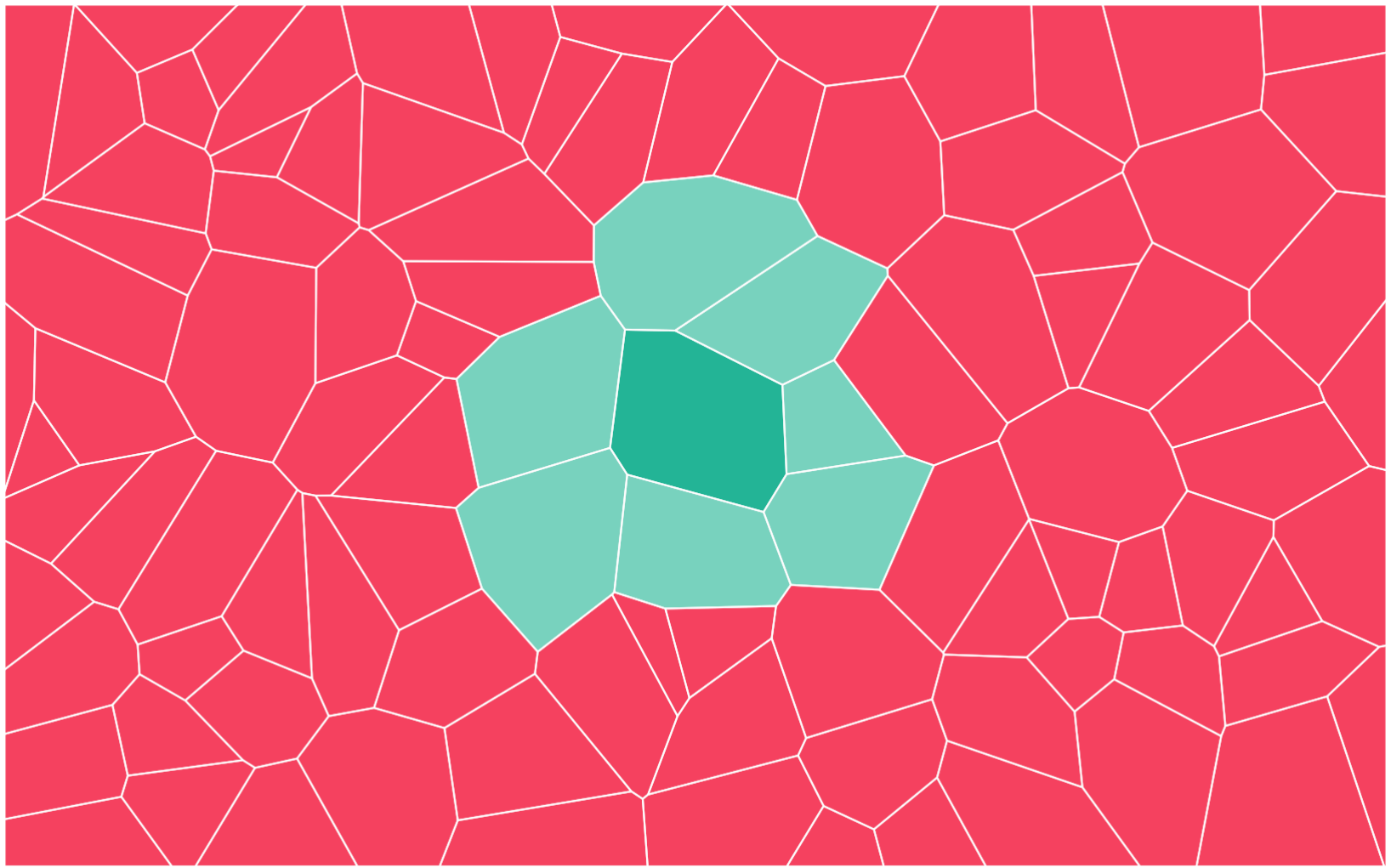
Graph forms isolated components.



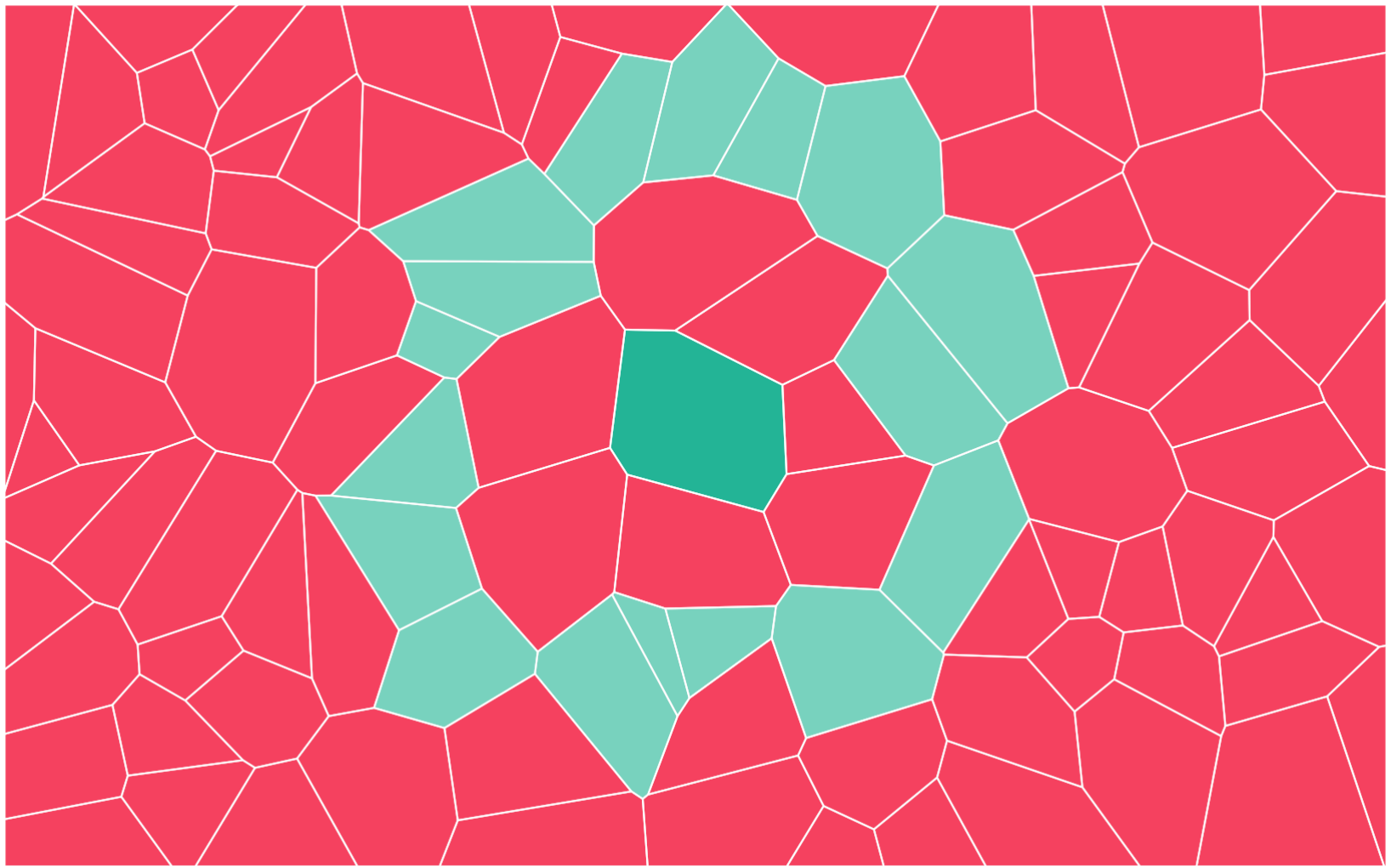
Graph forms isolated components.

# Graph routines

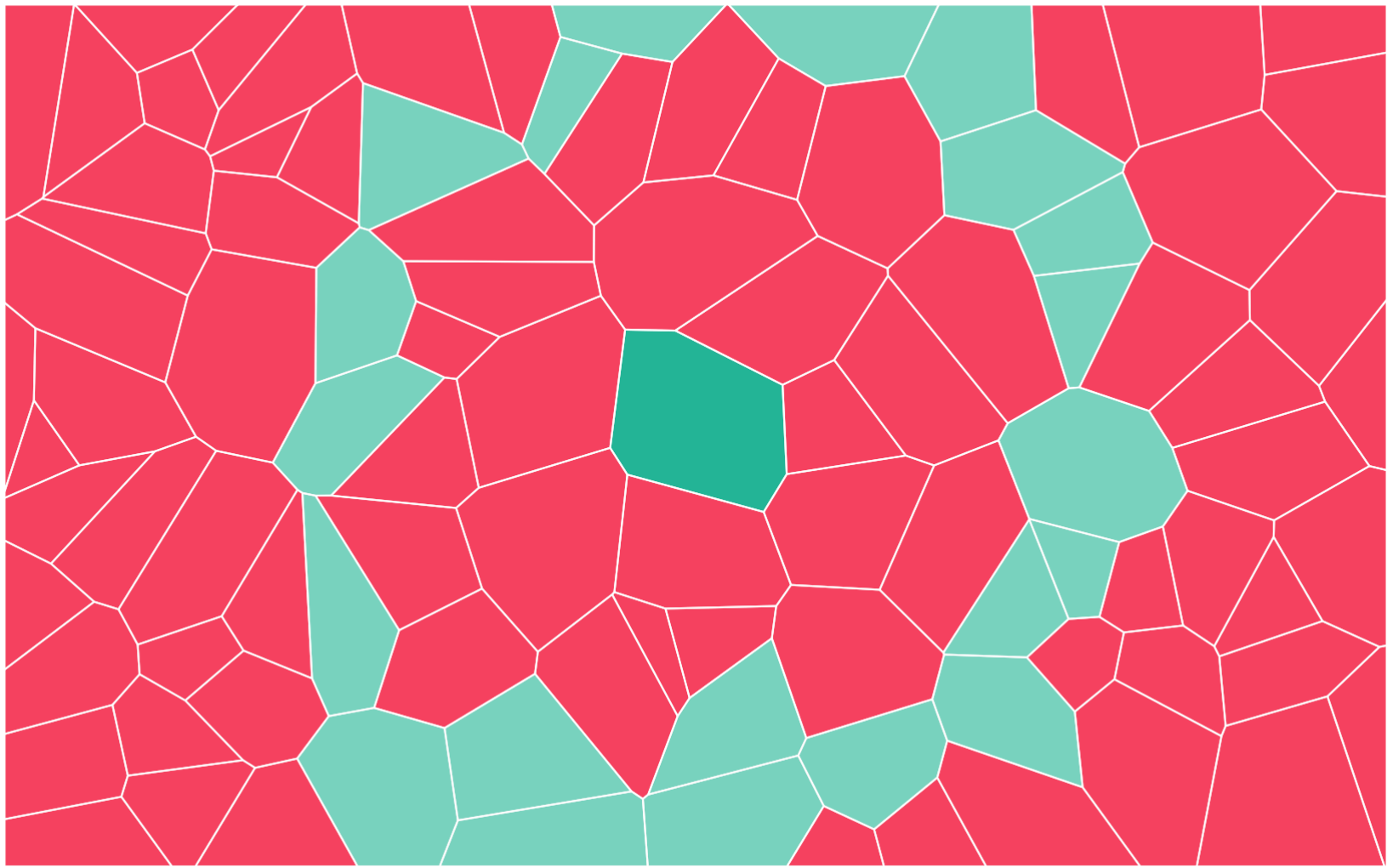
# Order of contiguity



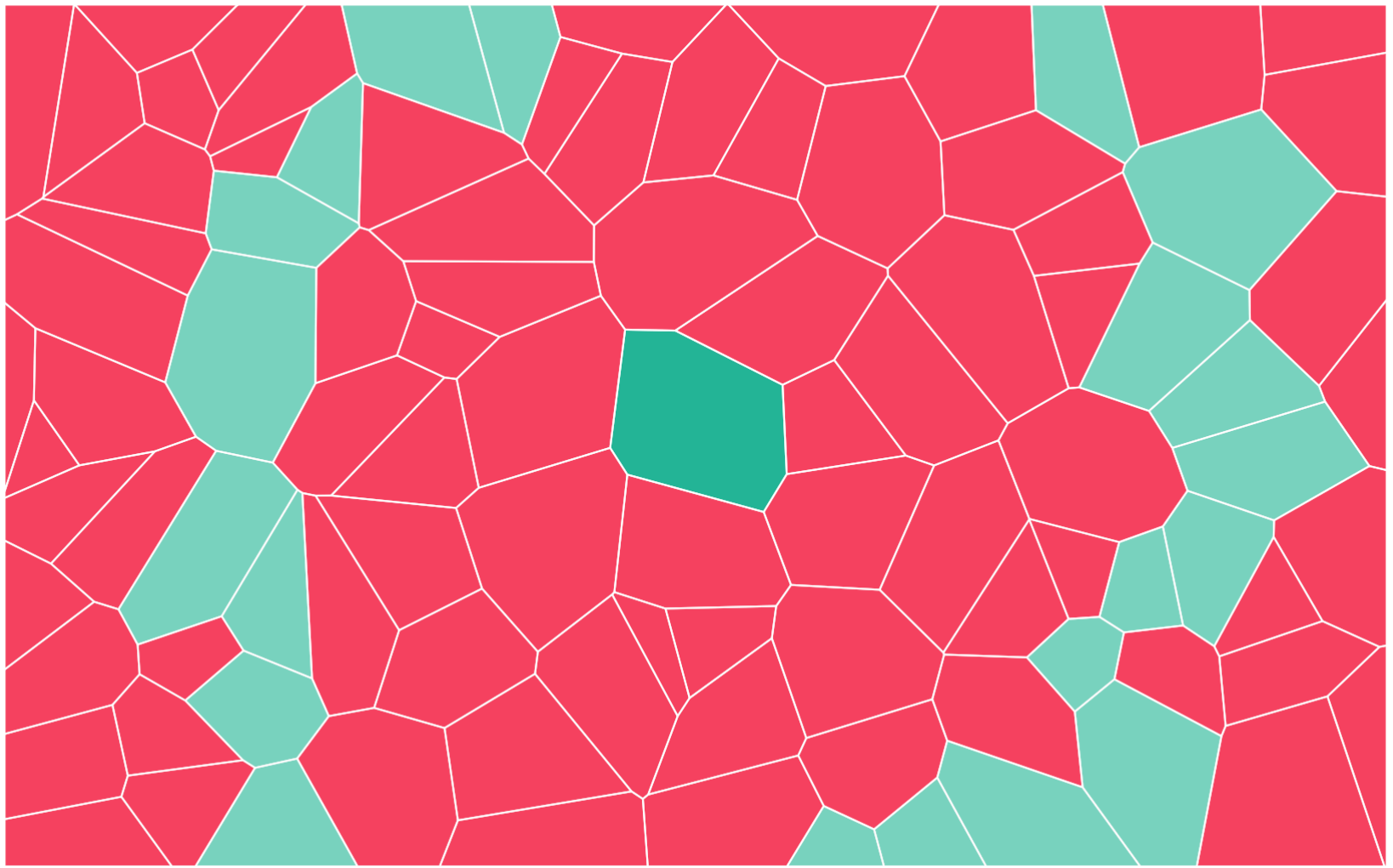
Let’s bring back contiguity.



Order 2 of contiguity

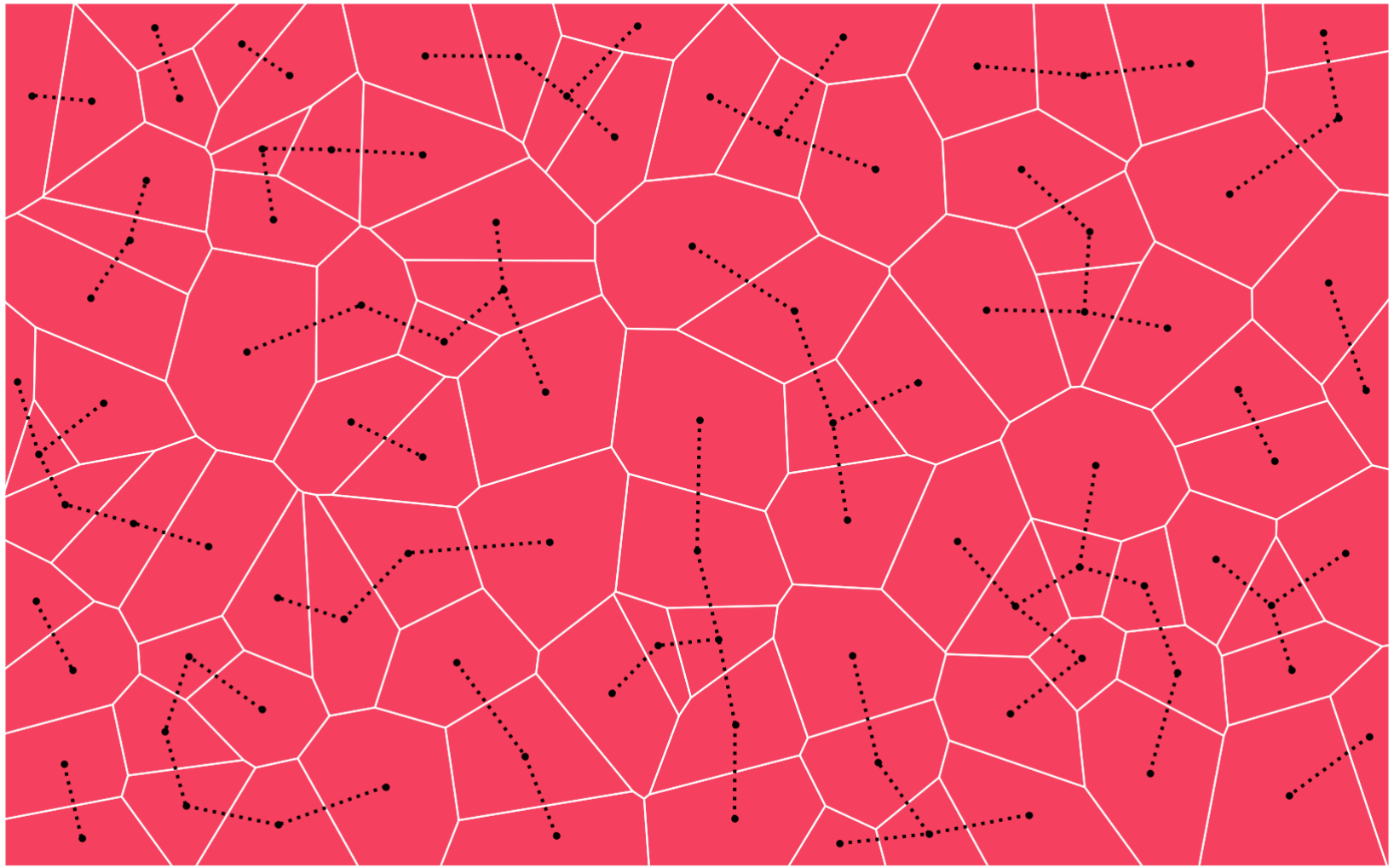


Order 3 of contiguity

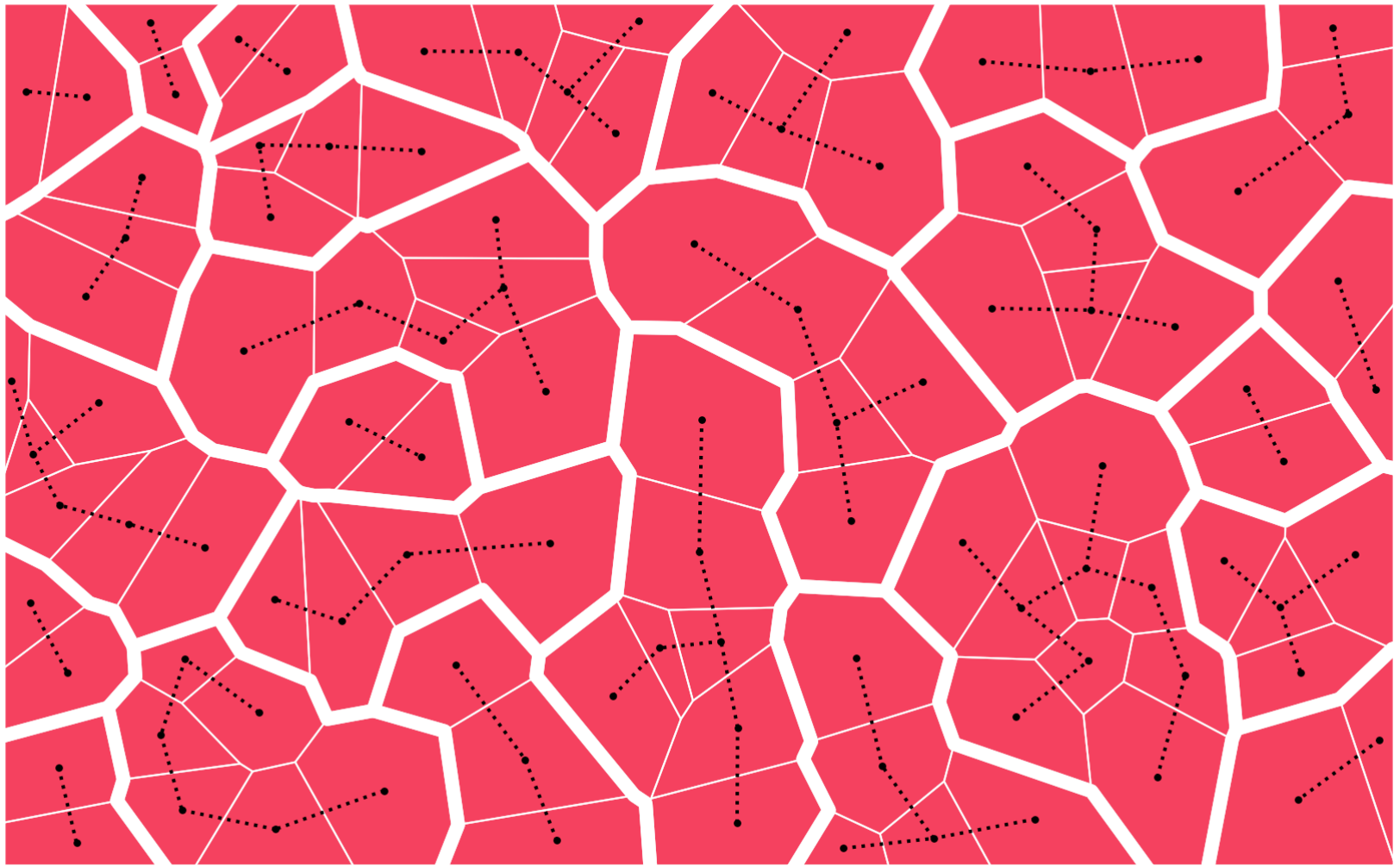


Order 4 of contiguity

# Component analysis

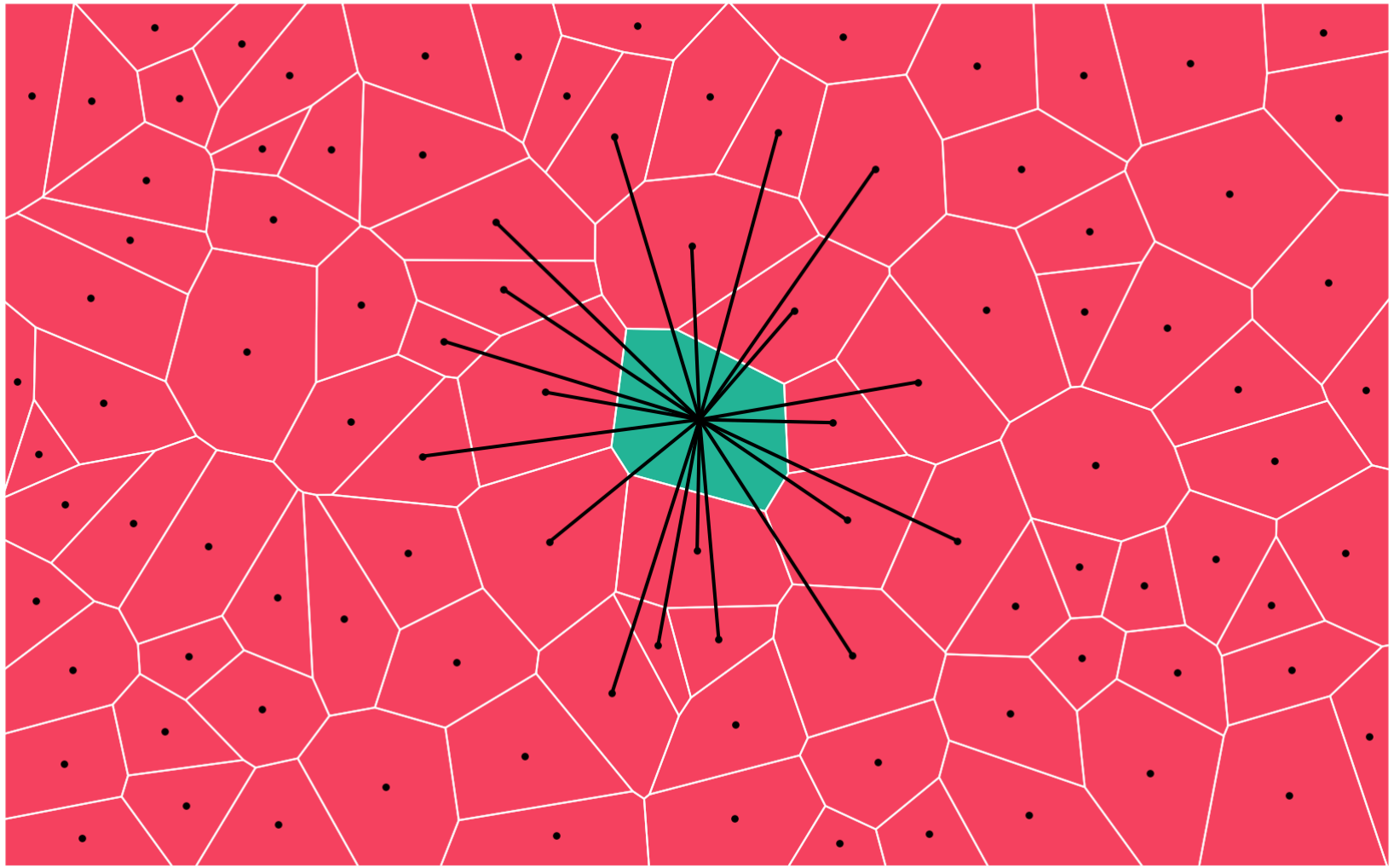


Let's bring back KNN with $k = 1$.

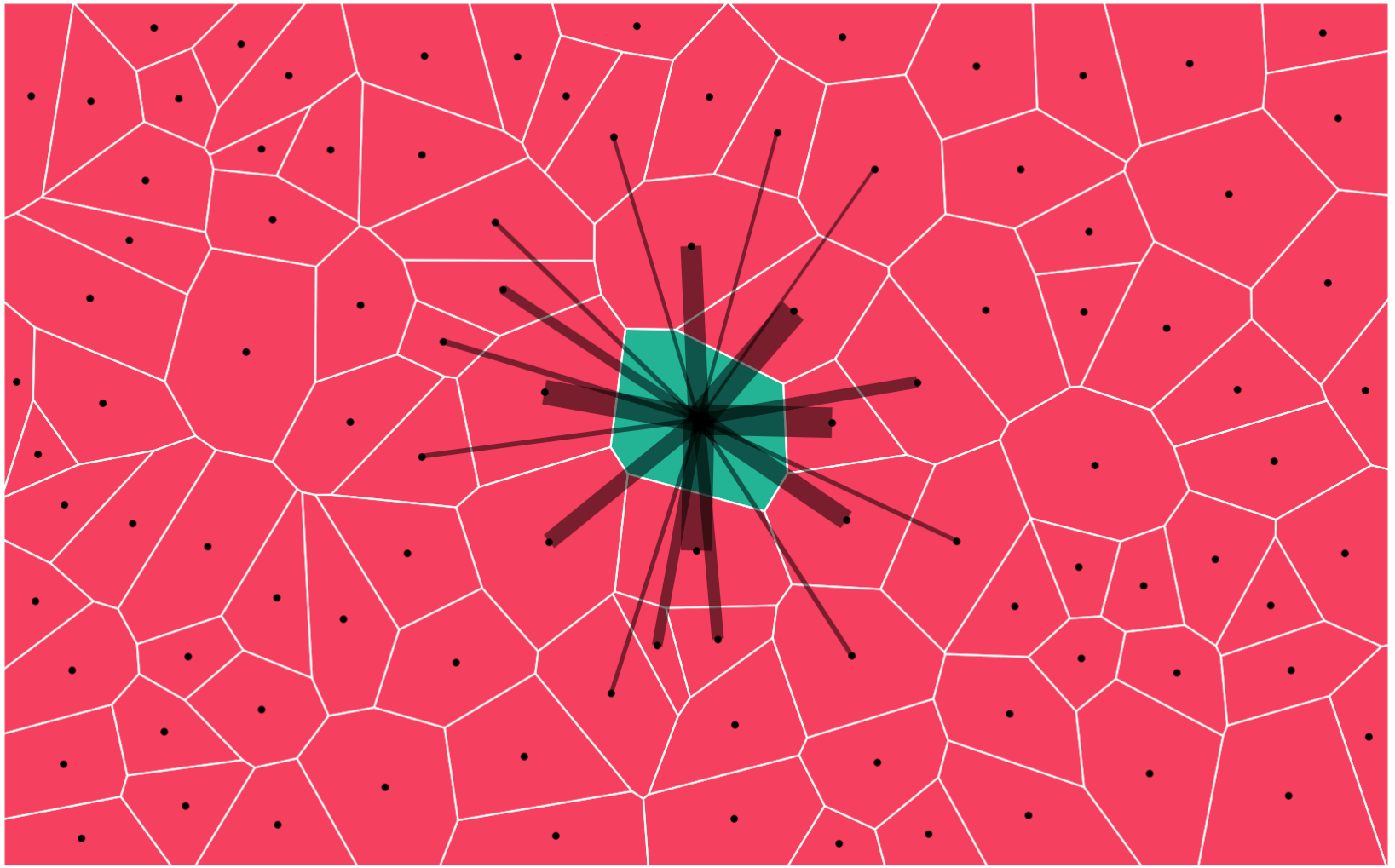


You can identify connected components.

# Weights are not always binary



Look at 20 nearest neighbours.



The closer a neighbour is, the higher the weight it has.

# Exploratory Spatial Data Analysis


ESDA

# Exploratory

## Focus on discovery and assumption-free investigation

# Spatial

**Patterns and processes that put *space* and *geography* at the core**

# Data Analysis

## Statistical techniques for exploration

# Questions that ESDA helps…

# Answer

## Is the variable I’m looking at concentrated over space?

## Do similar values tend to be located close by?

## Can I identify any particular areas where certain values are clustered?

## Ask

# What is behind this pattern?

## What could be generating the process?

## Why do we observe certain clusters over space?

# Spatial Autocorrelation

# Spatial Autocorrelation

## Statistical representation of Tobler's law
## Spatial counterpart of traditional correlation

## Degree to which similar values are located in similar locations

# Two flavours

**Positive: similar values → similar location (*closeby*)**
**Negative: similar values → disimilar location (*further apart*)**

# Two scales

**Global: do values tend to be close to other (dis)similar values?**

**Local: are there any specific parts of a map with an extraordinary concentration of (dis)similar values?**

# Global Spatial Autocorrelation

# “Clustering”

**Overall trend where the distribution of values follows a particular pattern over space.**

## Positive

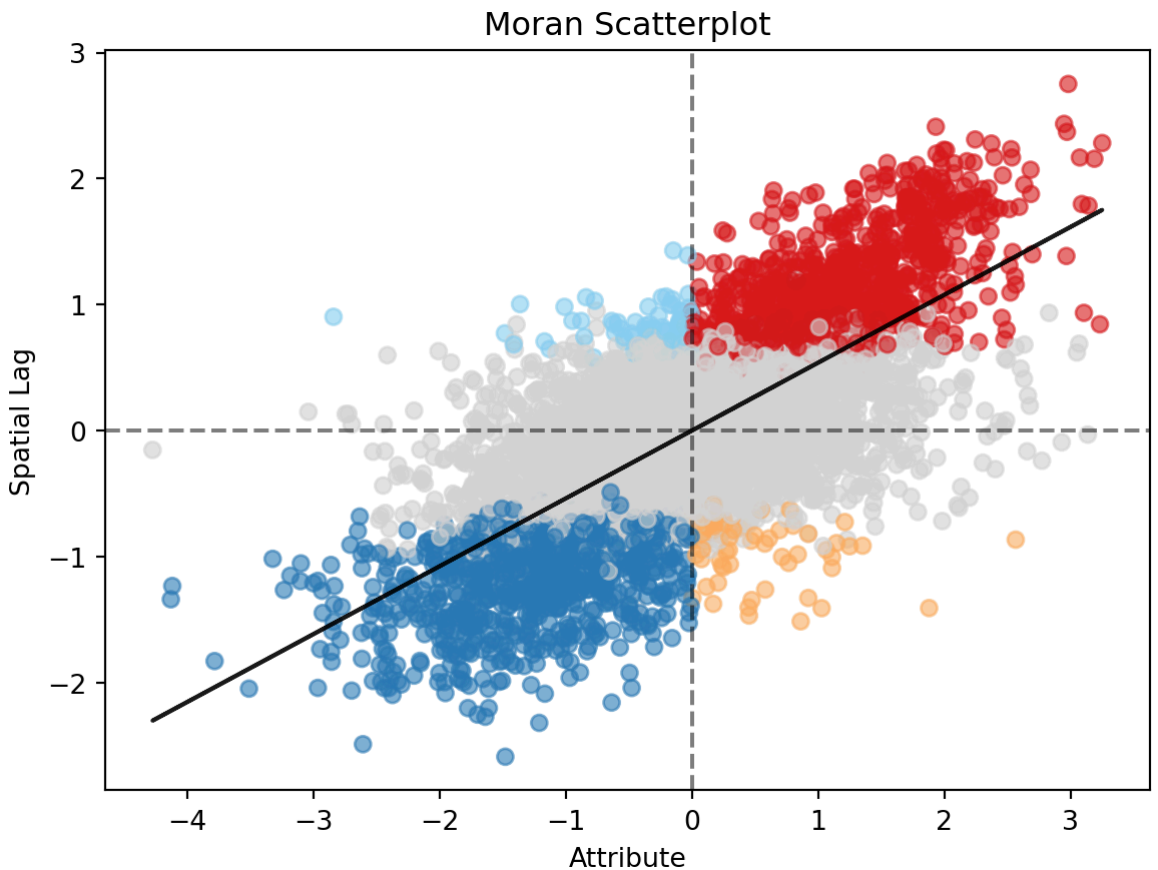
**Similar values close to each other (high-high, low-low)**

## Negative

**Similar values far from each other (high-low)**

# How to measure it?

# Moran Plot



# Moran’s I

**Formal test of global spatial autocorrelation**

**Statistically identify the presence of clustering in a variable**

**Slope of the Moran plot**

# Local Spatial Autocorrelation

# “Clusters”

***Pockets of spatial instability***

**Portions of a map where values are correlated in a particularly strong and specific way**

# Clusters

## High-High

## Positive SA of *high* values (*hotspots*)

## Low-Low

## Positive SA of *low* values (*coldspots*)

## High-Low

## Negative SA (*spatial outliers*)

## Low-High

## Negative SA (*spatial outliers*)

# LISAs

## Local Indicators of Spatial Association

## Statistical tests for spatial cluster detection

## Compares the observed map with many randomly generated ones to see how likely it is to obtain the observed associations for each location

# Lets code!