

JOURNALISM & MAPPING

AN INTRODUCTION TO GIS ANALYSIS

ERIC SAGARA

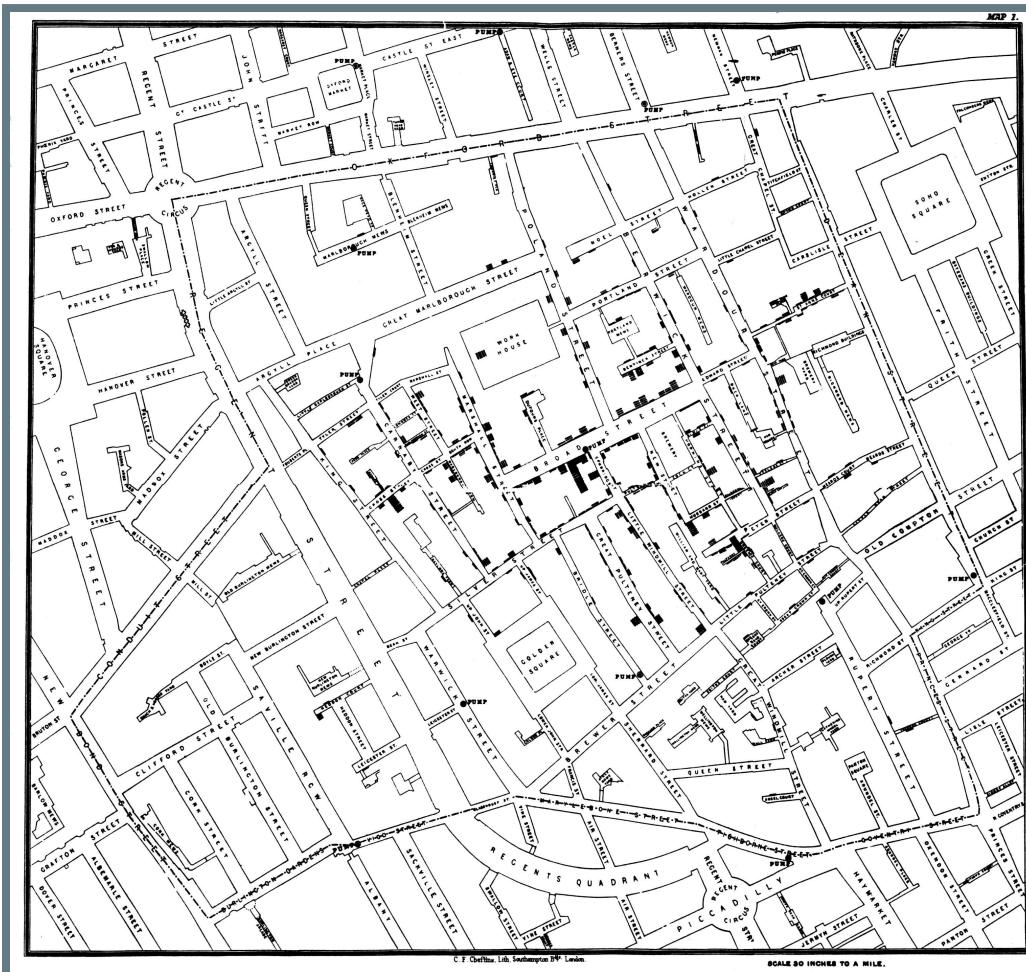
Reveal / Center for Investigative Reporting

esagara@cironline.org | @esagara

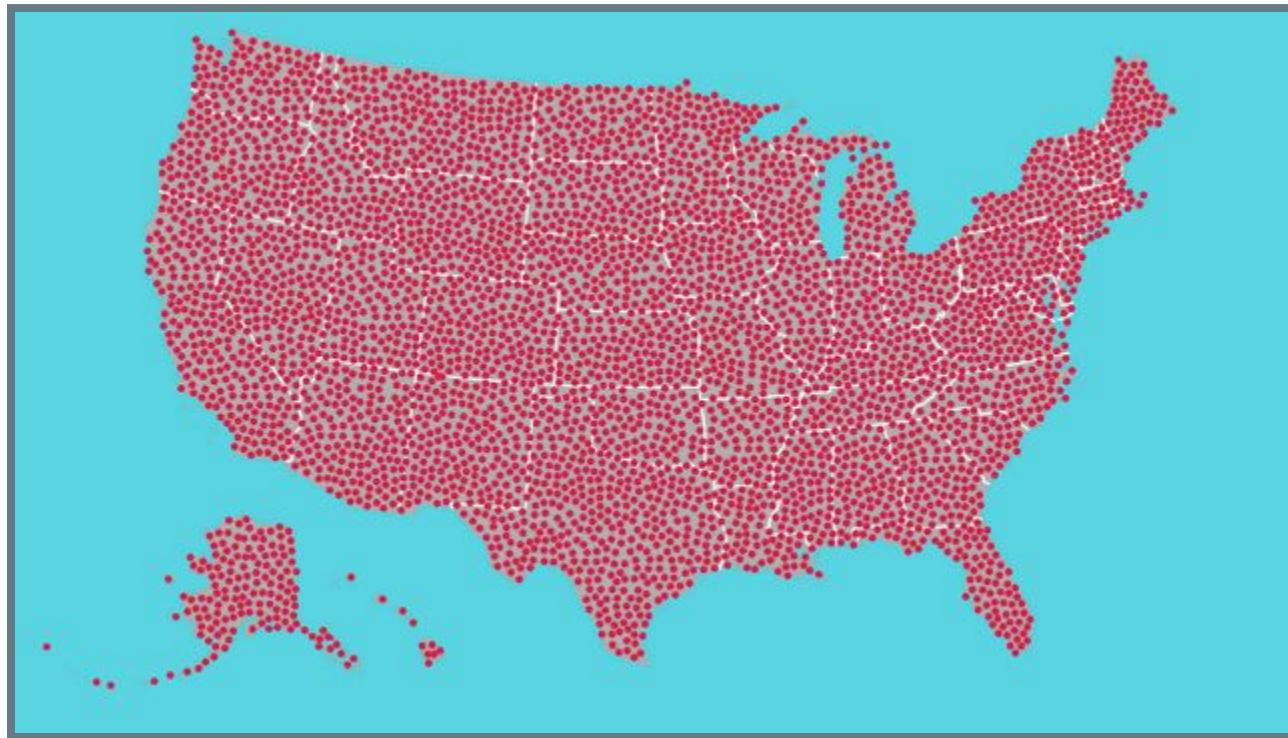
WHY MAP?

Mapping and GIS analysis is used to explore spatial relationships.

THIS IS A GOOD EXAMPLE OF MAPPING



THIS IS NOT



HOW DO WE MAP?

THERE ARE A VARIETY OF GIS SOFTWARE PACKAGES AVAILABLE

ARCGIS

Expensive, but fairly comprehensive in its features

QGIS

Free, not as rich in features (but getting better every day)

GDAL

A powerful, free command line tool. It has the highest learning curve and no graphical interface.

WHAT CAN WE MAP?

Many datasets today have a geographic component to it,
whether it be:

- An address
- Latitude and longitude
- County, city or state
- Census tracts
- Congressional districts
- Or other custom boundaries

**SOME DATASETS ARE DESIGNED SPECIFICALLY
FOR GIS**

SHAPEFILES

The most common format for geographic data. It is actually a set of at least three files.

- .shp - the actual shapefile
- .shx - an index of the feature geometry
- .dbf - the data attached to features

KML FILES

Also very common. It is a single file used to work with Google map products such as Fusion Tables, Google Maps and Google Earth. This file format is typically generated by someone wishing to display data online.

GEOJSON

Also meant to display geographic data online. It's a single file for use with JavaScript mapping libraries such as Leaflet and D3.

GEODATABASES

These are basically shapefiles on steroids. They can be very large in size and difficult to load into open source software such as QGIS (but it is possible).

WHAT TYPES OF GEOGRAPHIC DATA ARE THERE?

There are two basic types of geographic data

VECTOR

- **Points**

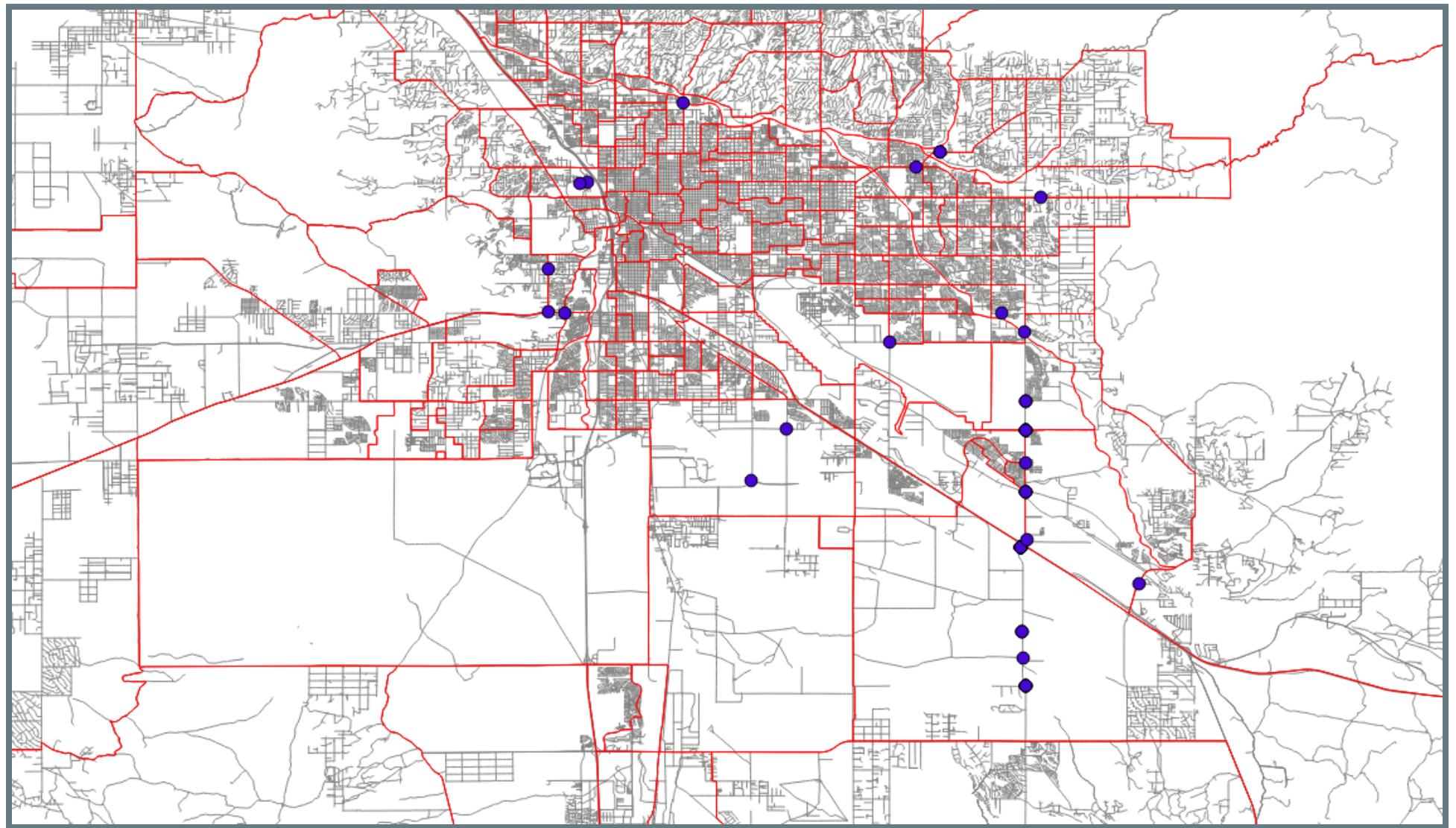
Each represents a single location, such as an address.

- **Lines**

Typically used to represent roads, waterways, trails, etc.

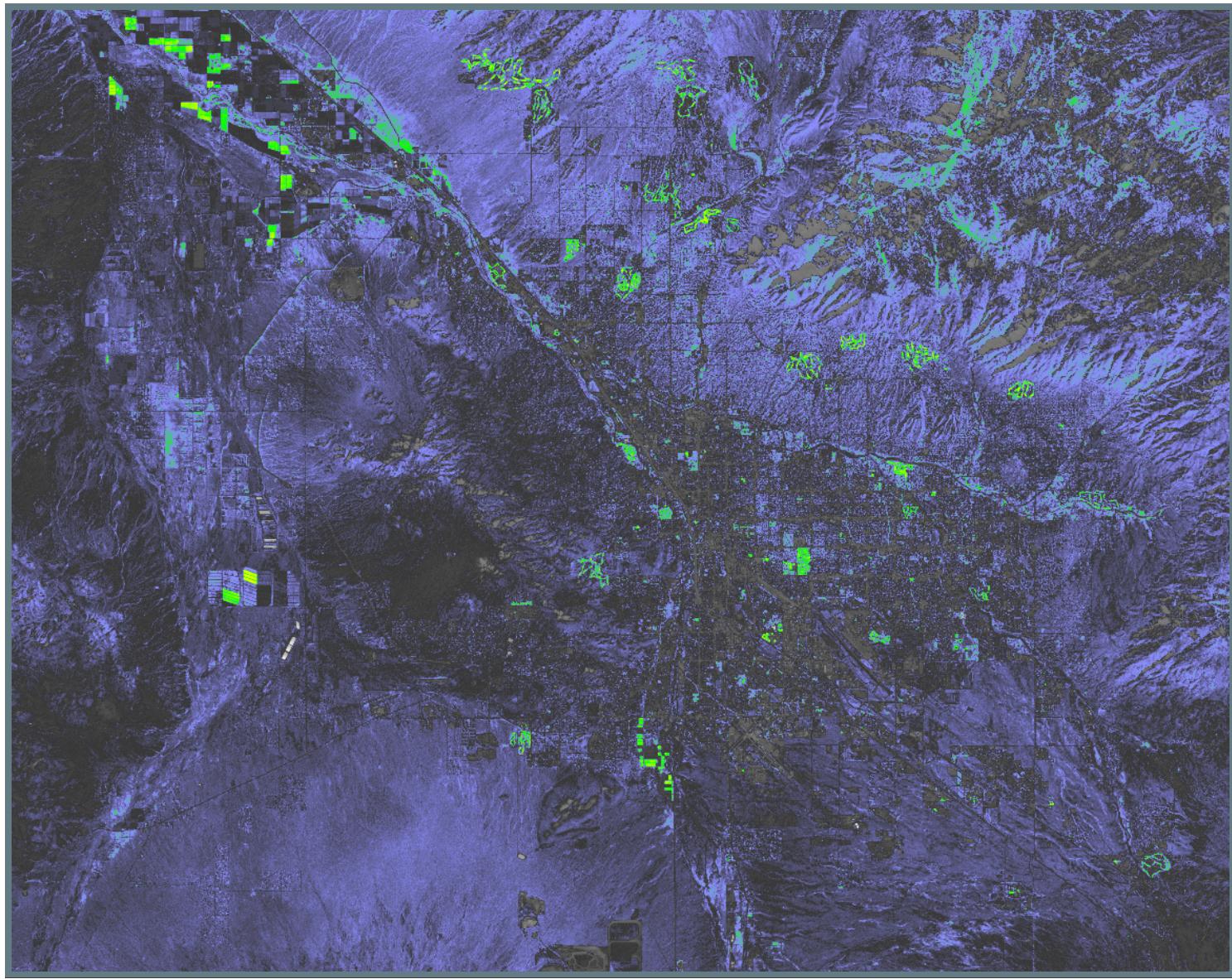
- **Polygons**

A boundary or a zone (such as a county, city or state)



RASTER

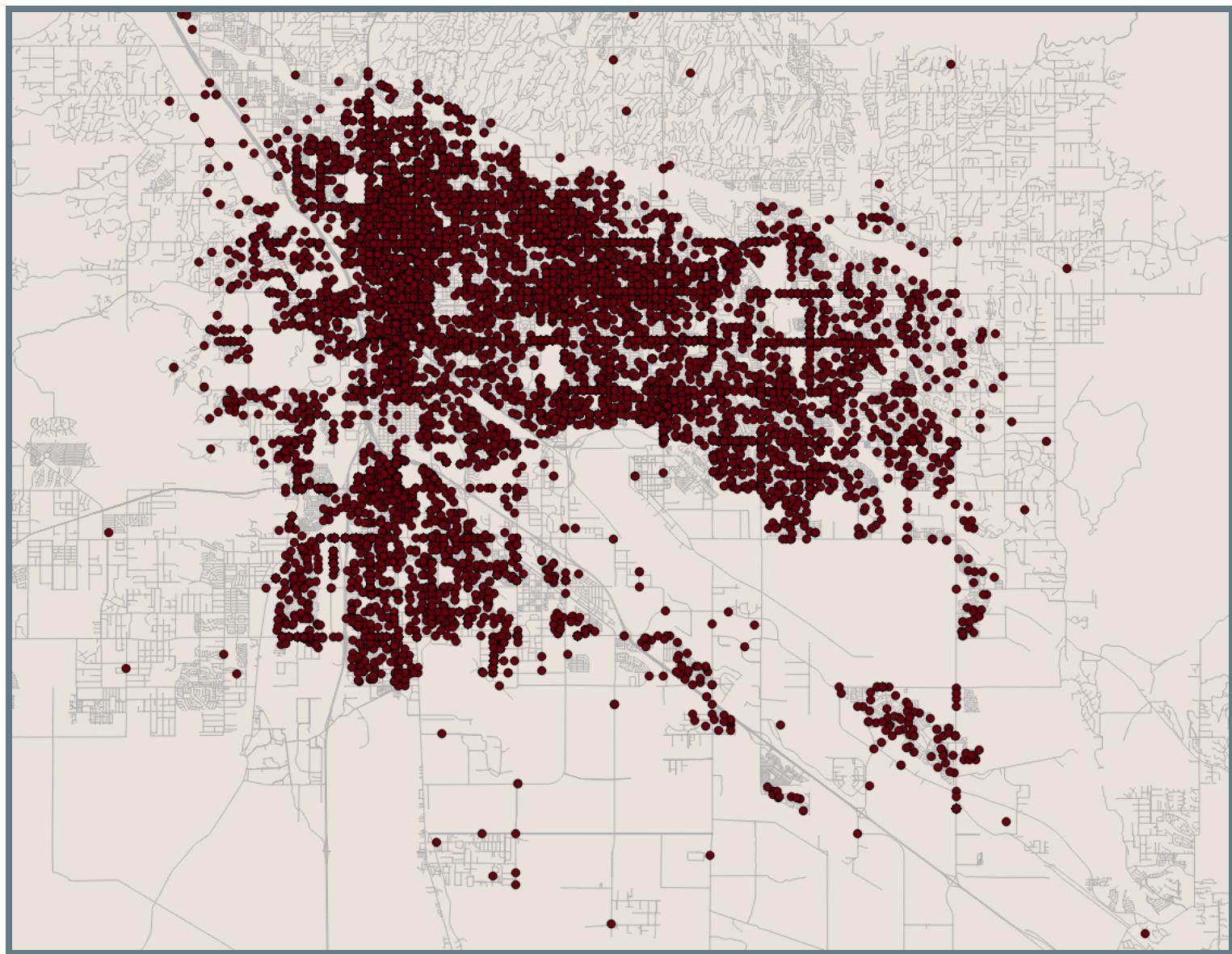
- Elevation data
- Satellite imagery
- Aerial imagery



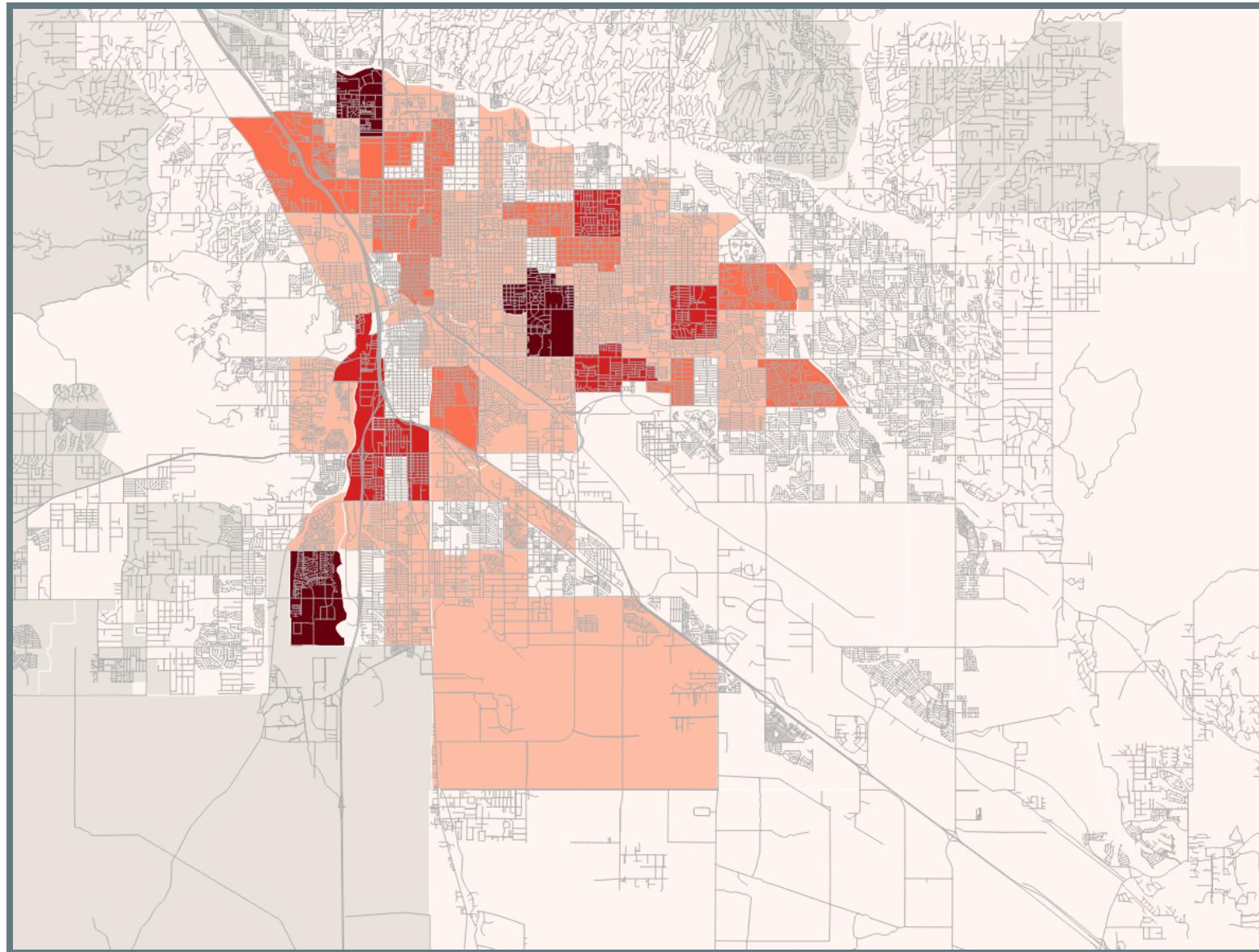
WE FIND STORIES COMBINING DIFFERENT DATASETS AND TECHNIQUES

THE TECHNIQUES WE USE CAN DRIVE THE STORY WE TELL

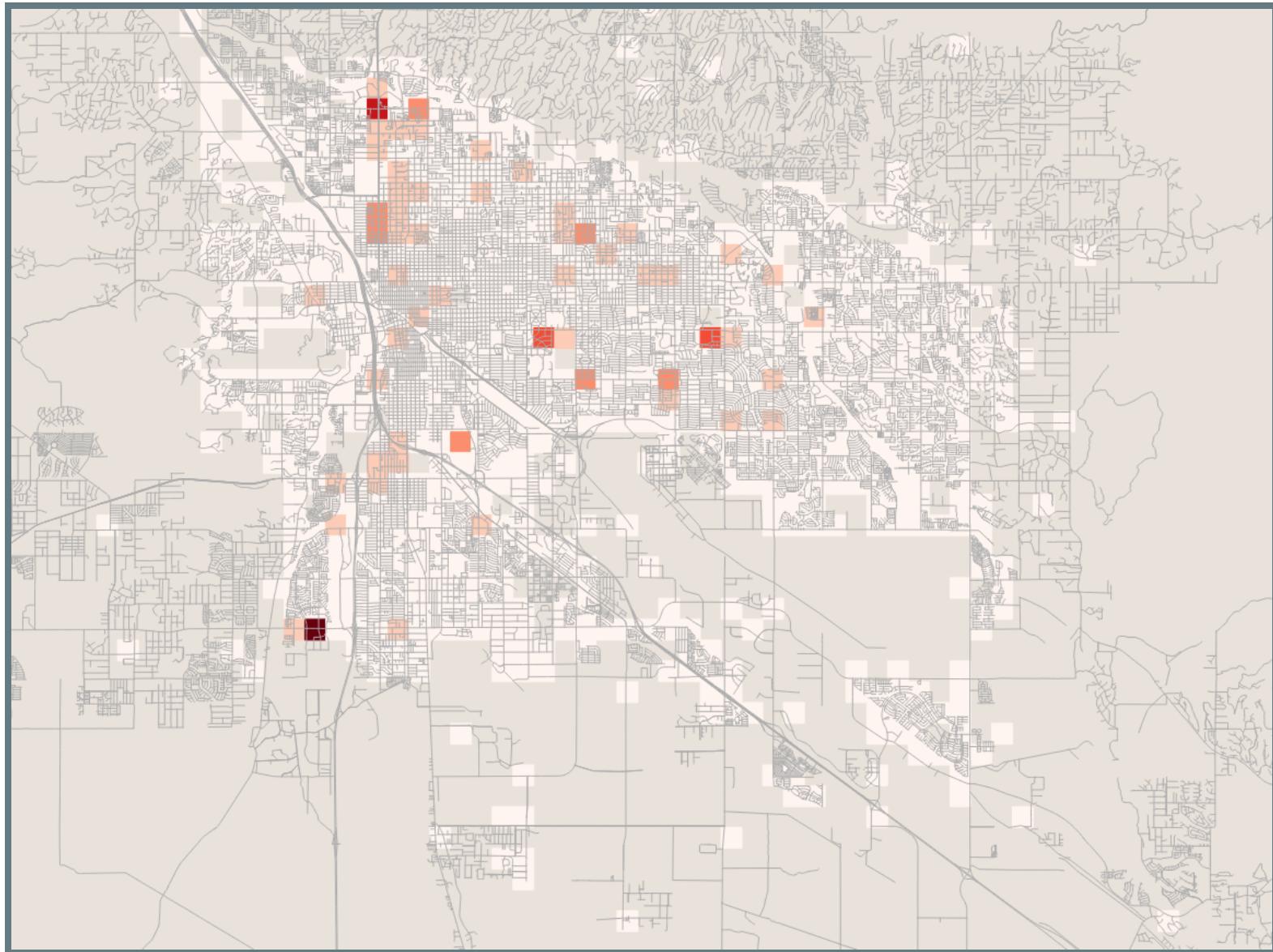
BASIC POINT MAP



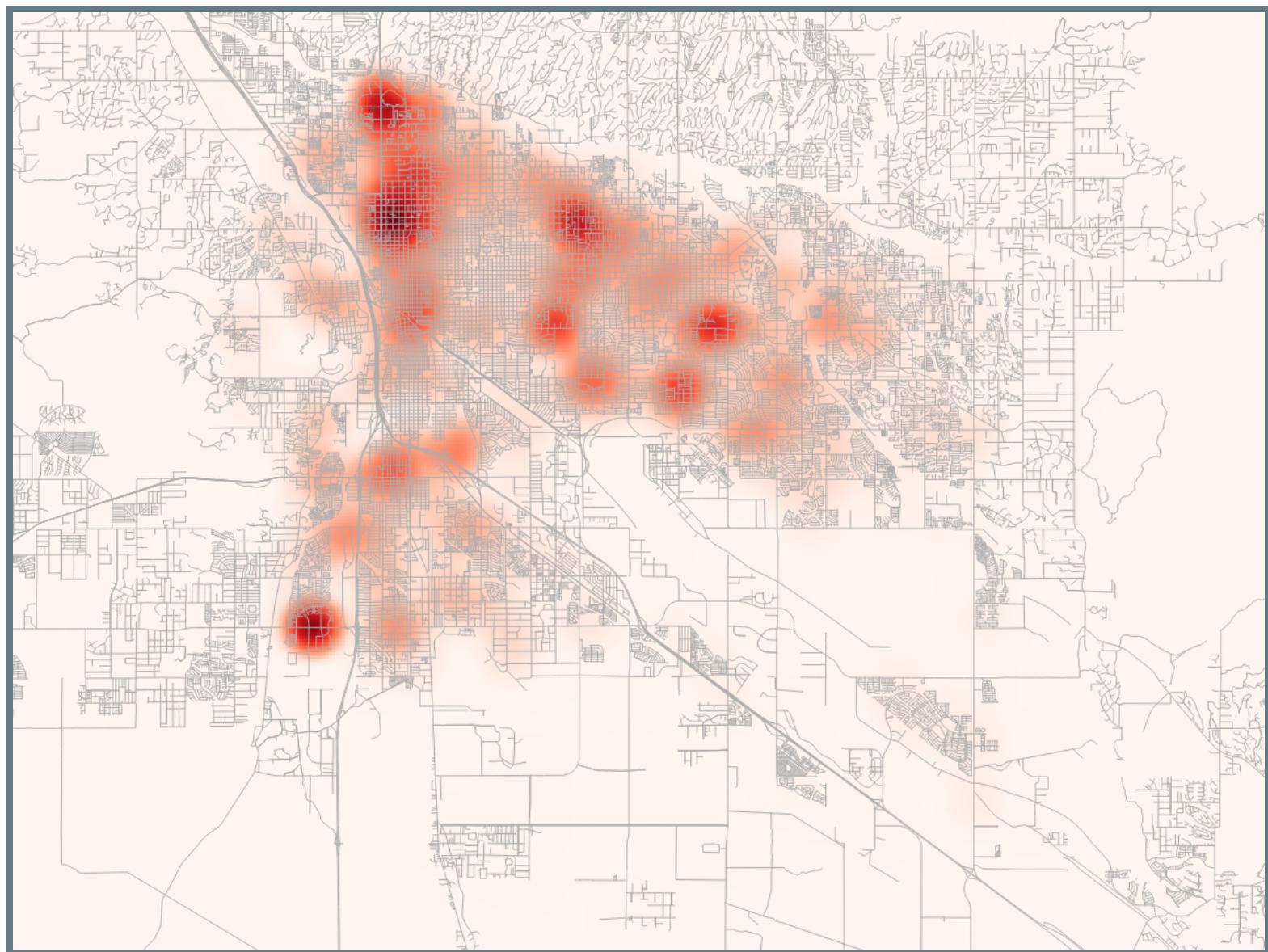
CHLOROPLETH MAP



BINNED MAP



HEAT MAP



A WORD ABOUT PROJECTIONS

**PROJECTIONS ARE HOW WE DRAW A THREE DIMENSIONAL
SPHERE ON A TWO DIMENSIONAL SURFACE**

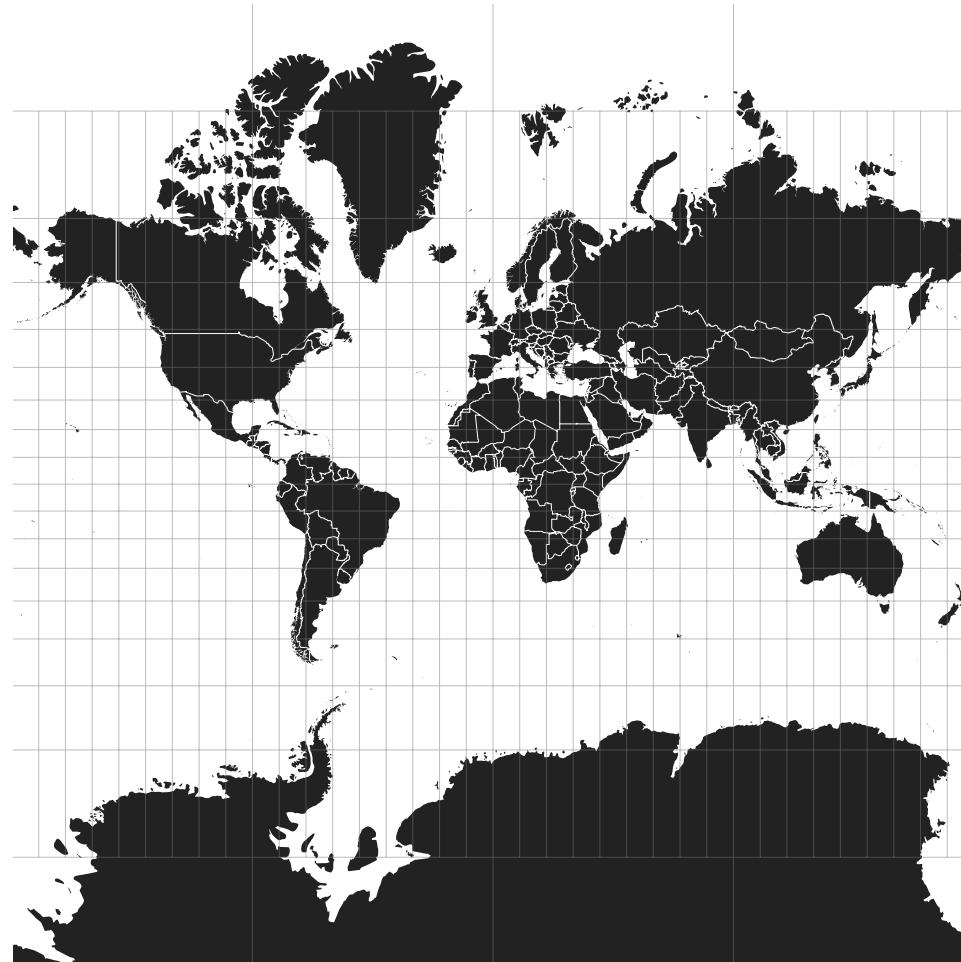


DIFFERENT PROJECTIONS ARE USED TO PRESERVE DIFFERENT PROPERTIES:

- Distance
- Direction
- Area
- Shape

SOME COMMON PROJECTIONS

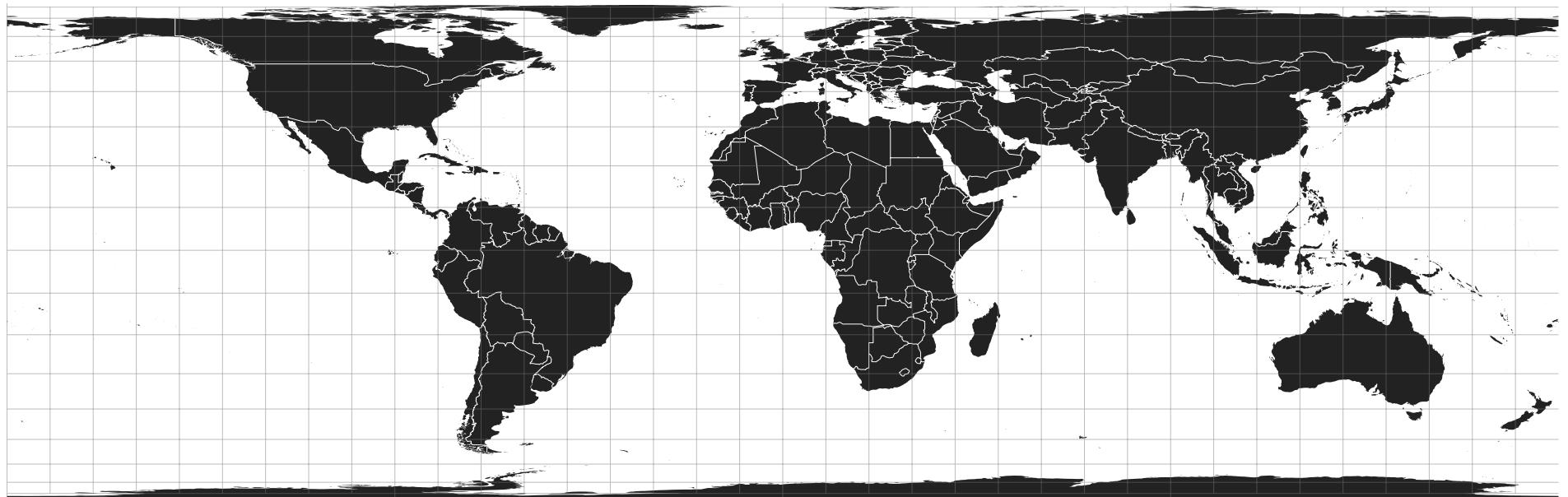
MERCATOR



ALBERS EQUAL AREA CONIC



LAMBERT CYLINDRICAL EQUAL-AREA



WHY DO WE NEED TO KEEP AN EYE ON PROJECTIONS?

All projections distort the world to some extent. Because of this we need to make sure datasets share the same projection when compared.

A WARNING

Mapping software will reproject "on the fly." This is only for viewing purposes; the underlying data is not changed. Your analysis is most likely incorrect if projections do not match.

AND THEN THERE IS DATUM

Datum is the coordinate reference system. In other words, it is how we measure locations on Earth.

SOME COMMON DATUM ARE:

- North American Datum 1927 (NAD27)
- North American Datum 1983 (NAD83)
- World Geodetic System 1984 (WGS84)

WHY DO WE MATCH DATUM BETWEEN FILES?

NAD27 and NAD83 sound similar. However they are based on different mathematical representations of the Earth's shape. The same location in two different datum can be off by a few hundred meters.

AND ALSO MAP UNITS

Many geographic datasets have units set to degrees. Degrees are not a measure of distance or area, but rather the angle of a point relative to the center of the earth.

If your analysis requires measurement of distance or area, convert your data to a projection and datum that support map units in meters or feet.