

EDS 223: Geospatial Analysis & Remote Sensing

Week 2



USGS via Unsplash

Welcome!

- Recap on week 1

Welcome!

- Recap on week 1
- Spatial data models

Welcome!

- Recap on week 1
- Spatial data models
- Vector data models

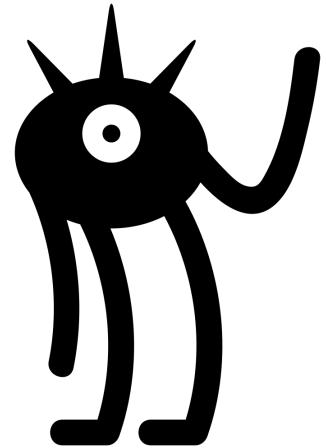
Welcome!

- Recap on week 1
- Spatial data models
- Vector data models
- Intro to ‘sf’

Welcome!

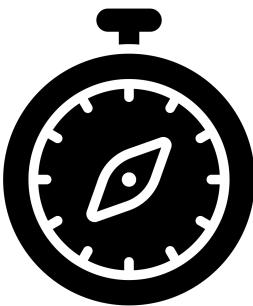
- Recap on week 1
- Spatial data models
- Vector data models
- Intro to ‘sf’
- Spatial data science stories

Week 1 recap



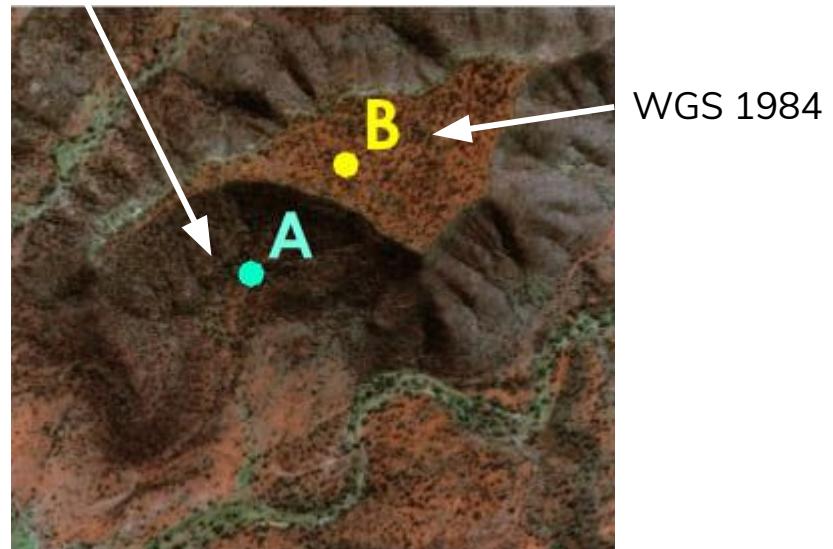
Coordinate reference systems

What does this look like in the real world?

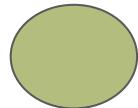


134.577°E, 24.006°S

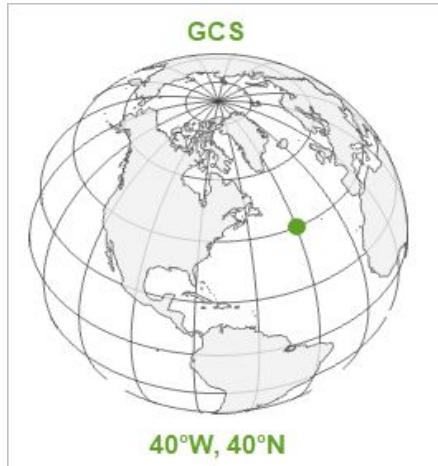
Australian Geodetic Datum 1984



Geographic vs. projected coordinate system



Geographic	Projected
Defines where the data is located on Earth	Provides instructions on how to draw the data onto a flat surface
3D	2D
Describes locations as angles	Describes locations in linear units



Geographic vs. projected coordinate system

- A PCS is a GCS that has been flattened using a map projection
- You can store data in a GCS, but you can't draw it on a flat map without a PCS
- Picking a GCS depends on where you are mapping
- Picking a PCS depends on where you are mapping AND the nature of the map you want to make

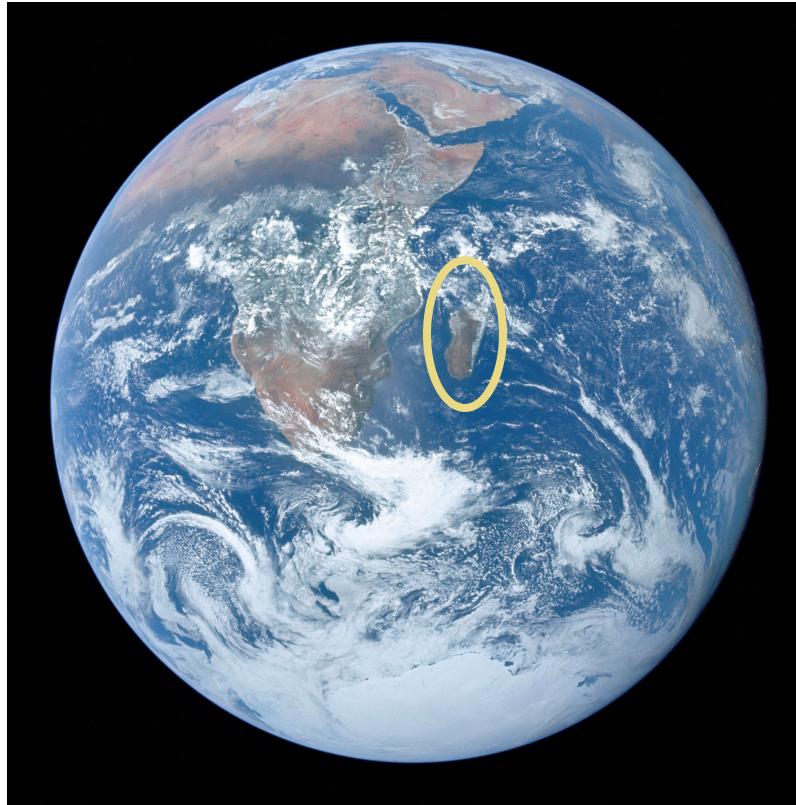
Spatial data models



Spatial data models



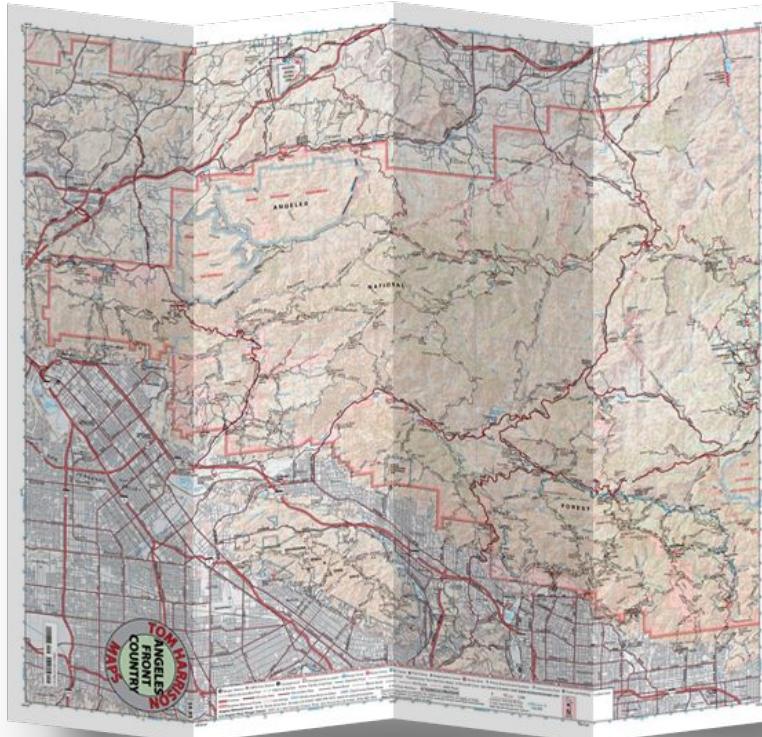
Spatial data models



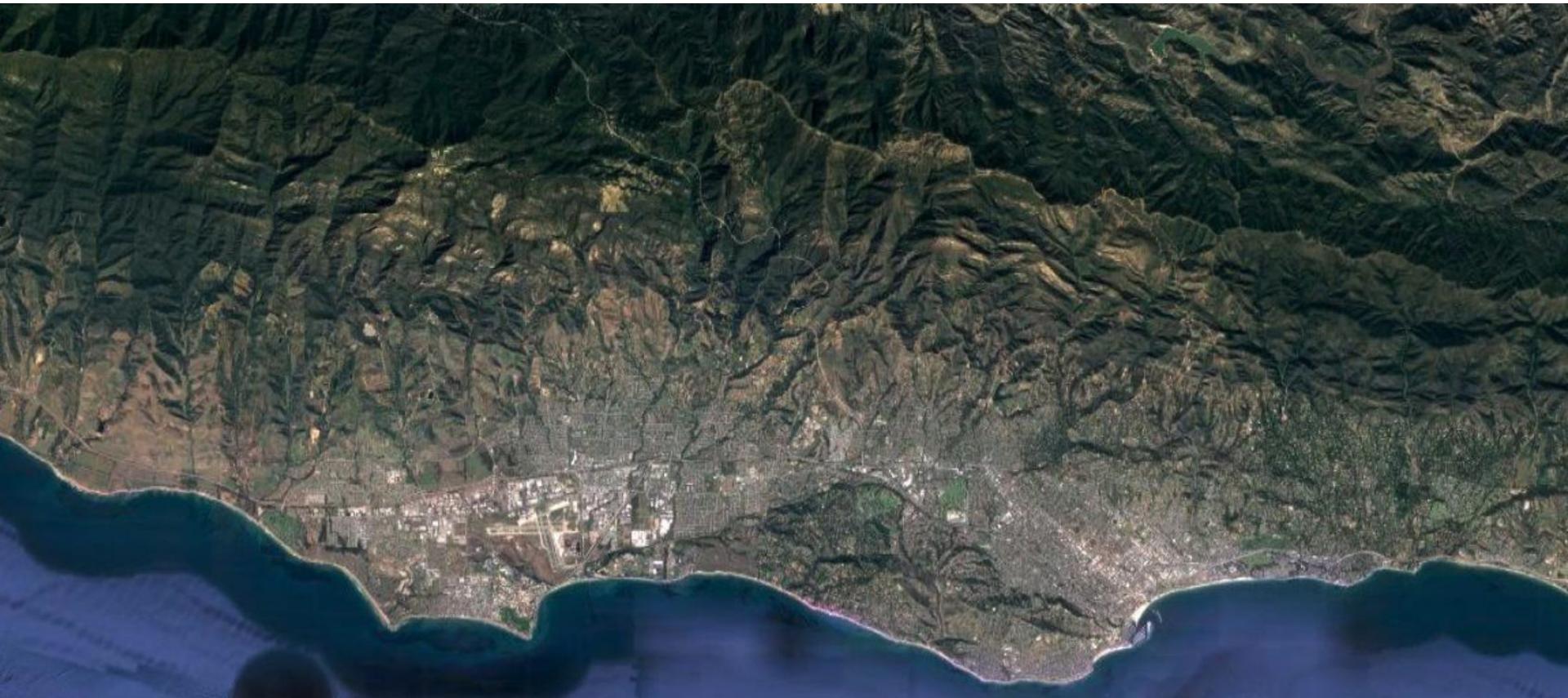
All models are wrong, but
some are useful.

- George E.P. Box

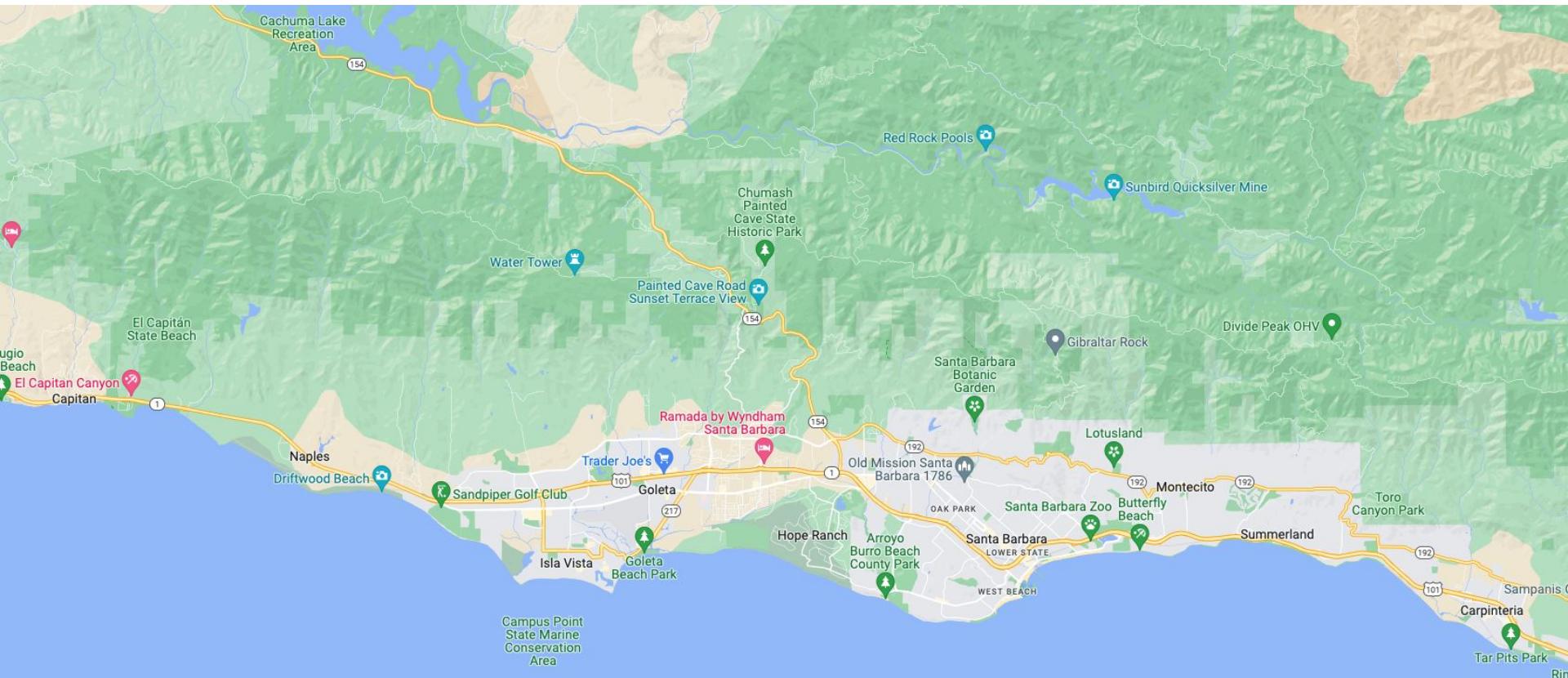
Maps are models!



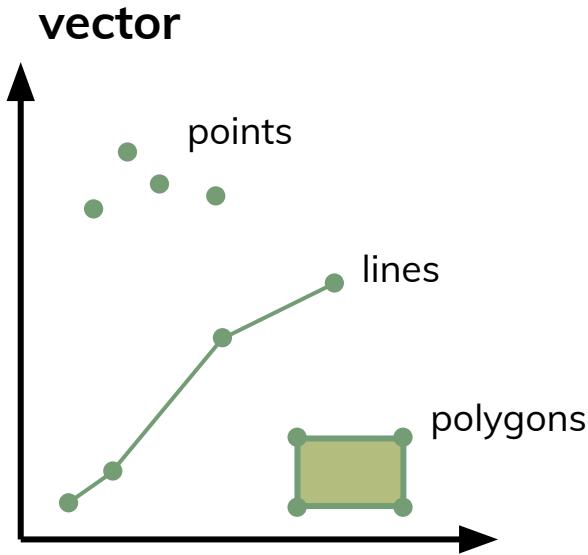
Spatial data models



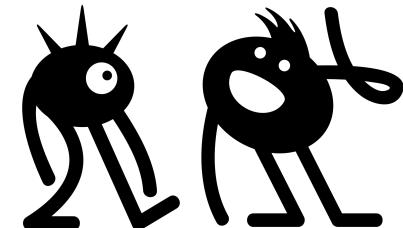
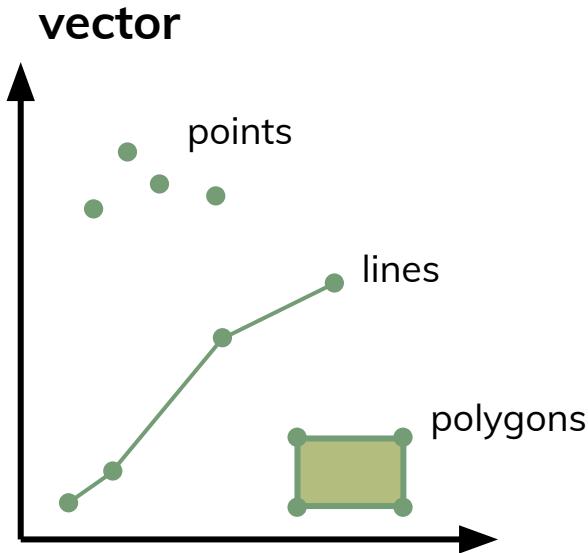
Spatial data models



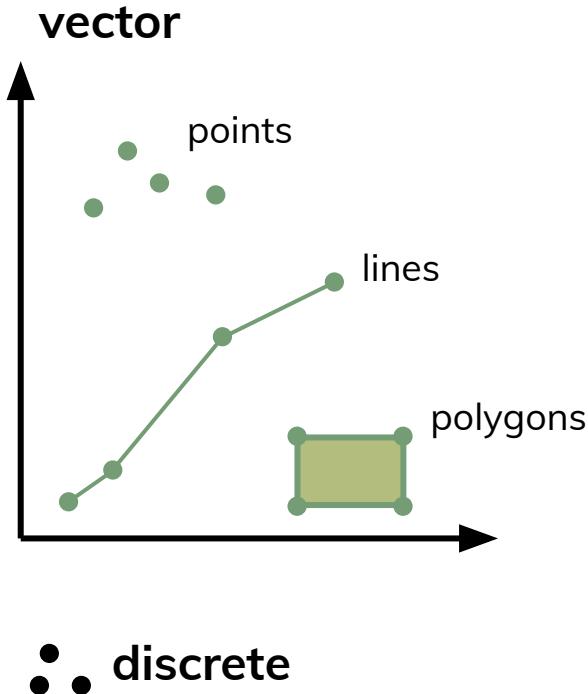
Spatial data models



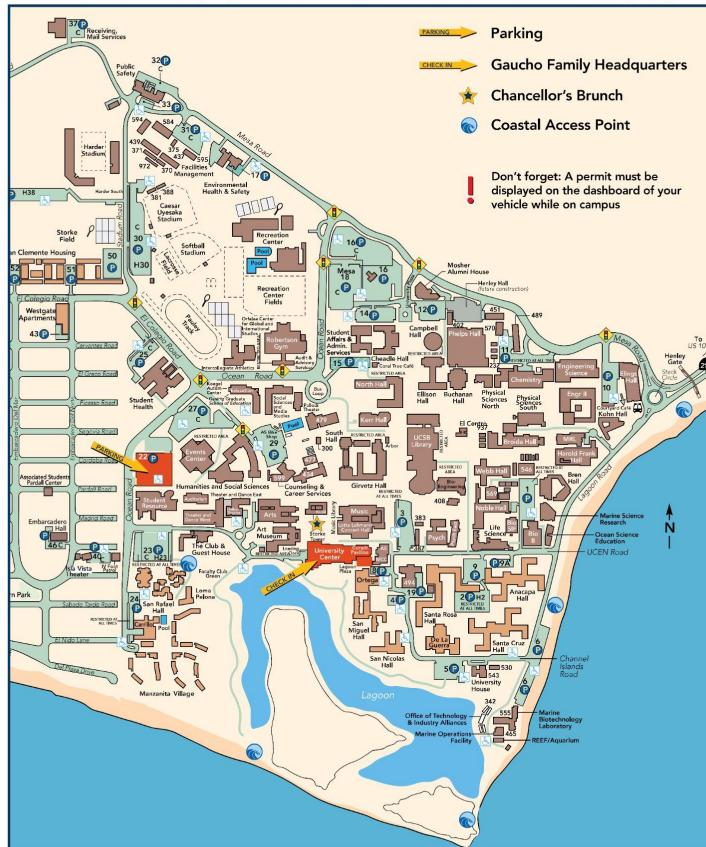
Spatial data models



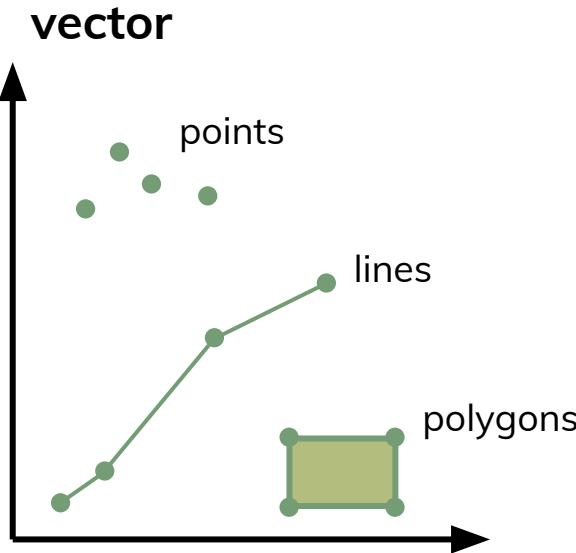
Spatial data models



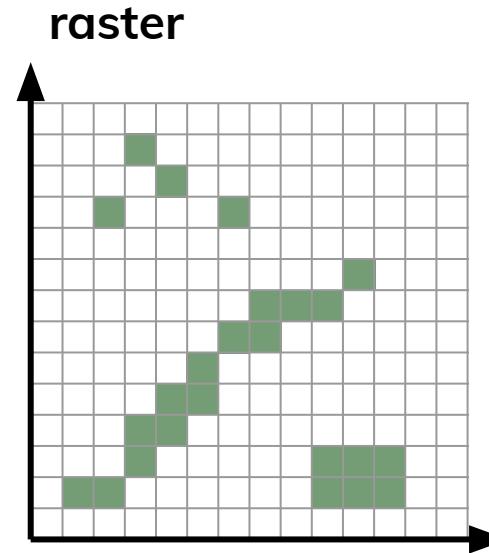
Spatial data models



Spatial data models

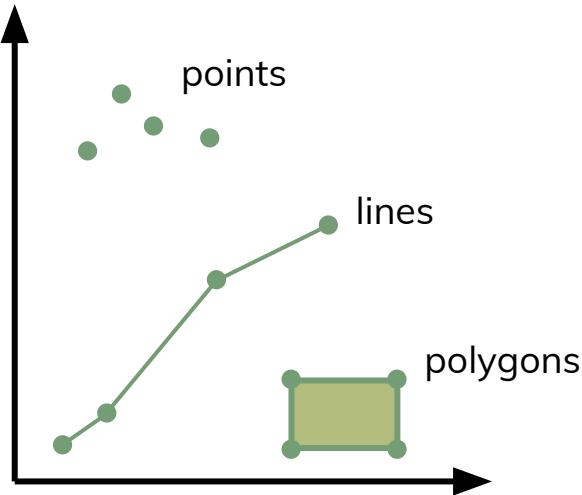


• • discrete



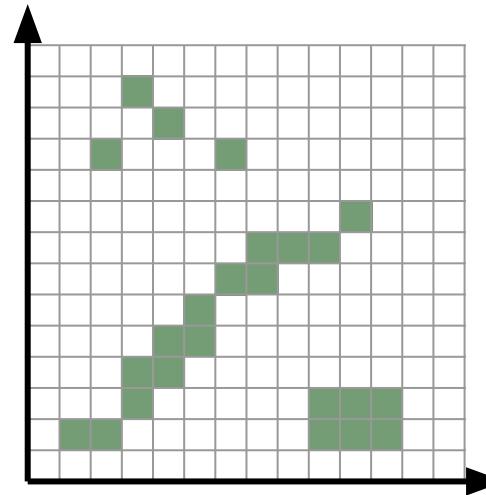
Spatial data models

vector



• discrete

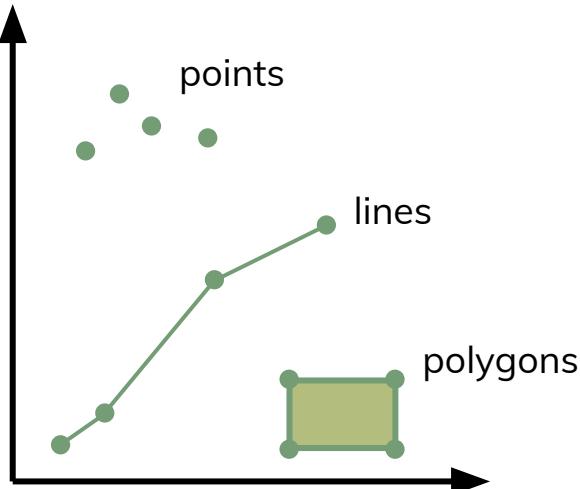
raster



continuous

Spatial data models

vector



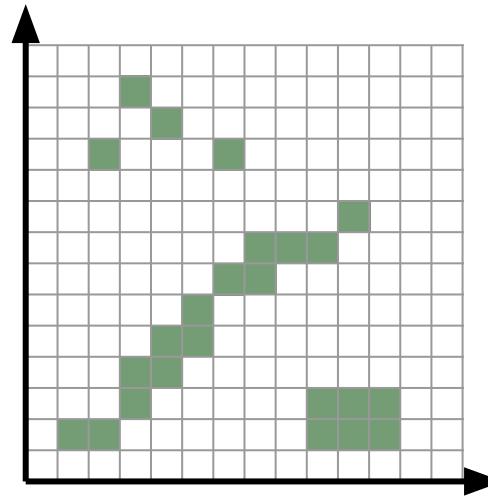
points

lines

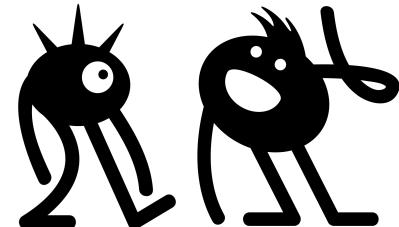
polygons

• discrete

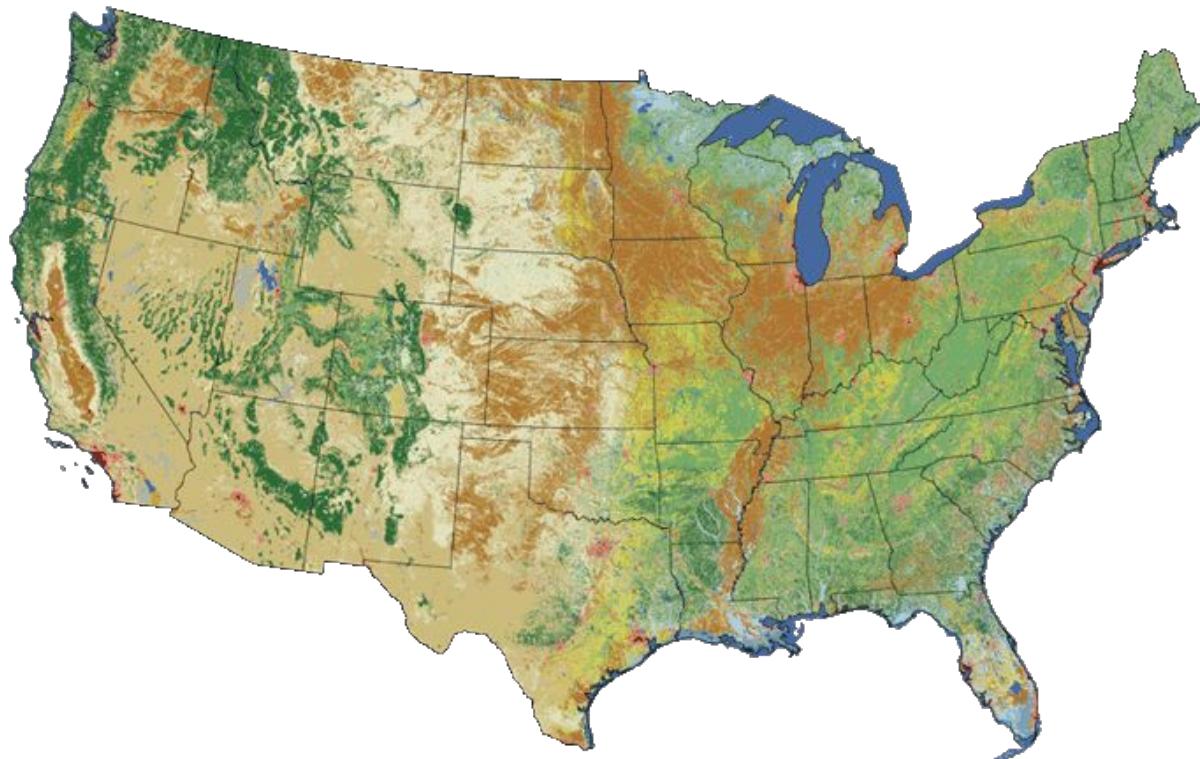
raster



continuous

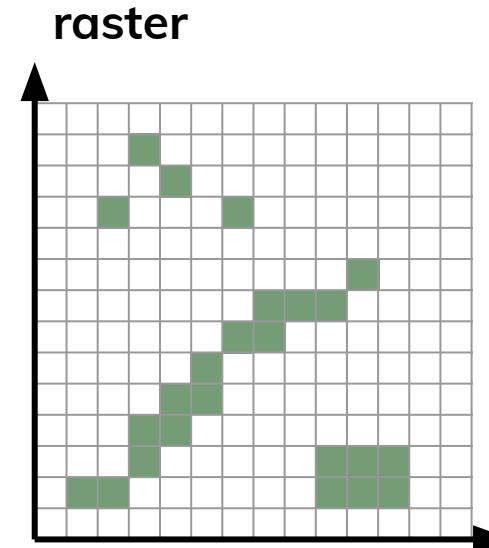
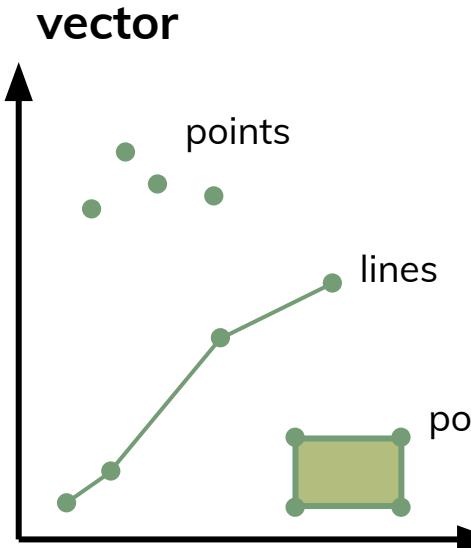


Spatial data models



Source: Multi-Resolution Land Characteristics Consortium

Spatial data models



• discrete

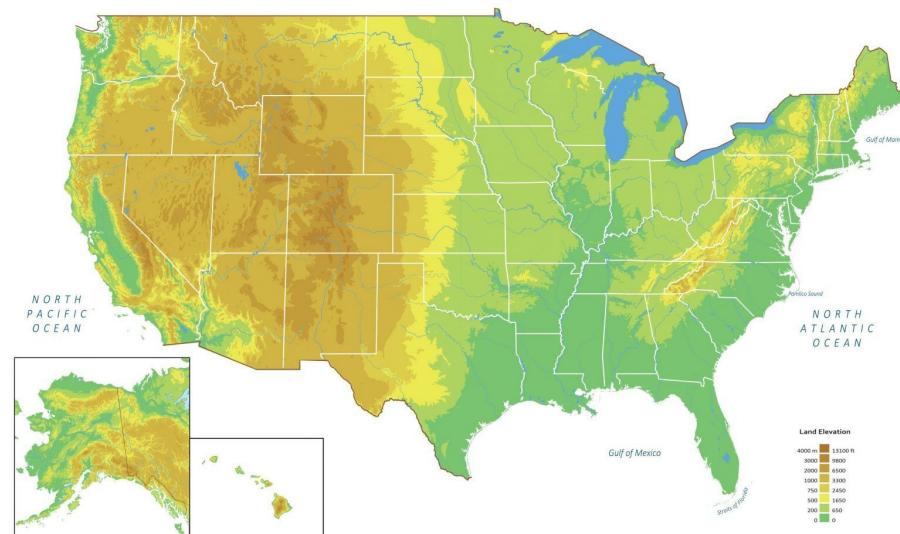
continuous

Spatial data models

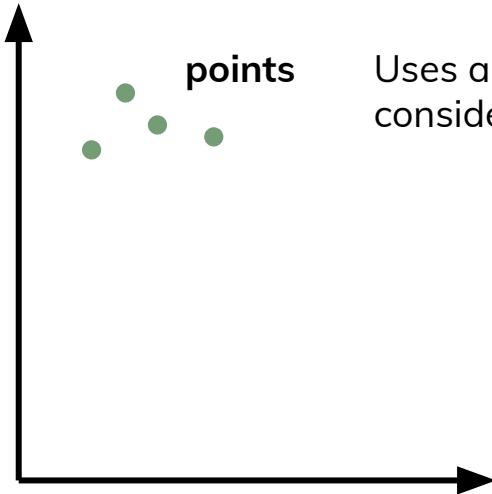
vector



raster

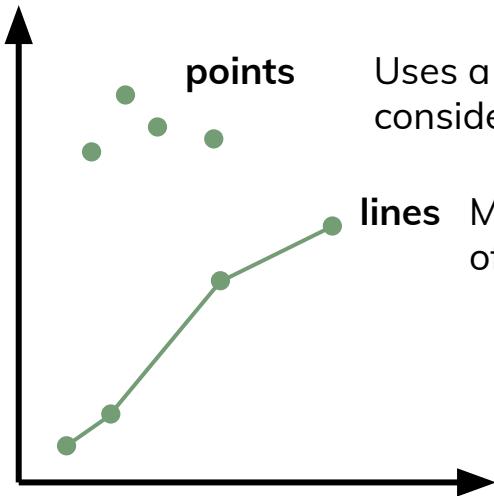


Vector data models



Uses a single coordinate pair to represent the location of an entity that is considered to have no dimension.

Vector data models



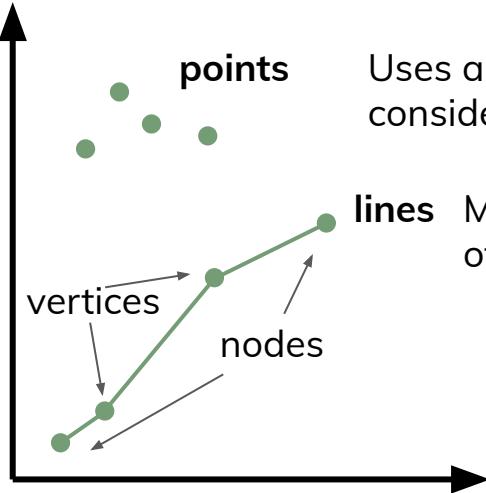
points

Uses a single coordinate pair to represent the location of an entity that is considered to have no dimension.

lines

Made up of line segments that run between adjacent, ordered sets of coordinate pairs.

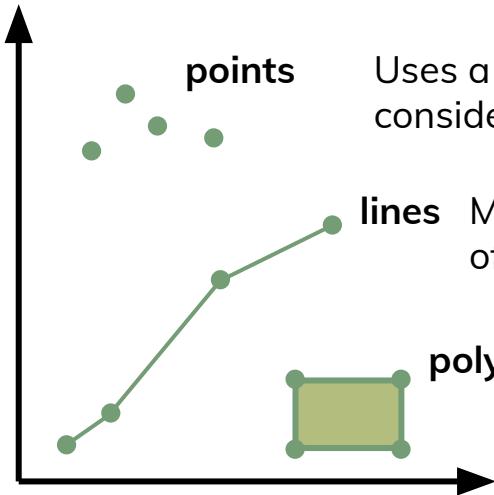
Vector data models



points Uses a single coordinate pair to represent the location of an entity that is considered to have no dimension.

lines Made up of line segments that run between adjacent, ordered sets of coordinate pairs.

Vector data models



points

Uses a single coordinate pair to represent the location of an entity that is considered to have no dimension.

lines

Made up of line segments that run between adjacent, ordered sets of coordinate pairs.

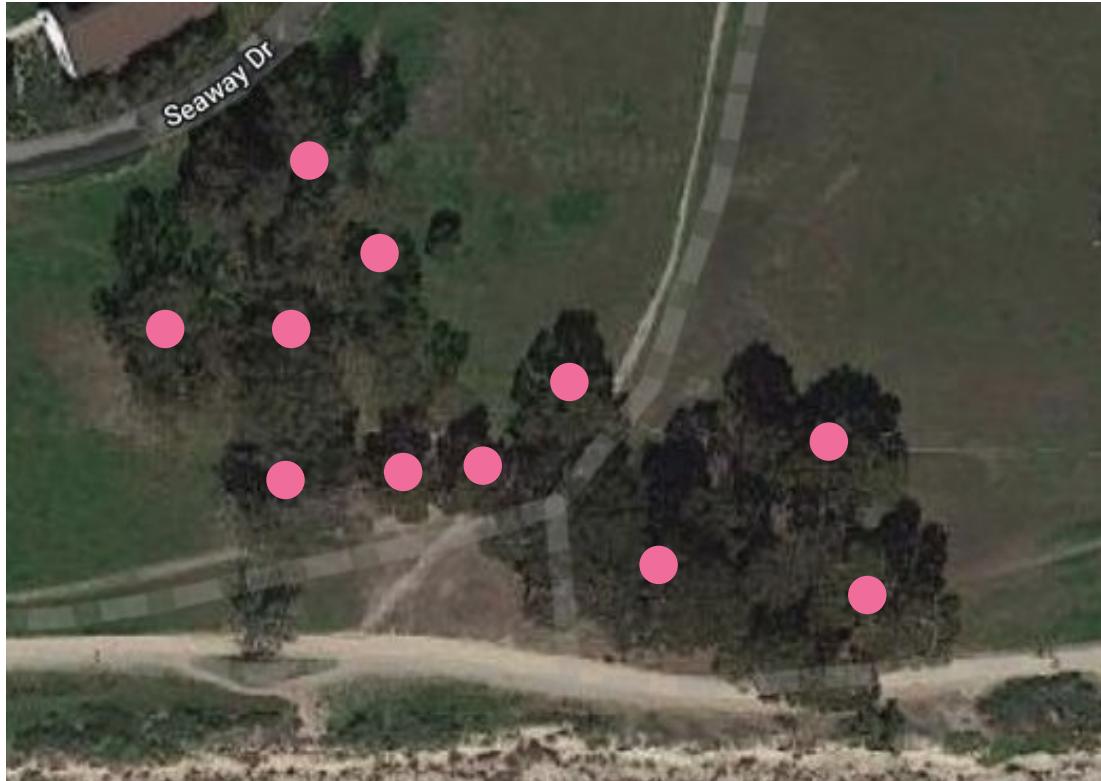
polygons

Formed by a set of connected lines.

Vector data models



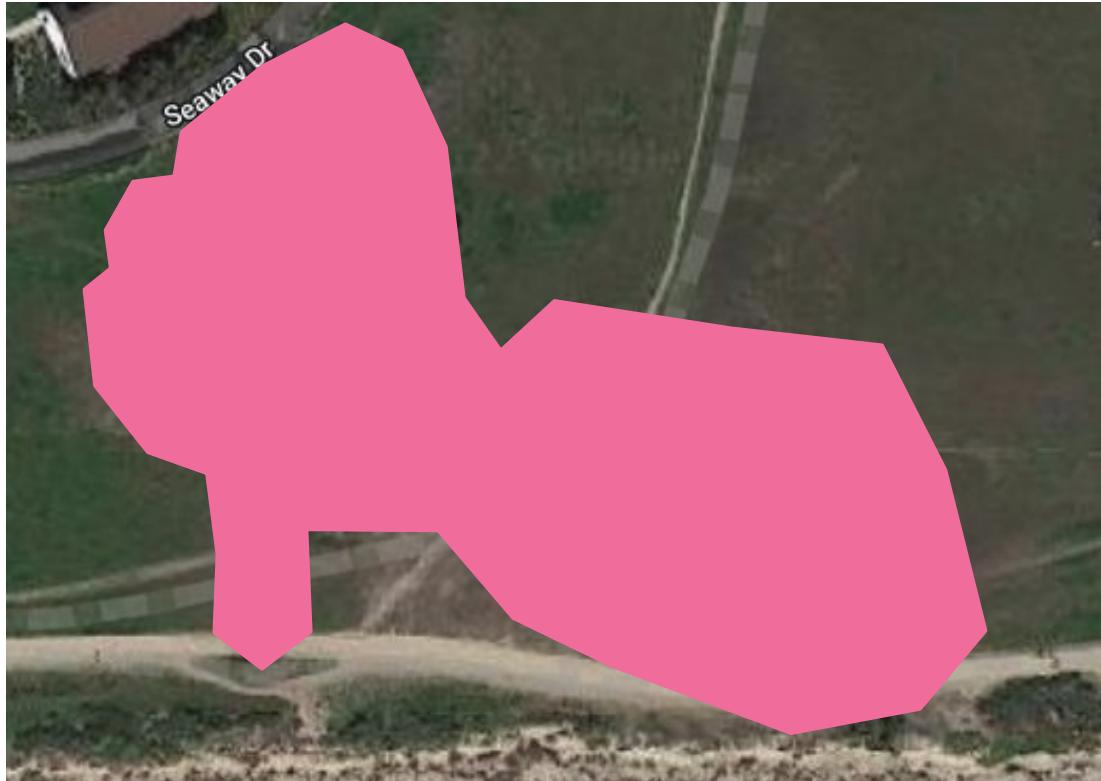
Vector data models



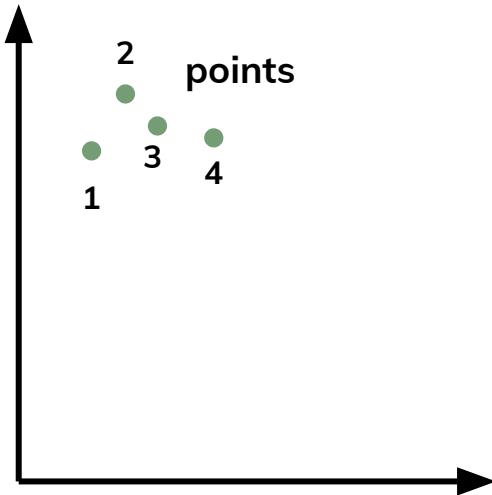
Vector data models



Vector data models



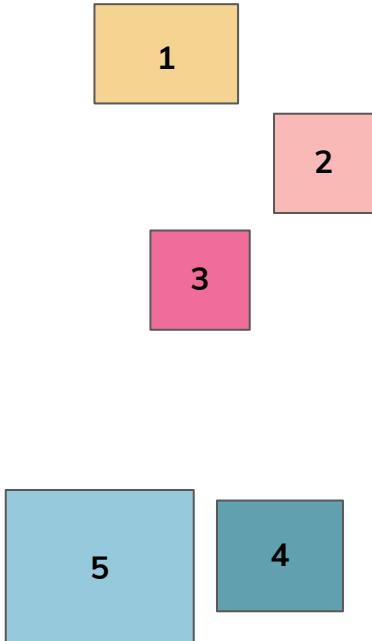
Vector data models



ID	Species	Age
1	Poplar	11
2	Oak	2
3	Beech	12
4	Cedar	15

Single vs. multi-part features

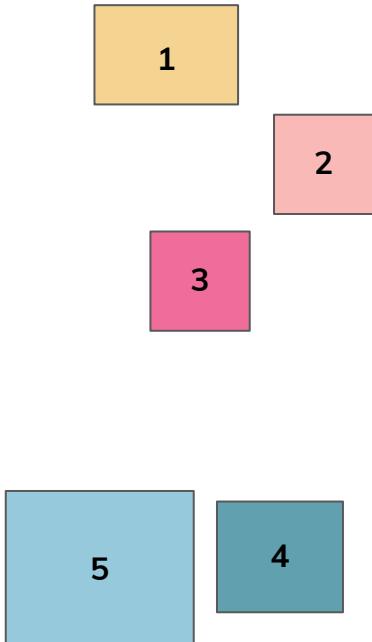
single-part



ID	Building
1	A
2	B
3	C
4	D
5	E

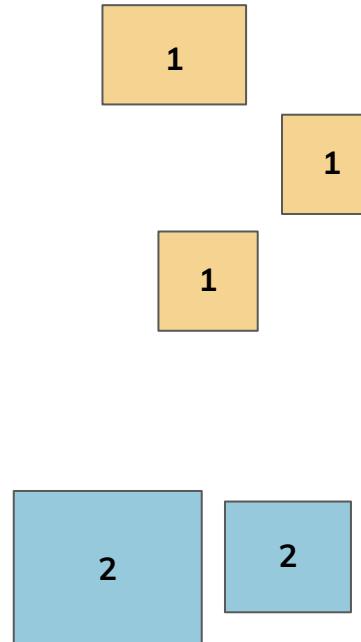
Single vs. multi-part features

single-part



ID	Building
1	A
2	B
3	C
4	D
5	E

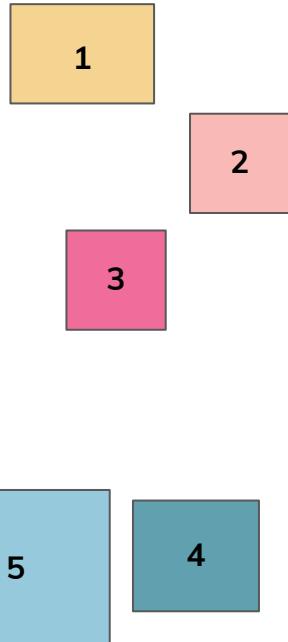
multi-part



ID	Campus
1	Main
2	Downtown

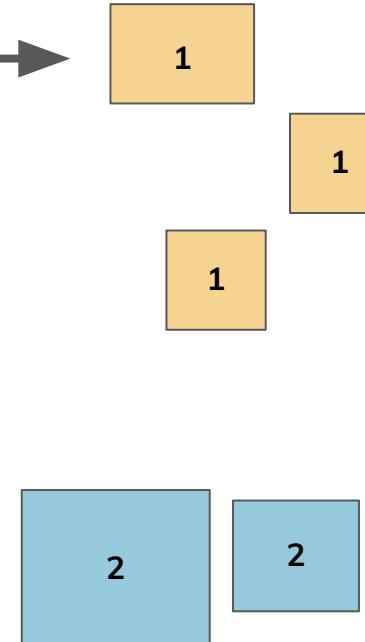
Single vs. multi-part features

single-part



ID	Building
1	A
2	B
3	C
4	D
5	E

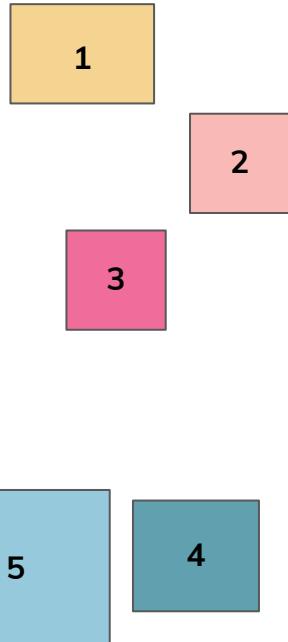
multi-part



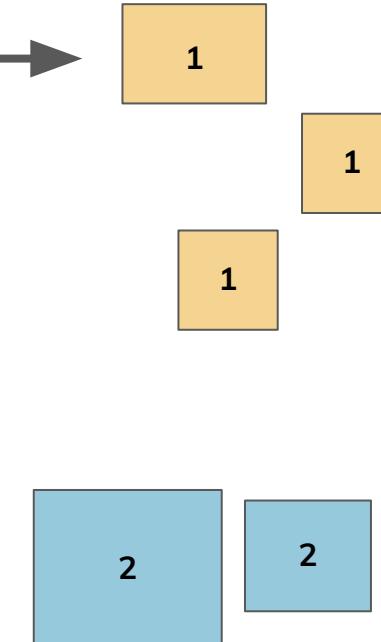
ID	Campus
1	Main
2	Downtown

Single vs. multi-part features

single-part



multi-part



ID	Building
1	A
2	B
3	C
4	D
5	E

ID	Campus
1	Main
2	Downtown

Notes of caution



Polygon inclusions



ID	Cover type	Species
1	Tree	Oak
2	Tree	Eucalyptus
3	Tree	Eucalyptus

Boundary generalizations



Spatial tools



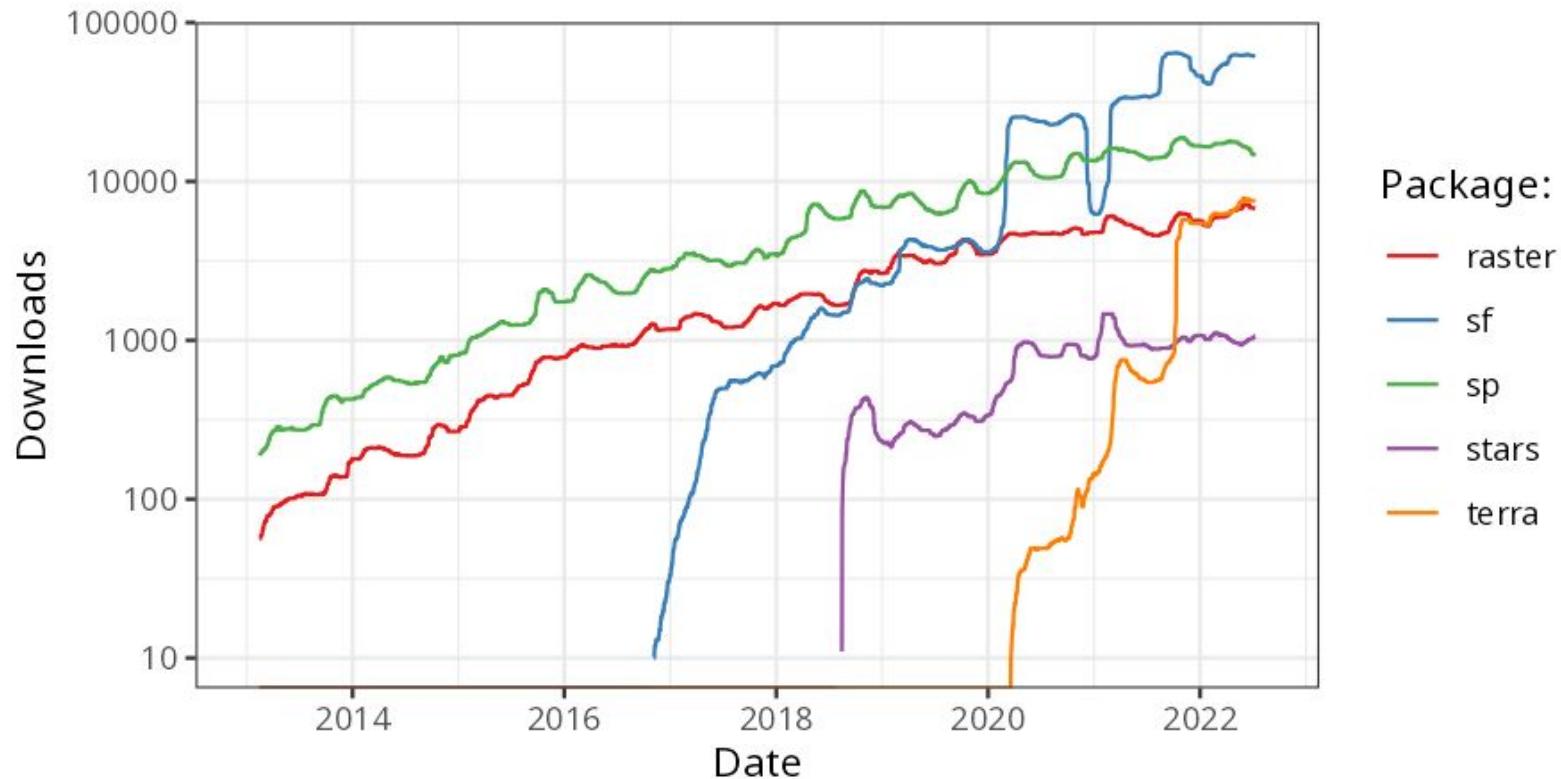
Spatial analysis with R



Spatial analysis with R



R's spatial ecosystem



Simple features: sf



Simple features: **sf**

Simple feature geometry
(sfg object)



(-36, 68)



Simple features: **sf**

Simple feature geometry
(sfg object)



(-36, 68)

Simple feature geometry column
(sfc object)



(-36, 68)
NAD38



Simple features: **sf**

Simple feature geometry
(sfg object)



(-36, 68)

Simple feature geometry column
(sfc object)



(-36, 68)
NAD38

Simple feature
(sf object)



(-36, 68)
NAD38



ID	Name	State
1	Toolik Field Station	Alaska

Advantages of **sf**

- Fast data reading and writing

Advantages of **sf**

- Fast data reading and writing
- Enhanced plotting performance

Advantages of **sf**

- Fast data reading and writing
- Enhanced plotting performance
- Consistent function names

Advantages of **sf**

- Fast data reading and writing
- Enhanced plotting performance
- Consistent function names
- Behave like data frames

Advantages of **sf**

- Fast data reading and writing
- Enhanced plotting performance
- Consistent function names
- Behave like data frames

Advantages of **sf**



Advantages of sf

dplyr : go wrangling

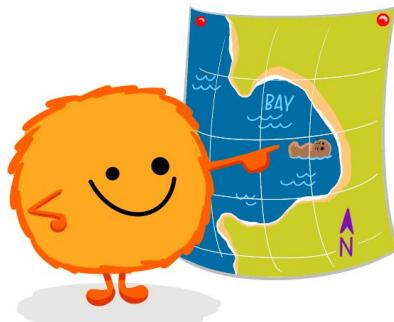


Advantages of sf

dplyr::filter()

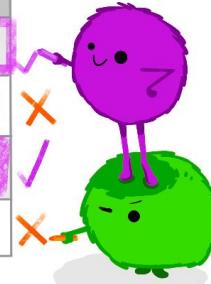
KEEP ROWS THAT
satisfy
your CONDITIONS

keep rows from... this data... ONLY IF... type is "otter"
filter(df, type == "otter" & site == "bay")
AND site is "bay"



	type	food	site
1	otter	urchin	bay
2	Shark	seal	channel
3	otter	abalone	bay
4	otter	crab	wharf

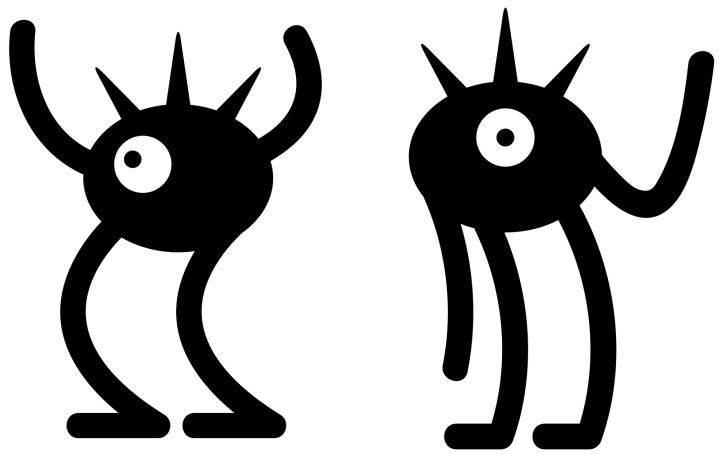
@allisonhorst



Advantages of sf



How are we feeling?



Spatial data science stories

Biodiversity data coverage

- International targets on biodiversity need metrics to track progress



Convention on
Biological Diversity

Biodiversity data coverage

- International targets on biodiversity need metrics to track progress
 - Target 19: By 2030, ensure that quality information, including traditional knowledge, is available to decision makers and public for the effective management of biodiversity through promoting awareness, education and research.



Convention on
Biological Diversity

Biodiversity data coverage



2,227,668,244

Occurrence records

76,365

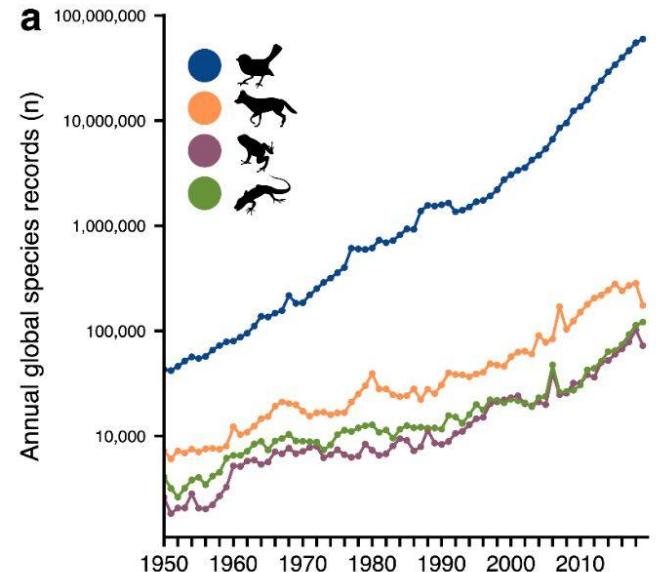
Datasets

1,913

Publishing institutions

7,765

Peer-reviewed papers using data



Oliver et al. 2021

Biodiversity data coverage



2,227,668,244

Occurrence records



76,365

Datasets



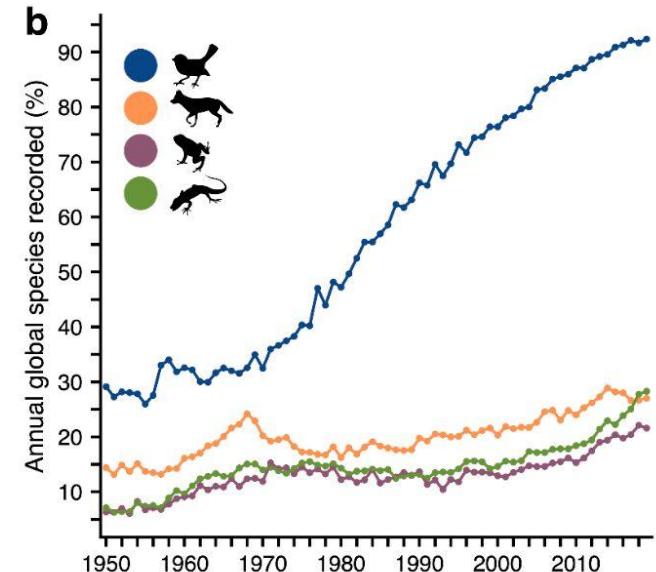
1,913

Publishing institutions



7,765

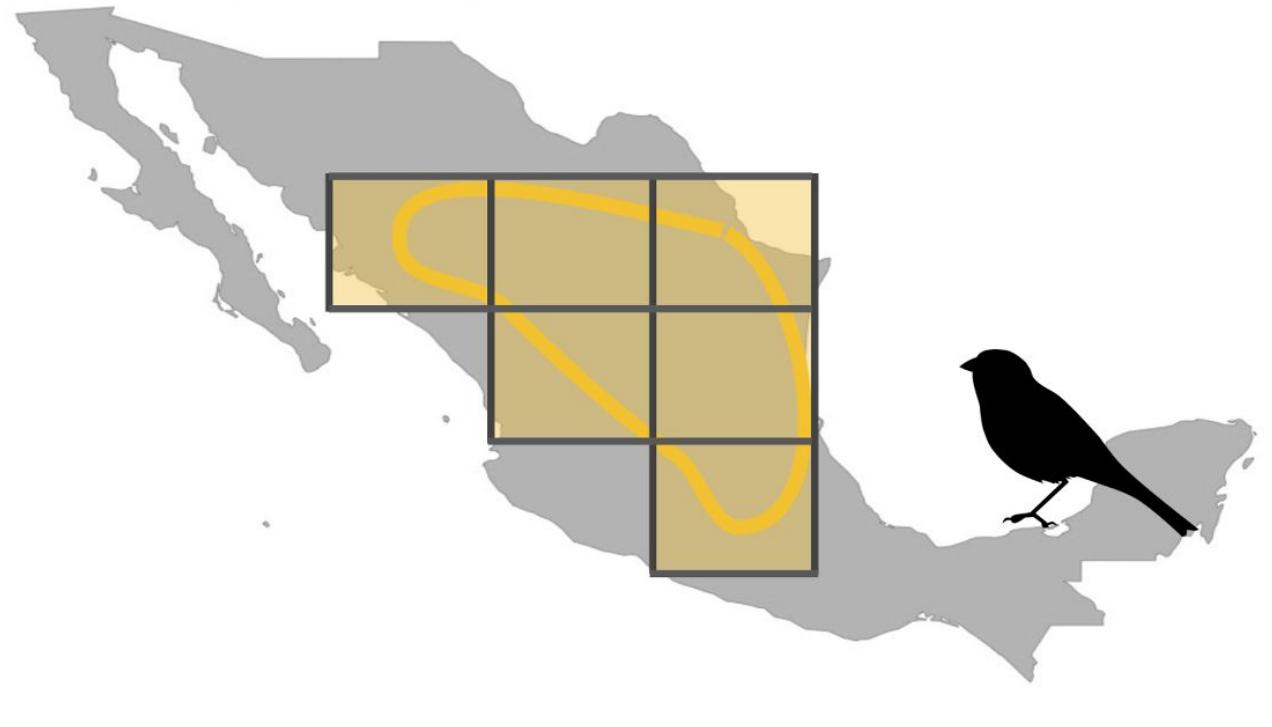
Peer-reviewed papers
using data



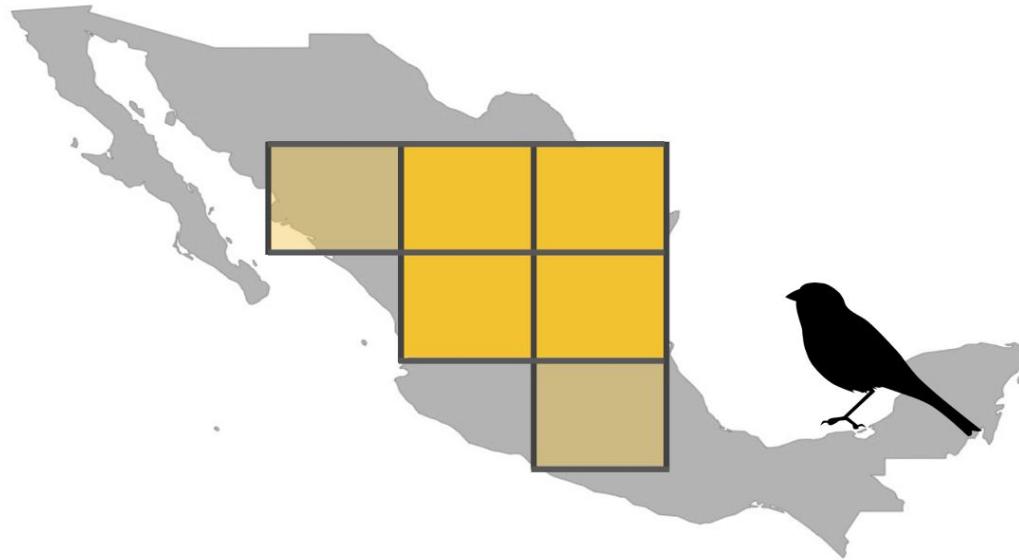
Oliver et al. 2021

Biodiversity data coverage

Gridded species ranges and national boundaries

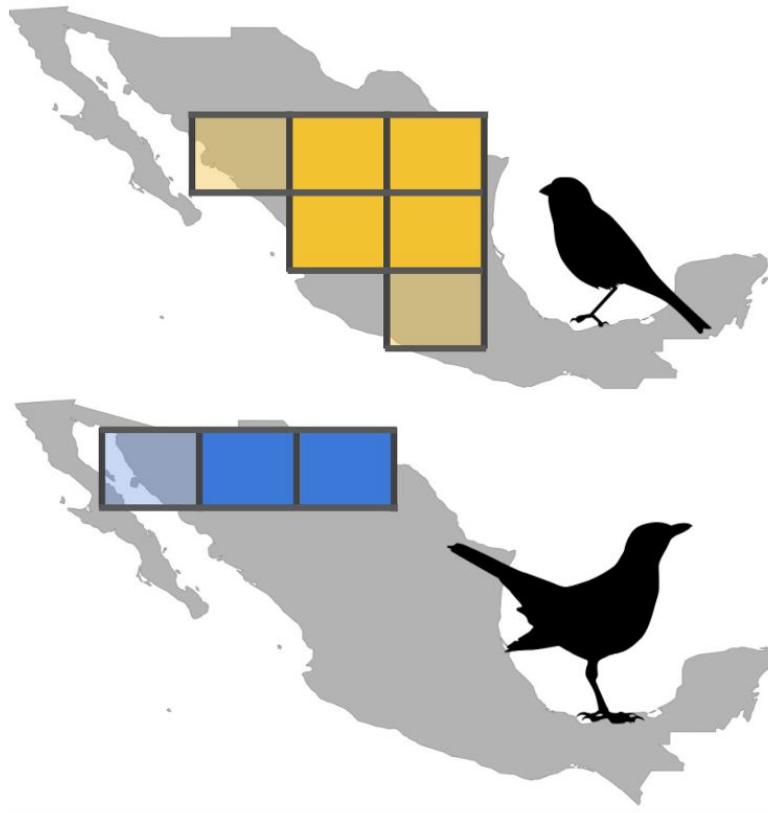


Biodiversity data coverage

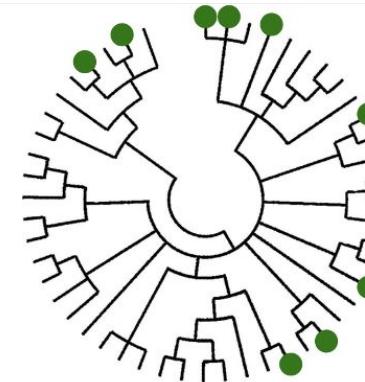


$$\text{Data coverage} = \frac{\sum \text{cells with records}}{\sum \text{expected cells}}$$

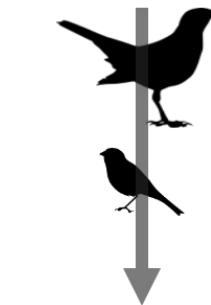
Biodiversity data coverage



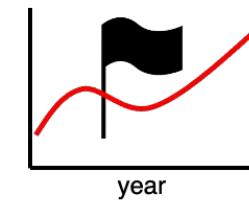
taxonomic diversity



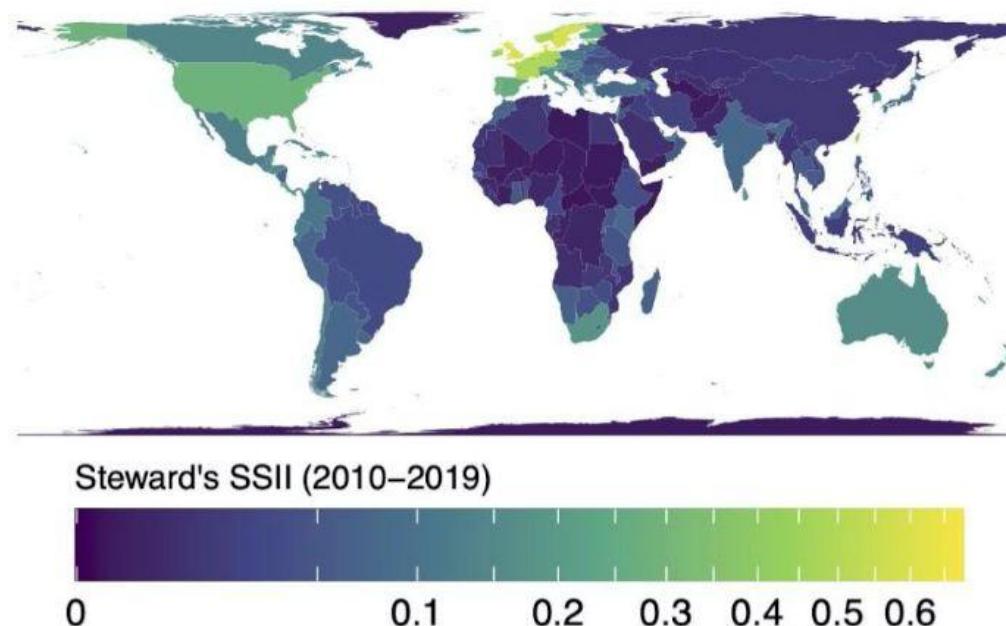
species-level



national-level



Biodiversity data coverage



Course logistics

ryoliver.github.io/EDS_223_spatial_analysis