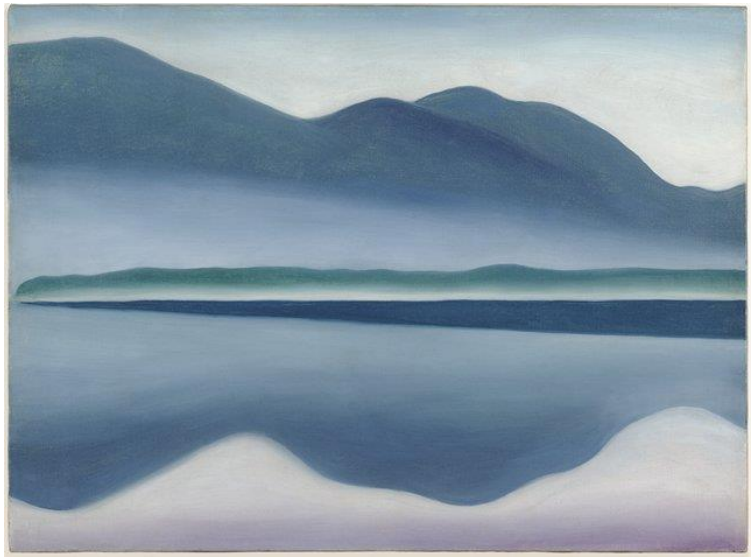
Generalisation

"the reduction of detail or simplification of reality"

Lawren Harris

"Nothing is less real than realism. Details are confusing. It is only by selection, by elimination, by emphasis that we get at the real meaning of things". *Georgia O'Keefe (1922)*



Lake George, Georgia O'Keefe

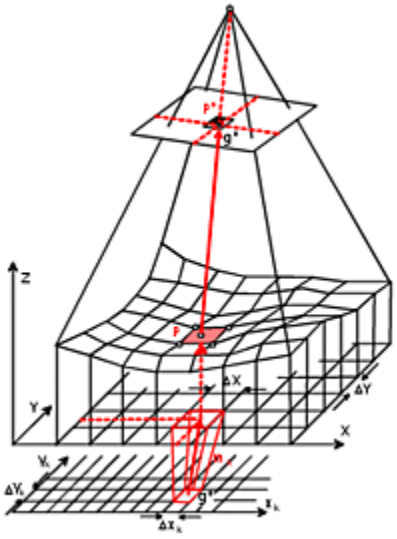
Generalisation

Watercolor map of Greater Vancouver rendered by Stamen Designs using OpenStreetMap data.
https://stamen.com/

Aerial Photos -> ORTHOPHOTOS

Once corrected, and **georeferenced,** photos can be used for (topographic) mapping and also as a base layer, with map data overlain on top.

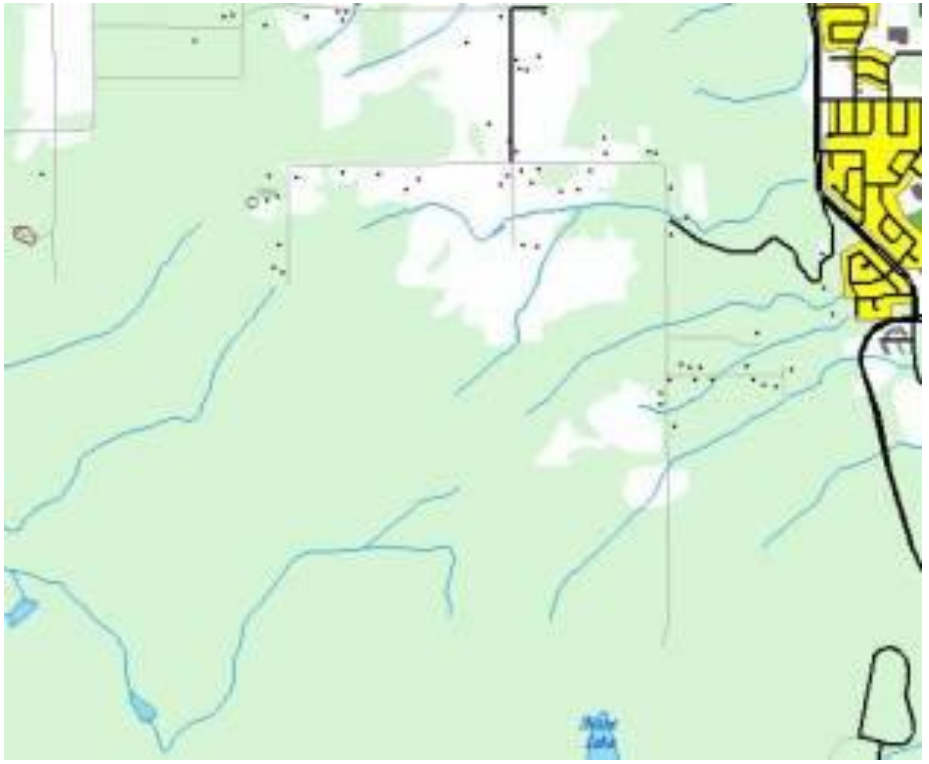
e.g. google maps, pgmap or BC- imap



"A map is a scaled, 2-D, generalized representation of a planetary surface"

Photos/images are not generalized: they need interpretation to be readable by users.

Air photo interpreted (= generalised) -> map



The process of **generalization** consists of these (three +) steps:

1. simplification -> **2. selection** -> **3. classification** (4. -> symbolisation)

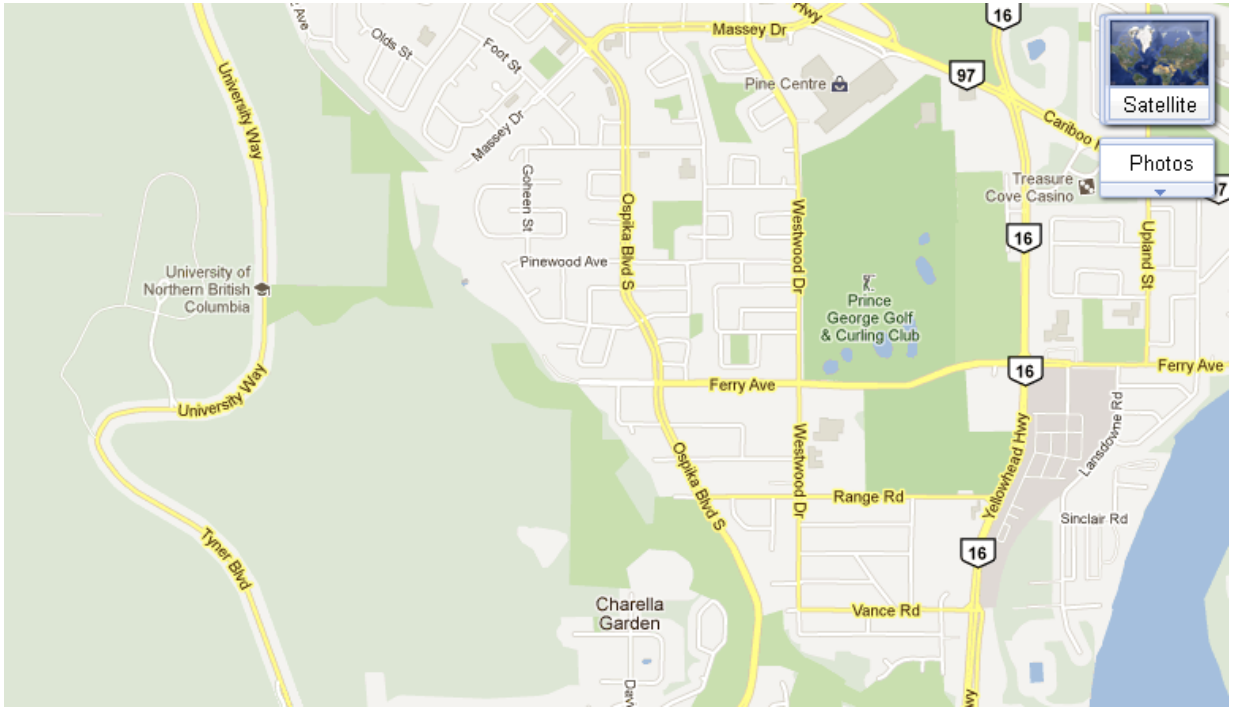
1. Simplification - spatial

All world features have three dimensions: length, width and height

They lose the third dimension as map displays are usually flat.

Simplification determines the most important characteristics (dimensions)

- 3D Volumes (reality)
- 2D Areas / polygons
- 1D Lines / arcs
- 0D Points



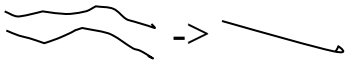
http://maps.google.ca

As SCALE decreases, generalisation increases

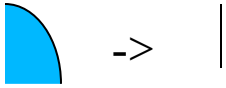
Simplification: features lose dimensionality

This is called **collapsing**

Examples :

areas -> lines: e.g. rivers, roads (left/right bank/edge-> single lines) 

areas -> points: e.g. cities (complex shape -> circle, dot) 

lines -> points: ? (few examples – dams ?) 

2. Selection - spatial

- maps /data can NOT retain all features/details

- Which feature types depend on the map **purpose**

  e.g. topographic/thematic map, parks map, city map

- How many features/details depends mostly on map **scale**

  e.g. more details on 1:50,000 than 1:250,000

<u>Selection</u> can be subdivided into:

a. Aggregation (merging of several elements, most common with **areas**)

b. Elimination (removal of certain elements: **points, lines and areas**)

c. Smoothing (removal of details in shape or outline: **lines and areas**)

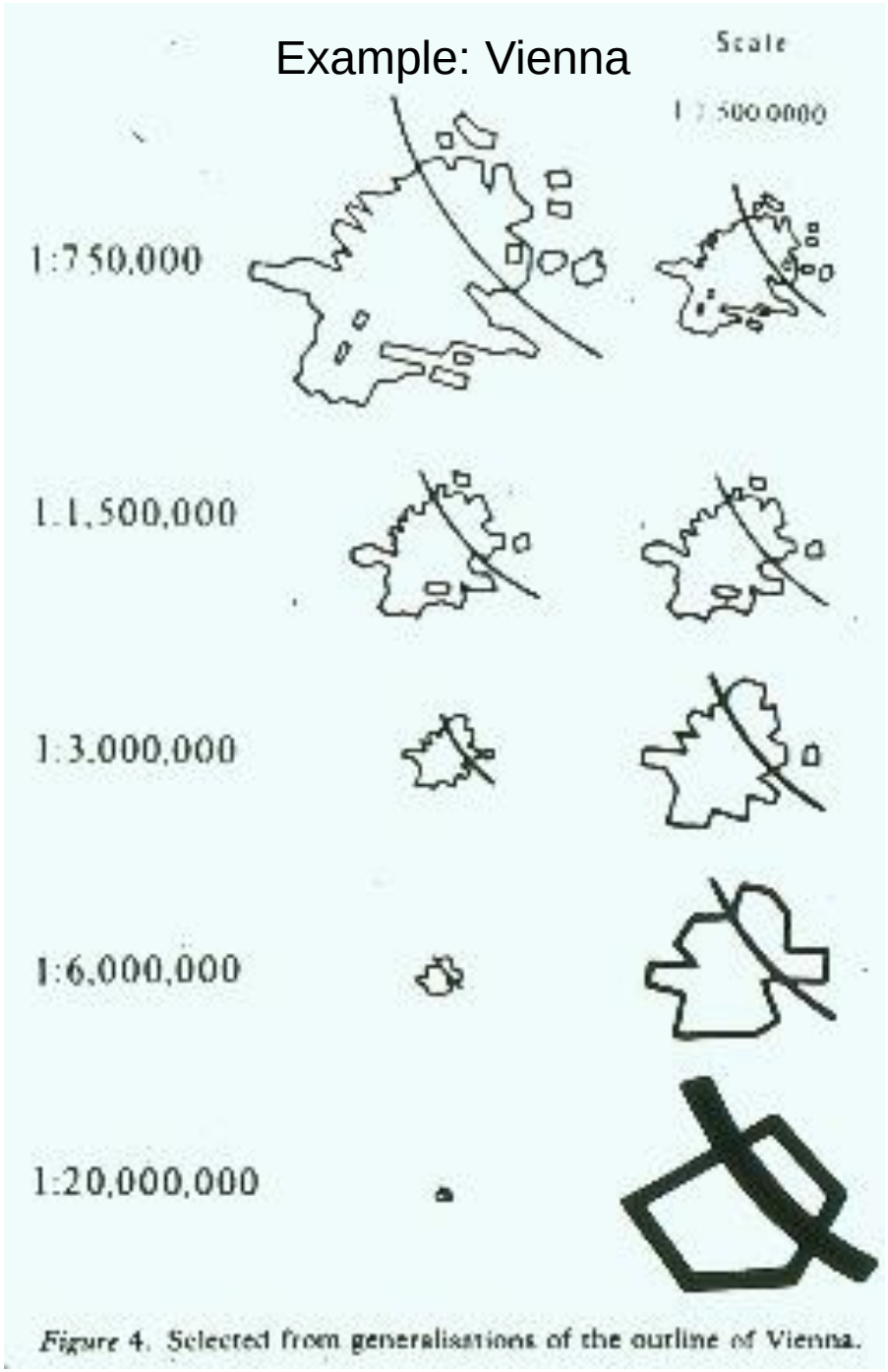
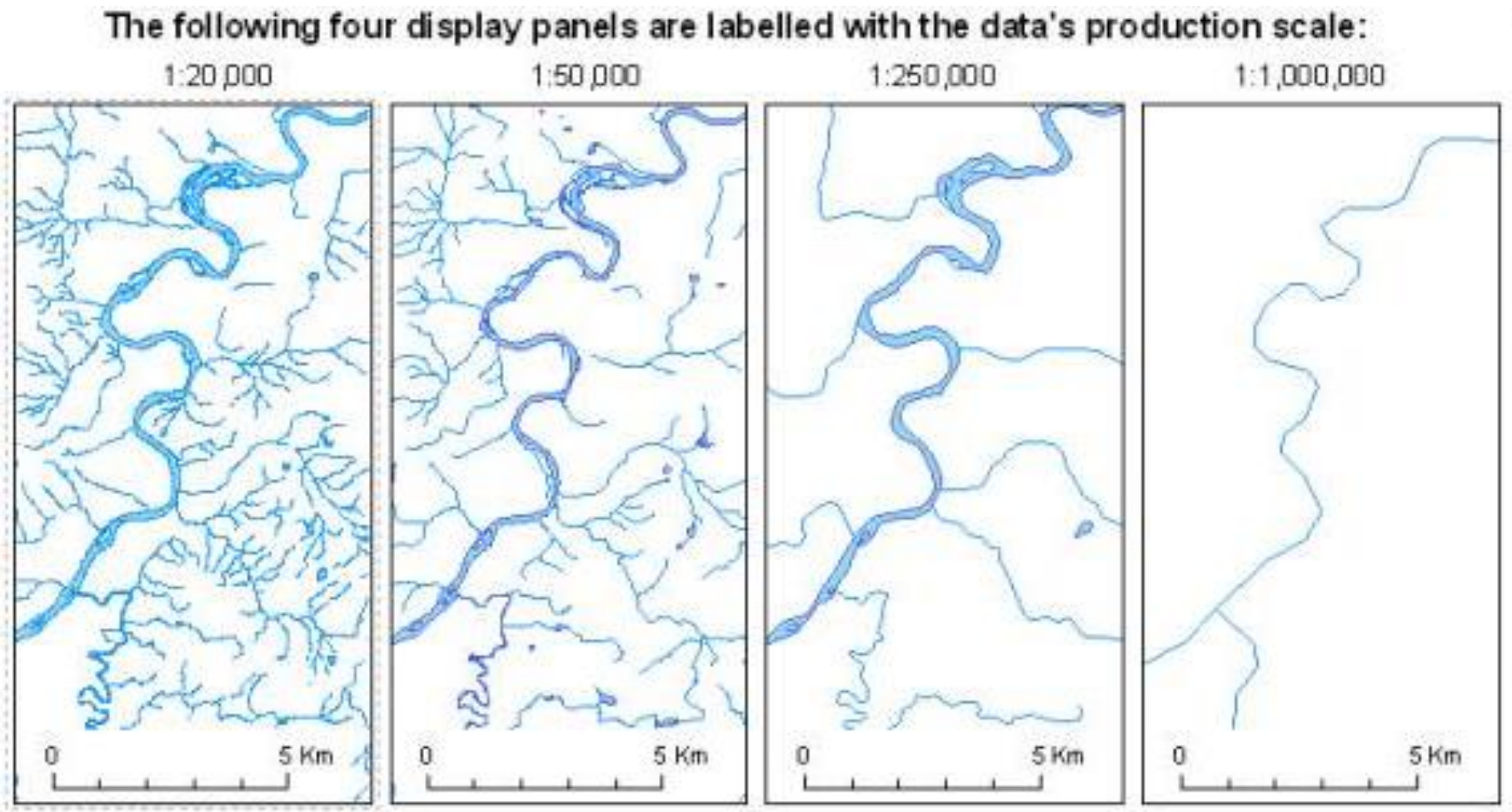


Figure 4. Selected from generalisations of the outline of Vienna.

Selection example: (elimination, smoothing) plus collapsing



Benoit Mandelbrot: "How Long Is the Coast of Britain? "

'Fractals' *Statistical Self-Similarity and Fractional Dimension (1967)*

Unit = 200 km, length = **2400 km**

Unit = 50 km, length = **3400 km**

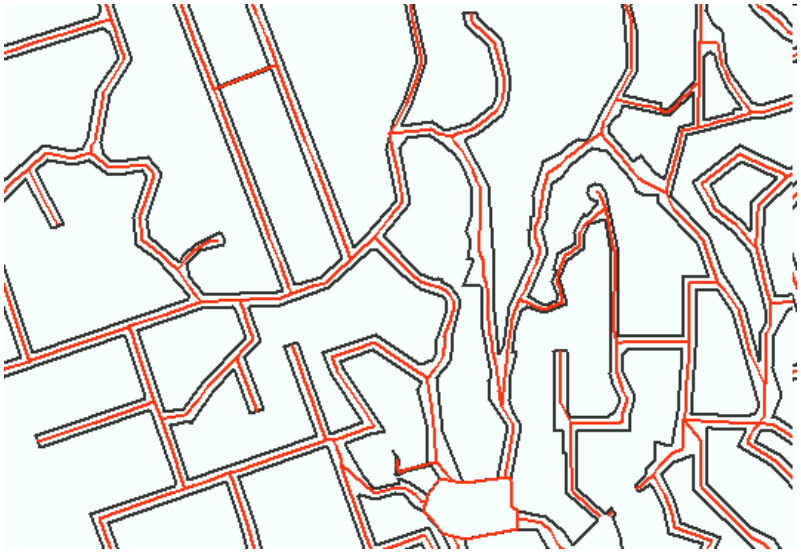
CIA: 12,429 km (1:100,000)

Ordnance Survey: 17,820 km (1:50,000)

Infinite? ->

Length of Coastline by Country

Canada – we're #1 (at any scale)
Length = 202,080 km or 243, 042 or 265,523 km or …..
Norway 58,113; Indonesia: 54,720; Russia: 37,653

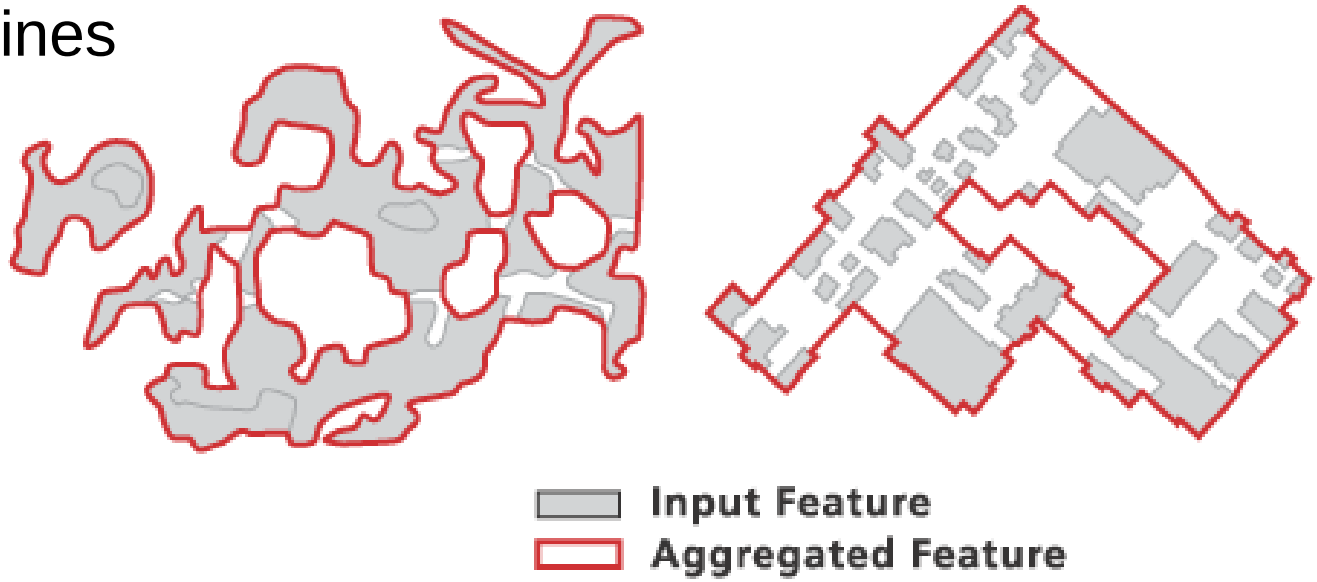


ArcGIS examples: Simplification

(for smaller scales)

Collapsing dual lines to a centre-line

Aggregating polygons



http://resources.esri.com/help/9.3/arcgisengine/java/gp_toolref/data_management_toolbox/an_overview_of_the_generalization_toolset.htm

ArcGIS generalization tools – 'GIS analysis' – assuming you need less detail ….

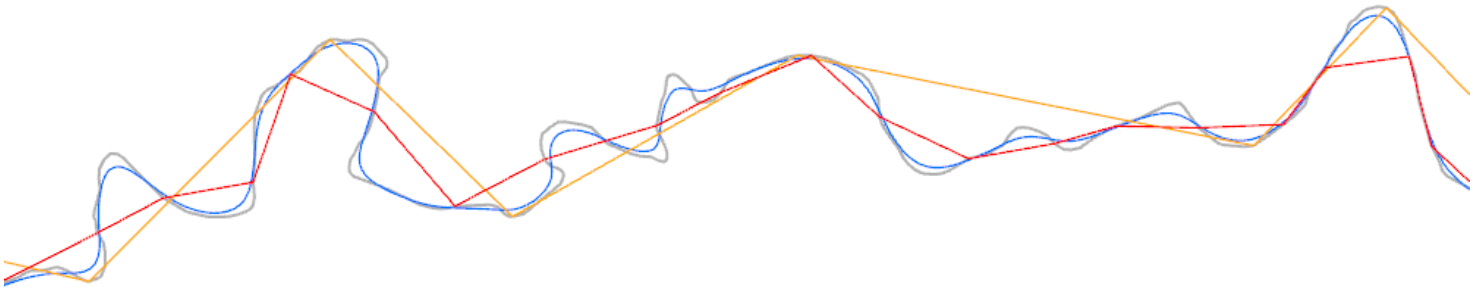
Aggregate Points	Creates polygon features around clusters of proximate point features.
Aggregate Polygons	Combines polygons within a specified distance of each other into new polygons.
Collapse Dual Lines To Centerline	Derives centerlines from dual-line (or double-line) features, such as road casings, based on specified width tolerances.
Collapse Road Detail	Collapses small, open configurations of road segments that interrupt the general trend of a road network, such as traffic circles, for example, and replaces them with a simplified depiction.
Delineate Built-Up Areas	Creates polygons to represent built-up areas by delineating densely clustered arrangements of buildings on small-scale maps.
Create Cartographic Partitions	Creates a mesh of polygon features that cover the input feature class where each polygon encloses no more than a specified number of input features, determined by the density and distribution of the input features.
Merge Divided Roads	Generates single-line road features in place of matched pairs of divided road lanes.
Simplify Building	Simplifies the boundary or footprint of building polygons while maintaining their essential shape and size.
Simplify Line	Simplifies lines by removing extraneous bends while preserving essential shape.
Simplify Polygon	Simplifies polygons by removing extraneous bends while preserving essential shape.
Smooth Line	Smooths sharp angles in lines to improve aesthetic or cartographic quality.
Smooth Polygon	Smooths sharp angles in polygon outlines to improve aesthetic or cartographic quality.
Thin Road Network	Generates a simplified road network that retains connectivity and general character for display at a smaller scale.

http://resources.arcgis.com/en/help/main/10.1/index.html#//00700000002r000000

Line generalization

Mapping -> as scale reduces

The aim of the algorithm is to produce a simplified polyline that has fewer points than the original but keeps the original's characteristics/shape - > ….. in cartooning as well as mapping



David H. Douglas and Thomas K. Peucker (Poiker) published a paper titled "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature" (Douglas & Peucker, 1973)

https://cartography-playground.gitlab.io/playgrounds/douglas-peucker-algorithm

Larger scale data (more detail) = costs more; Small scale = mostly free

3. Classification (by attributes)

We can't keep all features unique – e.g. every road or building might have special characteristics, but they have to be grouped. (so there is 'less detail in the data)

This is done in one of 3 ways:

Nominal: qualitative (type) - categorical

Ordinal: in sequence 'ranked' - hierarchical

Interval: quantitative (size) - numerical

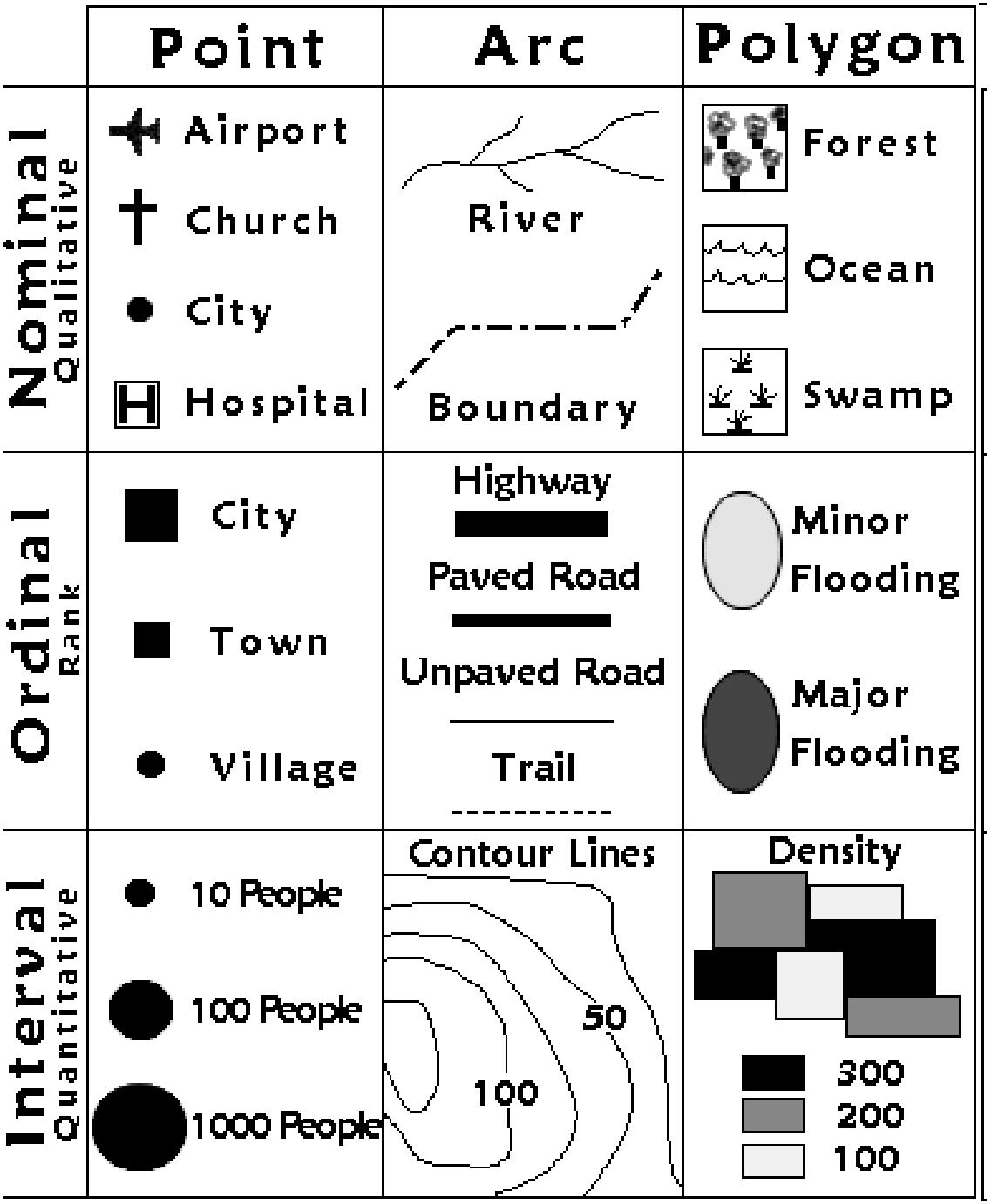
(Ratio): a sub-type of interval with an absolute 0 base

Classification

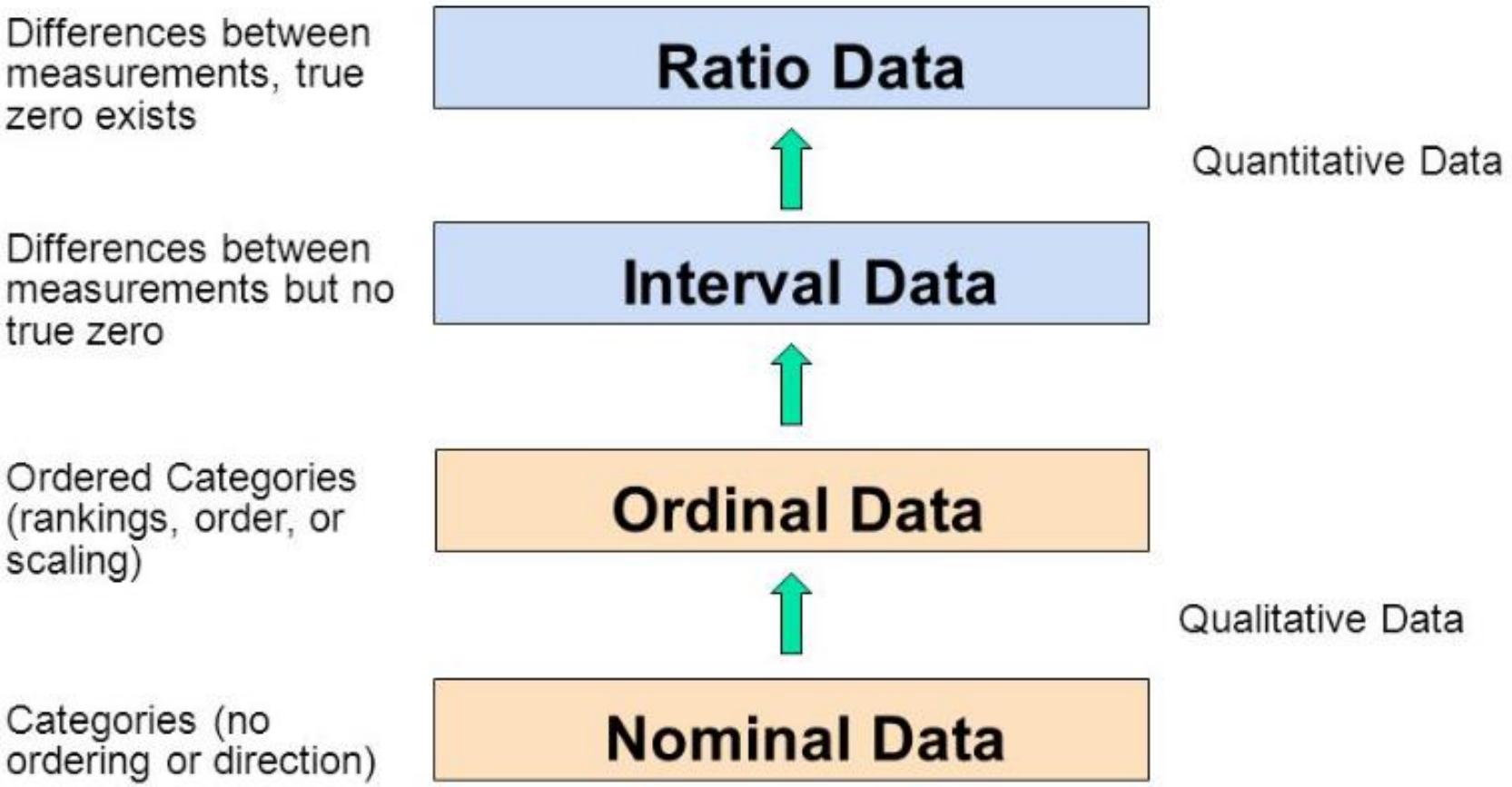
generalisation into groups or classes

Reminder:

*with **classification**, feature attributes are generalised; **simplification / selection** more applied to spatial elements*

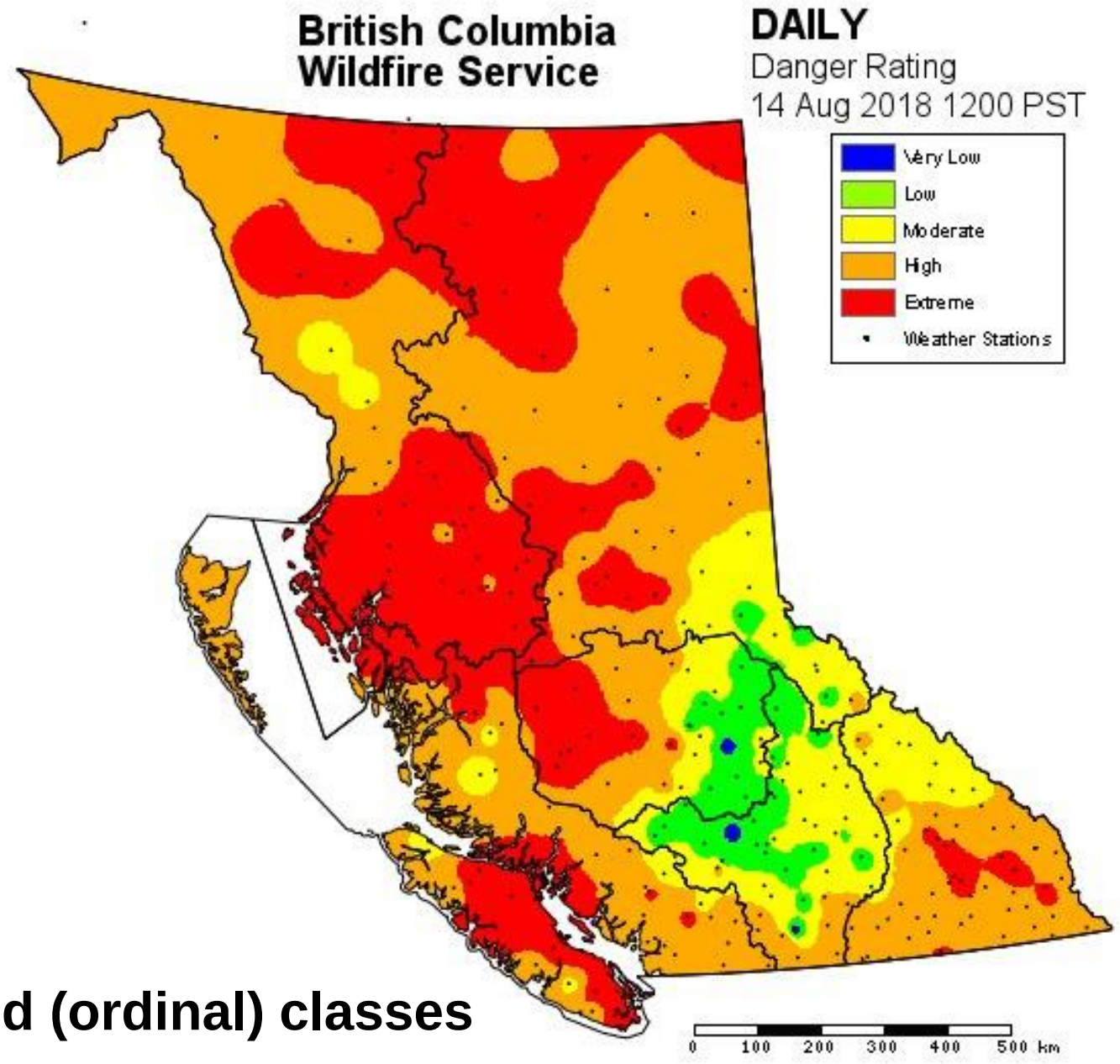


Grouping classes of data



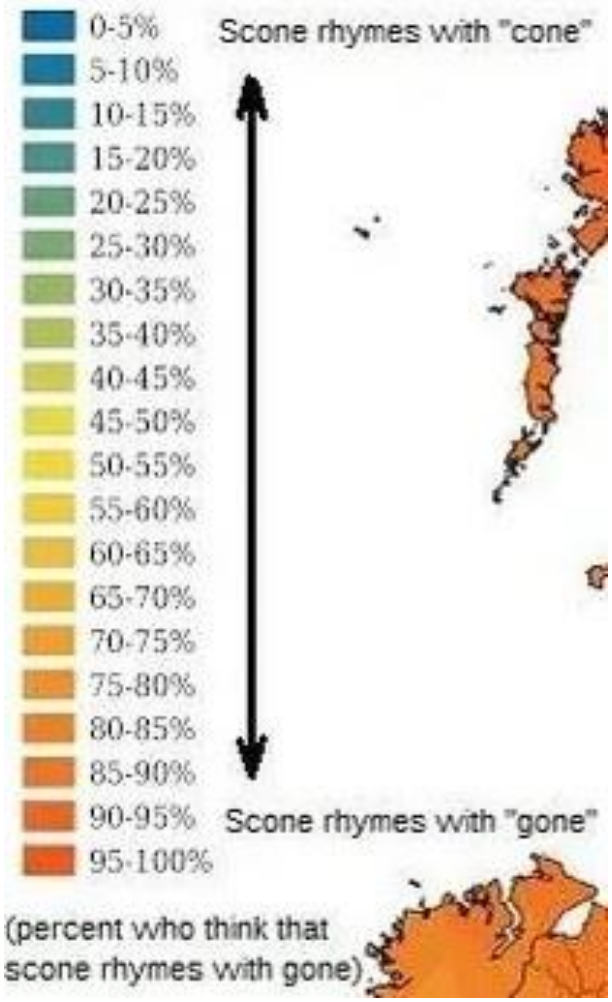
https://www.graphpad.com/support/faq/what-is-the-difference-between-ordinal-interval-and-ratio-variables-why-should-i-care

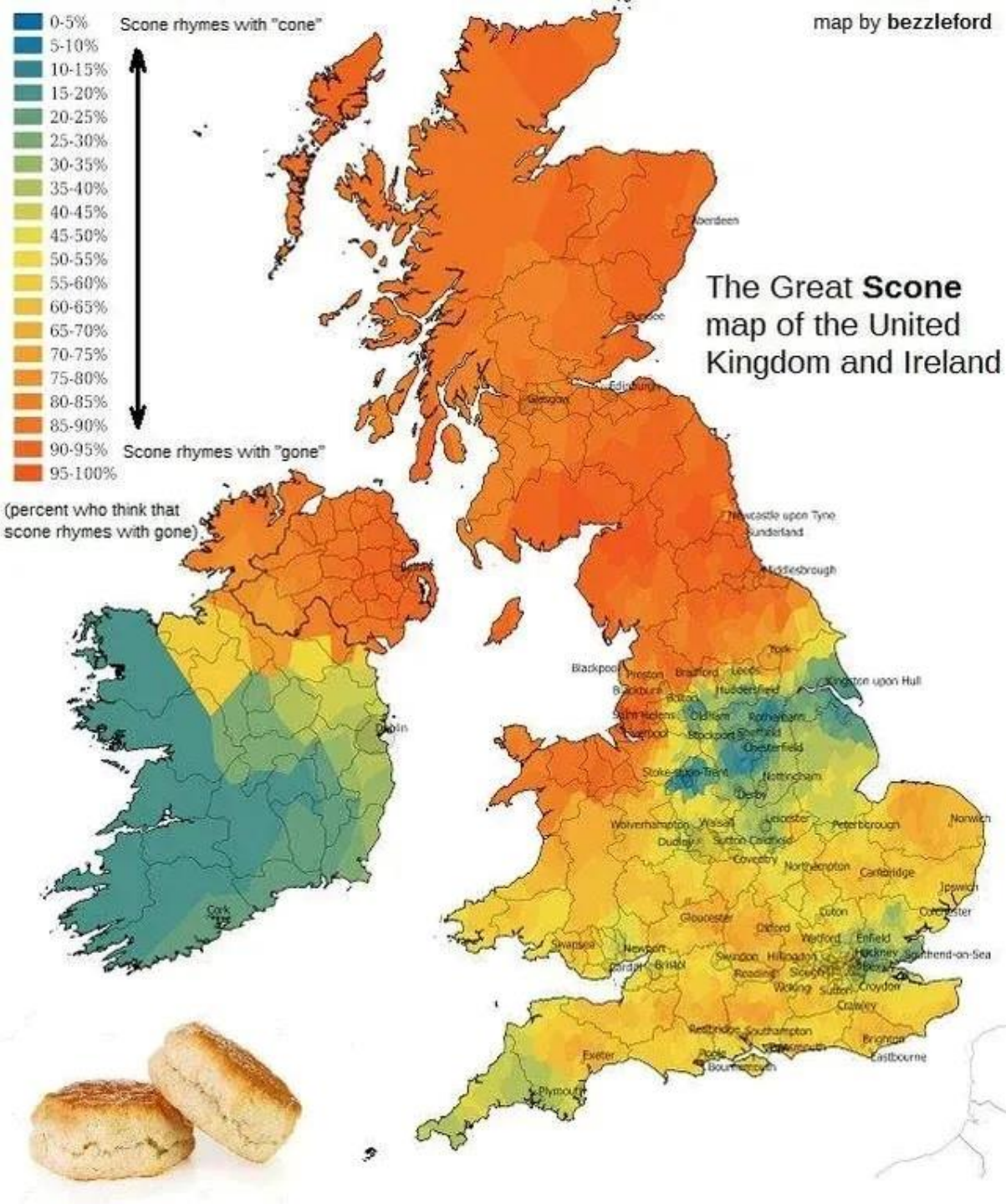
Categorical (nominal) classes



Ranked (ordinal) classes

Quantitative (interval) classes





Scale and Resolution

As scale decreases, both spatial and attribute details decrease

There is a practical level of detail associated with any scale

One could identify a visual 'minimal resolvable unit' (MRU) of ~ 0.5 mm.

This translates for common display scales in metres:

1:20,000	10m
1:50,000	25
1:250,000	125
1:1,000,000	500

This is also synonymous with **resolution** or **precision**

It is fixed with printed maps, variable with digital displays e.g. Google Maps

Online viewer: New Zealand http://www.topomap.co.nz

Summary of scale and generalization

Scale is the prime control on generalization

Data captured at one scale are <u>not</u> transferable to all scales:

- **data from a larger scale are too detailed for smaller scales,**
- **data from a small scale are too generalised for larger scales.**

Other Controls on Generalization (apart from scale)

Map / output purpose: how much detail, and what types of features

Space: is there room for these elements (+ lettering)

Data quality: survey methods may limit the details possible

Perceptual limits: the human eye / brain likes <= 7 classes

http://en.wikipedia.org/wiki/The_Magical_Number_Seven,_Plus_or_Minus_Two

Too much ? … there would be fewer place names in the west

Data quality: Spatial generalization, accuracy and precision: GPS / GIS

Trans Allegheny Trails

GPS Coordinates for Our Trailheads

6 to 10 Trail

2 trailheads

Allegheny Portage Railroad NHS: 40.458654,-78.54909

Foot of Ten Trailhead: 40.408106,-78.460739

6 decimal places

1° =111km; .1 = 11km, .01 = 1km .001 = 100m,
.0001 = 10m, .00001 =1m, .000001=10cm

1’ = ~ 2km; 1” = ~30m

Data quality: Spatial generalization, accuracy and precision: GPS and GIS

Trailhead Johnston Canyon Resort, 17.5 km (11 mi) northwest along the Bow Valley Parkway from the Trans-Canada Hwy.

GPS Coordinates

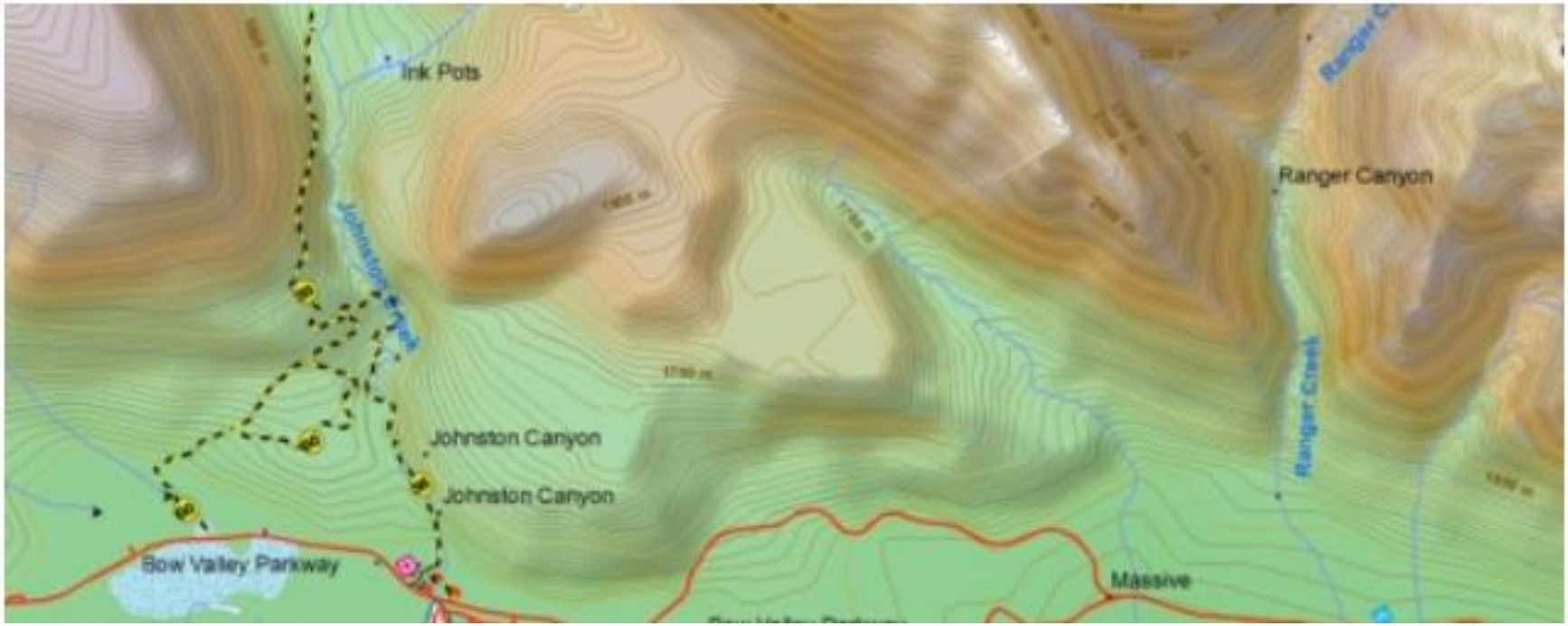
Latitude: 51.24542307241623

Longitude: -115.83992958068848

14 decimal places !
too precise; NOT accurate
not enough **generalization**
Last digit = millionths of a mm
(human hair = .1mm)

Detailed Map

Don't just repeat decimal places from GPS / GIS software !
General rule: we mostly don't need sub-metre precision



Summary review questions

- Why is generalisation necessary ?
- How is it related to map scale ?
- What happens to points, lines, areas as scale decreases ?
- What is meant by aggregation, elimination and smoothing ?
- What are the three main types of data classification ?

Reporting map locations: no need for sub-metre precision = it's wrong !

UTM: no need for (sub-metre) decimal values for x,y (eastings, northings)

DD/DMS/DDM: (1" =~ 30m), decimals - only need 5 (for °) / 3 (for ") / 1 (for ')