Lesson 8: Visualising and Analysing Geographic Data

Dr. Kam Tin Seong
Assoc. Professor of Information Systems (Practice)
School of Computing and Information Systems,
Singapore Management University

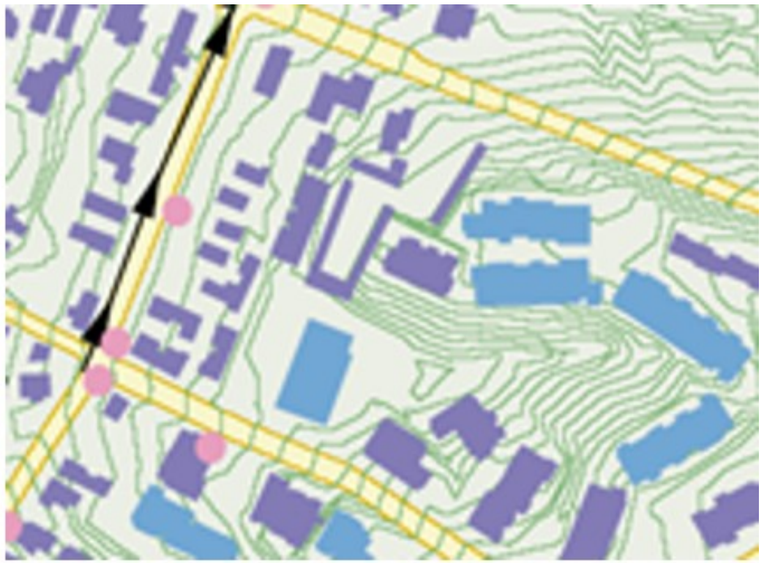
09 Jun 2023

What will you learn from this lesson?

- Introducing map
- Properties of geographical data
  - Geographical data models
  - Map projections
- Typology of maps
  - Reference maps
  - Thematic maps
- Proportional Symbol Map
- Choropleth Map]

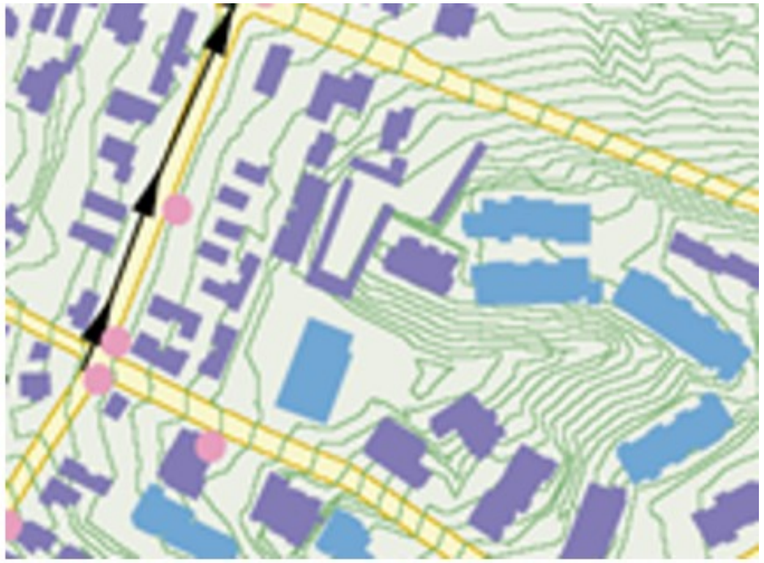
What is a Map?

A map is a model of real world depict by a collection of cartographic symbols or/and visual abstraction.

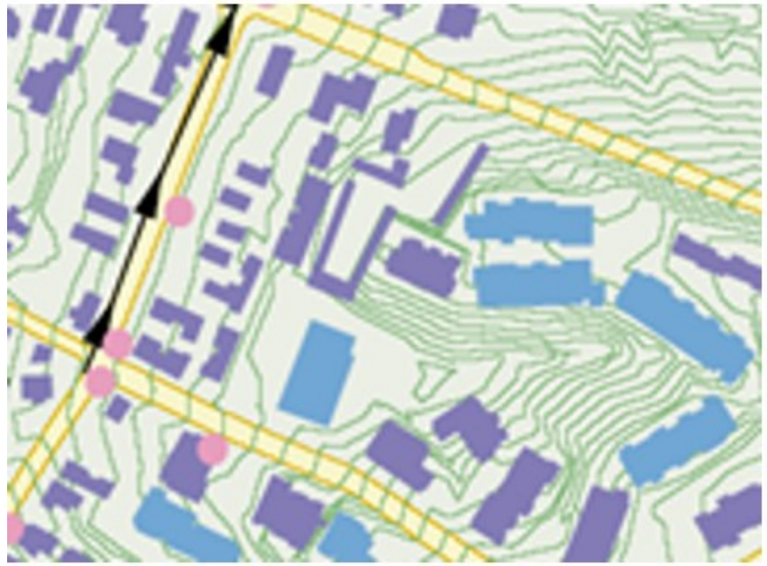


What is a Map?

An interface between geographical data and map users.



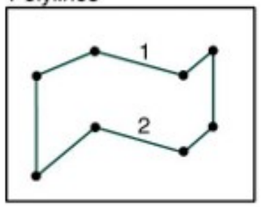
A Map and Geographical Data



Points

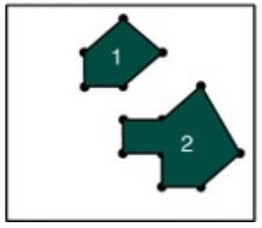
Point number	(x,y) coordinates
1	(2,4)
2	(3,2)
3	(5,3)
4	(6,2)

Polylines



Polyline number	(x,y) coordinates
1	(1,5) (3,6) (6,5) (7,6)
2	(1,1) (3,3) (6,2) (7,3)

Polygons



Polygon number	(x,y) coordinates
1	(2,4) (2,5) (3,6) (4,5) (3,4) (2,4)
2	(3,2) (3,3) (4,3) (5,4) (6,2) (5,1) (4,1) (4,2) (3,2)

Types of geographical data

XCOORD	YCOORD	NAME
103.8589	1.2840	Livewire (Marina Bay Sands)
103.8574	1.2947	Singapore Pools Suntec City
103.8469	1.2835	Kis Store
103.8493	1.2829	Singapore Turf Club China Square
103.8521	1.2836	Tay Sui Lan Agency
103.8492	1.2861	7-Eleven Circular Rd
103.8440	1.2858	Singapore Pools People's Park Centre
103.8464	1.2850	Feng Yuen Agency Enterprise

Geospatial data

SUBZONE_C	TOTAL_POP	YOUNG
AMSZ01	4990	850
AMSZ02	30390	3950
AMSZ03	28650	3410
AMSZ04	24160	3260
AMSZ05	19190	2620
AMSZ06	24550	3050
AMSZ07	6860	1030
AMSZ08	8370	1200

Aspatial data

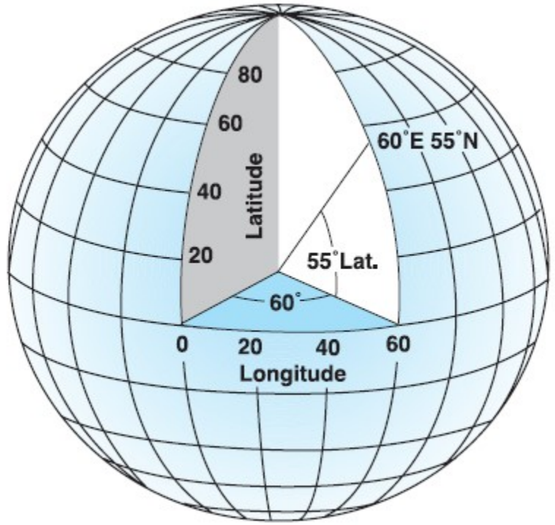
How location information are registered?

- Geographic Coordinates Systems
- Projected Coordinates Systems

How location information are registered?

Geographic Coordinates Systems

- A geographic coordinate system defines two-dimensional coordinates based on the Earth’s surface. It has an angular unit of measure, prime meridian and datum (which contains the spheroid).
- For example, WGS84, NAD27 and NAD83.

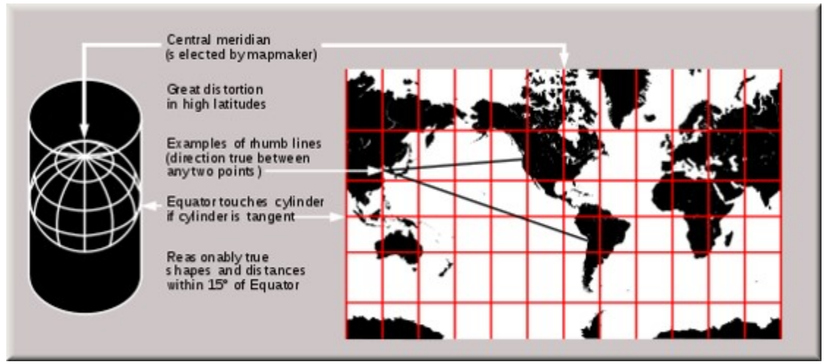


Source: Geographic Coodinate Ststem@wiki

How location information are registered?

Projected Coordinates Systems

- A projected coordinate system is defined on a flat, two-dimensional surface.
- A projected coordinate system, unlike a geographic one, has the advantage that lengths, angles, and areas are constant across the two dimensions. This is not true when working in a geographic coordinate system.
- A projected coordinate system is always based on a geographic coordinate system that can use a sphere or spheroid.



Source: Map Projection@wiki

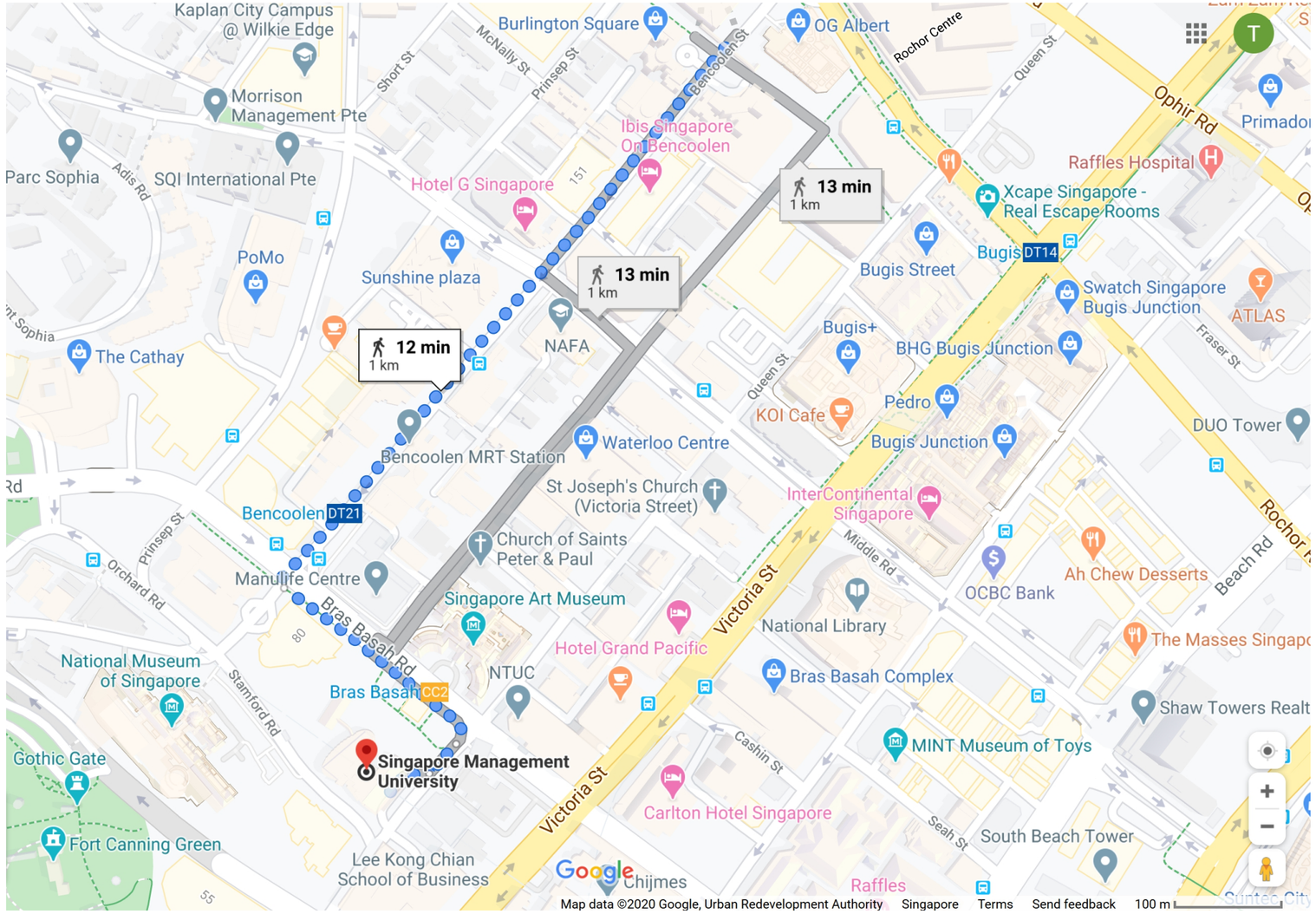
The Purpose of a Map

- To serve as a piece of historical archive.

Source: Battle of Singapore

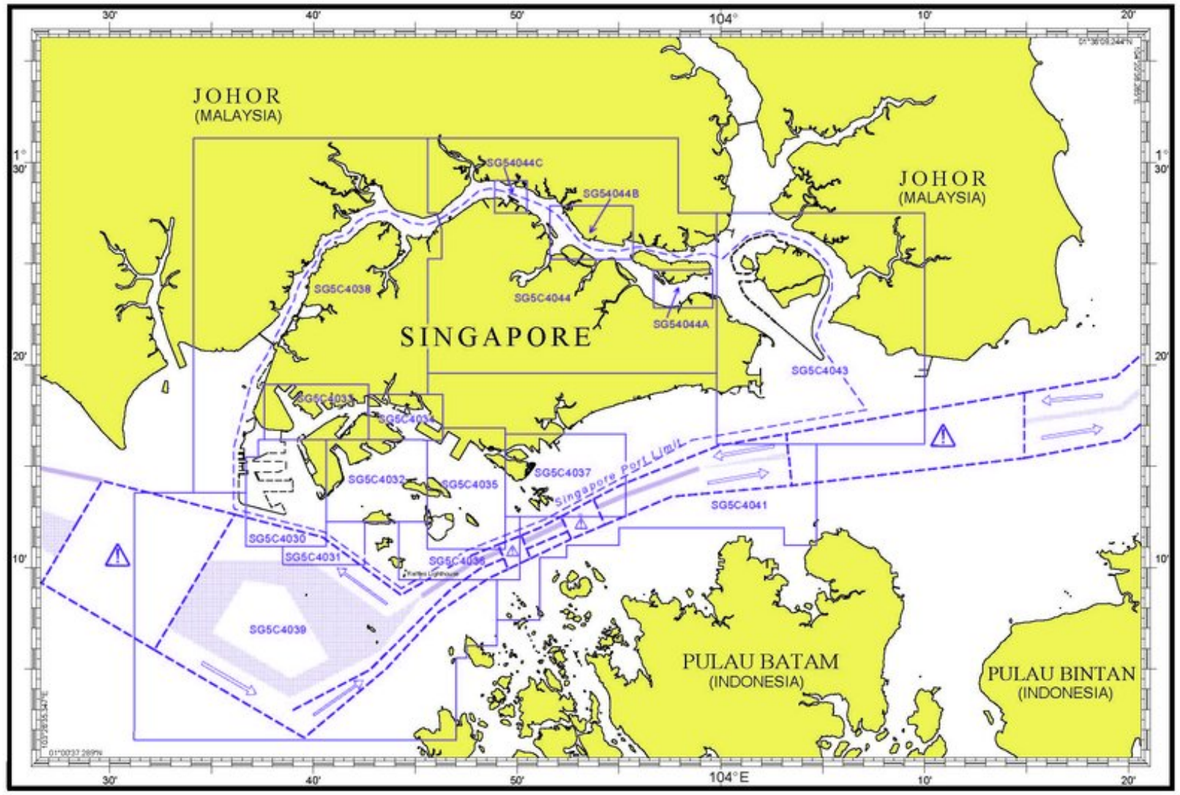
The Purpose of a Map

- To support navigation (i.e. Google Map).



The Purpose of a Map

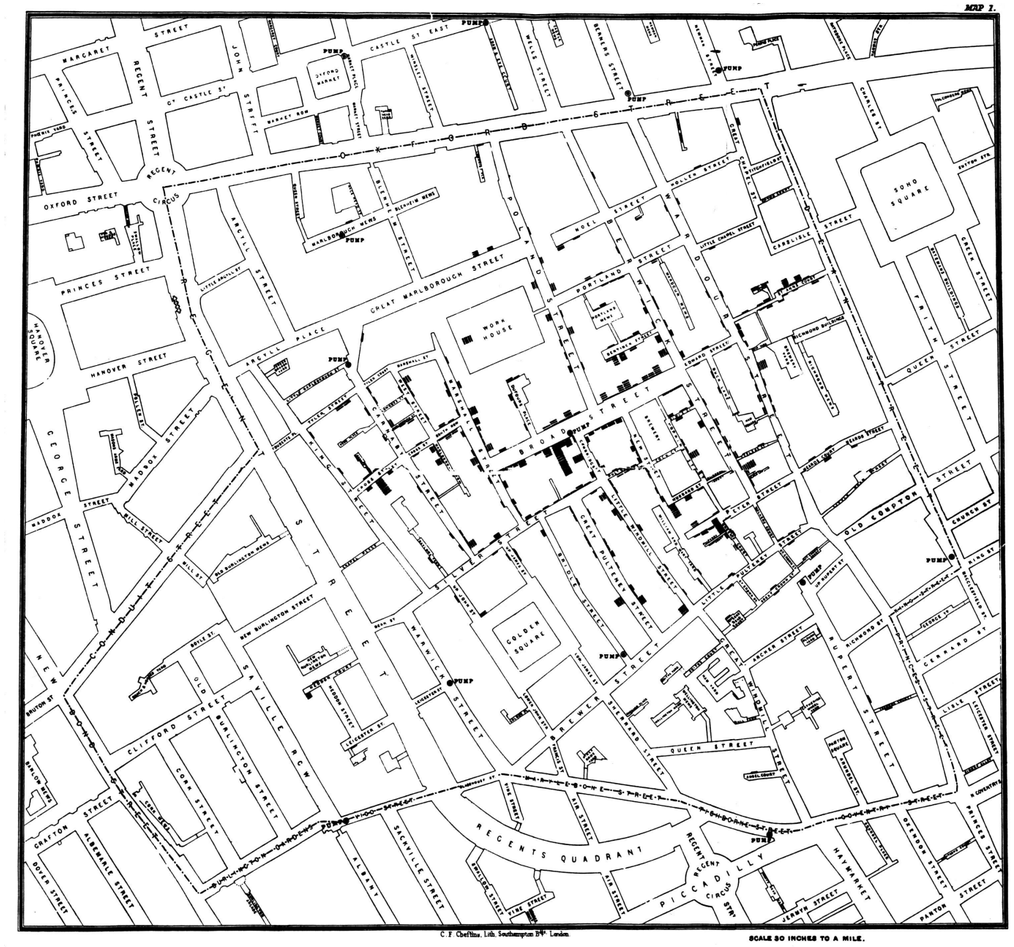
- To support navigation (i.e. sea chart of Singapore)



Source: Singapore electronic navigational chart

The Purpose of a Map

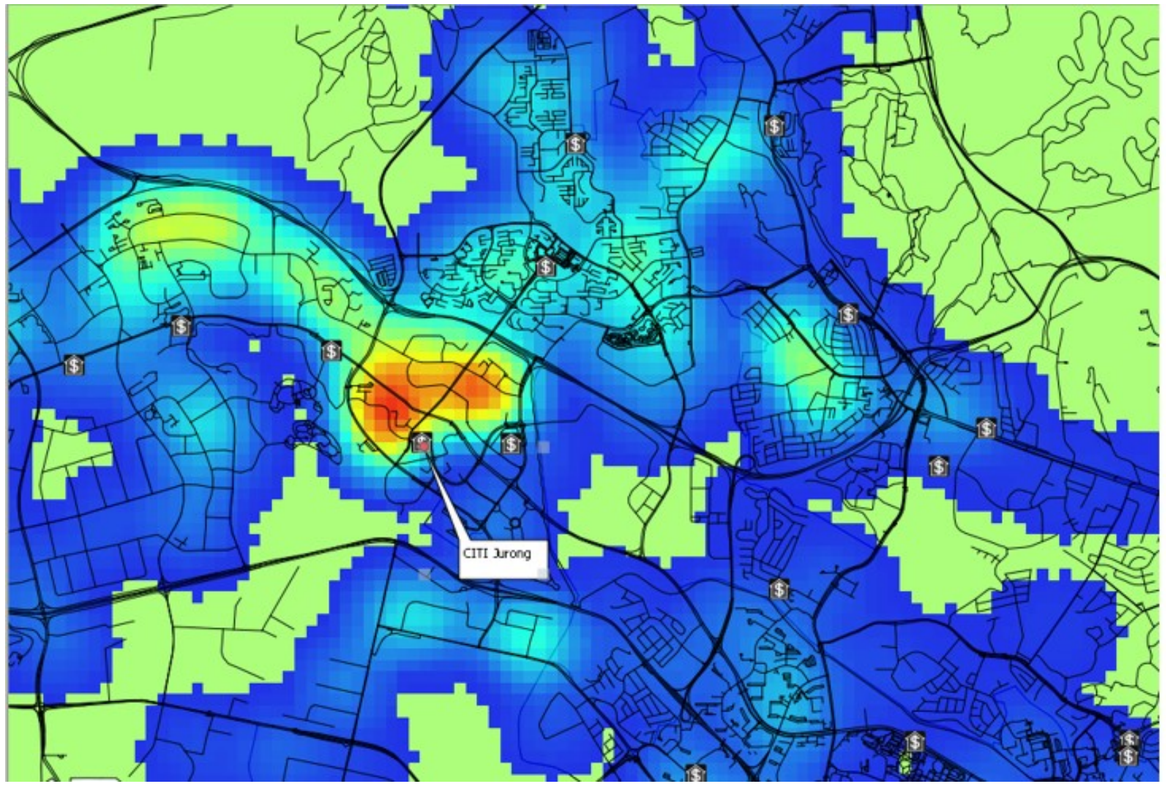
- To reveal spatial patterns.

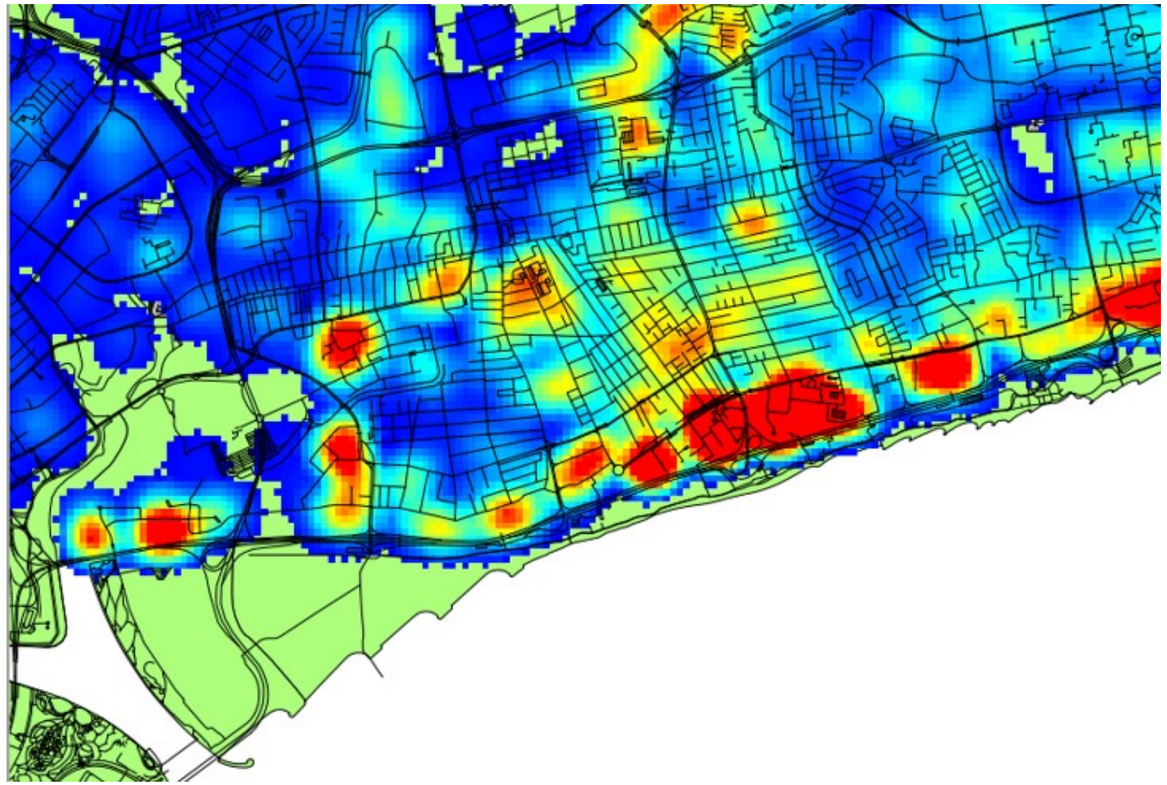


Reference: John Snow Cholera Map

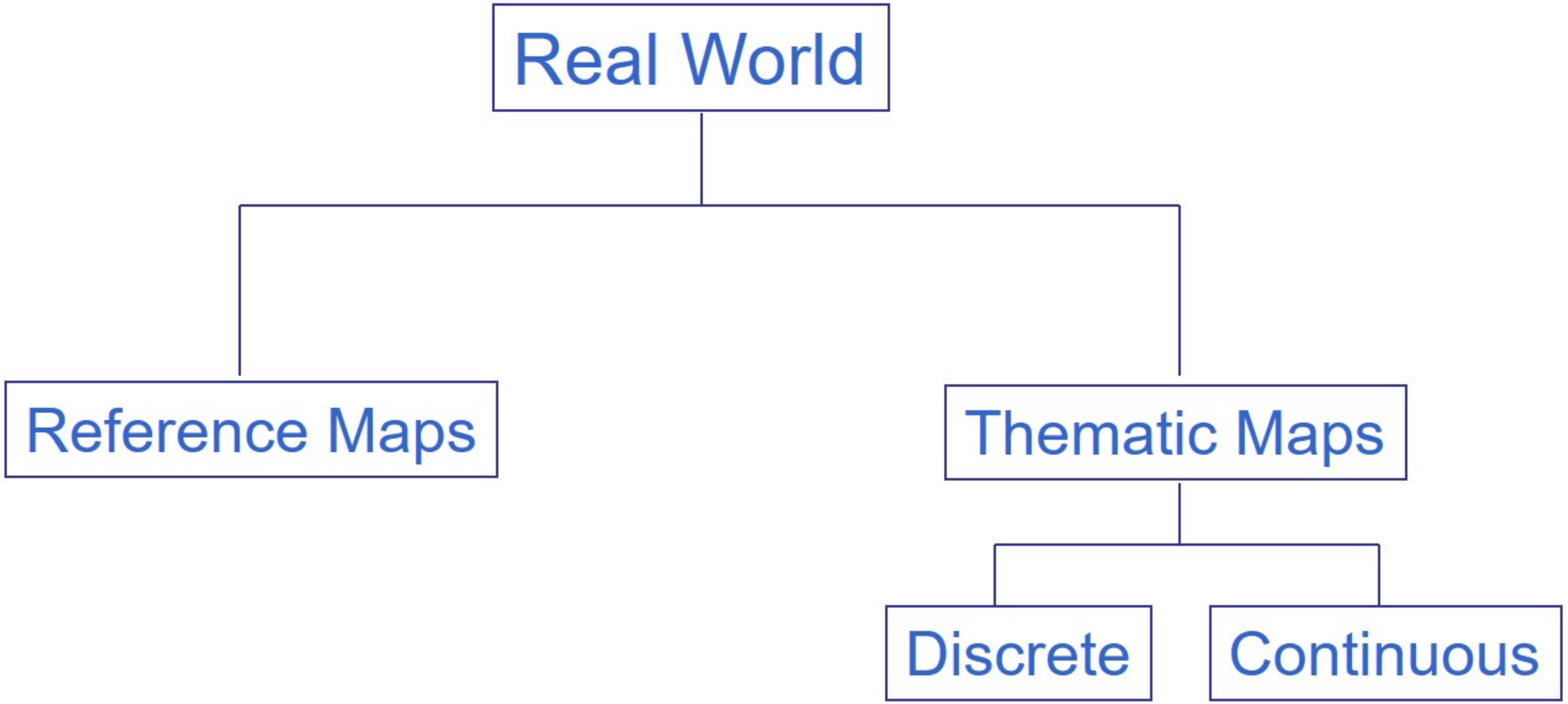
The Purpose of a Map

- To describe spatial relationship.



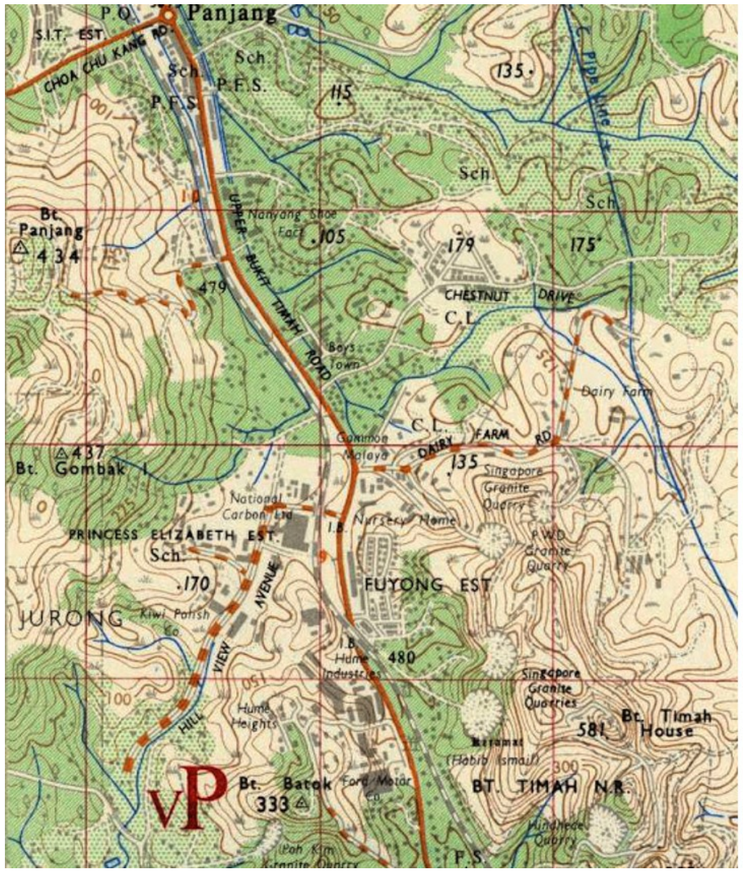


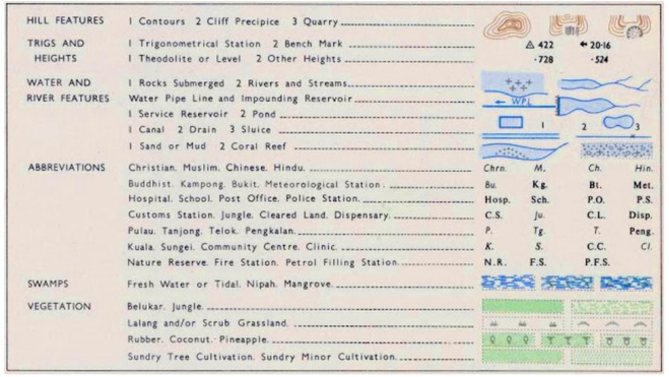
Typology of Maps



Reference maps

- Reference maps emphasize the location of spatial phenomena, such buildings, roads, vegetations, rivers, etc. Some of popular reference maps are **topographical map** and internet map such as Google map.

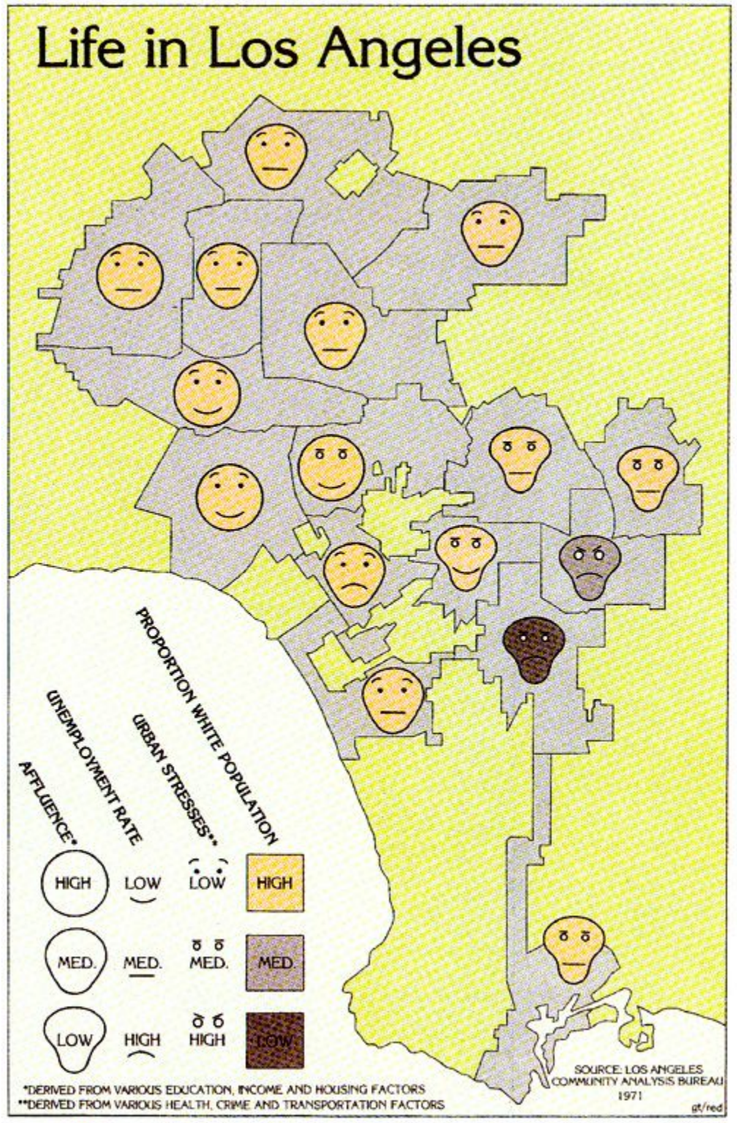




Thematic Map

- Thematic maps emphasize the spatial pattern of geographic attributes or statistics about places and relationships between places such as Life in Los Angeles.

Source: Life in Los Angeles



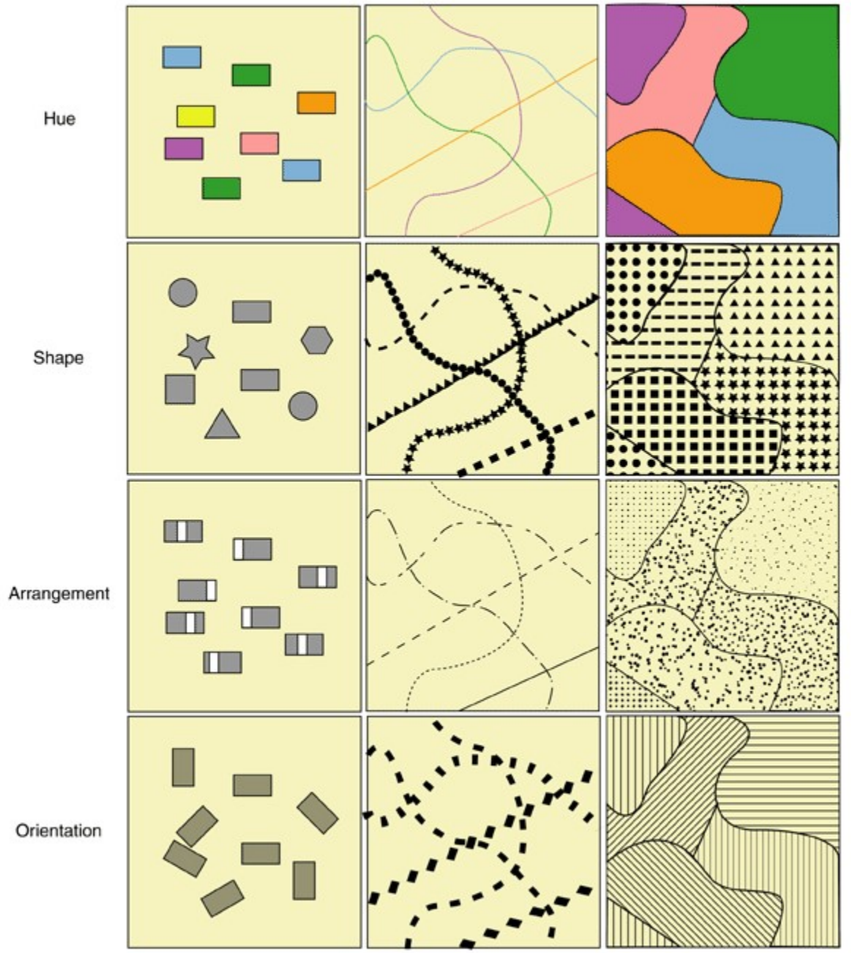
Thematic Mapping: Principles and Methods

- Displaying
  - Qualitative data
  - Quantitative data
- Choosing -Appropriate classification method for displaying data
  - Appropriate number of classes
- Techniques in data analysis
  - Using the classification histogram
  - Normalizing data

Qualitative Thematic Maps

Visual Variables and Cartographic Symbols

- Qualitative visual variables are used for nominal scale data.
- The goal of qualitative visual variables is to show how entities differ from each other.
- The visual variables that do a good job of showing ordinal differences are: colour value, colour saturation, size and texture/grain.
- Figure on the right for examples of these four ordinal visual variables used each in point, linear and areal symbols.



Qualitative Thematic Map

Point symbol map

- Different point symbols are used to represent school types.

Qualitative Thematic Map

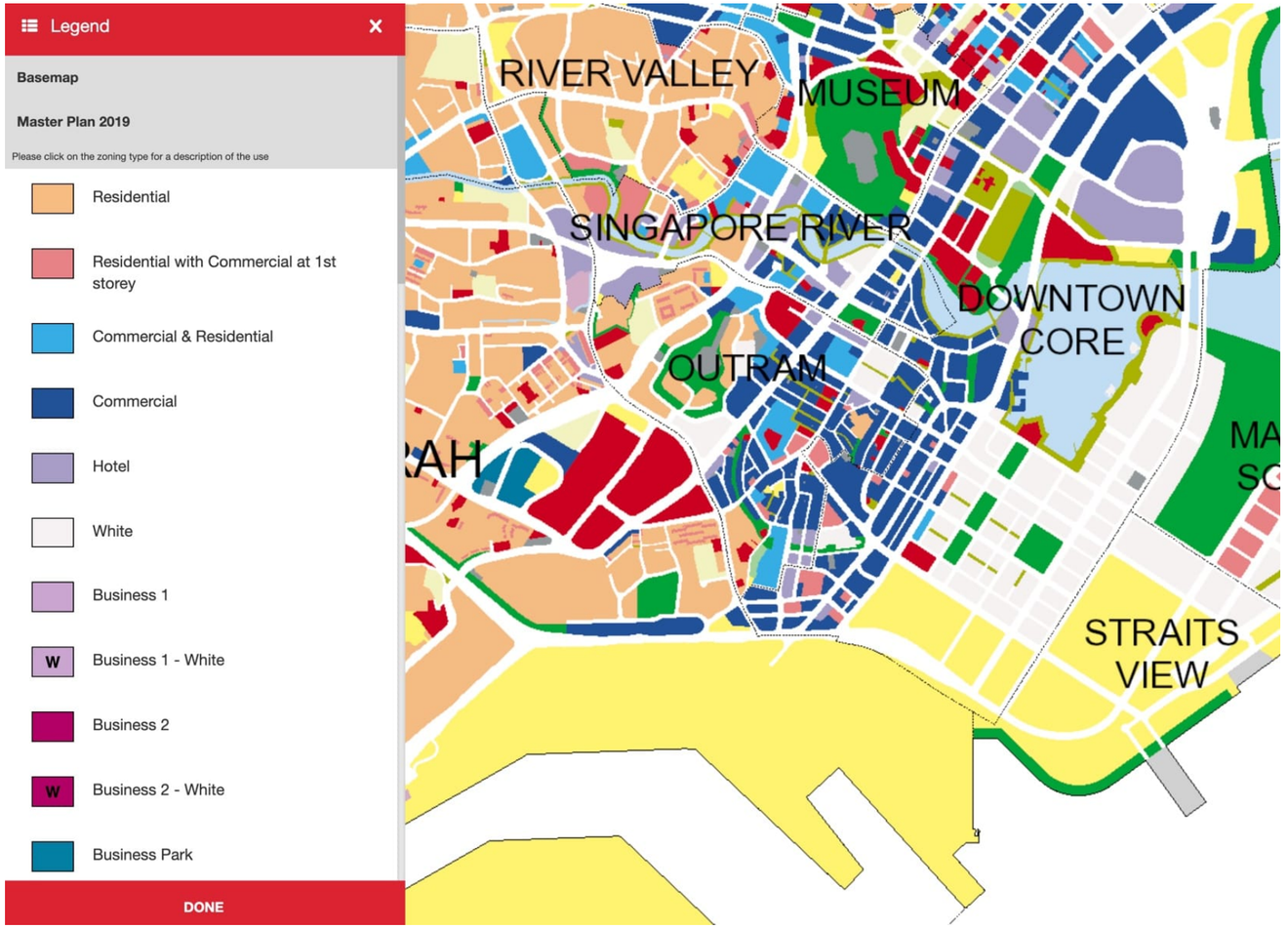
Line symbol map

- A road map is an example of a thematic map. It shows the road network of an area. In this map, lines with different colour intensity and tickness are used to differentiate hierarchy of roads.

Qualitative Thematic Map

Area map

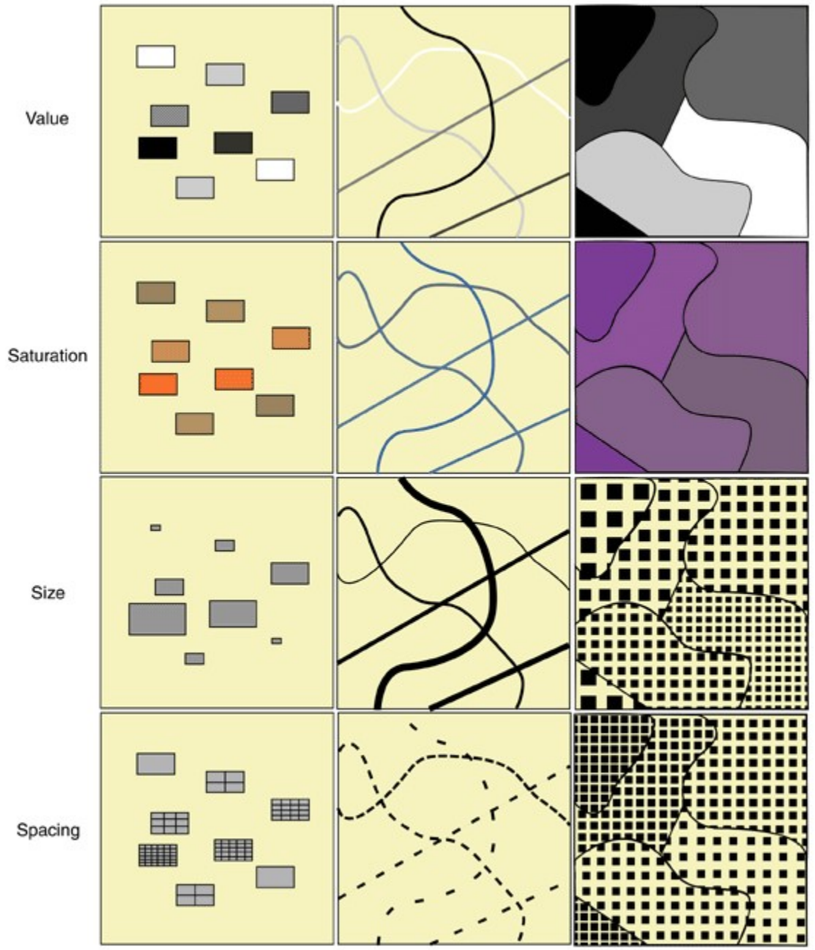
- Land use map below is a good example of a discrete thematic map. In this map, different colours are use to represent different land use types.



Quantitative Thematic Map

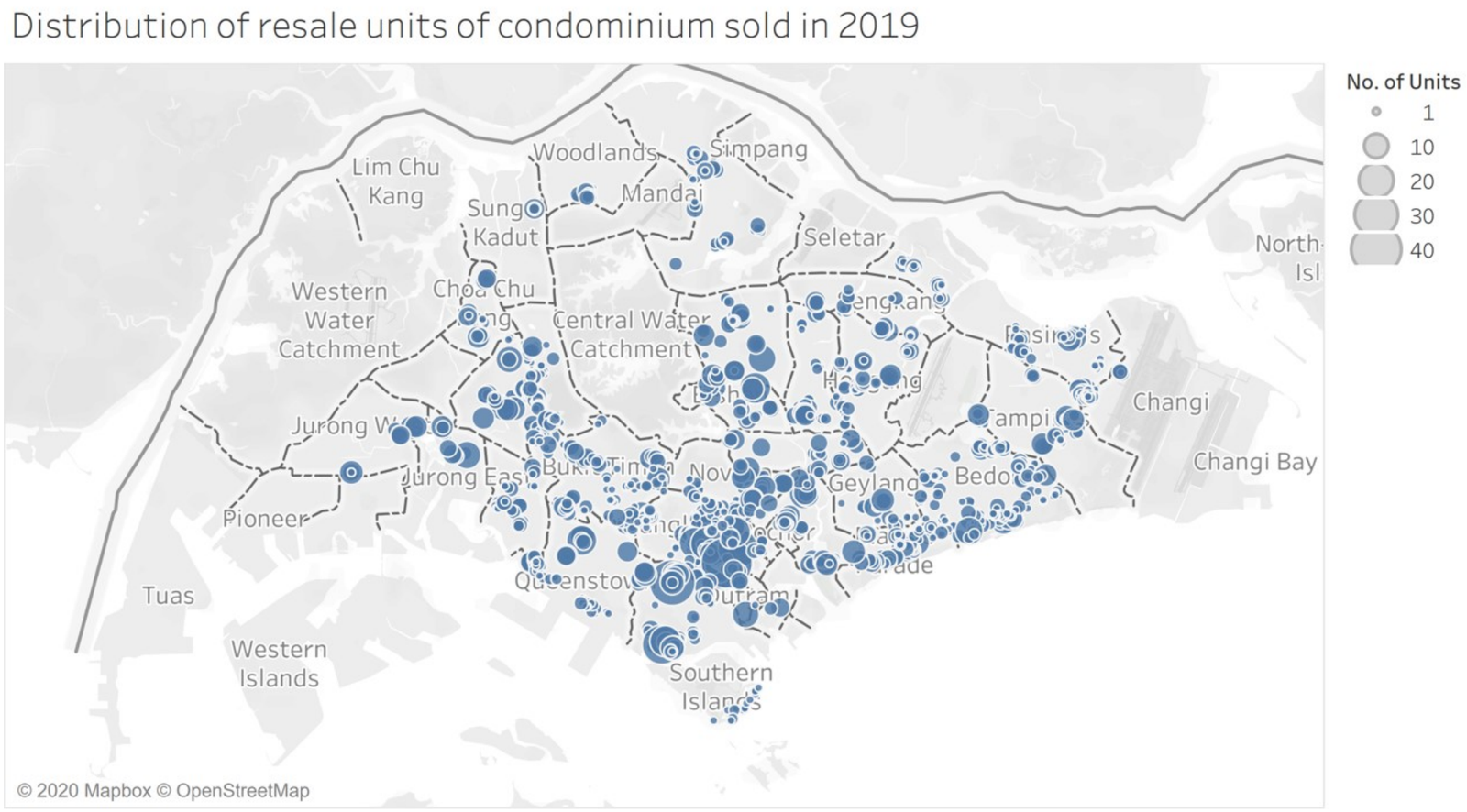
Visual Variables and Cartographic Symbols

- Quantitative visual variables are used to display ordinal, interval or ratio scale data.
- The goal of the quantitative visual variable is to show relative magnitude or order between entities.
- The visual variables that do a good job of showing ordinal differences are: colour value, colour saturation, size and texture/grain.
- Figure on the right shows of these four ordinal visual variables used each in point, linear and areal symbols.



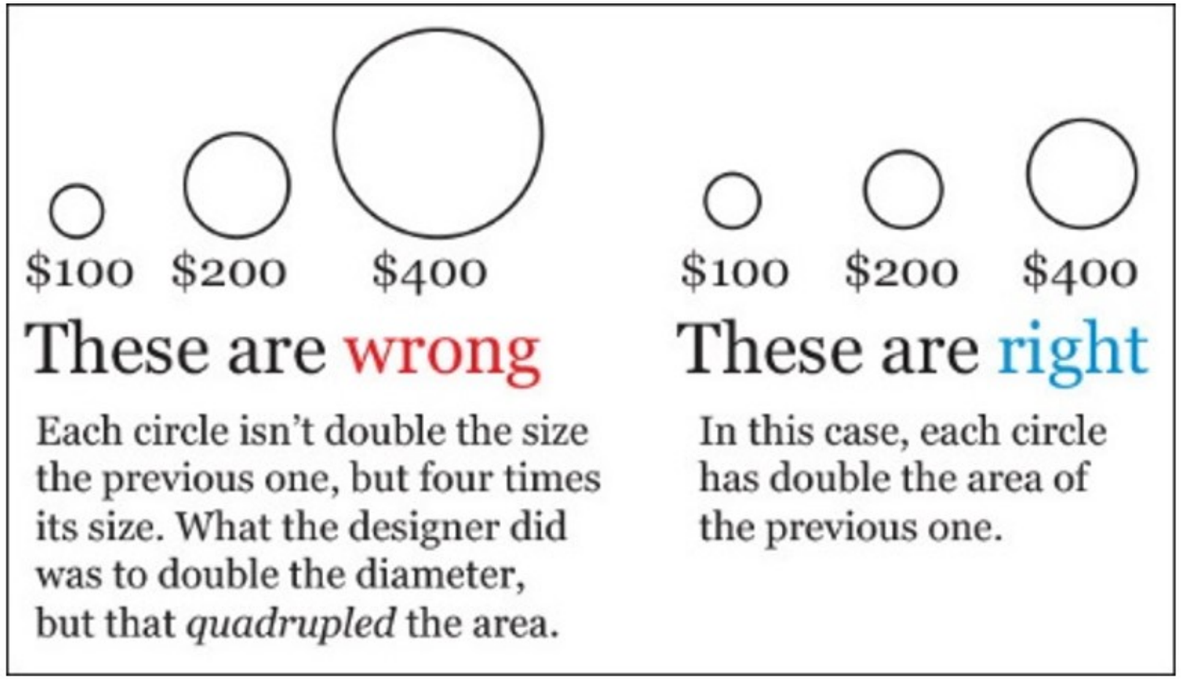
Proportional Symbol Map

- The proportional symbol technique uses symbols of different sizes to represent data associated with different areas or locations within the map.



How to size circles incorrectly and correctly?

Figure below explains how attribute values are map on circles. One strategy to overcome this problem is to use square instead of circle.

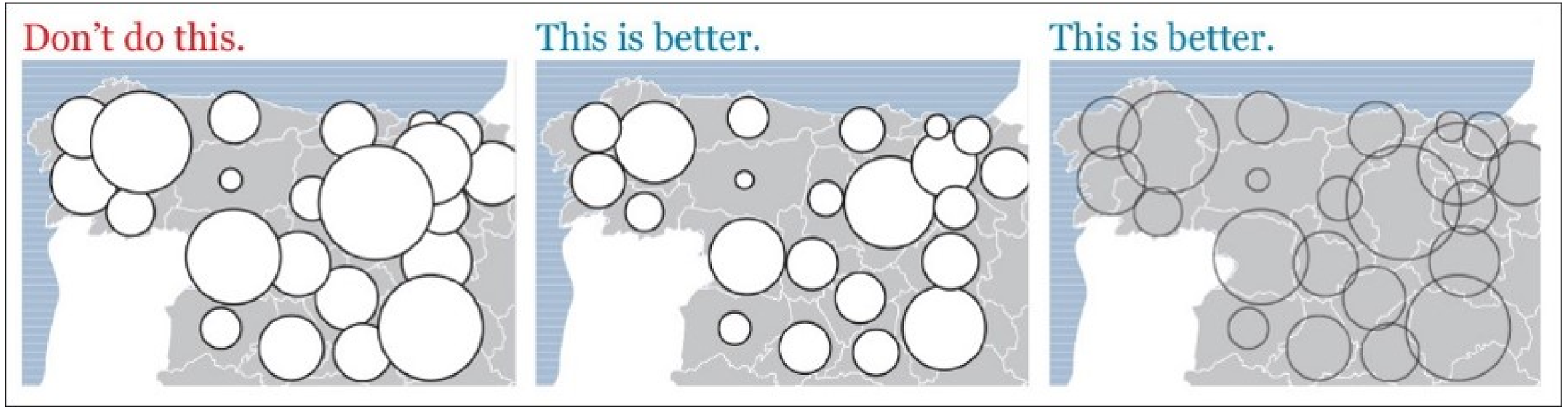


Imagine that the largest circle on a map of monthly household income equals $2,600.
The radius of this circle (R1) is 1.1 inches. How to calculate the radius (R2) of a circle representing $1,100?

$$R2 = \sqrt{\frac{\text{New value (1,100)}}{\text{Largest value (2,600)}}} \times R1 \rightarrow R2 = \sqrt{0.42} \times 1.1 = 0.71 \text{ inches}$$

How to make proportional symbol maps clearer?

When designing a proportional symbol map, one of the important consideration is to avoid the circles blocking each other. One possible strategy to overcome this problem is to applying opacity (also known as transparency).

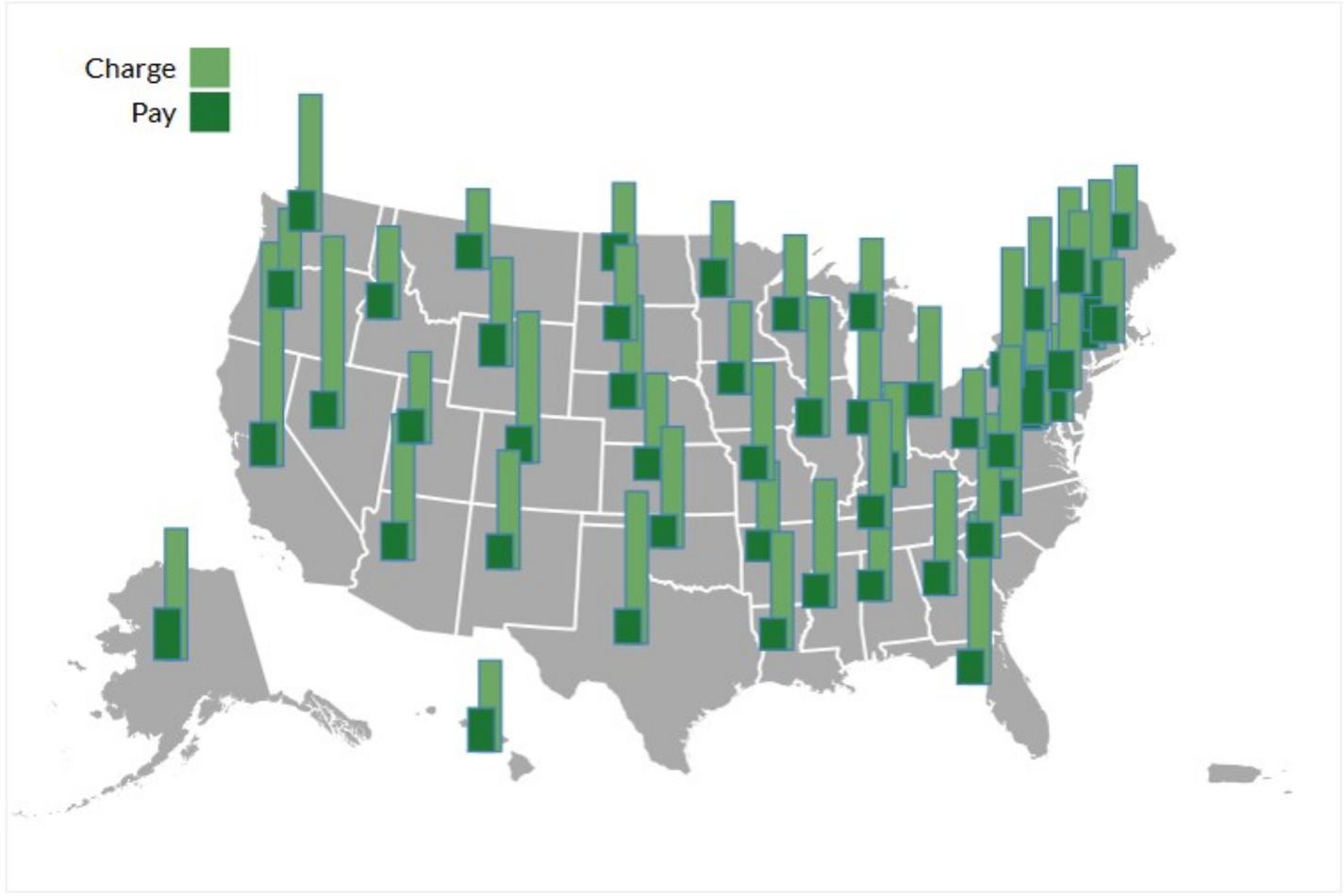


Proportional Symbol Map: Pie Chart Map

One important characteristic of proportional symbols is that they can easily be designed to represent more than one data value per location. Among the most common example is a **pie chart map** in which a circle is scaled proportionally to some total, and the size of wedges within the circle is scaled to depict a proportion of a total for two or more sub-categories.

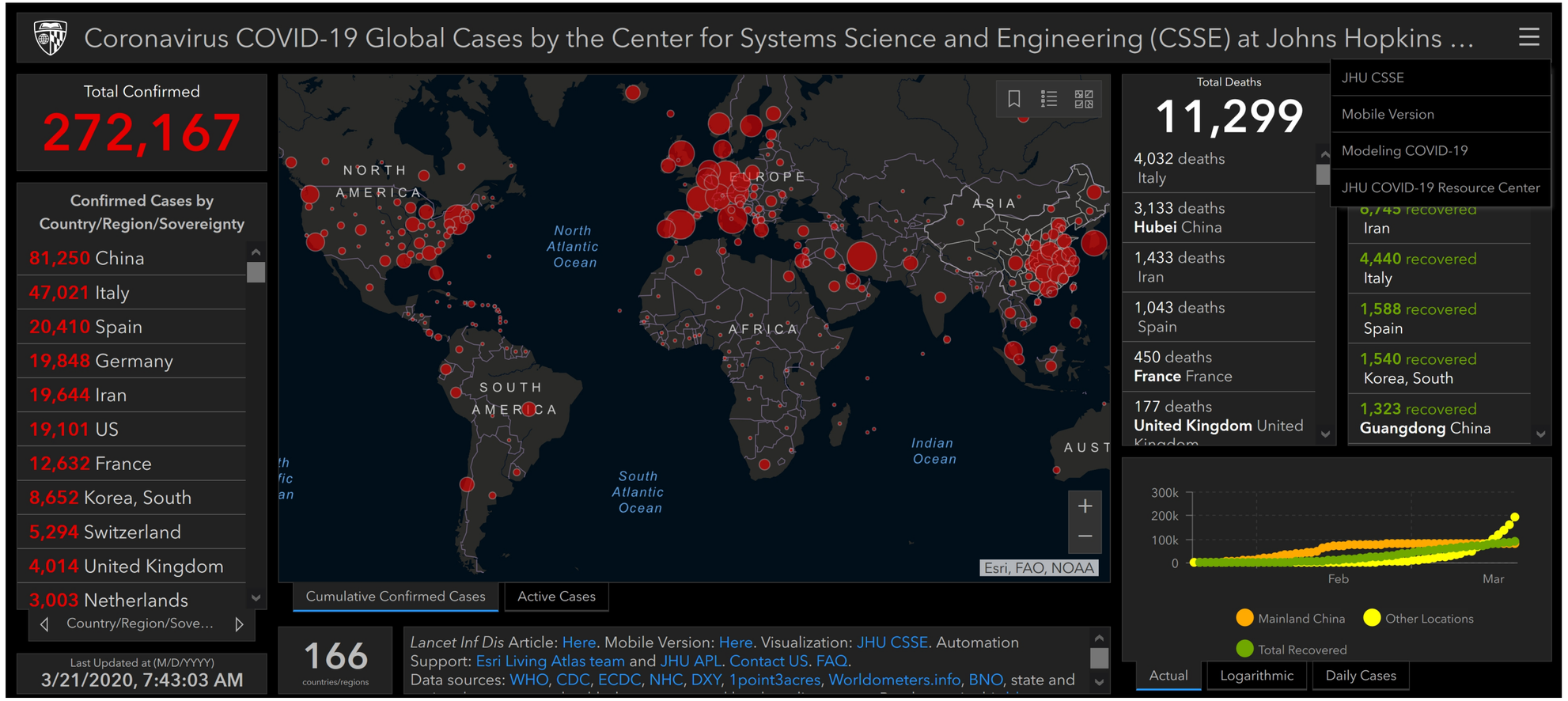
Proportional Symbol Map: Bar Chart Map

Proportional side-by-side bar map places side-by-side bar chart on a map based on the location of the observations.



Proportional Symbol Map - Junk Map

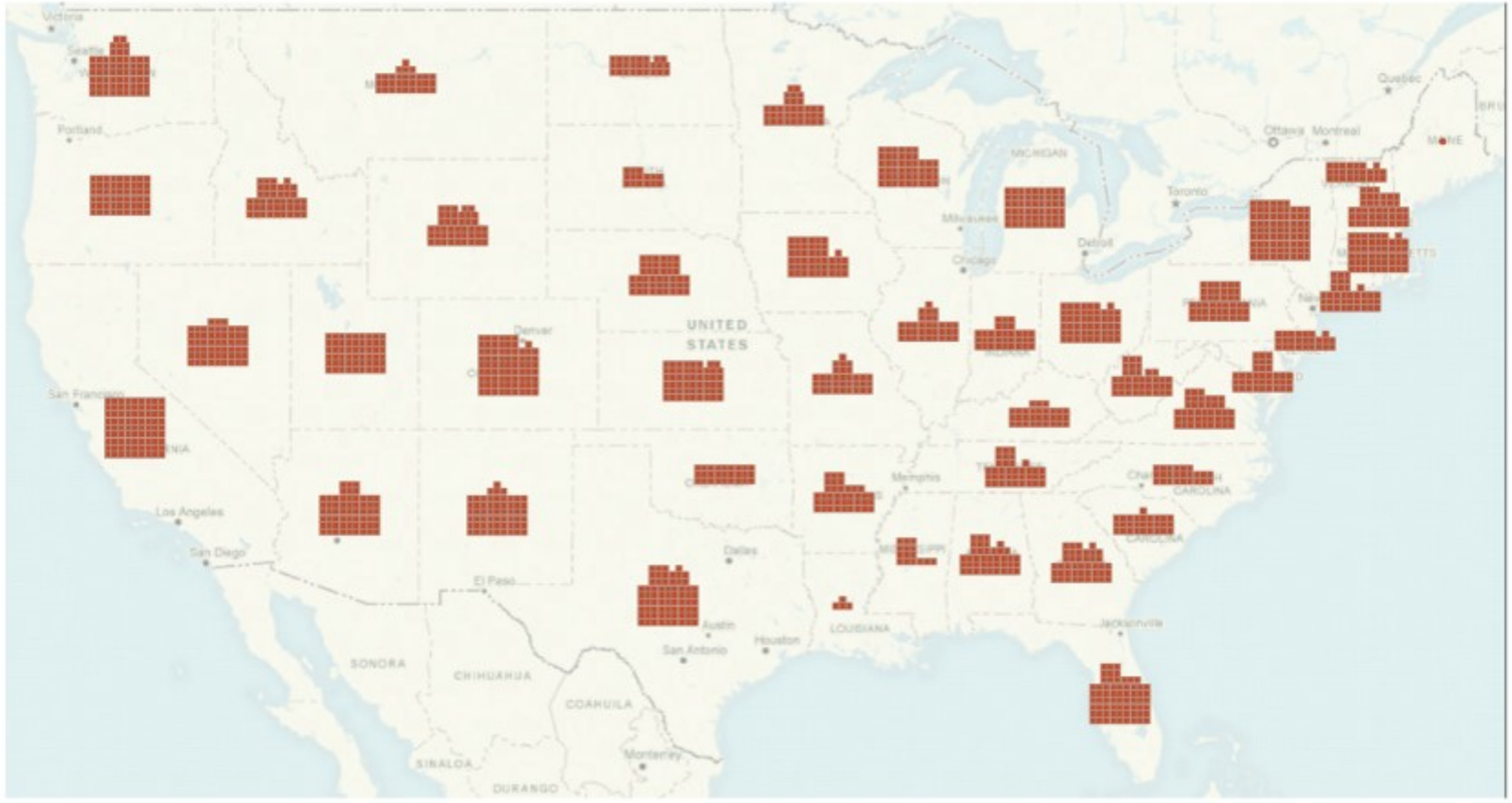
When preparing a proportional symbol map, it is important to ensure that the geographical level used is the same.



Source: COVID-19 Dashboard

Altertive of Proportional Symbol Map: Bricks Map

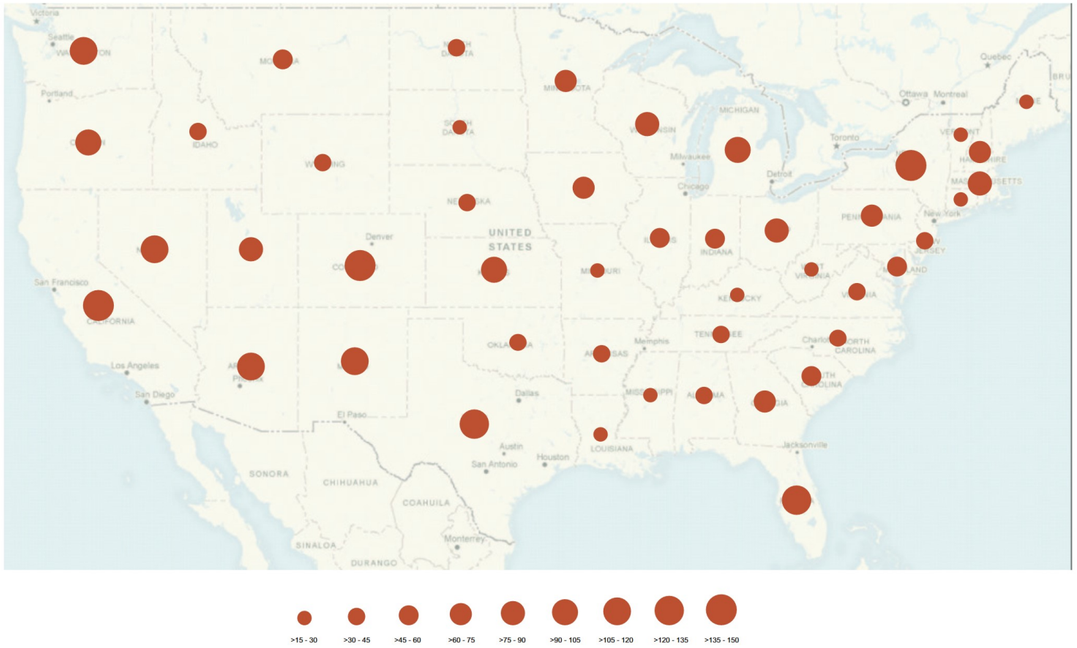
Bricks Map aims to provide better ways to encode quantitative information graphically—ways that can be easily, efficiently, and accurately decoded.

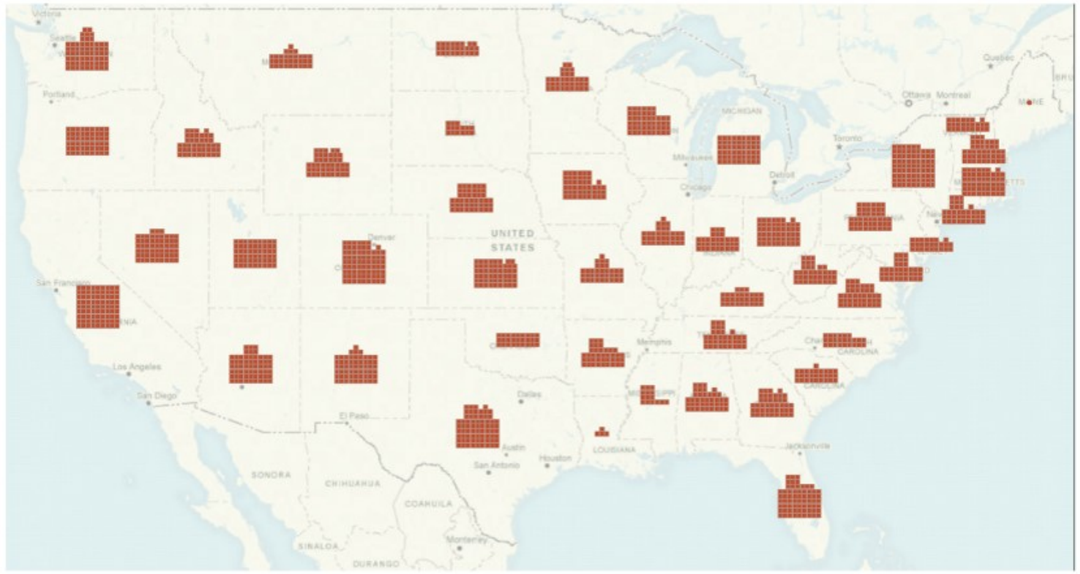


Source: Building insight with bricks.

Bricks Map versus Proportional Symbol Map

By plotting a proportional symbol map and a bricks map side-by-side as shown below, it is clear that bubbles of graduated sizes proportional to the value ranges that they represent can be difficult to distinguish than the bricks.

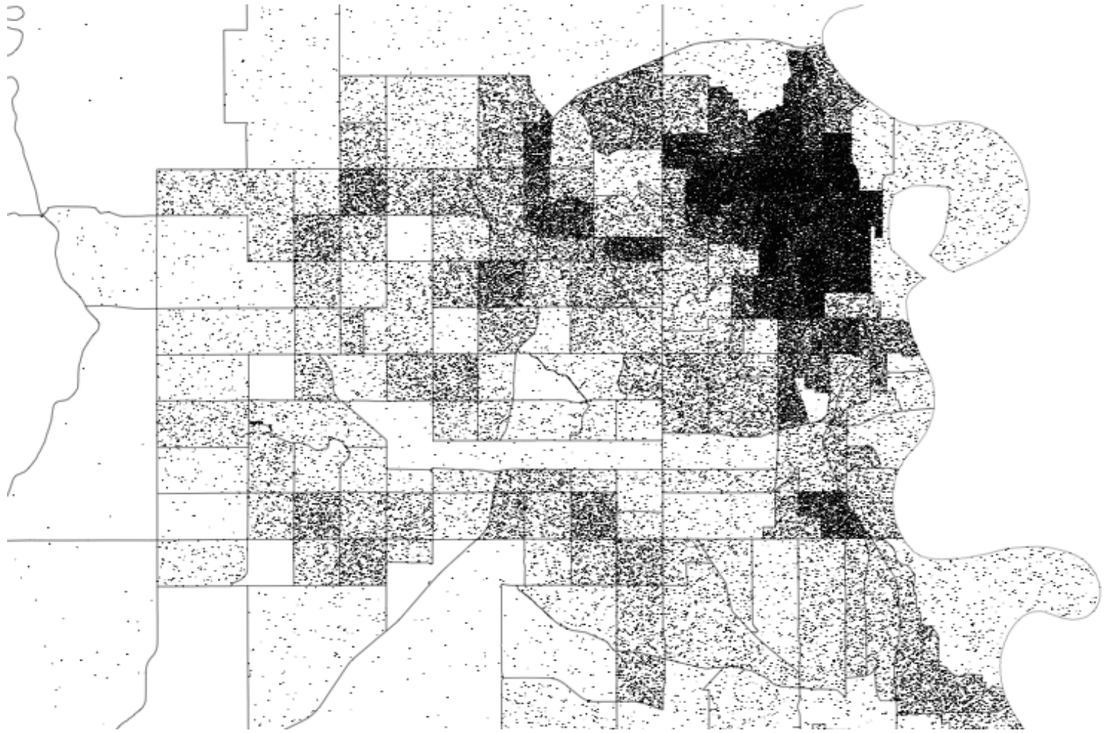




Copyright © 2013 Stephen Few, Perceptual Edge

Dot Density Map

A dot-density map is a type of thematic map that uses dots or other symbols on the map to show the values of one or more numeric data fields. Each dot on a dot-density map represents some amount of data.

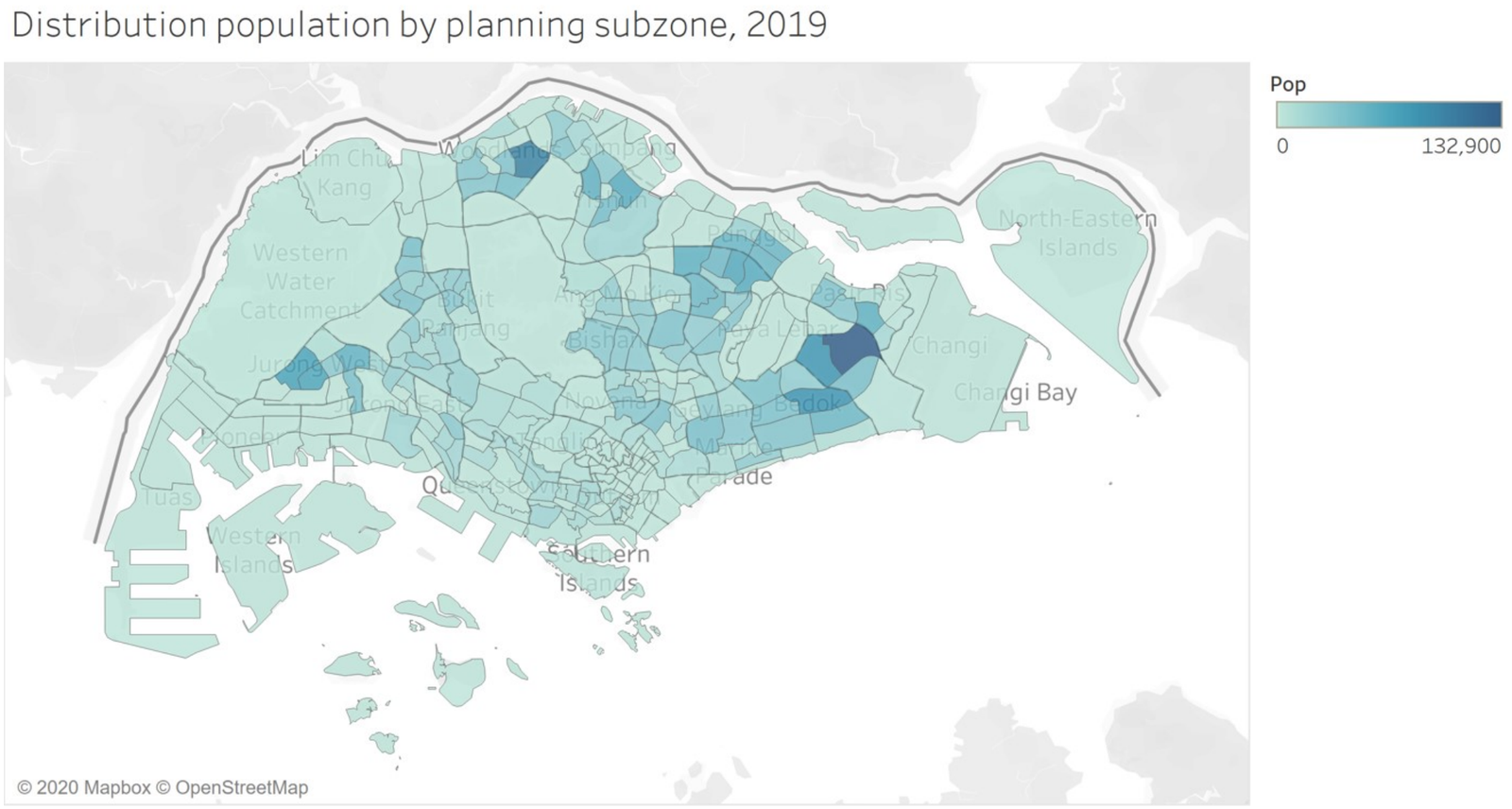


One dot represent 100 households.

Reference: Dot distribution map at wiki and Dot Density Maps

Choropleth Map

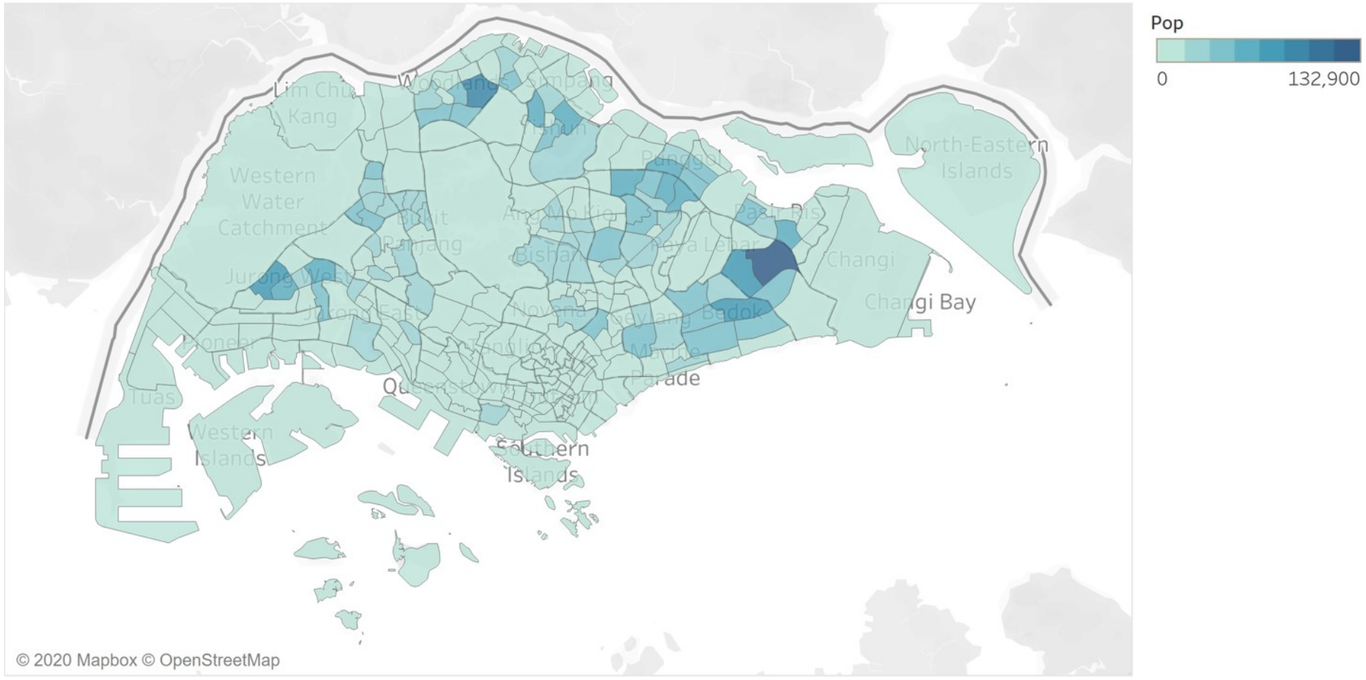
A choropleth map is a type of thematic map in which areas are shaded or patterned in proportion to a statistical variable that represents an aggregate summary of a geographic characteristic within each area, such as population or per capita income.



Classified choropleth map

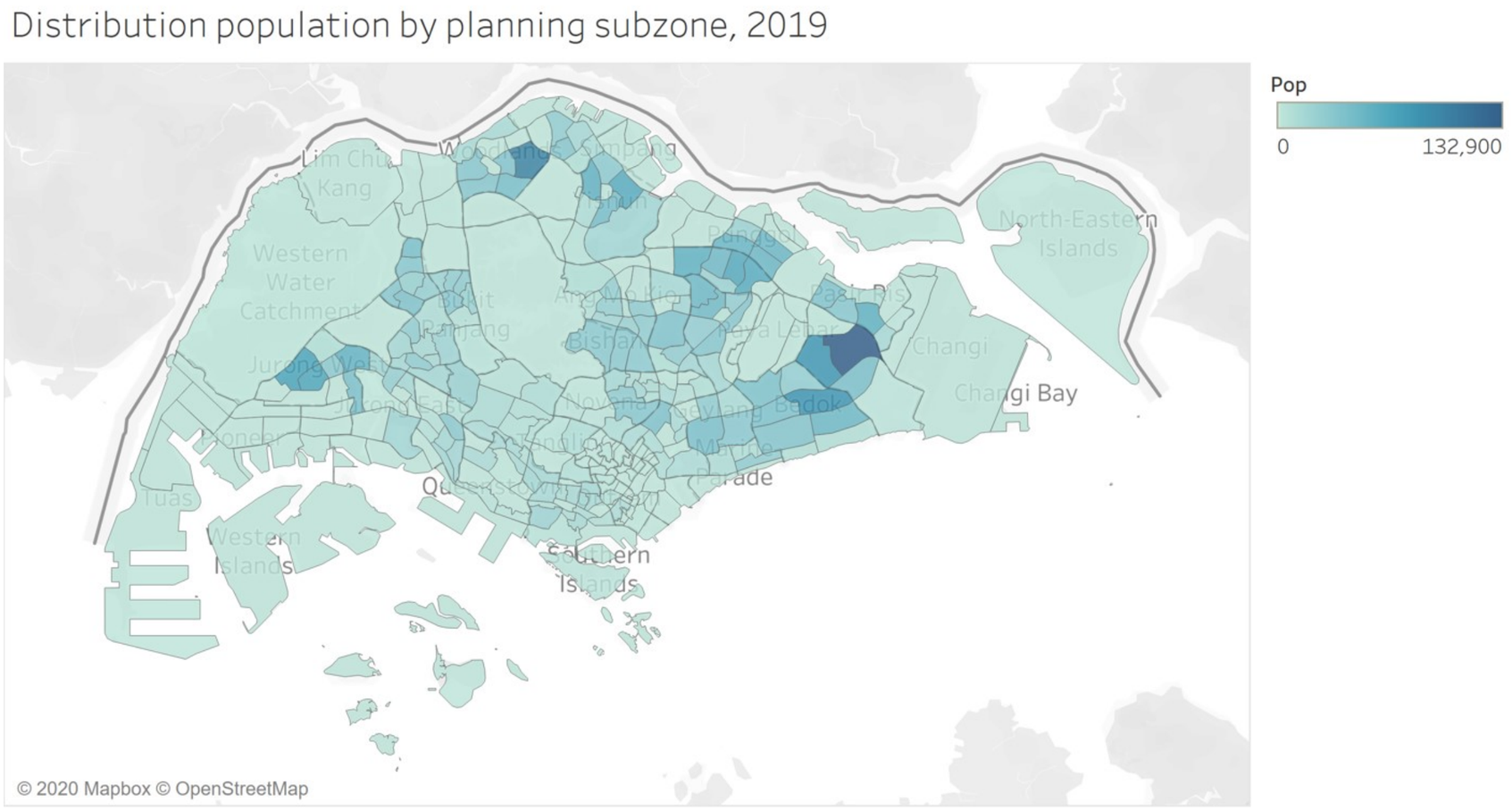
- A choropleth map can be either classified or unclassified.
- A **classed choropleth map** combines areal units into a smaller number of groups. Interval levels may vary, but typically 4 to 7 are used in a map. There is different classification techniques used to divide up the intervals.

Distribution population by planning subzone, 2019



Unclassified choropleth map

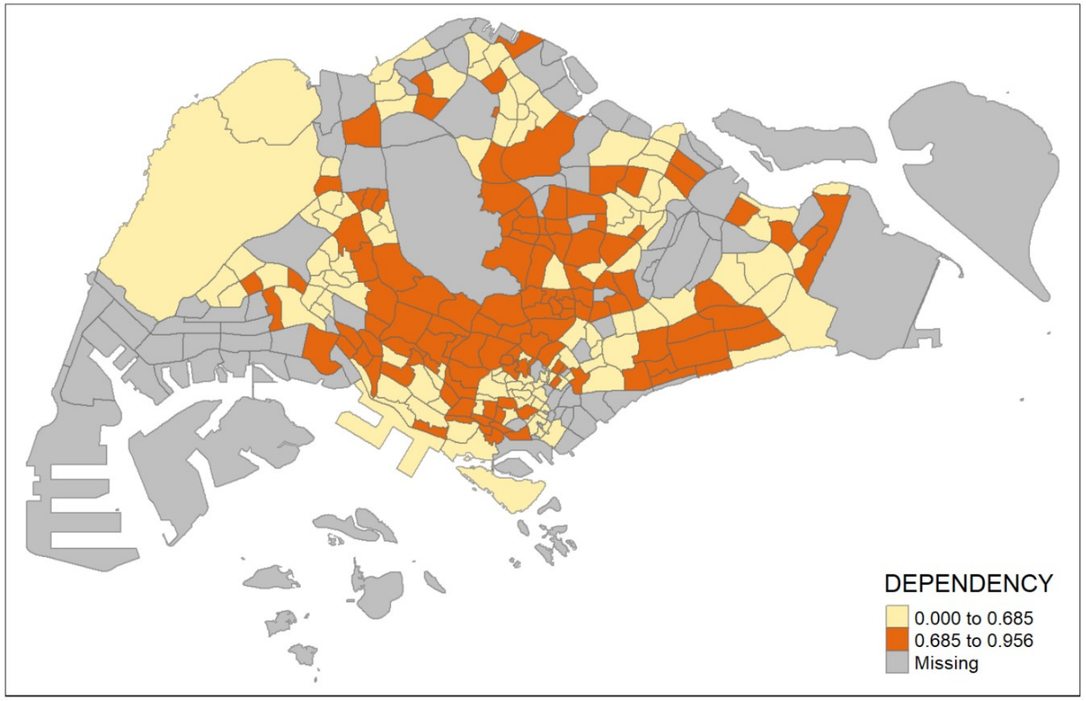
- An **unclassified choropleth** map is similar to a classed choropleth map; however, an unclassed choropleth maps do not have an averaged statistic towards each particular colour.

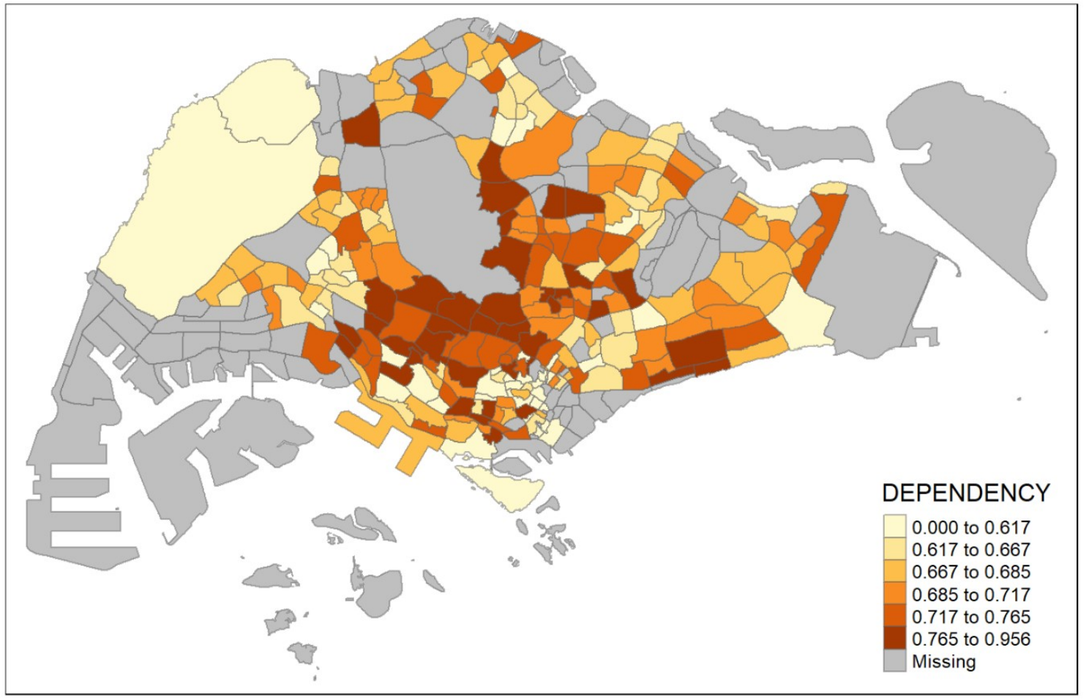


Choropleth mapping technique - Number of classes

The readability of a choropleth map will be affected by the number of classes used.

- When a small number of classes is used, the spatial distribution reveals will be very limited.
- When an appropriate number of classes is used, the spatial distribution reveals will be clear.





Choropleth mapping technique: Number of classes

Figure below provided a mathematical method to determine number of classes used for preparing a choropleth map.

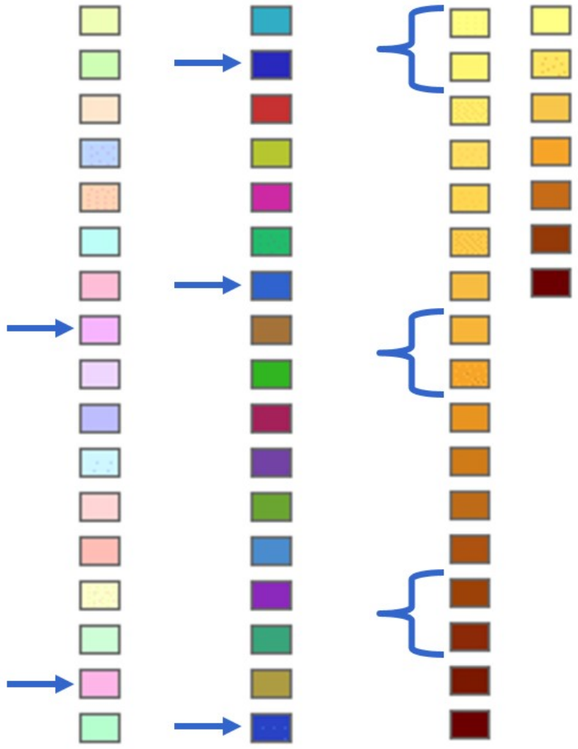
- Sturges' formula

$$k = 1 + 3.32 * \log n$$

n = number of values
k = number of classes

If n = 36
k = 1 + 3.32 * log n
k = 6 approx.

Example



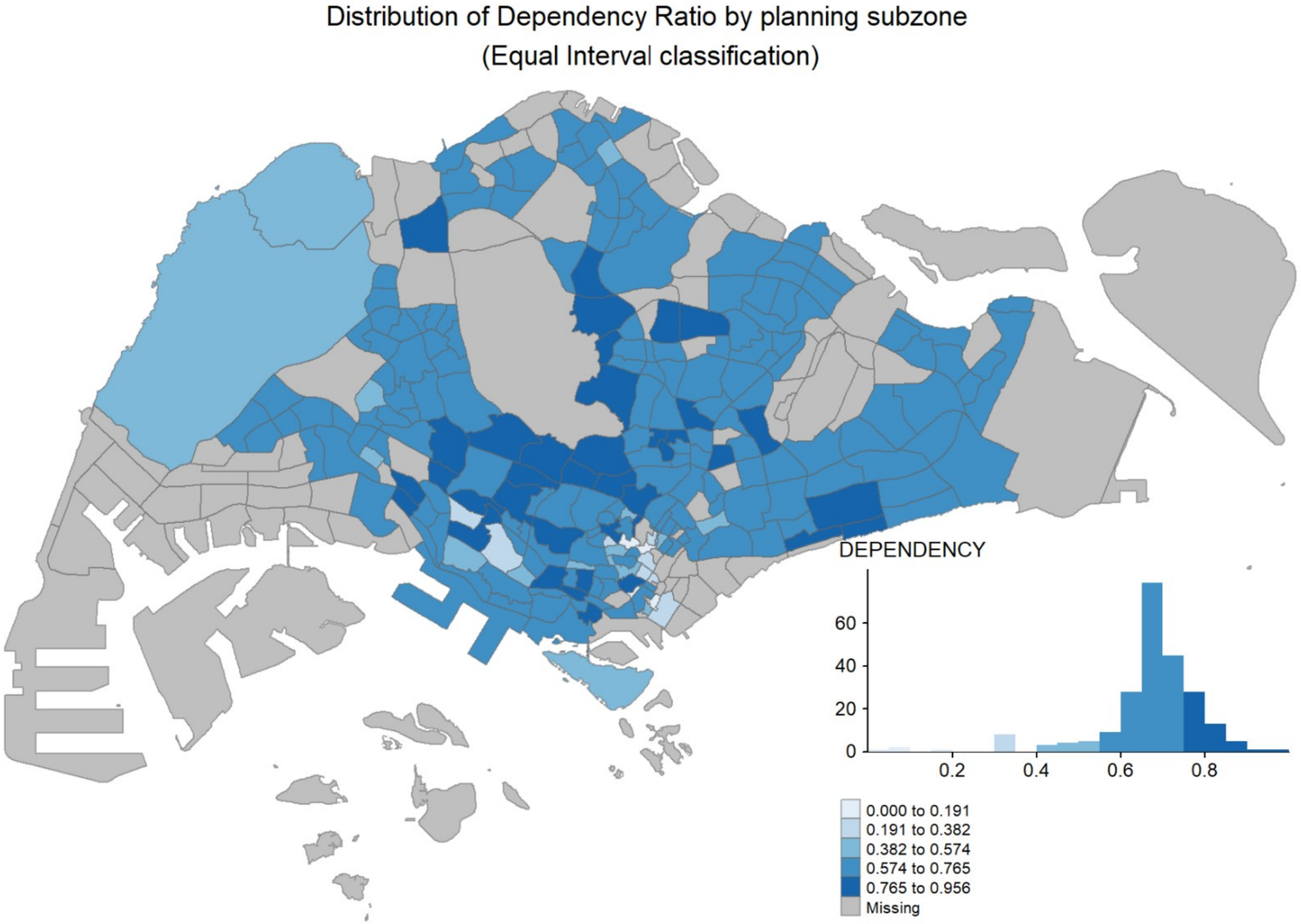
- Number of intervals?
  - Not less than 4
    - To avoid an overly generalized map
  - Not more than
    - 12 colors
    - 7 or 8 shades of the same color

Choropleth mapping technique: Methods of choosing classification scheme

- Based on the nature of the distribution
  - quantile, equal interval, natural breaks, standard deviations, defined interval
- Arbitrary
  - Can be based on round numbers.
  - Examples: Grouping according to age or census housing categories
  - Can result in empty categories

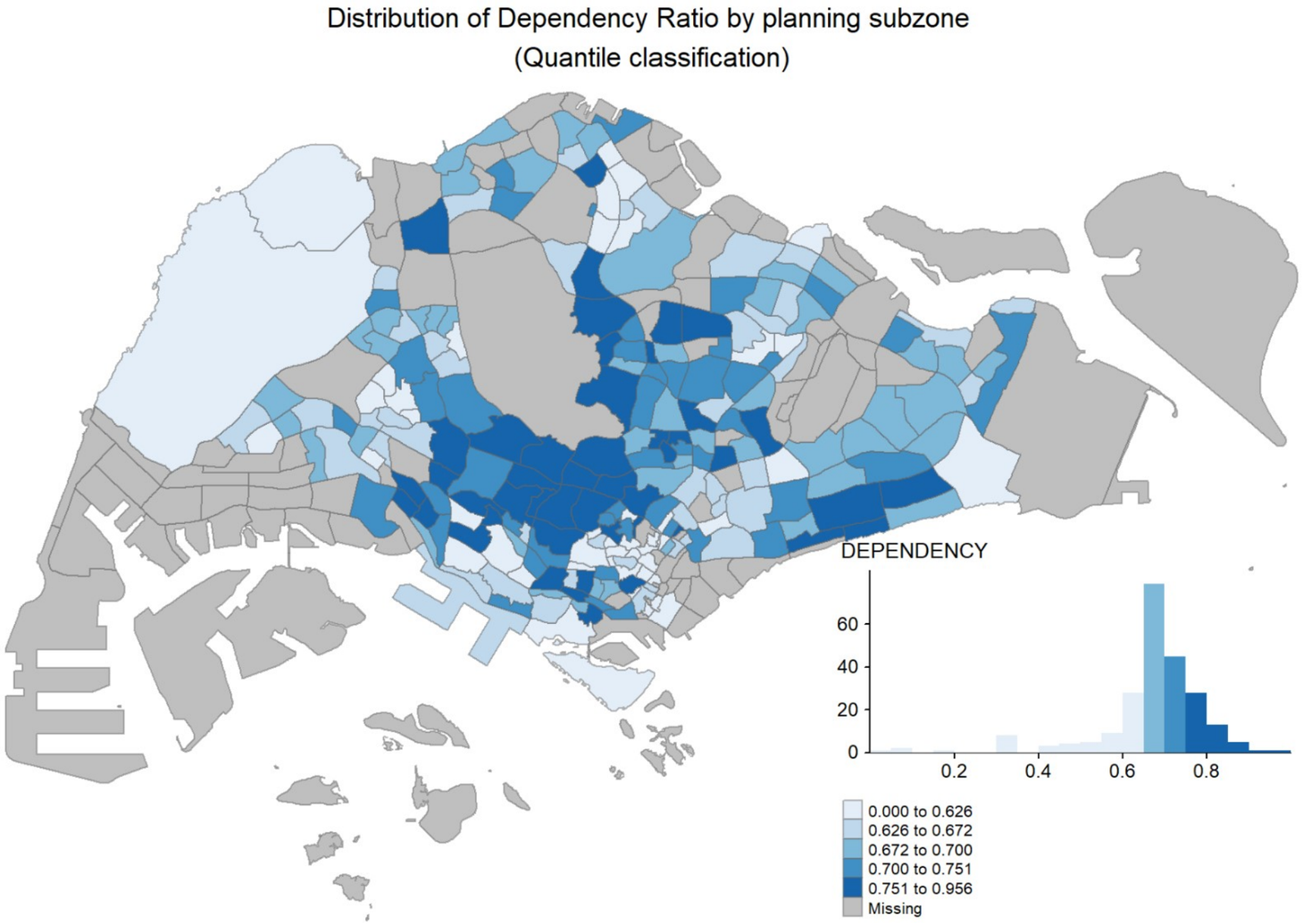
Data classification method: Equal interval

- Divides the range of attribute values into equally sized classes.



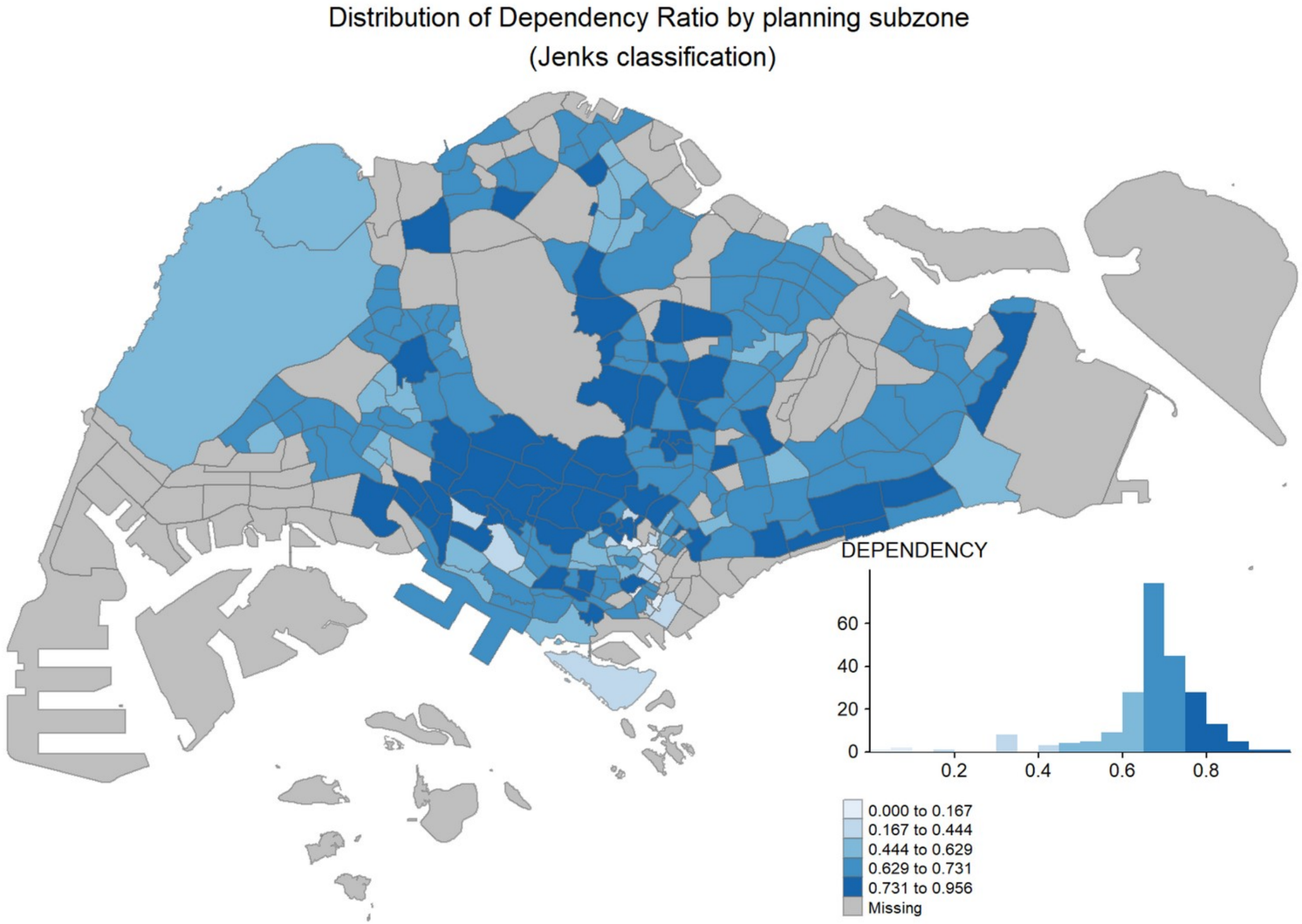
Data classification method: Quantitle

- Same number of features per class.



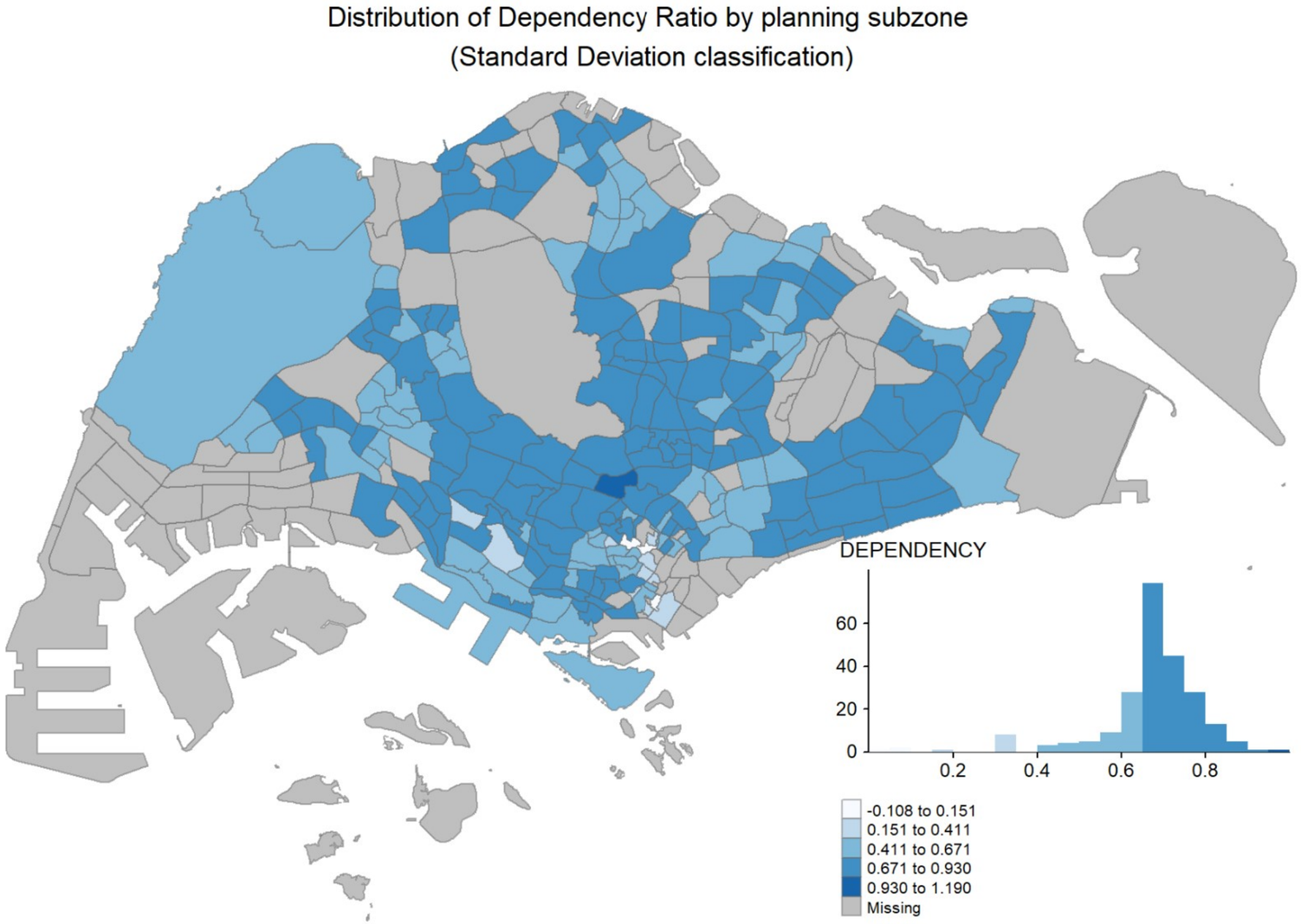
Data classification method: Jenks (also known as Natural breaks)

- Default Jenk's statistical optimization by finds natural groupings in the data.



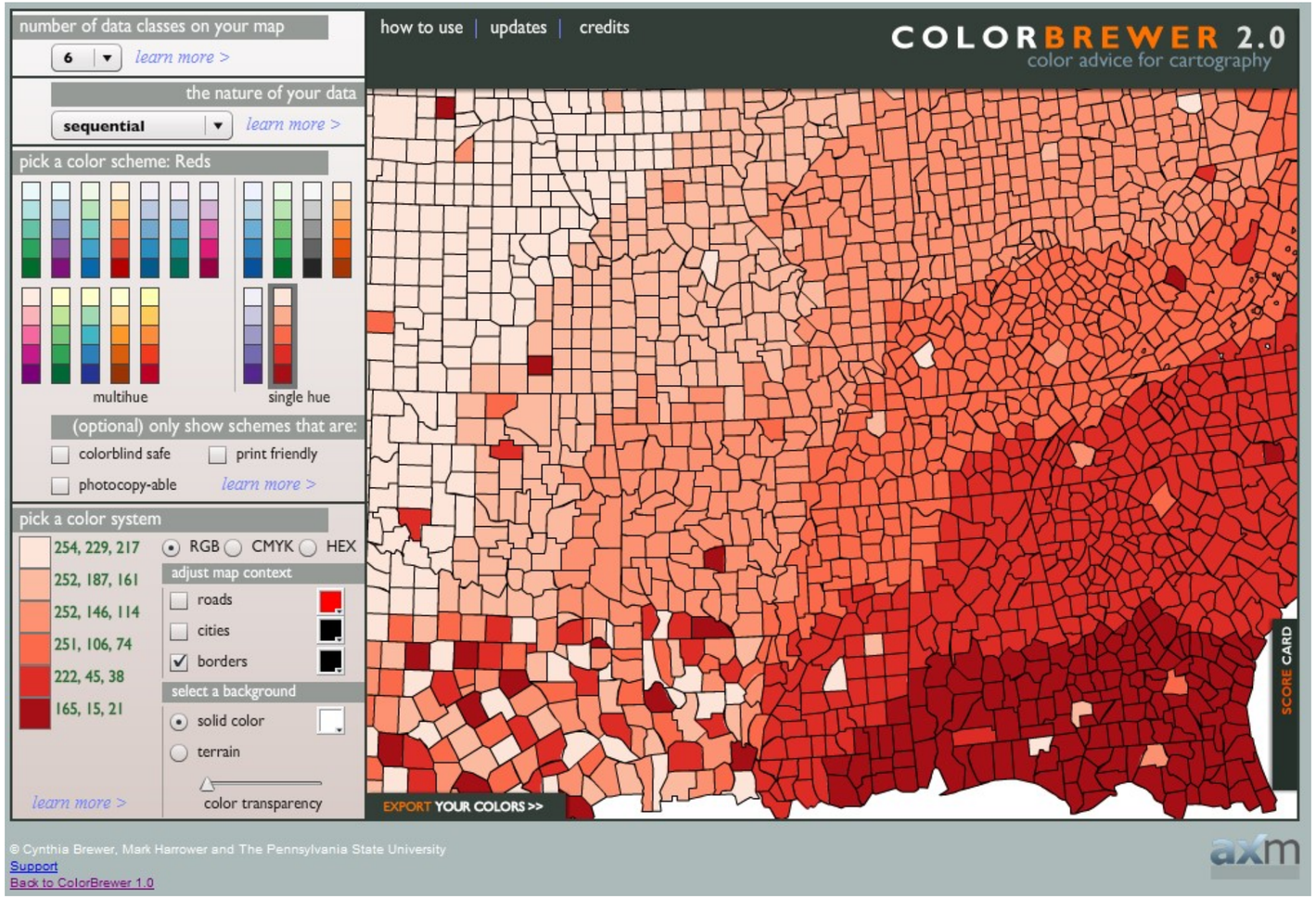
Data classification method: Standard deviation

- A measure of dispersion. Use if the distribution approximates a normal distribution.



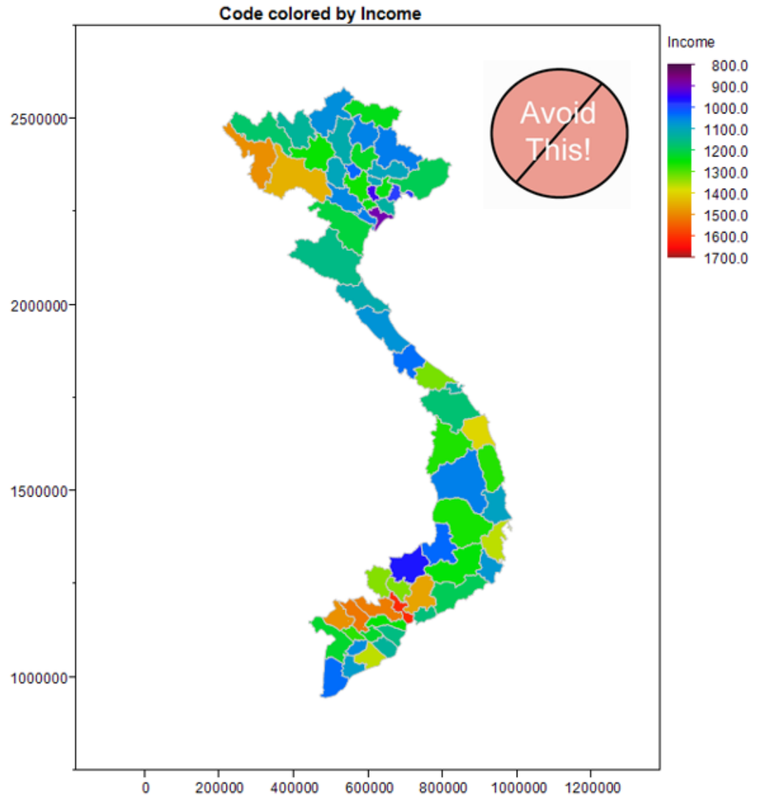
Choropleth map - Colour scheme

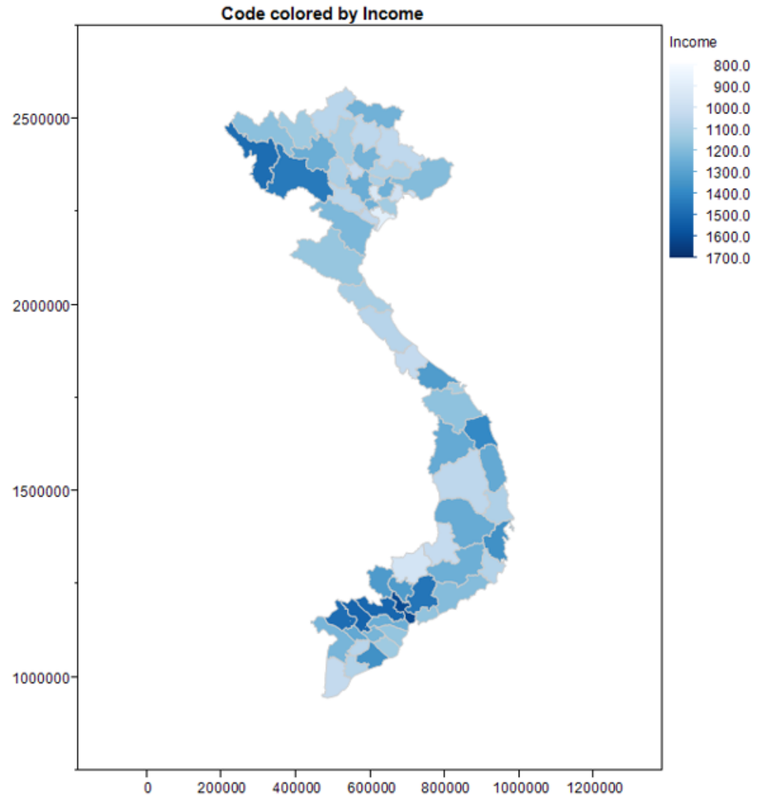
ColorBrewer is an online tool designed to help people select good color schemes for maps and other graphics.



Colour Scheme Tip

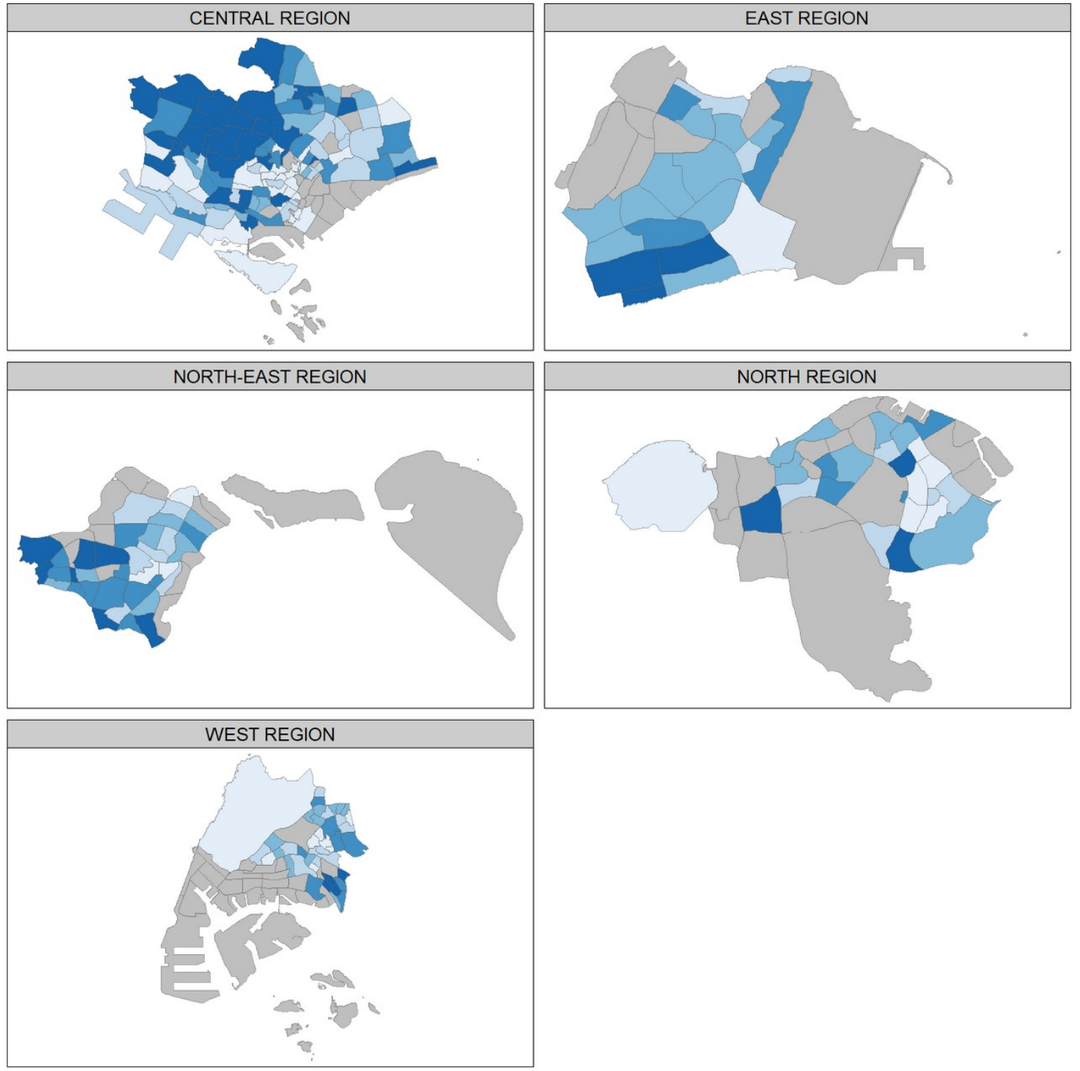
- Avoiding multiple colours in choropleth mapping
- Stick with a single hue (or a small set of closely related hues) and vary intensity from pale colours for low values to increasingly darker and brighter colours for high values.





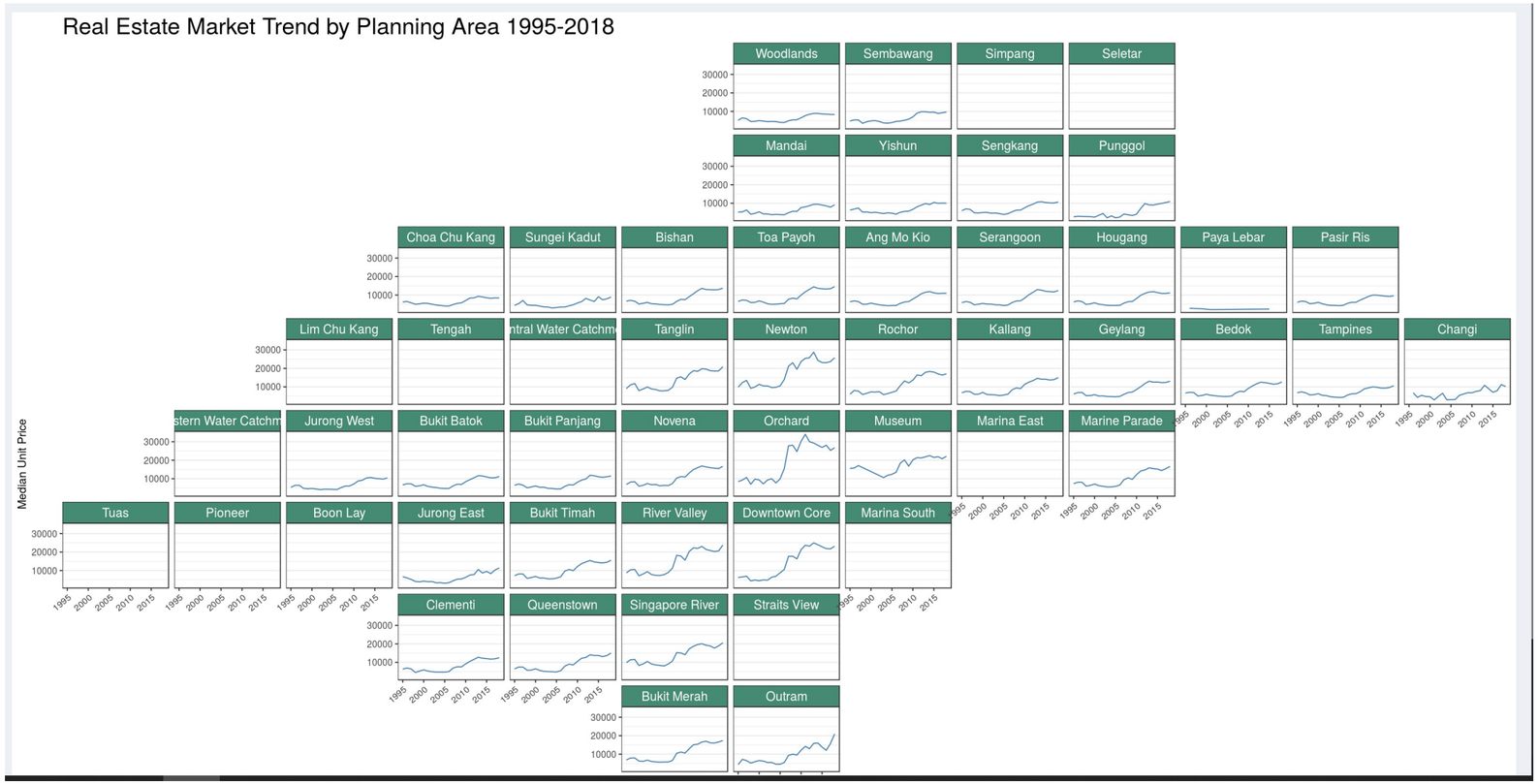
Trellis choropleth map

- **Trellis maps** , also known as **small multiple maps**, are composed of many maps arrange side-by-side, and sometimes stacked vertically.
- Small multiple maps enable the visualisation of how spatial relationships change with respect to another variable, such as time.



Geofacet

Geofacet takes data representing different geographic entities and apply a visualization method to the data for each entity, with the resulting set of visualizations being laid out in a grid that mimics the original geographic topology as closely as possible.



Source: Singapore Property Market Watch

Reference

Proportional Symbols

Choropleth Maps

Dot Density Maps

The Basics of Data Classification

Should a map be interactive?

Map interaction