AN ILLUSTRATED GUIDE

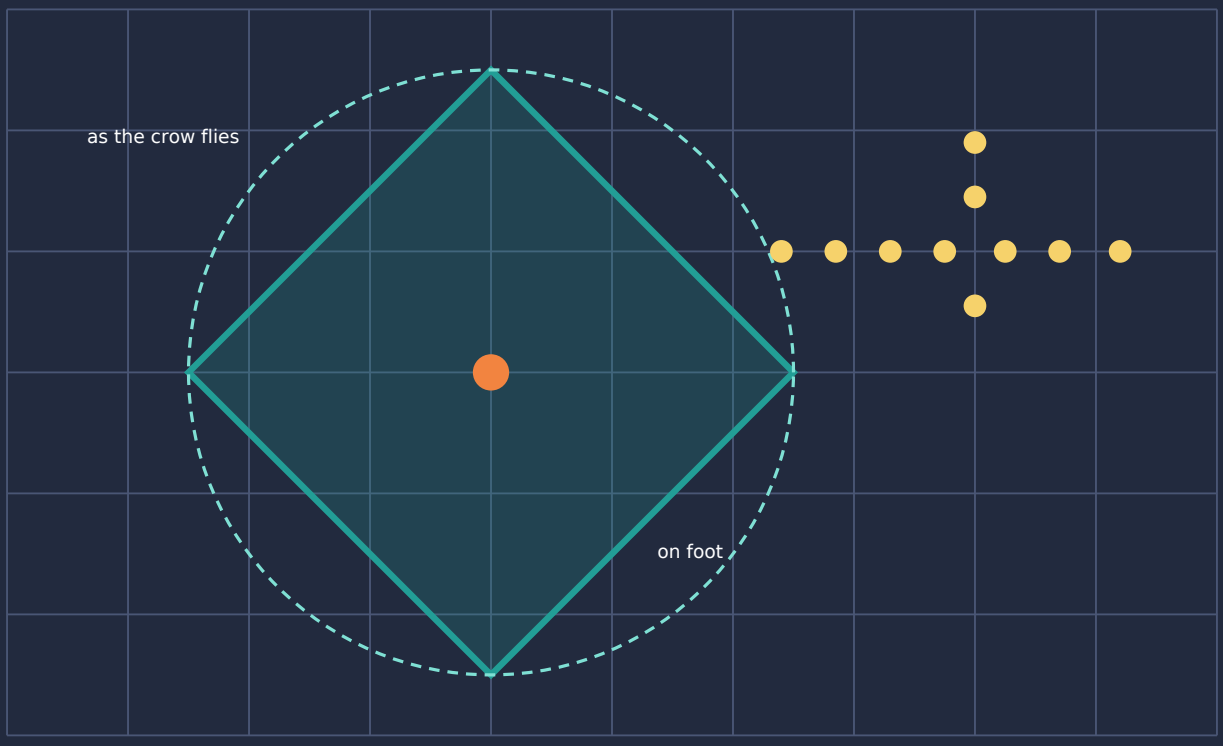
Spatial Analysis for City Design

A friendly guide to reading and measuring the city with maps and spatial data

Plain English, a picture for every idea, worked examples on a toy city, and a glossary, a quiz and an eight-week plan.

Sixteen short chapters · Quiz · Glossary · Eight-week plan

Prompted and directed by Mahmoud A.
Written and illustrated by Claude (Anthropic)



How to Use This Guide

Read it in order, or jump to what you need.

This guide explains **spatial analysis**, the craft of measuring places with maps and data, for people who shape cities or want to. You need no maths beyond school level and no coding. Every idea gets a plain explanation, a picture and a small worked example.

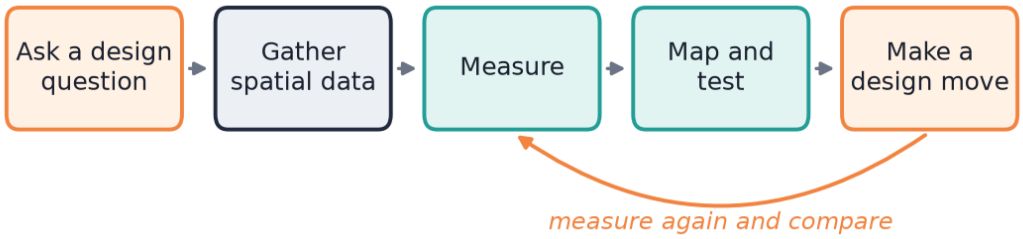


Figure 1. The whole guide in one picture. Ask, gather, measure, map, design, then measure again.

Who it is for

- Architects, urban designers, landscape architects and planners who want evidence behind their sketches.
- Students and curious readers who keep hearing words like **GIS**, **isochrone** or **space syntax** and want to know what they mean.
- Anyone who reviews a plan and wants to ask sharper questions about it.

How each chapter works

- A **banner** gives the topic in one sentence.
- Plain explanation first, then a picture, then a small example.
- Boxes: **TIP** (a shortcut), **SAY IT LIKE THIS** (a phrase you can use out loud), **WATCH OUT** (a common trap) and **IN ONE LINE** (the chapter in a sentence).
- At the back: myths and facts, a quiz with answers, a glossary, a one-page cheat sheet and further reading. Chapter 16 adds an eight-week plan.

A note on the numbers

Almost every example uses one **toy city**. It is a made-up town of 2 km by 2 km with a street grid, a railway, a park, 900 homes and about 120 shops. The Python script that builds this guide generates the city with fixed random seeds and then runs every calculation itself: distances, densities, statistics, isovists and the rest. Every number in the worked examples comes from that script. Because the city is invented, the numbers teach the methods. They say nothing about any real place.

Contents

Sixteen short chapters and a toolbox at the back.

1 **Why “Where” Matters**
Location is part of the answer, and near things tend to be alike.

2 **Maps, Coordinates and Projections**
How a round Earth becomes a flat map you can measure on.

3 **Vector and Raster**
Two ways to store the world: shapes or grids.

4 **Where Spatial Data Comes From**
Open maps, official records, satellites and sensors.

5 **Scale and the Modifiable Areal Unit Problem**
Draw the zones differently and the story changes.

6 **Buffers, Overlay and Selection**
Draw a zone, stack the layers, keep what qualifies.

7 **Distance and Accessibility**
Measure along the streets people actually walk.

8 **Density and Land-Use Mix**
How much is here, and how varied is it?

9 **Spatial Statistics: Clusters and Hot Spots**
Seeing patterns in points, and testing them against chance.

10 **Filling the Gaps with Interpolation**
Estimating values between the places you measured.

11 **Street Networks, Centrality and Space Syntax**
Which streets are central, and why movement follows layout.

12 **Visibility and Isovists**
What can you see from here?

13 **Measuring Urban Form**
Coverage, floor area, blocks and junctions in numbers.

14 **From Analysis to Design Moves**
Measure, change, measure again.

15 **Pitfalls and Honest Limits**
Maps persuade, so check them before they persuade you.

16 **Try It Yourself**
A short recipe, open tools and an eight-week plan.

At the back: Myths and Facts · Quick Quiz · Glossary · One-Page Cheat Sheet · Further Reading

1 Why “Where” Matters

Location is part of the answer, and near things tend to be alike.

Most questions about a city have a hidden second half. How many people can walk to the park? *Which* people, and from *where*? **Spatial analysis** is the set of methods that answers the “where” half with measurement instead of guesswork.

A map shows what is where. Spatial analysis goes one step further. It measures distances, counts what falls inside an area, finds clusters and tests whether a pattern could be chance. Almost every design decision is about location, such as where to put an entrance, a shop, a crossing or a tree. Most of this work happens in a **geographic information system** (GIS), software that stores, analyses and maps spatial data.

Tobler’s first law

In 1970 the geographer Waldo Tobler stated what he called the **first law of geography**: “Everything is related to everything else, but near things are more related than distant things.” Neighbouring houses tend to have similar prices. Neighbouring streets tend to have similar temperatures.

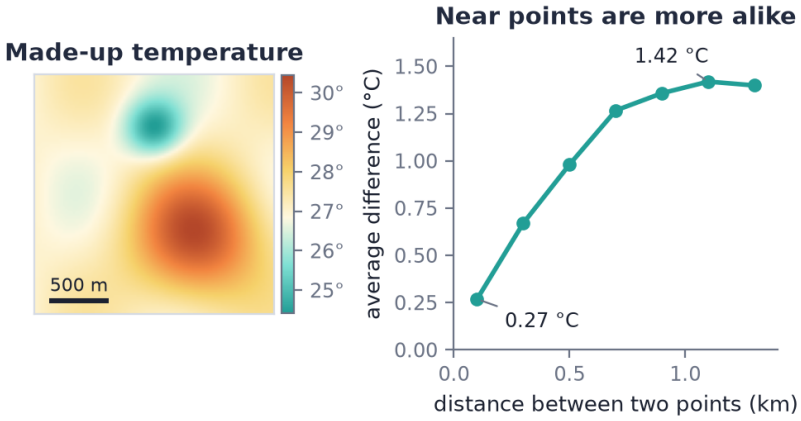


Figure 2. Tobler’s law in the toy city. Left: a made-up afternoon air temperature. Right: for 6,000 random pairs of points, the average temperature difference rises with distance, then levels off beyond about 1 km. Computed for this guide.

Points less than 200 m apart differ by only 0.27 °C on average. Points 1.0 to 1.2 km apart differ by 1.42 °C, about 5 times as much. This is good news and bad news. It is good because it lets us estimate values between measurements (Chapter 10). It is bad because many standard statistics assume that each observation is independent of the others, and nearby places are not (Chapter 9).

IN ONE LINE

Spatial analysis measures the “where” of a design question, starting from the idea that near things are related.

2 Maps, Coordinates and Projections

How a round Earth becomes a flat map you can measure on.

Spatial data stores every location as numbers called **coordinates**. The most familiar pair is **latitude** (north or south of the equator) and **longitude** (east or west of a reference line through Greenwich, London). Both are angles in degrees.

Angles are fine for finding a place but poor for measuring it. A degree of latitude is always about 111 km. A degree of longitude shrinks away from the equator, because the lines of longitude meet at the poles. At 60° it is only about 56 km.

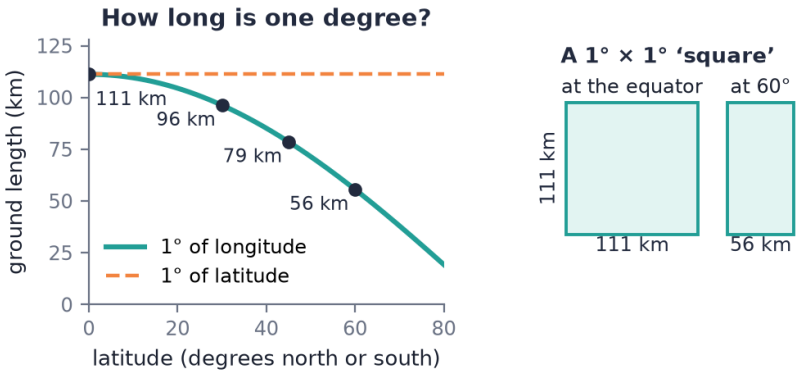


Figure 3. Longitude shrinks with latitude, so a 1° “square” narrows. Computed for this guide.

Projections and coordinate reference systems

A **map projection** is a recipe for flattening the curved Earth onto a plane. Every projection stretches something, and none keeps areas, shapes and distances all true. For city work, use a **projected coordinate system** in metres whose distortion is tiny over your study area, such as a national grid or a zone of the **Universal Transverse Mercator** (UTM) system, which splits the world into 60 narrow zones.

The full description of how coordinates relate to the Earth is a **coordinate reference system** (CRS). EPSG:4326 (a code from a standard catalogue of CRSs) is latitude and longitude on the WGS 84 datum (a model of the Earth’s shape) used by GPS. **Web Mercator**, used by most web maps, inflates areas towards the poles. Its coordinates look like metres but are stretched, to double at 60°, so never measure in it.

WATCH OUT

A buffer as big as a continent means your data is in degrees. Reproject it.

IN ONE LINE

Store data in degrees if you like, but measure in a local projected CRS.

3 Vector and Raster

Two ways to store the world: shapes or grids.

Spatial data comes in two main forms. **Vector** data stores features as geometry. A **point** is one coordinate pair, such as a tree or a bus stop. A **line** is a chain of points, such as a street centreline. A **polygon** is a closed line that encloses an area, such as a building or a park. Each feature has a row of facts in an **attribute table**, such as its number of storeys.

Raster data stores the world as a grid of equal square cells, like the pixels of a photo. Each cell holds one value per layer, such as a land-use code, a temperature or a height. Satellite images stack several layers, called **bands**. The size of the cells is the **resolution**. A 10 m raster has cells 10 m across.

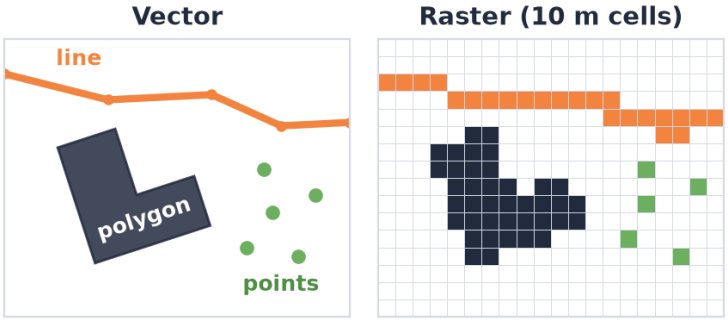


Figure 4. The same small scene stored two ways. Computed for this guide.

The script measured the L-shaped building both ways. The polygon covers 3,500 m². Counting the 10 m cells whose centres fall inside it gives 3,900 m², 11% too much. With 5 m cells the count is 3,575 m² (2% too much) and with 20 m cells it is 4,000 m² (14%). Finer cells follow the edges better but make bigger files.

Which one when?

	Vector	Raster
Best for	objects with clear edges: buildings, plots, streets, stops	continuous surfaces: heat, noise, elevation, satellite images
Weak at	smooth surfaces	sharp edges, which turn into staircases

Vector data can also record **topology**: which lines meet and which polygons share an edge. Network analysis depends on it. A street that looks connected on screen but does not share a node with its neighbour is, for the computer, a dead end.

IN ONE LINE

Vectors are shapes with attributes; rasters are grids of cells.

4 Where Spatial Data Comes From

Open maps, official records, satellites and sensors.

Good analysis starts with knowing who made the data, when, how and why. The table lists common families of data for city work.

Source	What it gives	Watch for
OpenStreetMap	Streets, paths, buildings, shops and more, mapped by volunteers; open licence	Coverage varies; check footpaths and heights
National mapping and land registry agencies	Base maps, plot boundaries, addresses, elevation	Licence terms; update cycles
Census and statistical offices	Population, households, age and jobs by small area	Data come in zones (Chapter 5); small counts hidden
Satellite imagery	Land cover, vegetation, surface heat, change; Landsat and Sentinel-2 are free	10 to 30 m pixels are coarse for streets; heat (thermal) data are coarser, about 100 m; clouds
Aerial photos and lidar	Roofs, tree canopy, terrain, building heights; lidar is laser scanning, usually from aircraft	Large files; may be costly
Public transport feeds	Stops, routes and timetables in the General Transit Feed Specification (GTFS), a standard format	Planned service can differ from actual service
Building footprint datasets	Outlines traced from imagery, often by machine learning	Missing or merged buildings; few heights
Sensors and mobile data	Counts, air quality, noise, movement traces	Who is missing; privacy

Before using a dataset, read its **metadata**, the data about the data. Five questions catch most problems. When was it collected? At what scale or resolution? In which coordinate reference system? What is missing, and why? Under what licence can I use and share it?

WATCH OUT

Missing data is rarely random. An empty area may mean that nobody mapped it.

IN ONE LINE

Know who made your data, when, how and why before you trust what it shows.

5

Scale and the Modifiable Areal Unit Problem

Draw the zones differently and the story changes.

Scale has two meanings. On a paper map it is the ratio between map and ground, such as 1:1,000. In analysis it usually means the size of the units you study: a plot, a block, a neighbourhood or a whole city. A pattern that is clear at one scale can vanish at another.

Much official data is published for zones, such as census areas or districts. Counting or averaging inside zones is called **aggregation**. The trouble is that people drew the zones, and other people could have drawn them differently. The **modifiable areal unit problem** (MAUP) names a simple fact. Change the zones and you change the results. The geographer Stan Openshaw, who named it with Peter Taylor in 1979, made it widely known with a 1984 booklet.

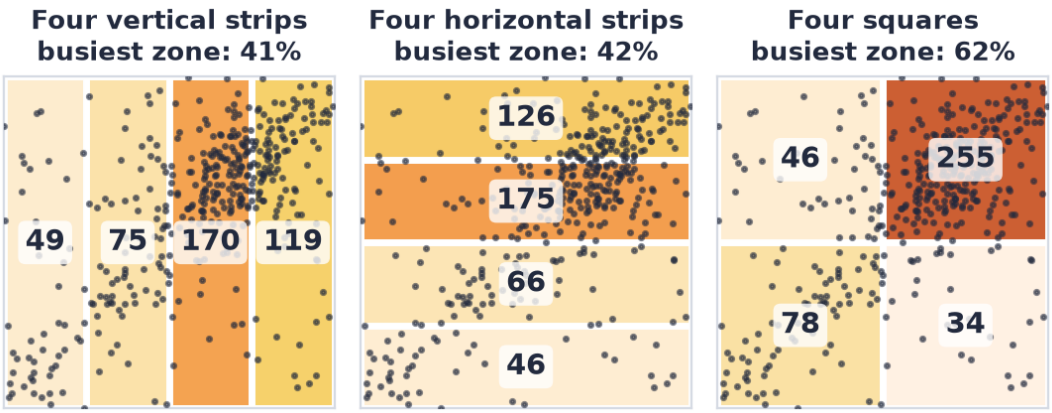


Figure 5. The same 413 points counted in three sets of four equal zones. The busiest zone holds 41%, 42% or 62% of the points, depending only on where the lines are drawn. Computed for this guide.

Two effects

- **The zoning effect.** Same number and size of zones, different boundaries, different answers (Figure 5).
- **The scale effect.** Bigger zones average away local variation, so relationships often look stronger than they are at street level.

To see the scale effect, the script made two variables, A and B, on 50 m cells. They share a gentle city-wide trend, but each also has a lot of local noise. The script then averaged them into larger cells and measured their **correlation**, a number from −1 to 1 that says how closely two variables rise and fall together.

Cell size	Number of cells	Correlation between A and B
50 m	1,600	0.20
100 m	400	0.53
250 m	64	0.85
500 m	16	0.96

At 50 m the link is weak (0.20). At 500 m it looks almost perfect (0.96). The data stayed the same in every row; the zone size alone made the link look stronger.

The ecological fallacy

A related trap is the **ecological fallacy**. It means assuming that what is true of a zone is true of each person or building in it. A district with a high average income can still contain many poor households.

> **TIP**
> Run the analysis at two or three zone sizes or layouts. If the conclusion survives, it is probably real. If it flips, say so in your report.

> **IN ONE LINE**
> Zones are choices, so test whether your result survives a different choice.

6 Buffers, Overlay and Selection

Draw a zone, stack the layers, keep what qualifies.

Three basic operations do a surprising amount of work. A **buffer** is the zone within a set distance of a feature, such as everything within 400 m of a tram stop. An **overlay** stacks two or more layers and combines them, such as a buffer and a flood map. A **selection** keeps only the features that meet a condition, either about their attributes (“more than four storeys”) or about their location (“inside the buffer”).

Worked example: homes near transit, outside the flood zone

The toy city has 900 homes, two tram stops and a low-lying strip in the south that floods. How many homes are close to a stop and also safe from flooding? The script buffered each stop by 400 m, overlaid the flood zone and selected the homes that pass both tests.

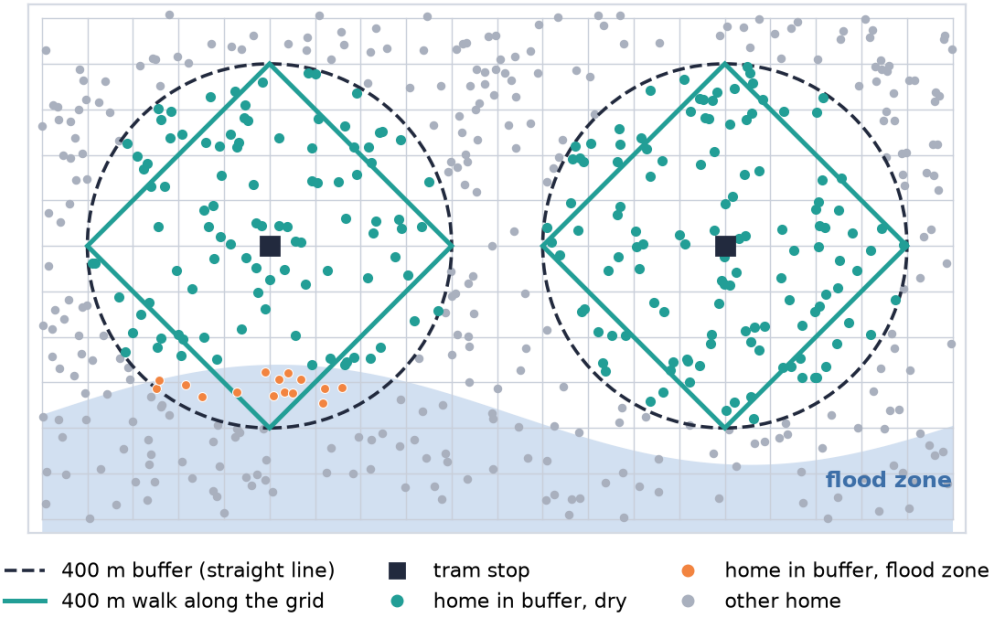


Figure 6. The southern half of the toy city. Dashed circles are 400 m buffers; the teal diamonds show 400 m measured along the street grid. Grid squares are 100 m across. Computed for this guide.

Step	Homes
All homes	900
Within a 400 m buffer of a stop	239
... of which inside the flood zone	15
Within the buffer and outside the flood zone	**224**
Within 400 m by actual walking route (Chapter 7)	153
Within a 400 m walk and outside the flood zone	**145**

The buffer gives a quick first answer. The last two rows use the street network instead of a circle, and the count falls by about 36%. A circle assumes that people can walk in straight lines through buildings and back gardens.

The core operations

Operation	What it does	Example question
Buffer	Makes the zone within a distance	What lies within 400 m of each stop?
Clip	Cuts one layer to the outline of another	Keep only the streets inside the study area.
Intersect	Keeps the parts where layers overlap	Which homes are near a stop and in the flood zone?
Union	Combines two layers, keeping every piece of both	Combine catchments and flood areas.
Dissolve	Merges features that share a value	Turn many plots into one zone per land use.
Spatial join	Copies attributes across by location	Give each home the name of its district.

SAY IT LIKE THIS

Within 400 m as the crow flies, 239 homes qualify. Along real streets, 153 do.

IN ONE LINE

Buffer, overlay and select: draw the zone, stack the layers and keep what passes every test.

7 Distance and Accessibility

Measure along the streets people actually walk.

Accessibility means how easily people can reach the things they need: shops, schools, parks and jobs. The simplest measure is the distance to the nearest one. Richer measures count how many destinations lie within a time limit, or weight each destination by how far away it is (a **gravity measure**). All of them depend on one choice: which distance?

Straight-line distance, also called Euclidean distance or distance “as the crow flies”, ignores everything in the way. **Network distance** follows streets and paths, so it respects blocks, rivers, railways and missing links. The ratio of the two is the **detour ratio**. On a perfect grid, the average detour ratio over all directions is 4/π, about 1.27. Barriers push it much higher.

Worked example: walking to the park

The script measured both distances from every home to the nearest edge of the park. The network distances come from **Dijkstra’s algorithm**, a classic method that spreads outwards from the destination, always extending the shortest known route first, until every street corner has its shortest distance.

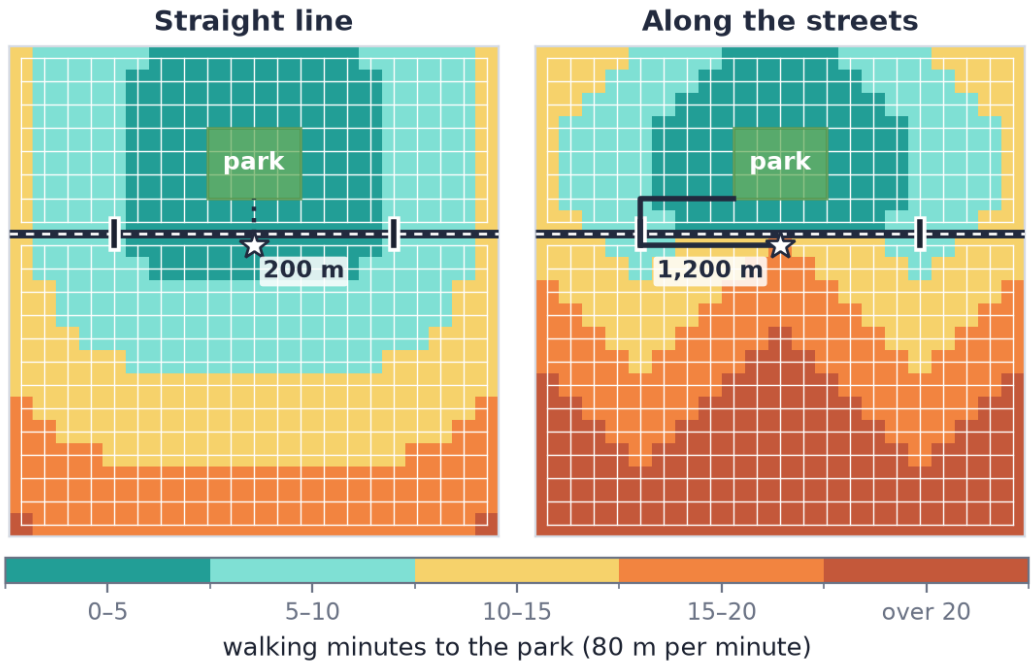


Figure 7. Walking time to the park at 80 m per minute. Left: straight-line distance. Right: along the streets. The railway (dark line) can only be crossed at two bridges, so the corner marked with a star faces a long detour. Computed for this guide.

Measure	Straight line	Along the streets
Homes within 400 m (5 minutes)	23%	16%
Homes within 800 m (10 minutes)	60%	33%
Median detour ratio, homes north of the railway	1	1.15
Median detour ratio, homes south of the railway	1	1.47

One corner shows the problem best. It sits 200 m from the park as the crow flies. On foot it is 1,200 m, because the walk must go round by a bridge.

Isochrones and the 15-minute idea

An **isochrone** (from the Greek for “equal time”) is a line joining all the places you can reach in the same travel time. The edges between the colour bands in Figure 7 (right) are isochrones. The “15-minute city” idea, popularised around 2020, asks that daily needs be within about 15 minutes on foot or by bike. Isochrones let you test that promise street by street. In the toy city only 53% of homes are within a 15-minute walk of the park.

WATCH OUT

Walking speed matters. This guide uses 80 m per minute (about 4.8 km/h), a common planning assumption. Many older people, young children and people with disabilities walk more slowly, so their 15-minute area is smaller. Hills, long waits at crossings and streets that feel unsafe shrink it too.

IN ONE LINE

Measure access along real routes and in minutes, and the barriers in a plan become visible.

8 Density and Land-Use Mix

How much is here, and how varied is it?

Density is an amount divided by an area. It sounds simple, but the answer depends on what you count and which area you divide by. One **hectare** (ha) is 10,000 m², the area of a square 100 m by 100 m.

Measure	Counts	Divided by
Dwellings per hectare	homes	land area
People per hectare	residents	land area
Floor area ratio (Chapter 13)	total floor area of all storeys	plot area
Net density	homes or people	only the plots that hold them
Gross density	homes or people	the whole area, including streets, parks and schools

Net and gross density can differ a lot for the same neighbourhood, so always say which one you mean.

Land-use mix and the entropy index

Land-use mix describes how varied the activities in an area are. A street with homes, shops, a clinic and workshops gives more reasons to walk than a street of homes alone. A common score for mix is the **entropy index**, borrowed from information theory. For an area with k possible uses, where p_i is the share of use i:

$$\textbf{entropy} = -\Sigma\, p_i \ln(p_i) / \ln(k)$$

The score runs from 0, when everything is one use, to 1, when all k uses have equal shares.

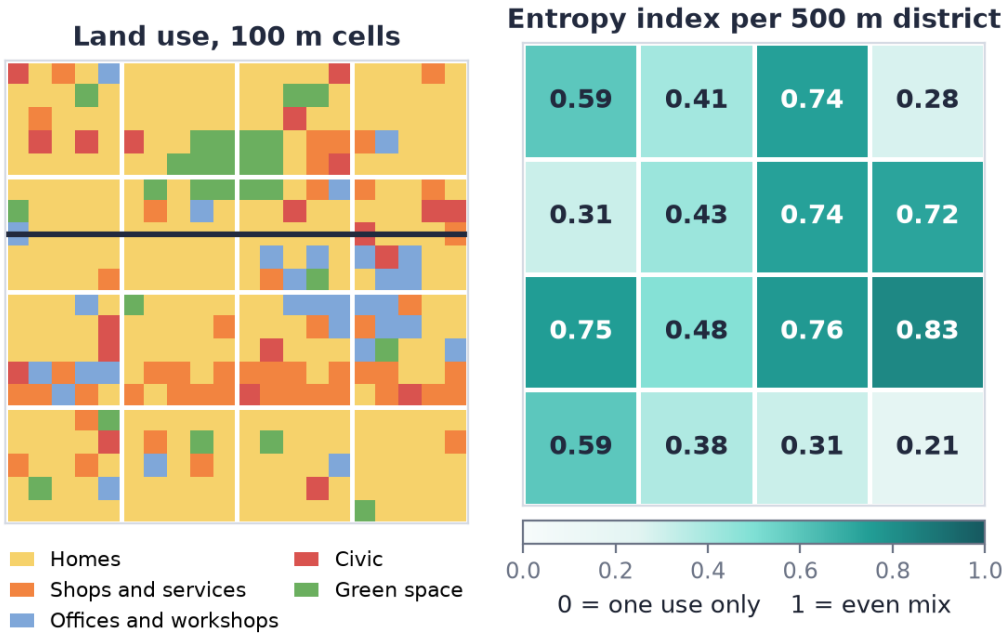


Figure 8. Left: land use in the toy city on a 100 m grid. Right: the entropy index for each 500 m district, with k = 5 uses. Computed for this guide.

Worked example: the most mixed district

Use	Cells	Share p	−p ln p
Homes	9	0.36	0.368
Shops and services	8	0.32	0.365
Offices and workshops	6	0.24	0.343
Civic	1	0.04	0.129
Green space	1	0.04	0.129
Sum	25	1.00	**1.332**

Divide the sum by ln(5) = 1.609 and the score is **0.83**. The least mixed district, almost all homes, scores 0.21. The whole city together scores 0.64. That is the scale effect again. A city can be mixed overall while many of its districts are not.

WATCH OUT

Entropy does not know what is next to what. A district with all its shops in one corner can score the same as one with shops on every street. Pair it with distance measures.

IN ONE LINE

Density says how much; the entropy index says how varied, from 0 for one use to 1 for an even mix.

9 Spatial Statistics: Clusters and Hot Spots

Seeing patterns in points, and testing them against chance.

Kernel density

The eye finds patterns even in random dots, so we measure instead. **Kernel density estimation** (KDE) turns points into a smooth surface. Imagine placing a small, soft hill over each shop and adding all the hills together. The shape of the hill is the **kernel**. Its width is set by the **bandwidth**; for the quartic kernel used here, the bandwidth is the hill's radius. The result is a surface of shops per hectare.

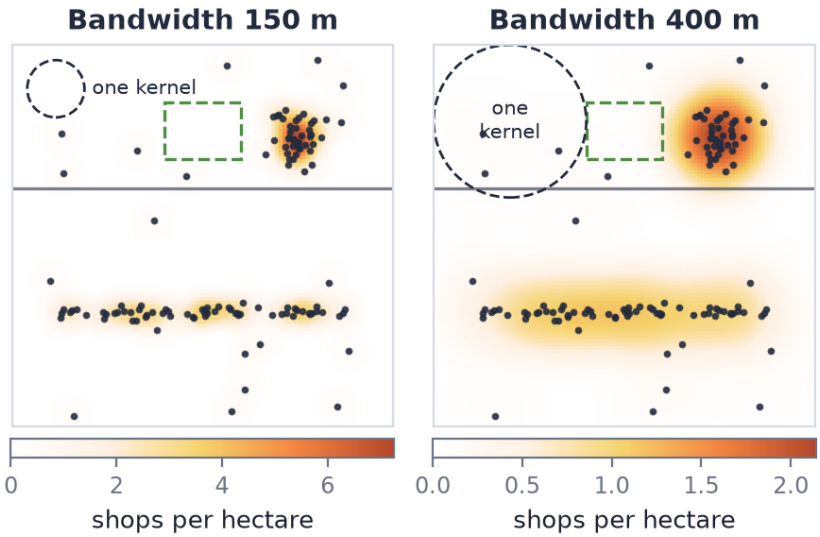


Figure 9. Kernel density of 122 shops (quartic kernel). Left: a 150 m bandwidth picks out the high street and a compact centre. Right: 400 m blurs them into broad zones. Computed for this guide.

With a 150 m bandwidth the peak is 7.3 shops per hectare; with 400 m it is 2.1. The bandwidth matters far more than the exact shape of the kernel, so match it to the question, such as a short walk for local shopping. Near the map edge, part of each hill falls off the map. With 400 m, 1.7% of the total is lost this way, an example of an **edge effect**.

Spatial autocorrelation and Moran's I

Spatial autocorrelation asks whether nearby values are more alike than chance (positive), less alike (negative) or neither (close to zero). Its best-known measure is **Moran's I**, introduced by the statistician Patrick Moran in 1950. It compares each value with its neighbours' values. A **spatial weights matrix** records who counts as a neighbour. Here, cells that share an edge are neighbours. A **p-value** is the chance of seeing a pattern at least this strong if the values were scattered at random; small values, such as below 0.05, suggest the pattern is real.

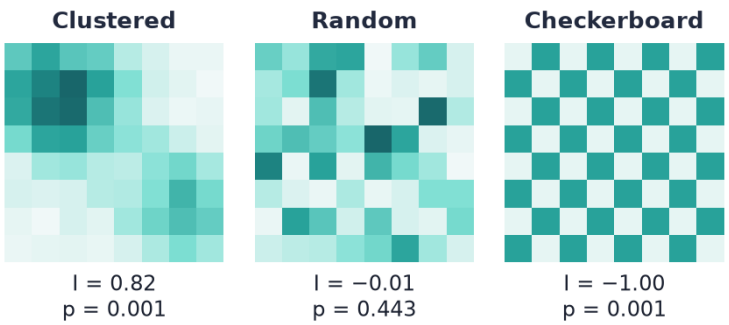


Figure 10. Moran's I for three 8 × 8 grids. With no pattern at all, the expected value is −1/63, about −0.016. The p-values come from 999 random shuffles of each grid. Computed for this guide.

Is the clustered grid's I of 0.82 just luck? The script shuffled its values 999 times and recomputed I. None reached 0.82 (the highest was about 0.31), so the **pseudo p-value** of this **permutation test** is (0 + 1) / (999 + 1) = 0.001. The toy-city shops, counted in 200 m cells, give I = 0.27 (p = 0.001). Shops cluster.

Hot spots

Local statistics give one value per place instead of one per map. Examples are the **local Moran's I**, one of the local indicators of spatial association (LISA) named by Luc Anselin, and the **Getis-Ord Gi***. Gi* compares each cell plus its neighbours with random scattering. It is a **z-score**, the number of standard deviations a value sits above or below what random scattering would give, so above about 1.96 marks a **hot spot** at the 95% level and below −1.96 a **cold spot**.

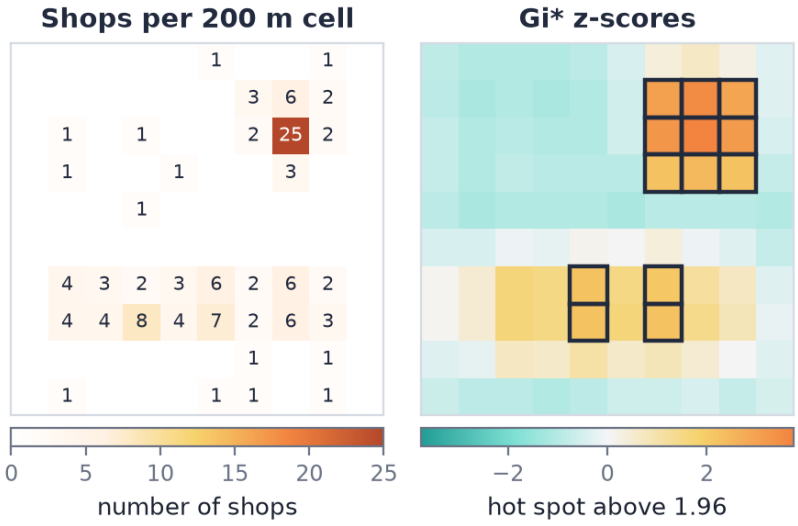


Figure 11. Left: shops per 200 m cell. Right: Gi z-scores using each cell and its eight neighbours; outlined cells are hot spots. Computed for this guide.*

13 cells are hot spots and none are cold spots, since empty patches are normal here. Testing many cells invites false alarms (Chapter 15).

IN ONE LINE

Kernel density shows clusters; Moran's I and Gi* test them against chance.

10 Filling the Gaps with Interpolation

Estimating values between the places you measured.

Sensors and survey points measure only a few places. **Spatial interpolation** estimates the values in between, trusting Tobler's law.

- **Nearest neighbour.** Each place takes the value of the closest sensor, inside its Thiessen (or Voronoi) polygon. Fast, but blocky.
- **Inverse distance weighting** (IDW). A weighted average in which closer sensors count more. With power 2, a sensor twice as far away gets a quarter of the weight.
- **Kriging.** A statistical method that fits a **variogram**, a curve of how much values differ as the distance between them grows (Figure 2 is a rough version of one), to set the weights and estimate uncertainty.

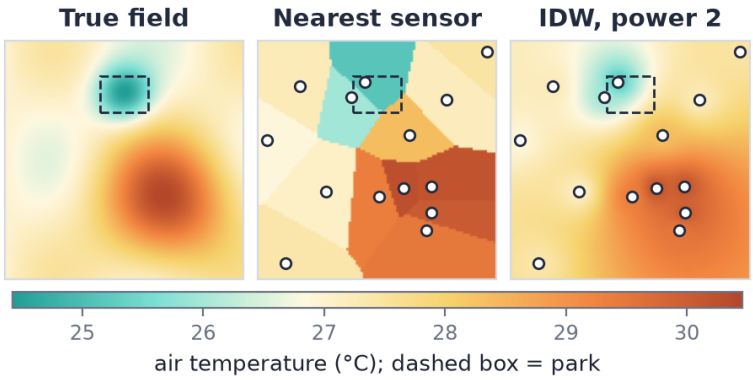


Figure 12. Afternoon air temperature in the toy city. Left: the made-up true field. Middle: nearest sensor. Right: IDW with power 2. Dots are the 14 sensors. Computed for this guide.

How good is the guess? In real life you never see the true field, so you use **cross-validation**. Hide one sensor, predict it from the others, compare, and repeat. The **RMSE** (root mean square error) gives the typical error. Leave-one-out gives 0.98 °C for nearest sensor and 0.74 °C for IDW. Against the true field the errors are 0.92 and 0.69 °C, so the cheap check ranks the methods correctly.

WATCH OUT

IDW never predicts outside the range of its sensors. The coldest sensor sits inside the park, about 50 m from its northern edge, and reads 25.2 °C. At the park's cool centre, IDW therefore estimates 26.0 °C against a true 24.4 °C. If a feature matters, put a sensor at its centre as well as near its edge.

IN ONE LINE

Interpolation fills gaps from nearby measurements; check it with cross-validation.

11 Street Networks, Centrality and Space Syntax

Which streets are central, and why movement follows layout.

A street network can be stored as a **graph**: **nodes** (junctions and dead ends) joined by **edges** (street segments). Once a city is a graph, we can ask which parts of it are central.

- **Closeness** measures how near a node is, on average, to all other nodes. High closeness marks a good place from which to reach everyone, such as a market.
- **Betweenness** measures how often a node or street lies on the shortest paths between all pairs of places. High betweenness marks lots of passing movement.

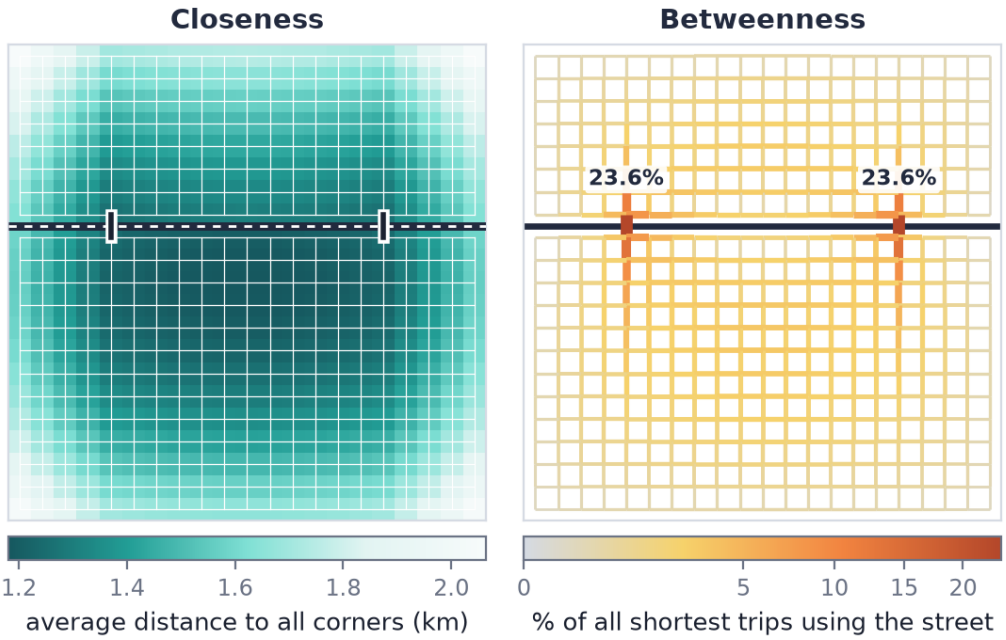


Figure 13. Left: closeness in the toy city, shown as the average distance from each corner to all others (darker is more central). Right: betweenness of each street, as its share of the shortest trips between all 97,020 pairs of corners. Computed for this guide.

The railway dominates. 47% of all pairs of corners lie on opposite sides of it, so every one of those trips must use a bridge. Each bridge carries 23.6% of all shortest trips (ties are split evenly), about 16 times the median street. In a real town, such links are likely places for crowds, shops and conflicts.

Space syntax in brief

Space syntax is a family of theories and methods, developed by Bill Hillier, Julienne Hanson and colleagues from the 1970s onwards, that relates spatial layout to how people move and meet. Its core claim is that the structure of the street network, by itself, shapes much of the pattern of movement. The classic tool is the **axial map**, the fewest and longest straight lines that cover all the open space of a layout. Each line becomes a node, and lines that cross or touch are linked.

Distance is counted in **steps**, meaning changes of direction, instead of metres. From each line the method counts the steps to every other line.

- **Integration** measures how shallow a line is, on average, from all the others. Highly integrated lines can be reached with few turns and tend to carry more movement.
- **Choice** counts how often a line lies on the shortest step-paths between all pairs of lines. It is betweenness on the axial map.

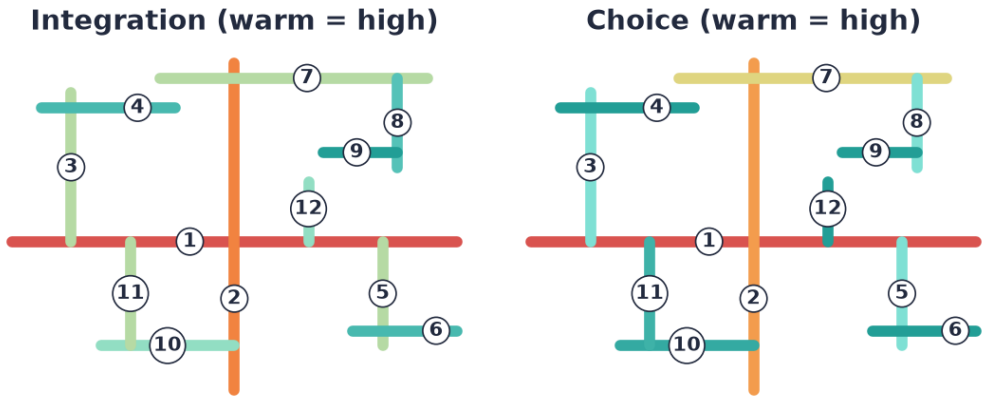


Figure 14. A twelve-line axial map. Left: integration. Right: choice. Computed for this guide.

The main street (line 1) is the most integrated (1.74) and has the highest choice, 40. Counting tied routes as fractions, it lies on the shortest paths between 40 of the 55 pairs of other lines. The north-east court (line 9), 4.2 steps deep on average, is the least integrated (0.45). The script uses Hillier and Hanson's standard normalisation, which compares mean depth with a reference value for a layout of the same size, so that higher always means more integrated. Modern studies often use street segments, count depth by the angle turned and look within a walking radius, but the logic is the same.

SAY IT LIKE THIS

Integration tells you which streets are easy to get to; choice tells you which streets people pass through.

IN ONE LINE

Treat streets as a graph, and closeness, betweenness, integration and choice show where movement is likely to concentrate.

12

Visibility and Isovists

What can you see from here?

Design is often about what people can see. Kevin Lynch's book *The Image of the City* (1960) showed how people build mental maps from paths, edges, districts, nodes and landmarks, much of it by sight.

An **isovist** is the area visible from one point, as if you shone a lamp in every direction and painted everything it lit. The architect Michael Benedikt set out isovist measures in 1979, such as **area** and **perimeter**. Later work added others, such as the longest line of sight. A large isovist feels open. A long, thin one pulls the eye in one direction.

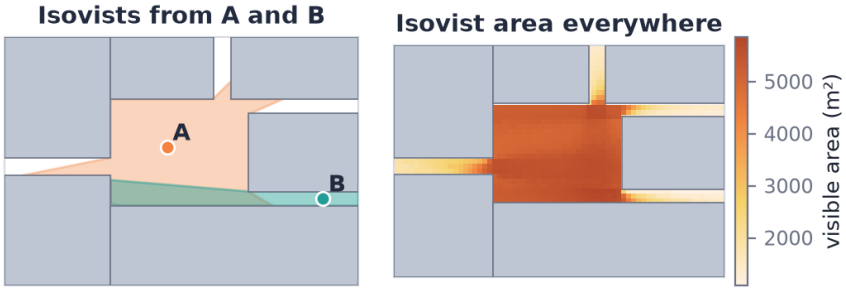


Figure 15. Left: isovists from a point in a square (A) and a point in a narrow street (B), each cast with 720 rays that stop at the first wall. Right: isovist area from every point on a 3 m grid. Computed for this guide.

Measure	A (square)	B (narrow street)
Visible area	5,122 m²	1,373 m²
Isovist perimeter	435 m	302 m
Longest line of sight	84 m	120 m

From A you see 5,122 m², almost four times the 1,373 m² seen from B. Yet B has the longer view, 120 m down the street against 84 m.

Visibility graph analysis

Repeating the isovist from every point of a fine grid gives an **isovist field**. In Figure 15 (right) the visible area ranges from 1,087 to 5,850 m². **Visibility graph analysis** (VGA), set out by Alasdair Turner and colleagues in 2001, goes a step further. It links every pair of points that can see each other and measures the resulting network, such as how many steps of sight each point is from all the others. Such maps are used to study where people gather. A plan ignores trees and level changes, so walk the site.

IN ONE LINE

One isovist shows what a person sees; a field of them reveals the plan.

13

Measuring Urban Form

Coverage, floor area, blocks and junctions in numbers.

Urban morphology is the study of the physical form of cities: plots, buildings, blocks and streets. A few simple ratios describe a lot of it.

Metric	How to compute it	What it tells you
Building coverage (also called ground space index, GSI)	building footprint ÷ site area	how much ground is built on
Floor area ratio (FAR; also floor space index, FSI, or plot ratio)	total floor area of all storeys ÷ site area	how much building there is
Average storeys	FAR ÷ coverage	how tall, on average
Open space ratio (OSR)	unbuilt ground ÷ total floor area	open ground per square metre of floor
Block size	area or perimeter of each street block	how often you meet a street
Intersection density	junctions of three or more streets per km²	how many route choices a walker has

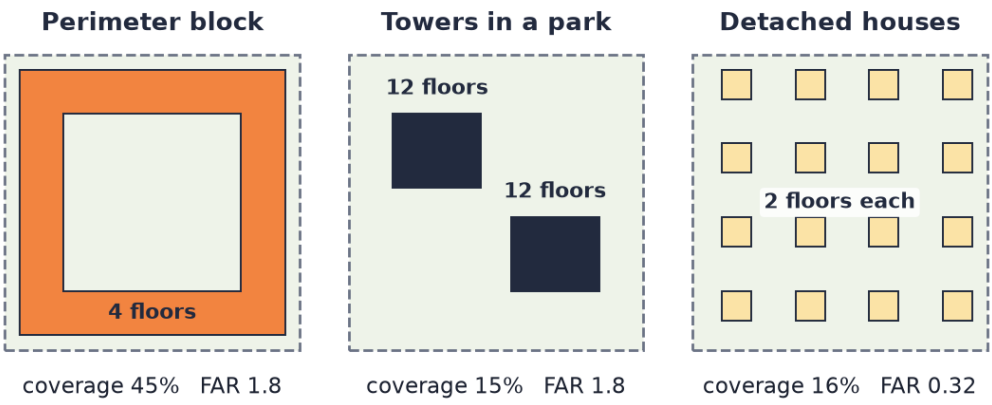


Figure 16. Three 1-hectare plots. The perimeter block and the towers have the same floor area ratio but very different coverage, height and open space. Computed for this guide.

Plot	Coverage	FAR	Average storeys	Open space ratio
Perimeter block	45%	1.80	4	0.31
Towers in a park	15%	1.80	12	0.47
Detached houses	16%	0.32	2	2.63

The perimeter block and the towers both hold 18,000 m² of floor space on one hectare. The block does it with 45% coverage and 4 storeys, the towers with 15% and 12. The same FAR can mean very different places. This is why Meta Berghauser Pont and Per Haupt combine several of these measures in one chart, which they call the Spacematrix.

The toy city’s street form

The toy city has 82.1 km of street on 4 km², or 20.5 km per km². Its blocks are 1 ha each (100 m by 100 m). Of its 441 street corners, 433 are junctions of three or more streets, about 108 per km². (The corners on the map edge are counted too, which is a small edge effect.) Small blocks and many junctions give walkers many route choices.

WATCH OUT

Ratios hide how things are arranged. A FAR of 1.8 can be one bulky building or a fine grain of small ones. Always look at the plan as well as the numbers.

IN ONE LINE

Coverage, FAR, height, open space, block size and junction density turn the look of a place into numbers you can compare.

14 From Analysis to Design Moves

Measure, change, measure again.

Analysis earns its place when it changes a design. The loop is simple. Measure the current situation, propose a move, measure again and compare. Because the toy city lives in a script, a move can be tested in seconds.

Worked example: one new footbridge

Chapter 7 showed homes south of the railway cut off from the park. Chapter 11 showed the two bridges carrying a heavy load. The move is a new footbridge in the middle of the railway, right below the park. The script added this one 100 m link to the network and ran the same calculations again.

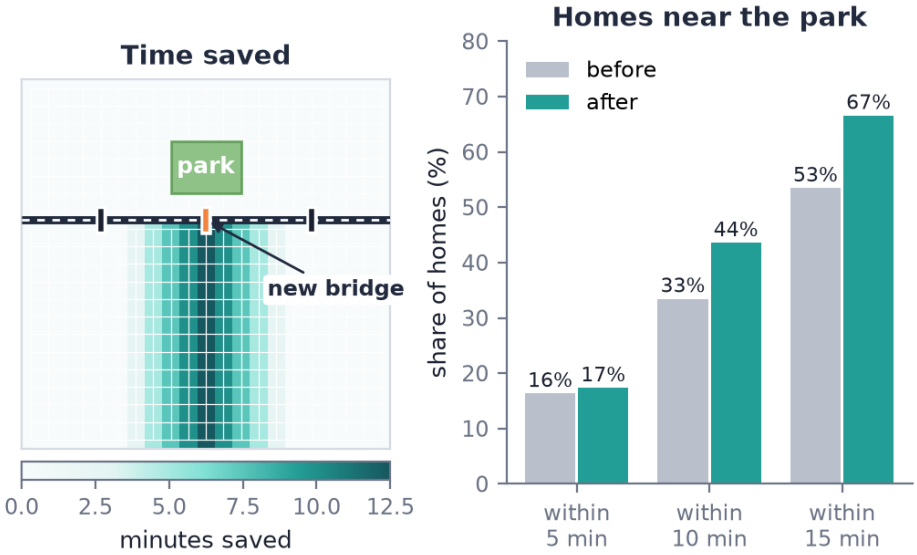


Figure 17. The effect of one new footbridge. Left: minutes saved on the walk to the park, at each street corner. Right: share of homes within 5, 10 and 15 minutes' walk of the park. Computed for this guide.

Measure	Before	After
Homes within a 10-minute walk of the park	33%	44%
Homes within a 15-minute walk of the park	53%	67%
Average walk to the park, south of the railway	18.1 min	15.2 min
Walk from the starred corner (Figure 7)	1,200 m	200 m
Share of shortest trips over each old bridge	23.6%	14.0%
Share of shortest trips over the new bridge	none	19.3%
Average shortest trip between two corners	1,474 m	1,428 m

One short structure lifts the share of homes within a 10-minute walk from 33% to 44% and takes about 41% of the load off each old bridge. The new bridge at once becomes one of the busiest links in the city, so it deserves width, light and things to see at both ends.

A menu of moves

If the analysis shows	Consider
Long detours, high detour ratios	New links: crossings, cut-throughs, permeable blocks
A street with very high betweenness or choice	Wide pavements, safe crossings, shop fronts, calmer traffic
Low land-use mix	Ground-floor shops and services; a school or clinic
A shop hot spot far from most homes	More homes nearby, or a second centre elsewhere
A heat hot spot	Shade trees, lighter surfaces, water, open paths for air
Small, hemmed-in isovists at a key entrance	Open sight lines, move walls, add a landmark
High FAR with little open space	Shared courtyards, roof gardens, a nearby pocket park

SAY IT LIKE THIS

We tested it. One new bridge raises the share of homes within ten minutes of the park from 33% to 44%.

IN ONE LINE

Measure, change, measure again: the before-and-after comparison turns analysis into a design argument.

15

Pitfalls and Honest Limits

Maps persuade, so check them before they persuade you.

A map looks authoritative. That is exactly why it needs care. These traps catch beginners and experts alike.

Colour classes change the story

Most maps group values into a few colour classes. **Equal intervals** split the range of values into equal steps. **Quantiles** put the same number of areas in each class. Figure 18 shows the same shop density both ways.

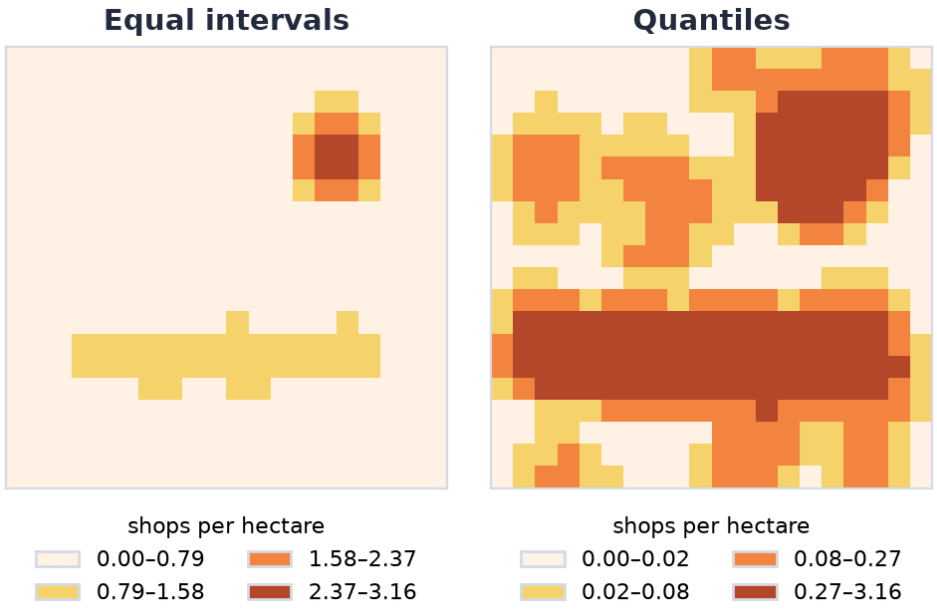


Figure 18. The same kernel density of shops (300 m bandwidth, 100 m cells) in four classes. Left: equal intervals. Right: quantiles. Computed for this guide.

With equal intervals only 1% of cells land in the top class and 87% in the bottom one, so the map looks quiet. With quantiles a quarter of the cells land in each class by design, and the whole city looks busy. Both are correct. Readers will still draw different conclusions, so say which method you used and why.

More traps

Pitfall	What goes wrong	A habit that helps
Mixed or wrong CRS	Layers misalign; distances come out in degrees	Check the CRS of every layer; measure in metres
Straight-line distance	Access looks better than it is	Use network distance for walking
Zones (MAUP) and the ecological fallacy	Results depend on the zones; zone averages are read as facts about individuals	Try two zonings; avoid claims about individuals
Edge effects	Places near the study boundary look emptier or less central	Include a margin around the study area
Missing or biased data	Unmapped places look empty	Check coverage; ask who is missing
False precision	Three decimals on rough data	Round sensibly; show uncertainty
Many tests at once	Even with no real pattern, about one test in twenty passes the 95% level by chance	Use stricter thresholds; distrust isolated hot cells
Pattern read as cause	Shops on busy streets: which came first?	Treat patterns as clues to test

WATCH OUT

Measures of movement such as betweenness and space syntax describe potential. Real flows also depend on land use, public transport, safety, slope and culture. Count people on site before you lean on a prediction.

IN ONE LINE

Check the CRS, the distance type, the zones, the classes and the gaps before you trust a map, including your own.

16

Try It Yourself

A short recipe, open tools and an eight-week plan.

A first project in seven steps

1. Pick one question about a place you know, such as “How many homes are within a 10-minute walk of a park?”
2. Download the streets, buildings and parks for that area from OpenStreetMap.
3. Reproject every layer to a local metric CRS.
4. Build the walking network and compute the network distance from homes, or a grid of points, to the nearest park entrance.
5. Map the result with clear classes and a legend that states the walking speed.
6. Propose one design move, such as a new path or a pocket park, and compute again.
7. Write half a page: question, data, method, result and limits.

Open tools

Tool	What it is	Good for
QGIS	Free, open-source desktop GIS	Viewing, editing, buffers, overlay and map making; many plugins
GeoPandas	Python library for vector data	Scripted overlay, joins and reprojection
PostGIS	Spatial extension for the PostgreSQL database	Large datasets and spatial queries in SQL
OSMnx	Python package for street networks from OpenStreetMap	Downloading networks, network distance, centrality and street statistics
PySAL	Python spatial analysis library	Spatial weights, Moran’s I, LISA and hot spots
depthmapX	Open-source space syntax software	Axial and segment analysis, isovists and VGA
momepy	Python library for urban morphology	Building, plot and block measures
R with sf, spdep and terra	Spatial packages for the R language	Vector data, spatial statistics and rasters
GDAL	Open-source library and command-line tools	Converting and reprojecting almost any format

TIP

Start small: one neighbourhood, one question, one evening. A finished small study teaches more than a half-finished big one.

An eight-week plan

Week	Focus	Do this
1	Coordinates and maps	Install QGIS. Load OpenStreetMap data for your area. Find its CRS and reproject it to metres.
2	Vector and raster	Measure a few building areas. Rasterise them at two cell sizes and compare.
3	Buffers and overlay	Count homes within 400 m of each stop. Overlay a hazard or green-space layer.
4	Networks and access	Compute walking distance to the nearest park. Compare it with straight-line distance.
5	Density and mix	Compute dwellings per hectare and an entropy index for a few districts.
6	Patterns	Make a kernel density map of shops. Compute Moran’s I and a hot-spot map.
7	Centrality and visibility	Compute betweenness or choice for your streets. Cast isovists from three key points.
8	Design and report	Test one design move. Write up before and after, with limits.

IN ONE LINE

Pick a question, measure along real routes, test a design move and write down your limits.

Myths and Facts

Common misunderstandings, cleared up.

Myth	Fact
A map is an objective picture of reality.	Every map involves choices: what to include, which projection, which zones and which colour classes. Good maps state their choices.
Straight-line distance is good enough for walking.	Barriers and street layout can make real walks much longer. In the toy city, a corner 200 m from the park was a 1,200 m walk.
Latitude and longitude are fine for measuring.	Degrees are angles, and a degree of longitude shrinks towards the poles. Measure in a projected CRS with metres.
More detailed data always gives better answers.	Detail helps only when it suits the question. Results can change with scale, so test more than one.
A hot spot proves a cause.	A hot spot shows where values cluster. Explaining why takes other evidence.
Space syntax predicts exactly how many people will use a street.	It estimates movement potential from layout. Real counts also depend on land use, transport, safety and more, so check them on site.

Quick Quiz

Ten questions. Answers are on the next page.

1. Put Tobler’s first law into your own words.
2. Why should you not measure distance in degrees of longitude?
3. What is the difference between vector and raster data?
4. What does the modifiable areal unit problem say?
5. A home is 200 m from a park as the crow flies. Why might it still be a 15-minute walk?
6. What does an entropy index of 0 mean? And 1?
7. What does the bandwidth control in kernel density?
8. Moran’s I is close to −1. What pattern does that suggest?
9. What is the difference between integration and choice in space syntax?
10. Two plots have the same floor area ratio. Must they look alike?

Quiz Answers

No peeking until you have tried.

1. Near things tend to be more alike than distant things.
2. A degree of longitude covers less ground as you move away from the equator: about 111 km at the equator and 56 km at 60°.
3. Vector stores points, lines and polygons with attributes. Raster stores a grid of cells, each holding one value.
4. Results based on zones change when the boundaries or sizes of the zones change.
5. Barriers such as a railway or a river, and the street layout, can force long detours. Network distance captures this.
6. 0 means a single land use. 1 means all the uses have equal shares.
7. How far each point’s influence spreads. A small bandwidth shows local detail; a large one smooths it away.
8. A checkerboard-like pattern, where neighbours are unlike each other (negative spatial autocorrelation).
9. Integration measures how easily a line is reached from all others, in few steps. Choice measures how often it lies on the shortest paths between other lines.
10. No. The same floor area can come as a low building with high coverage or as tall towers with lots of open ground.

Glossary

Plain definitions of the words in this guide.

Term	Meaning
Accessibility	How easily people can reach what they need.
Axial map	The fewest, longest lines covering all open space.
Bandwidth	The reach of each point’s hill in kernel density.
Betweenness	How often a place lies on shortest paths between others.
Buffer	The zone within a set distance of a feature.
Choice	Space syntax name for betweenness on an axial map.
Closeness	How near a place is, on average, to all others.
CRS	Coordinate reference system; ties numbers to places.
Ecological fallacy	Reading a zone average as a fact about individuals.
Entropy index	Land-use mix score from 0 (one use) to 1 (even mix).
FAR	Floor area ratio, total floor area ÷ site area.
Getis-Ord Gi*	A local statistic that flags hot and cold spots.
GIS	Geographic information system; maps and analyses spatial data.
IDW	Inverse distance weighting; nearer values count more.
Integration	Space syntax measure of how few steps away a line is.
Isochrone	A line joining places with equal travel time.
Isovist	The area visible from one point.
Kernel density (KDE)	A smooth surface built from points.
LISA	Local indicators of spatial association.
MAUP	Modifiable areal unit problem; zones change results.
Moran’s I	One number for spatial autocorrelation on a map.
Projection	A method for flattening the Earth onto a plane.
Raster	A grid of equal cells, each holding one value.
Spatial autocorrelation	Nearby values being alike, or unlike, more than chance.
Vector	Points, lines and polygons with attributes.

One-Page Cheat Sheet

Pin it next to your screen.

Idea	In short
Tobler’s law	Near things are more alike. Location is evidence.
CRS	Store in degrees if you must; measure in a local projected CRS.
Vector or raster	Shapes for objects, grids for surfaces.
Data	Ask who made it, when, how, why, and what is missing.
MAUP	Zones are choices. Test two zonings and two scales.
Buffer, overlay, select	Draw the zone, stack the layers, keep what passes.
Access	Use network distance and minutes. Watch the detour ratio.
Density and mix	Say net or gross. Entropy runs from 0 to 1.
Clusters	KDE to see; Moran’s I and Gi* to test. Bandwidth matters.
Interpolation	IDW or kriging; check with leave-one-out.
Centrality	Closeness for reach, betweenness or choice for passing movement.
Isovists	Area and longest view from where people stand.
Urban form	Coverage, FAR, storeys, open space, block size, junctions.
Design loop	Measure, change, measure again, and report the limits.

Further Reading

Where the ideas come from.

Later editions and reprints may differ in pages and publisher; check the version you use before citing.

- de Smith, M. J., Goodchild, M. F. and Longley, P. A. *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools*. Several editions; also free online.
- Longley, P. A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W. (2015). *Geographic Information Science and Systems*, 4th edition. Wiley.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46, 234–240.
- Openshaw, S. (1984). *The Modifiable Areal Unit Problem*. Concepts and Techniques in Modern Geography 38. Geo Books.
- Anselin, L. (1995). Local indicators of spatial association: LISA. *Geographical Analysis*, 27(2), 93–115.
- Getis, A. and Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24(3), 189–206.
- Hillier, B. and Hanson, J. (1984). *The Social Logic of Space*. Cambridge University Press.
- Benedikt, M. L. (1979). To take hold of space: isovists and isovist fields. *Environment and Planning B*, 6, 47–65.
- Turner, A., Doxa, M., O’Sullivan, D. and Penn, A. (2001). From isovists to visibility graphs: a methodology for the analysis of architectural space. *Environment and Planning B: Planning and Design*, 28, 103–121.
- Lynch, K. (1960). *The Image of the City*. MIT Press.
- Boeing, G. (2017). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65, 126–139.
- Berghauser Pont, M. and Haupt, P. (2010). *Spacematrix: Space, Density and Urban Form*. NAi Publishers.
- Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2), 17–23.