TU Delft

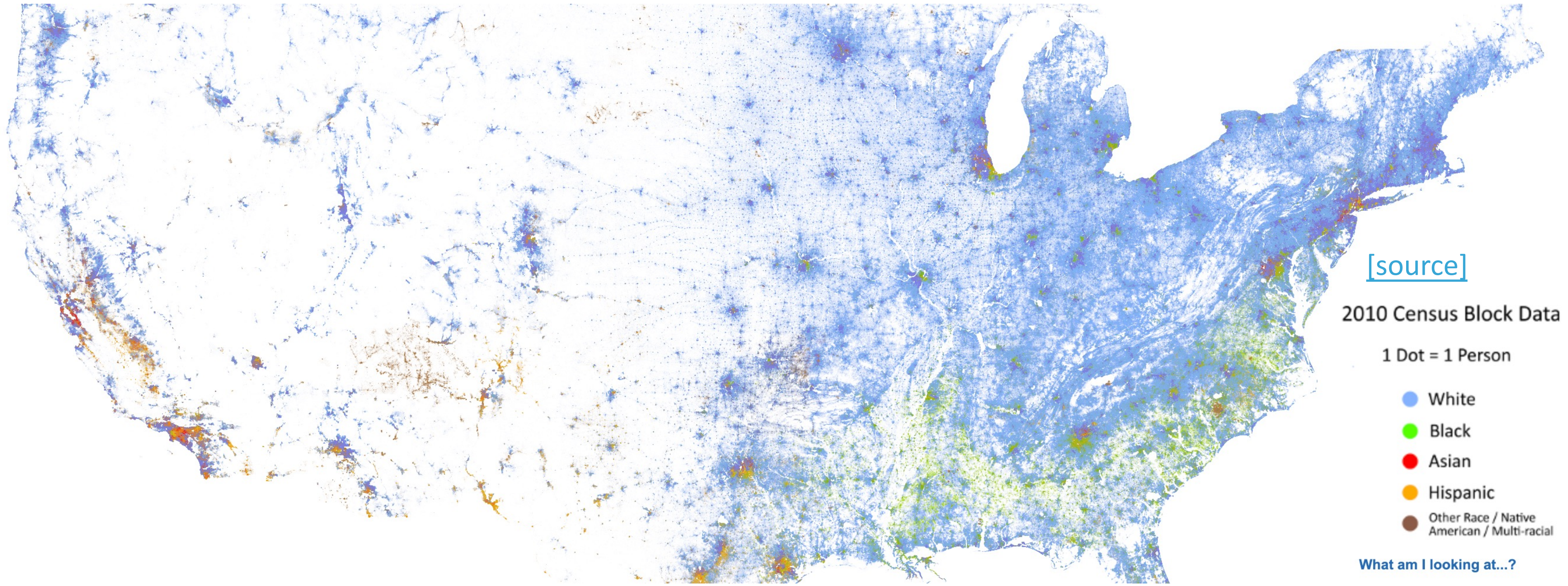
Spatial Data Science

Networks

& Space

(EPA 122A)

Lecture 7



Trivik Verma

The Racial Dot Map developed by the Demographics Research Group at University of Virginia

Last Time

- Geo-Visualisation
- Dangers of Geo-Vis
- Mapping Data
  - MAUP
  - Choropleths

Today

- Introduction to Networks
- The need to represent space formally
- Spatial weights matrices
  - What
  - Why
  - Types
- The spatial lag

Why are we studying ***Networks?***

Wide Range of Applications

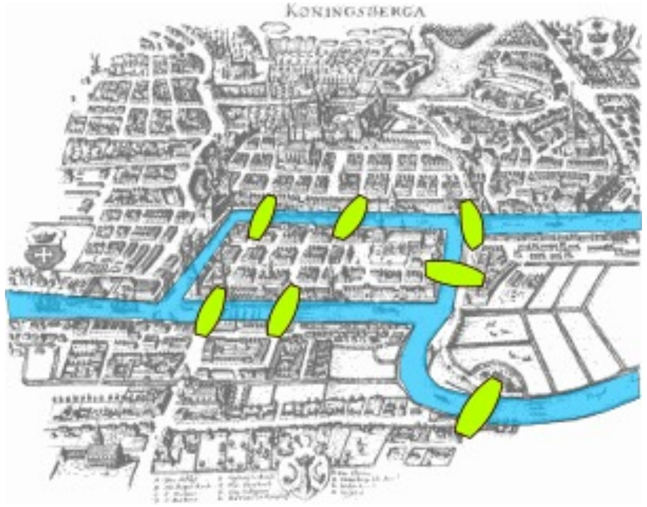
- Operations research: urban traffic, airport configuration, assembly lines …

- Physical systems / processes: PV array, battery, hydraulics, aerodynamics …

- Others: crowd behavior, family formation, wildfire spread, disease spread …

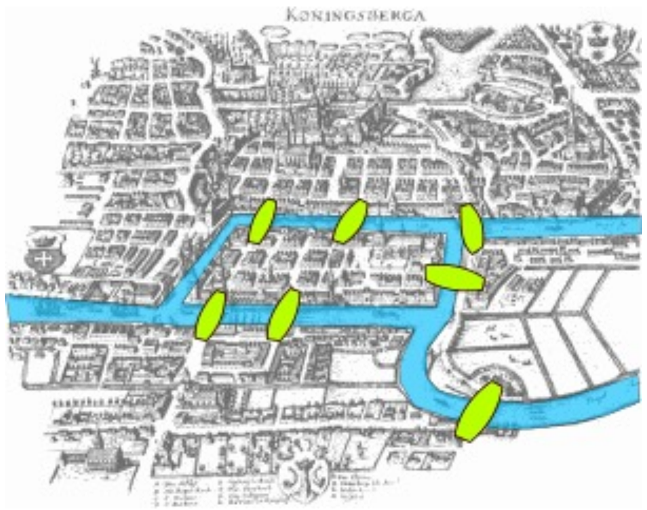
Introduction to Networks

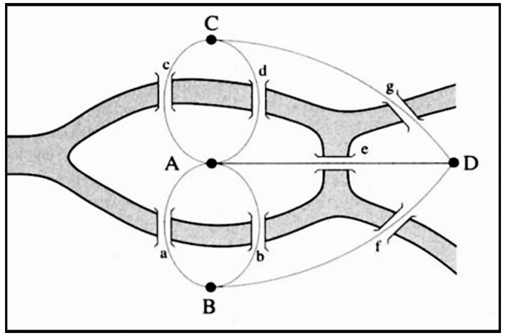
- Pattern of relationships/connections among a set of “components”
- Growing public fascination with connectedness of modern society
- Early days : Euler’s 1735 solution of the Königsberg bridge problem

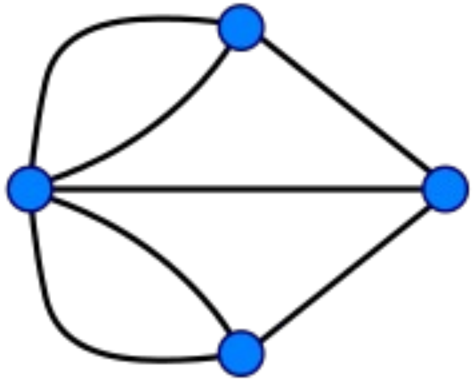


Can you cross each bridge exactly once in a walk?

Can you cross each bridge exactly once in a walk?

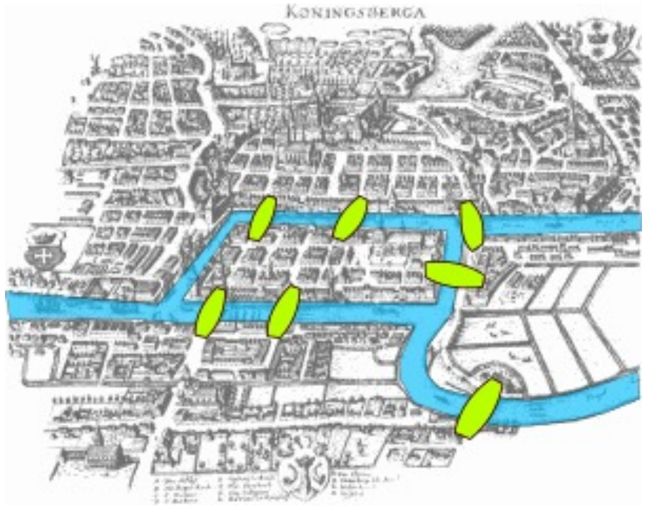






Introduction to Networks

- Pattern of relationships/connections among a set of “components”
- Growing public fascination with connectedness of modern society
- Early days : Euler’s 1735 solution of the Königsberg bridge problem



Can you cross each bridge exactly once in a walk?

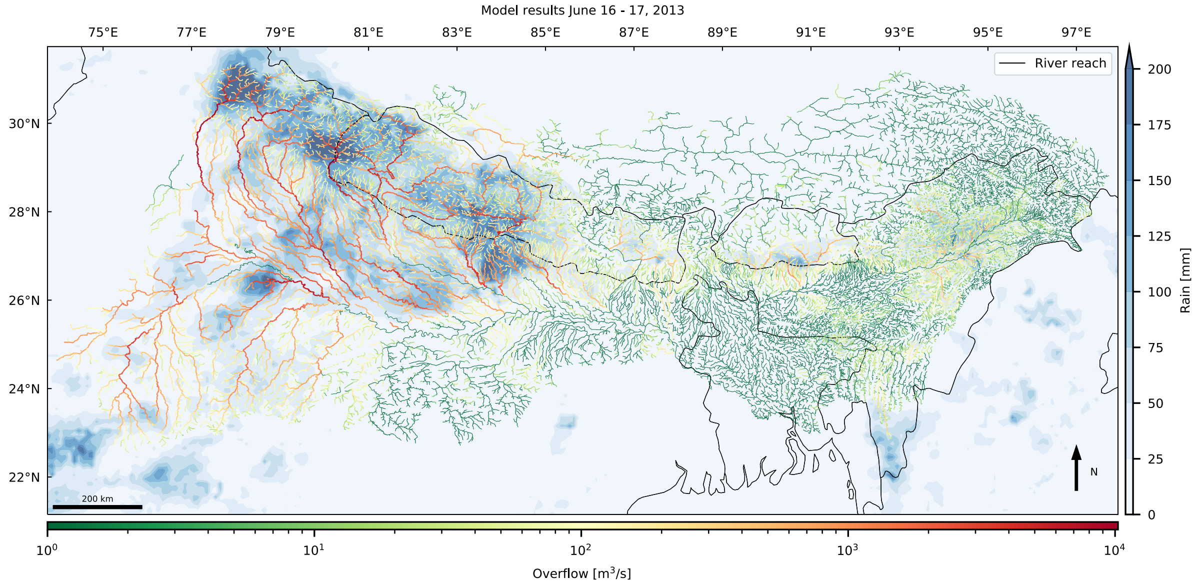
- Complex system modelled as a mathematical network (graph)
- Level of abstraction to understand systems (powerful across disciplines)

History of Networks

- Network-based analysis has a long-standing history
  - Study of Königsberg bridges (Euler, 1736)
  - Laws of electrical circuitry (Kirchoff, 1845)
  - Molecular structure in chemistry (Cayley, 1874)
  - Power grids (1910), telecommunications and the Internet (1960)
  - Complex power grids (Me, 2012)

History of Networks (Examples)

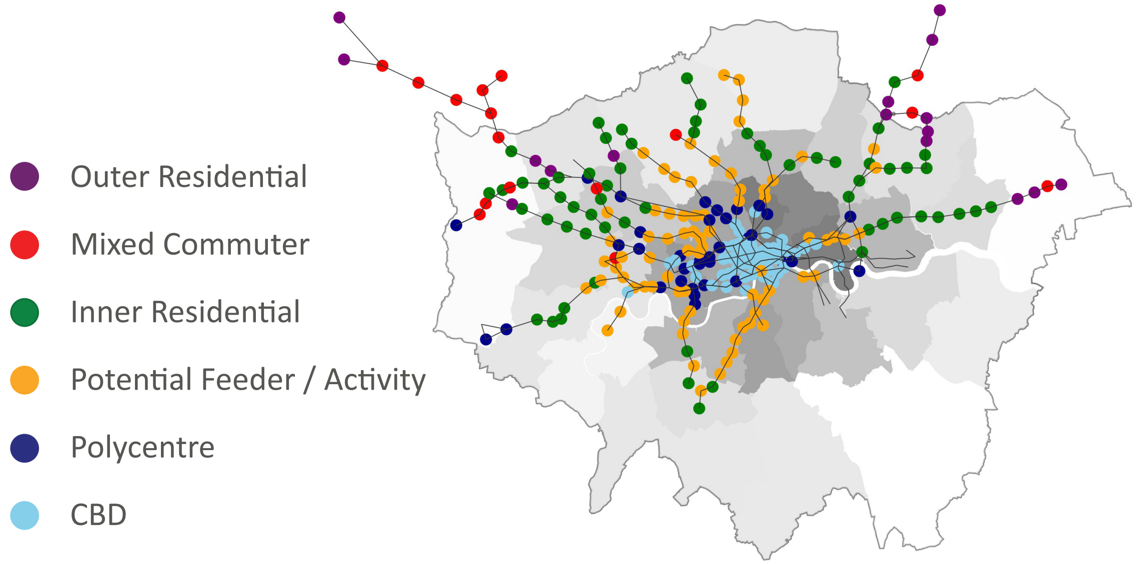
- Network-based analysis has a long-standing history
  - A multi-model simulation of river flooding in Northern India



van Meurs, B. (2019). Riverine flood risk screening with a simple network-based approach: A proof of concept in the Ganghes-Brahmaputra basin.

History of Networks (Examples)

- Network-based analysis has a long-standing history
  - No Simulation, but Network helps to understand the structure of a city
  - Benefits of simulating traffic data on this network?

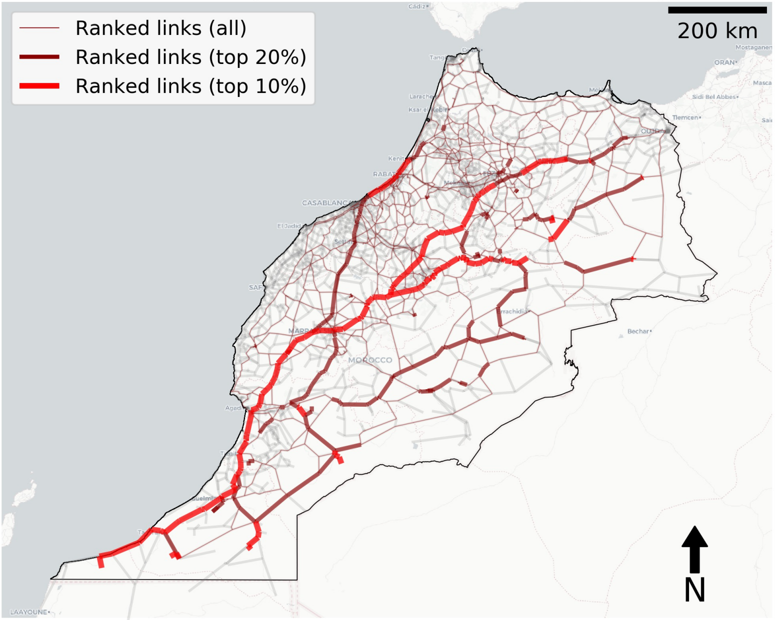


Verma, T., Sirenko, M., Kornecki, I., Cunningham, S., & Araújo, N. A. (2020). Extracting spatiotemporal demand for public transit from mobility data. *arXiv preprint arXiv:2006.03351*.

History of Networks (Examples)

- Network-based analysis has a long-standing history
  - Criticality of elements in a network changes depending on the underlying measure, or the underlying values behind the measure

Morocco: Population equalization principle



Yap, J.R., Jafino, Bramka., Verma, T. (2021). The Differentiating Effect of Choice of Moral Principles on National Road Transport Network Criticality Analysis. *In Review*.

History of Networks

- Network-based analysis has a long-standing history
  - Study of Königsberg bridges (Euler, 1736)
  - Laws of electrical circuitry (Kirchoff, 1845)
  - Molecular structure in chemistry (Cayley, 1874)
  - Power grids (1910), telecommunications and the Internet (1960)

- Interest exploded in the last two decades
  - Systems thinking in science
  - Data and Computation
  - Social and Physical Globalisation

Why are we studying ***Networks****?*

Usual trade-off between losing details in an idealized representation while gaining insights into the simplified problem

- Simple representations of complex systems in society
- Derive properties mathematically, computationally and analytically (systems thinking)
- Prediction of properties and outcomes
- Understanding common features of different networks

Introduction to Networks

Application Examples (Urban Data Science)

Routes and driving time

- Which is the fastest way to …?

Accessibility of objects

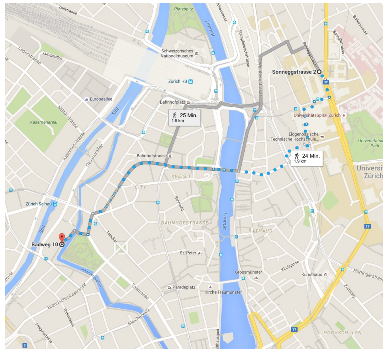
- Which facility is closest?

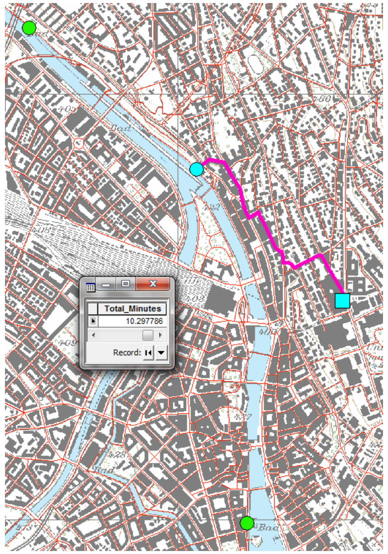
Vehicle-routing-problems

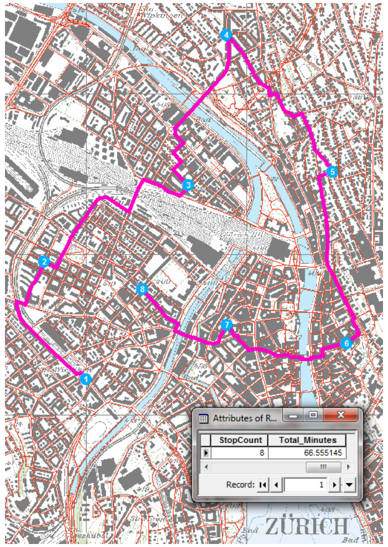
- Which sequence of stops is most effective?

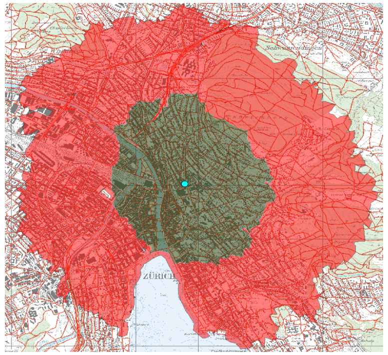
Accessibility of Zones

- What is the accessible area in a given time?









Introduction to Networks

Application Examples

Routes and driving time

- Which is the fastest way to …?

Accessibility of objects

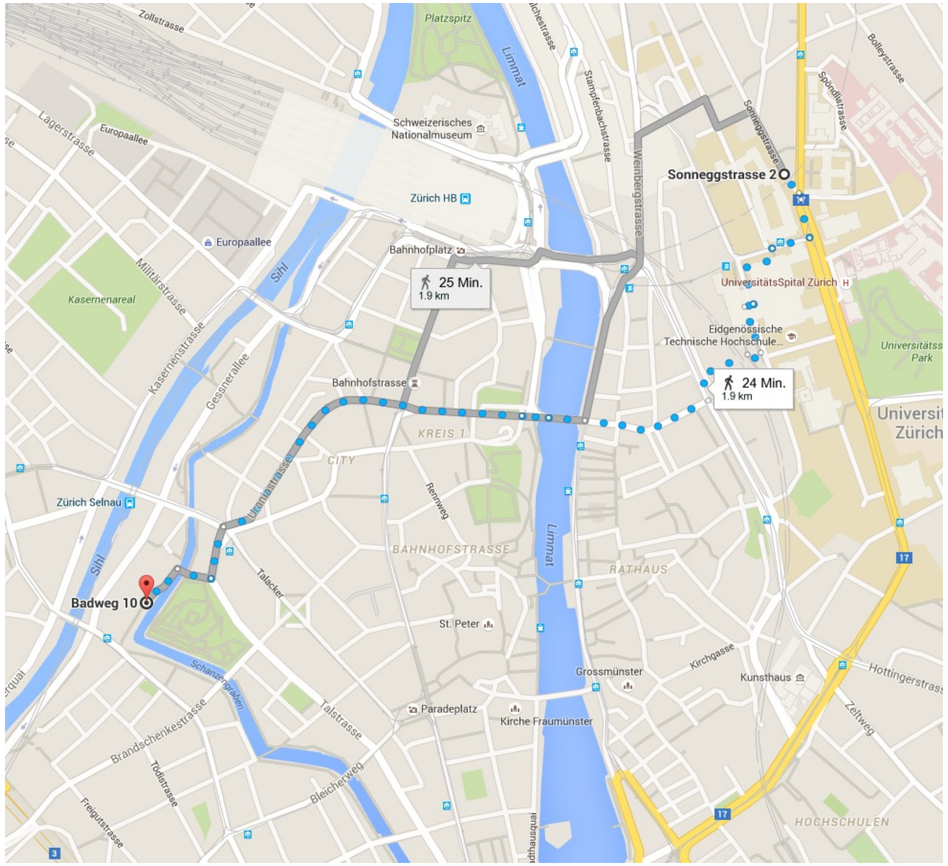
- Which facility is closest?

Vehicle-routing-problems

- Which sequence of stops is most effective?

Accessibility of Zones

- What is the accessible area in a given time?



© 2016 Google

Introduction to Networks

Application Examples

Routes and driving time

- Which is the fastest way to …?

Accessibility of objects

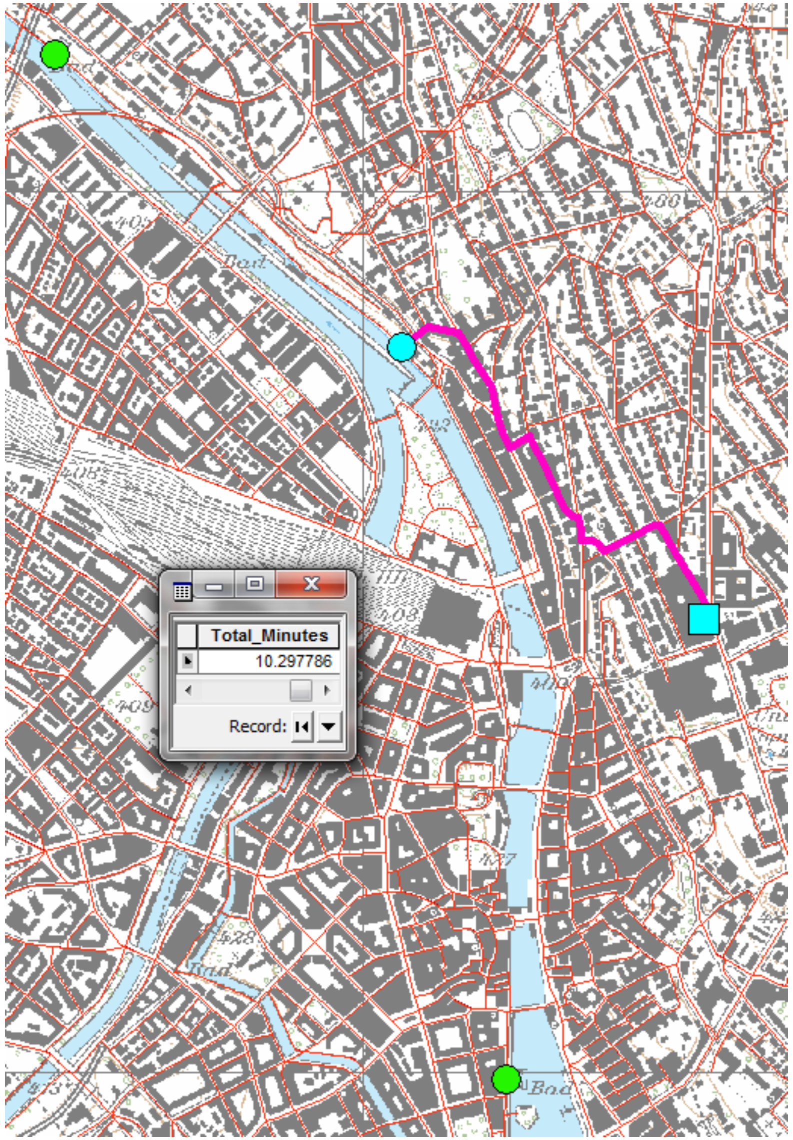
- Which facility is closest?

Vehicle-routing-problems

- Which sequence of stops is most effective?

Accessibility of Zones

- What is the accessible area in a given time?



Introduction to Networks

Application Examples

Routes and driving time

- Which is the fastest way to …?

Accessibility of objects

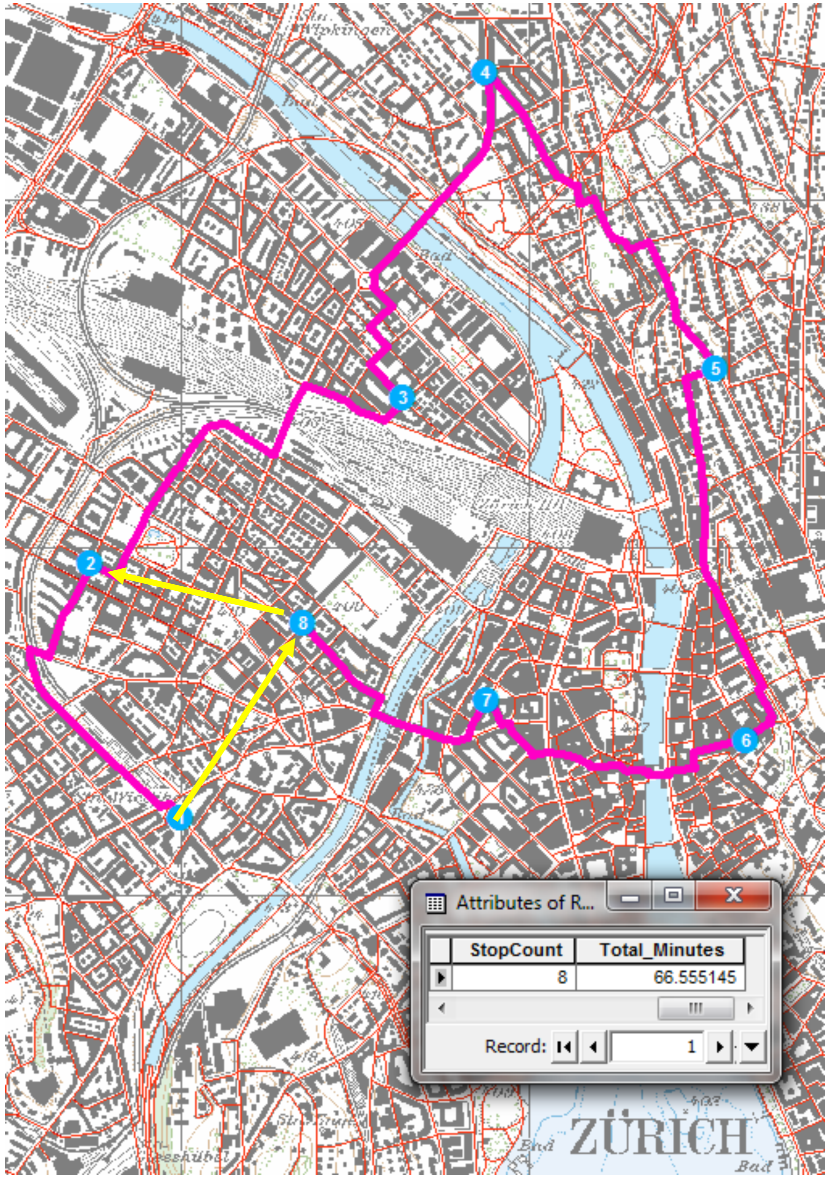
- Which facility is closest?

Vehicle-routing-problems

- Which sequence of stops is most effective?

Accessibility of Zones

- What is the accessible area in a given time?



Introduction to Networks

Application Examples

Routes and driving time

- Which is the fastest way to …?

Accessibility of objects

- Which facility is closest?

Vehicle-routing-problems

- Which sequence of stops is most effective?

Accessibility of Zones

- What is the accessible area in a given time?

© 2016 swisstopo (JA100120, JD100042) © 2016 Esri, ArcGIS

Overview of Networks

What is a network?

- Definitions
- Elements, representation
- Types of graphs
- Structural properties (not part of this course)

Definitions

- **Network**:

  *"A geometric-topological arrangement of **nodes** and **edges**, e.g. in the form of a **graph** […]"*

Definitions

- **Network**:

  "*A geometric-topological arrangement of **nodes** and **edges**, e.g. in the form of a **graph** […]*"

- **Network analysis**:

  "*A basic group of analysis functions […] based on **line-like phenomena** to calculate and determine relations […].*

Definitions

- **Network**:

  *"A geometric-topological arrangement of **nodes** and **edges**, e.g. in the form of a **graph** […]"*

- **Network analysis**:

  *"A basic group of analysis functions […] based on **line-like phenomena** to calculate and determine relations […].*

  This includes

  - *shortest path analysis,*
  - *searching for the nearest neighbour or the best location,*
  - *calculating a minimum spanning tree or*
  - *the solution of the travelling salesman problem.*

Definitions

- **Network**:

  *"A geometric-topological arrangement of **nodes** and **edges**, e.g. in the form of a **graph** […]"*

- **Network analysis**:

  *"A basic group of analysis functions […] based on **line-like phenomena** to calculate and determine relations […].*

  This includes
  - *shortest path analysis,*
  - *searching for the nearest neighbour or the best location,*
  - *calculating a minimum spanning tree or*
  - *the solution of the travelling salesman problem.*

  *Important for this group of analysis function is the correct representation of topological relations because mathematical methods of **topology** and **graph theory** are used."*

Definitions

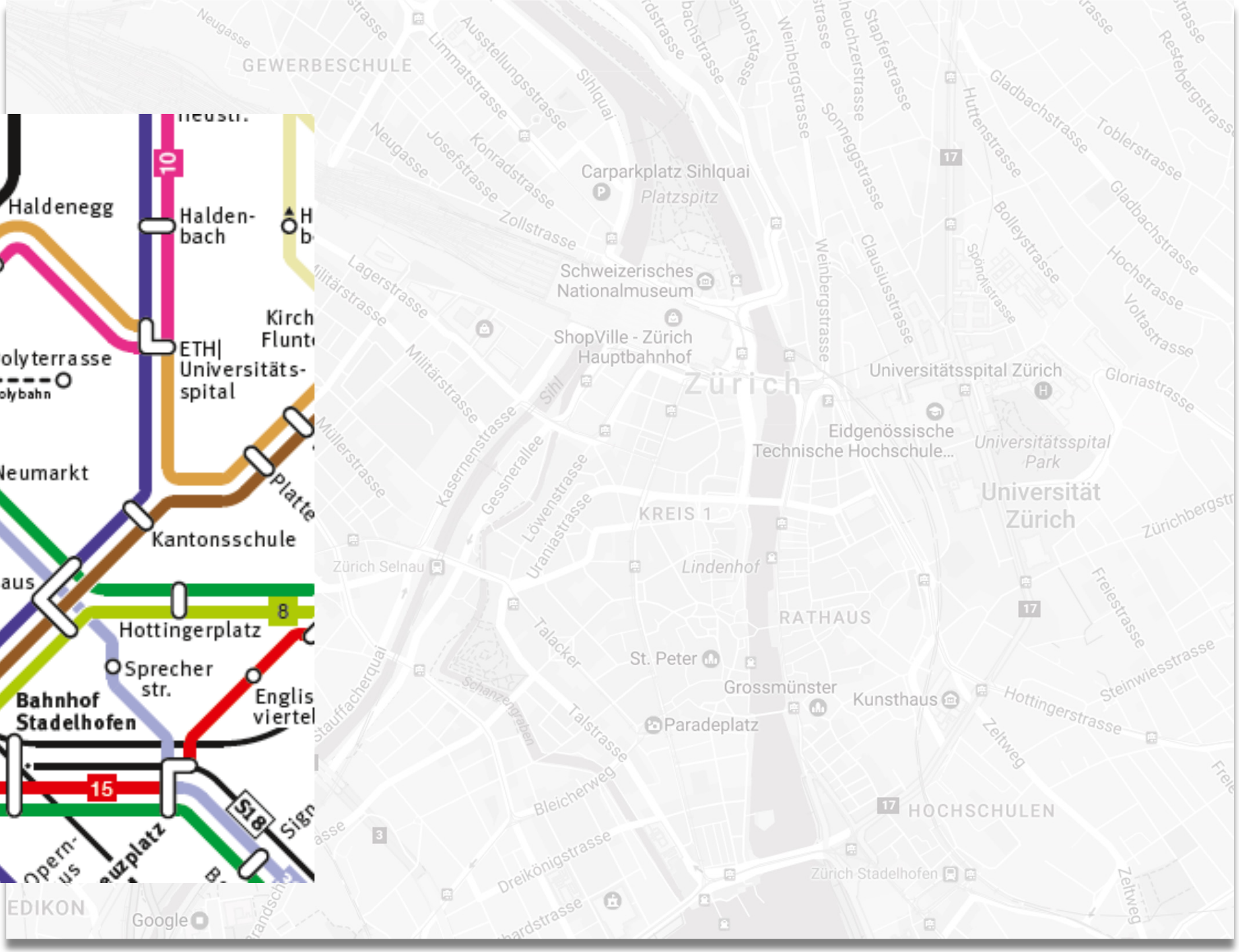
"***Topological characteristics** …*

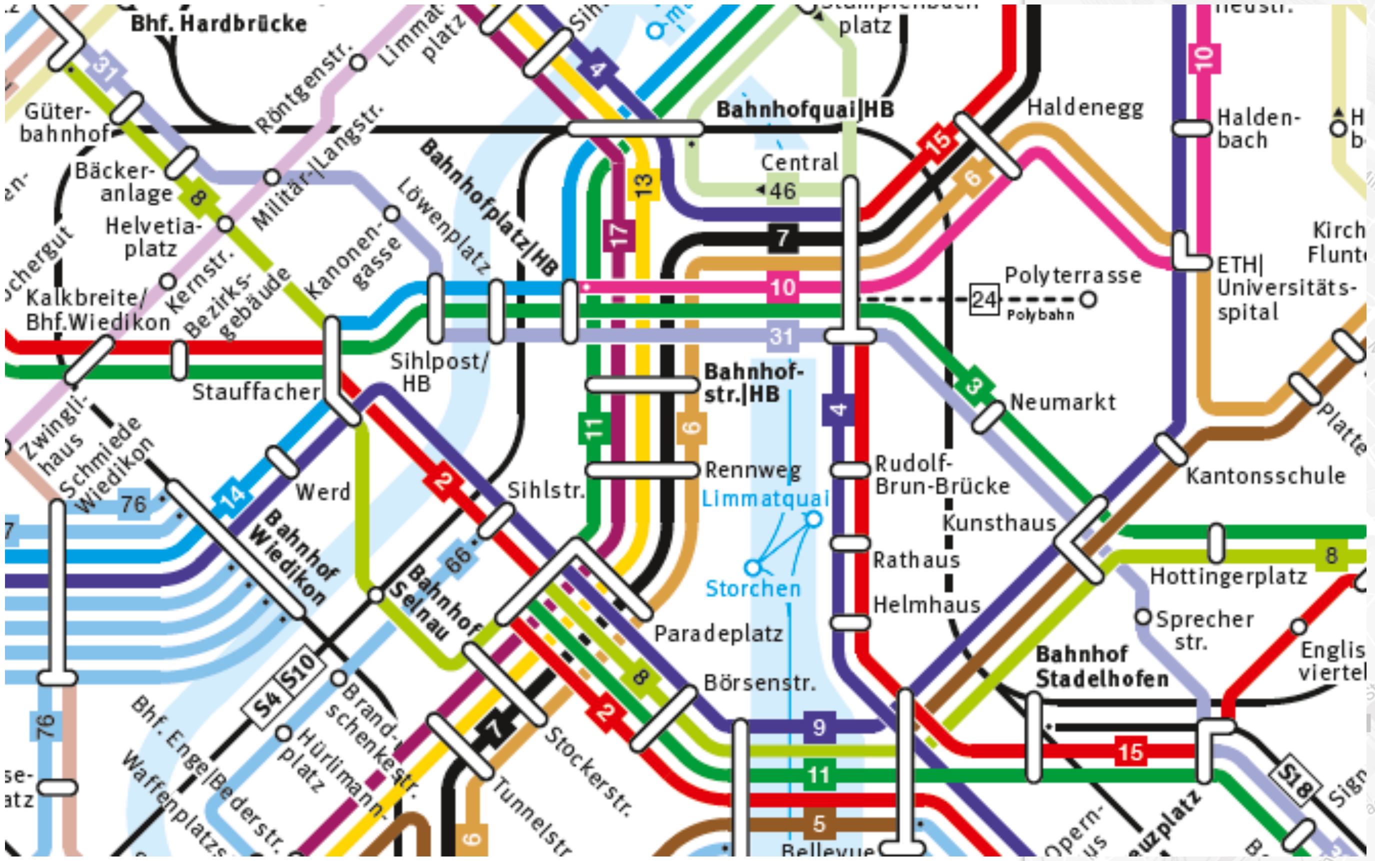
*… describe the **relative spatial relations between objects** […]. Typical topological relations are related to **adjacencies**, (e.g. if two areas are adjacent), **containedness** (e.g. if a house is located on a certain spatial unit) or the **intersection** (e.g. if two roads cross).*"

It's about **mutual positions** and **arrangement** of geometrical objects in space … not about metrical relations.

Definitions

"*Topological characteristics …*





What is a network?

Elements/representation

- Vertex, edge
- Adjacency matrix

Types

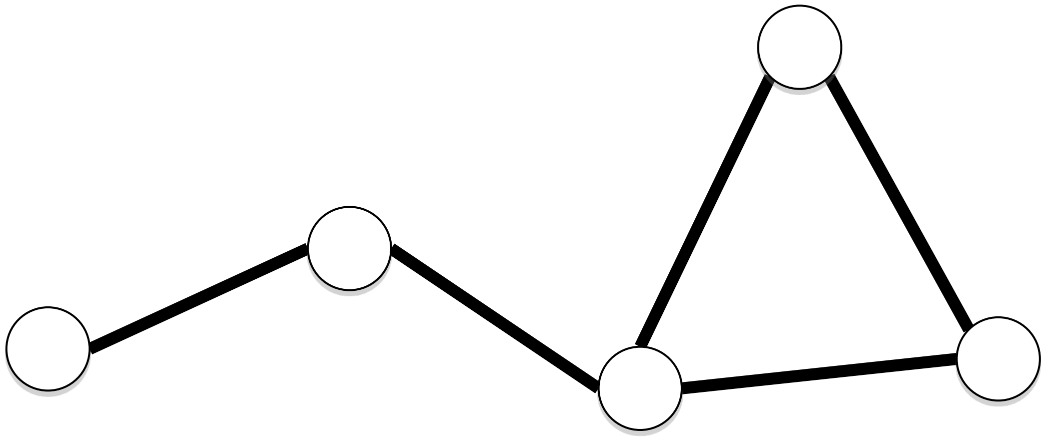
- Weighted vs. unweighted
- Directed vs. undirected
- Connected vs. disconnected
- Cyclic vs. acyclic
- Complete, tree, cubic, star

Elements/representation

Graph G = (V, E)

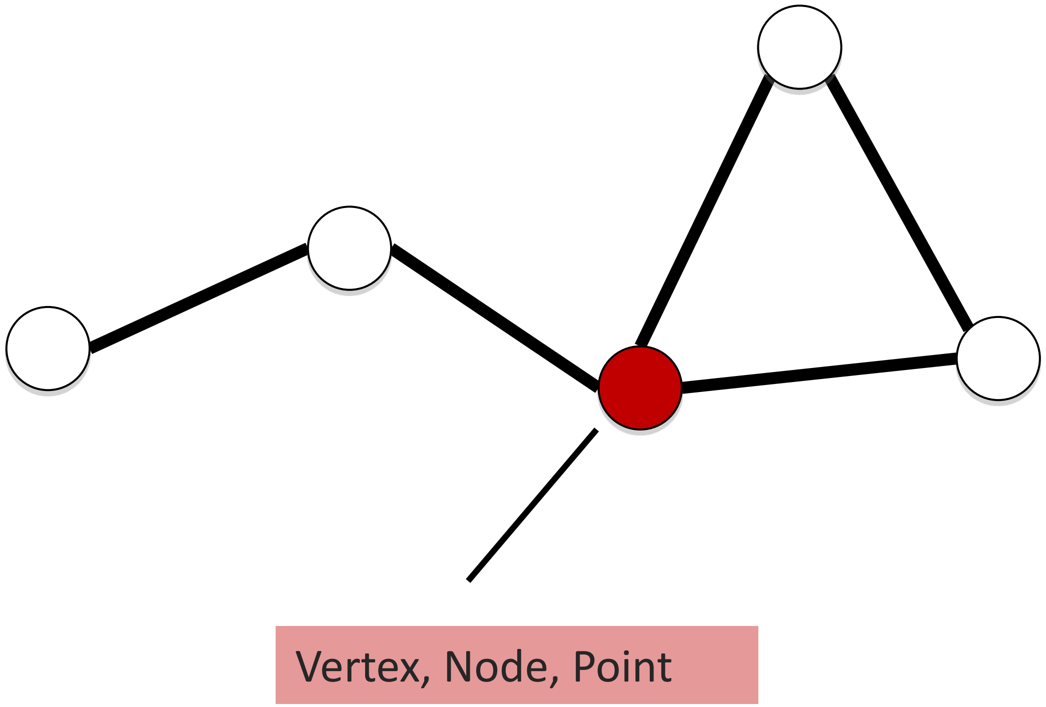
V = vertices (singular vertex)

E = edges



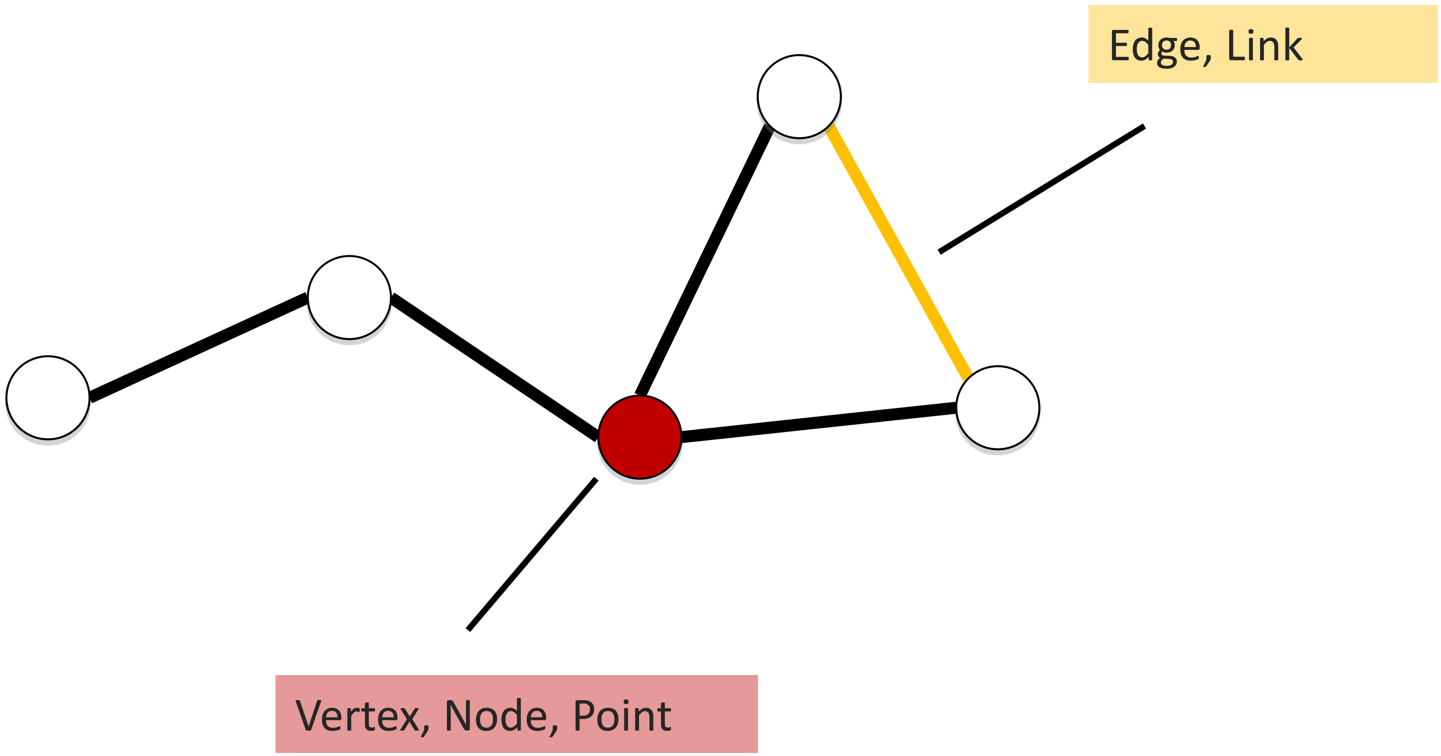
Elements/representation

Vertex V, edge E



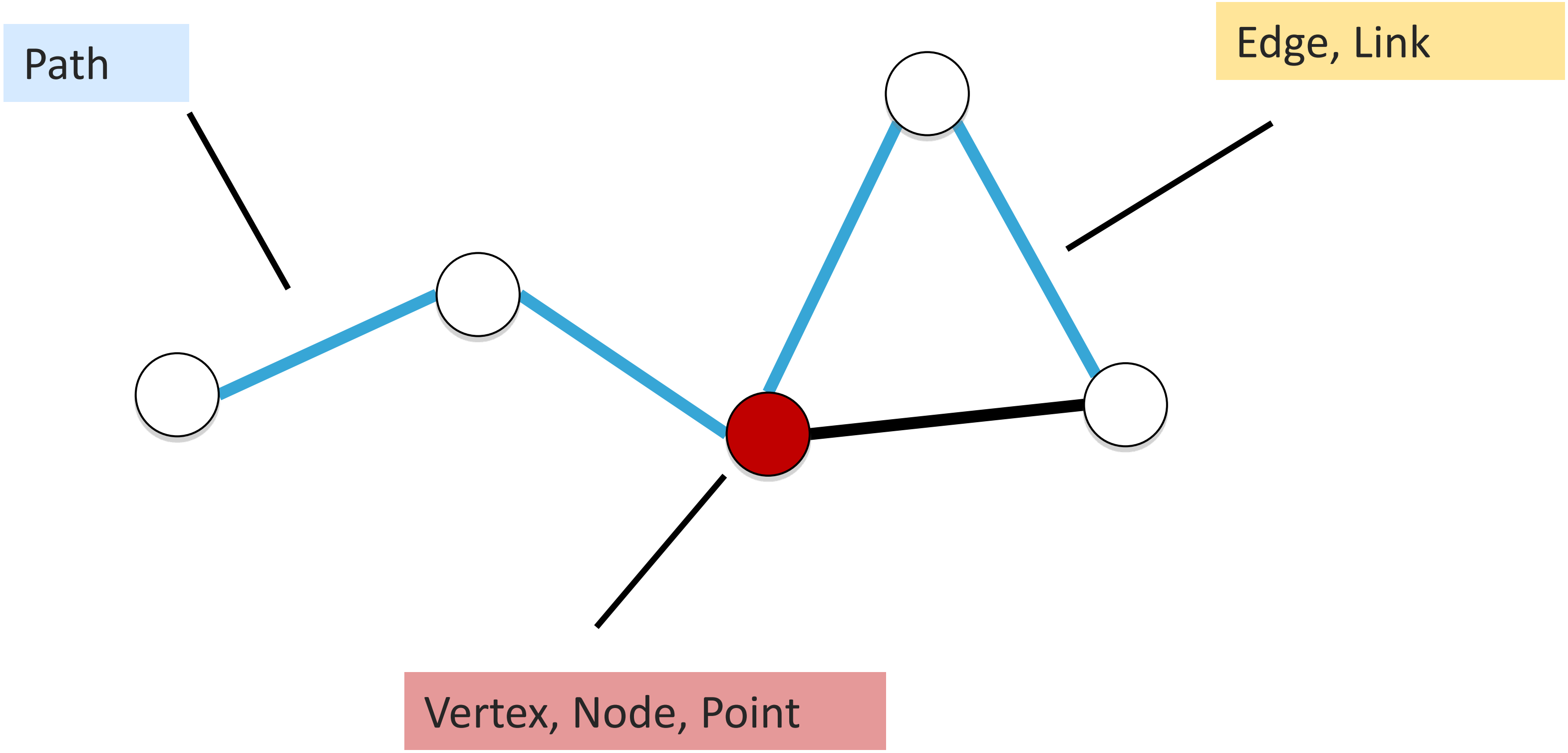
Elements/representation

Vertex V, edge E



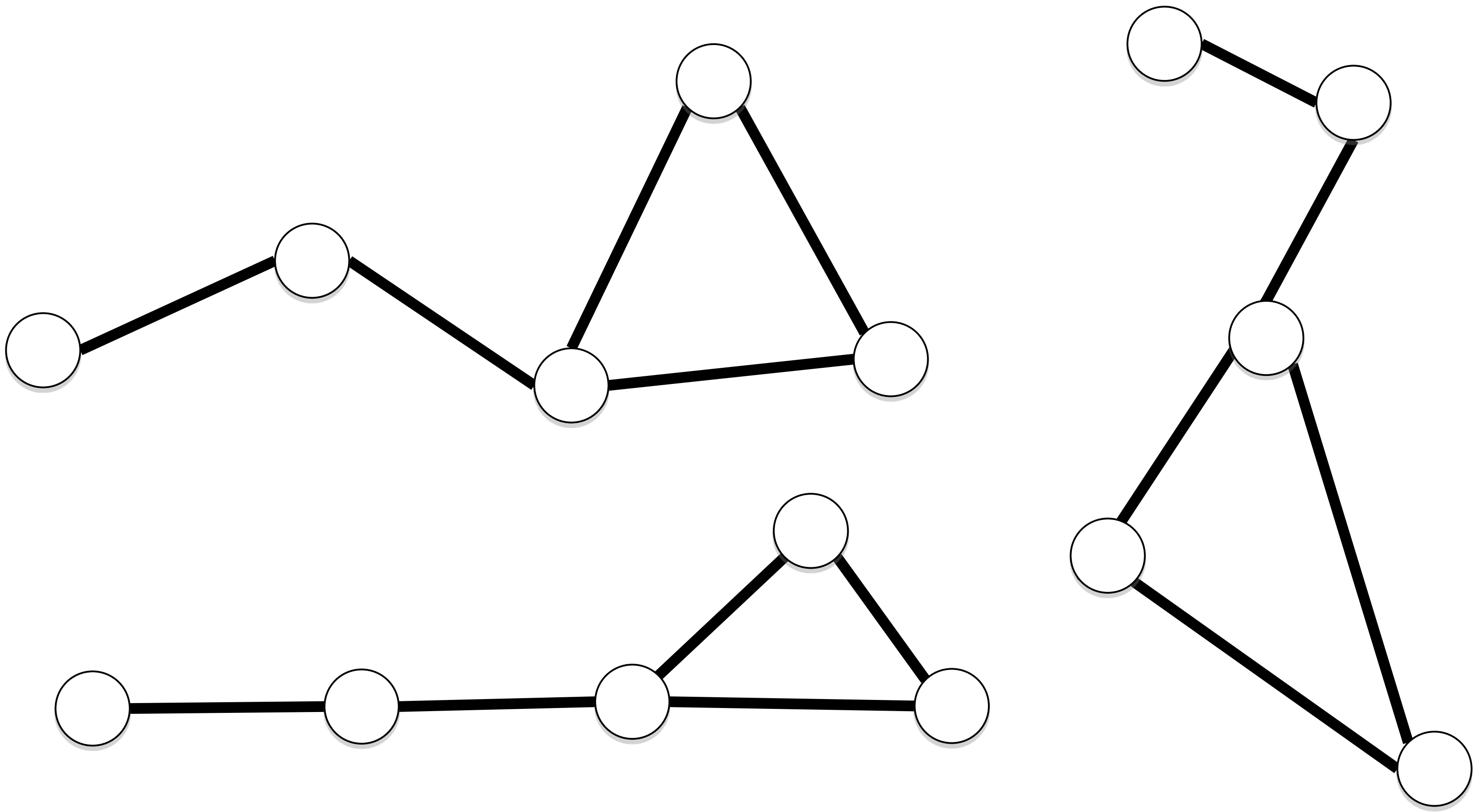
Elements/representation

Vertex V, edge E

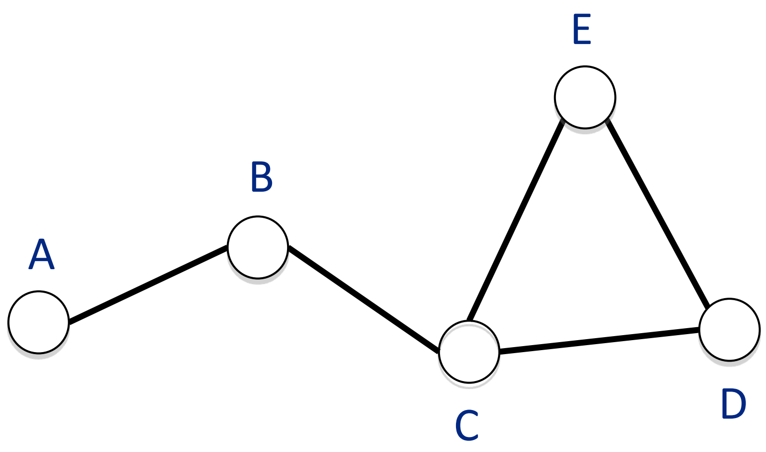


Elements/representation

Graph = independent from its visualisation



Elements/representation

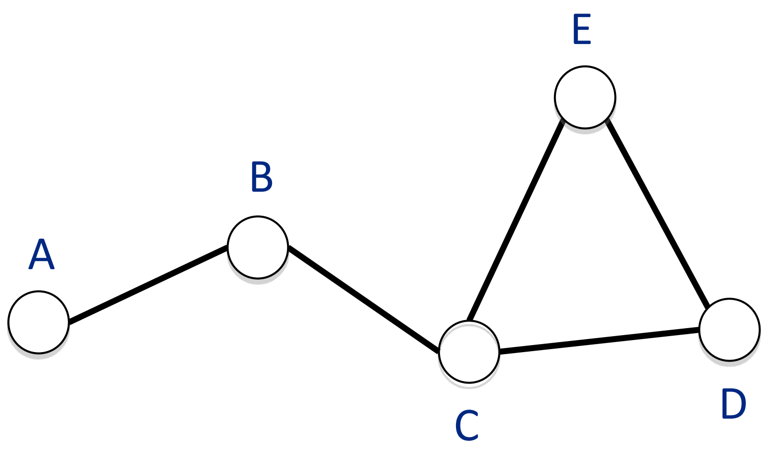


Adjacency matrix

	A	B	C	D	E
A					
B					
C					
D					
E					

TU Delft

Elements/representation

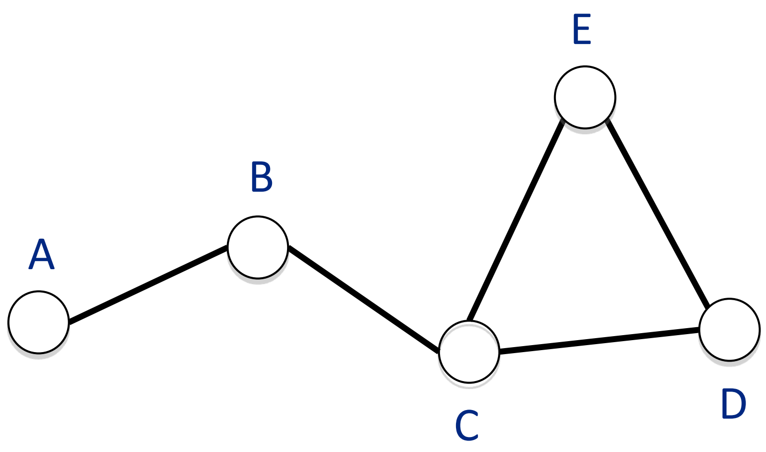


Adjacency matrix

	A	B	C	D	E
A	0	1	0	0	0
B	1	0	1	0	0
C	0	1	0	1	1
D	0	0	1	0	1
E	0	0	1	1	0

Elements/representation

Degree of a vertex:
= the number of edges connecting to a vertex

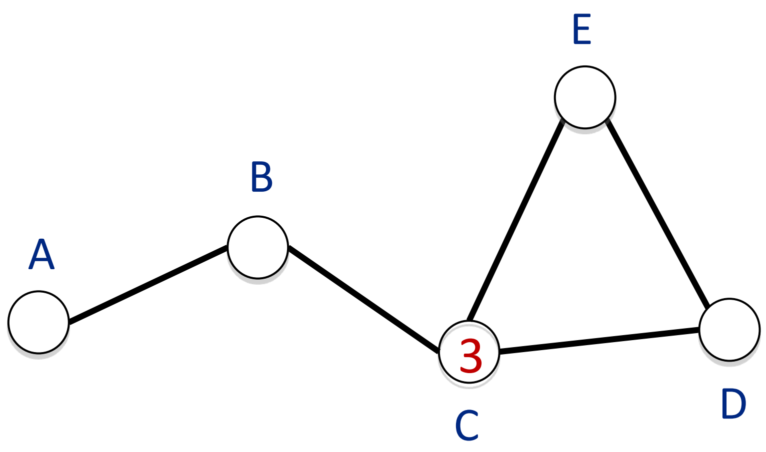


Adjacency matrix

	A	B	C	D	E
A	0	1	0	0	0
B	1	0	1	0	0
C	0	1	0	1	1
D	0	0	1	0	1
E	0	0	1	1	0

Elements/representation

Degree of a vertex:
= the number of edges connecting to a vertex



Adjacency matrix

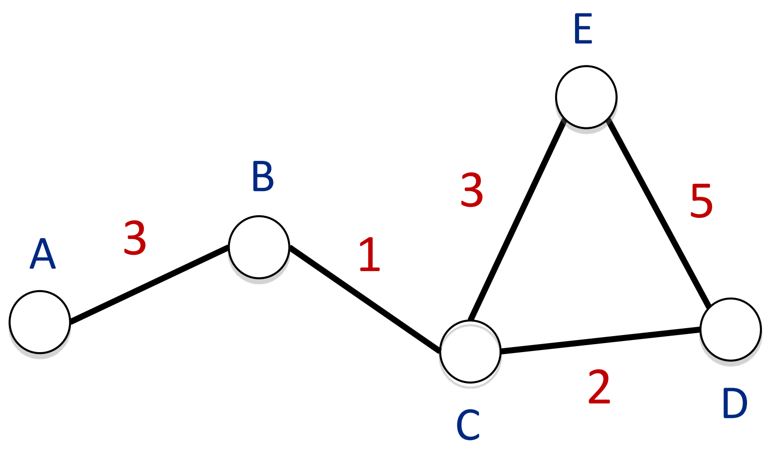
	A	B	C	D	E
A	0	1	0	0	0
B	1	0	1	0	0
C	0	1	0	1	1
D	0	0	1	0	1
E	0	0	1	1	0

3

3

Cost/Weight of Links

Weighted Degree of a node:
= the weighted sum of links connecting to a node



Adjacency matrix

	A	B	C	D	E
A	0	3	0	0	0
B	3	0	1	0	0
C	0	1	0	2	3
D	0	0	2	0	5
E	0	0	3	5	0

6

6

What is a network?

Elements/Representation

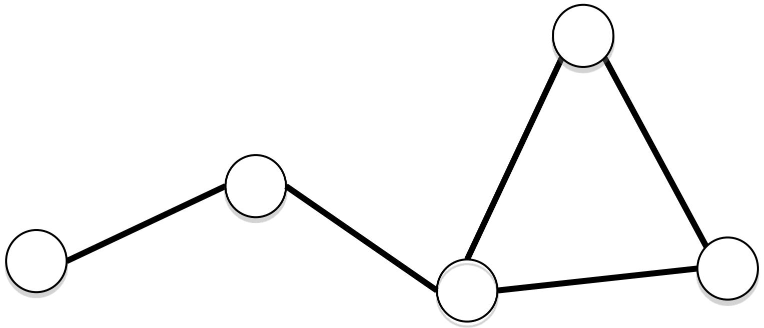
- Vertex, Edge
- Adjacency matrix

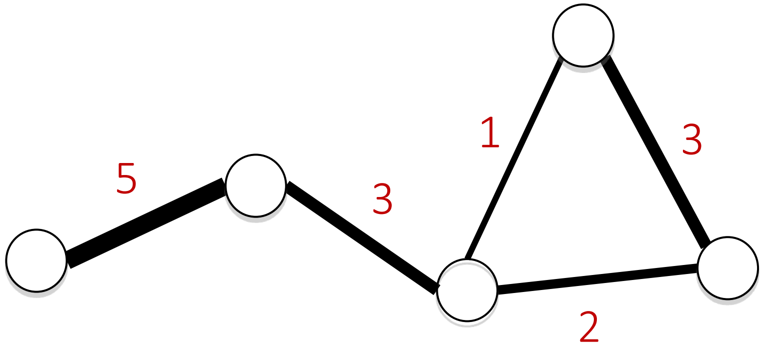
Types

- Weighted vs. unweighted
- Directed vs. undirected
- Connected vs. disconnected
- Cyclic vs. acyclic
- Complete, tree, cubic, star

Types of graphs

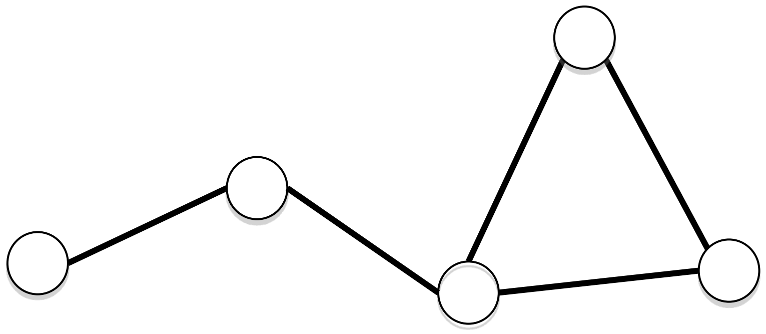
unweighted graph vs. weighted graph

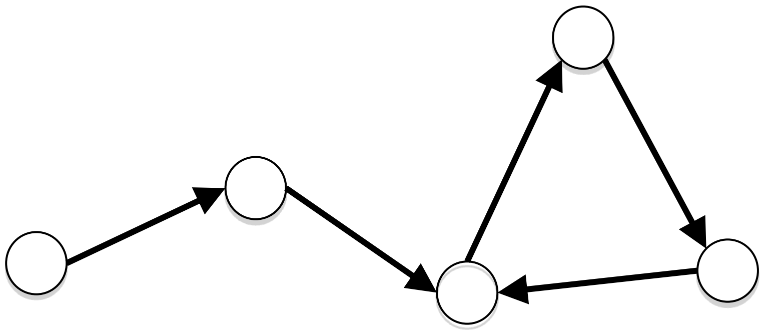




Types of graphs

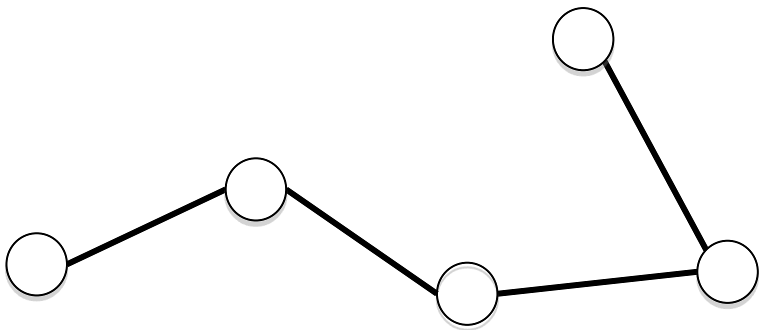
undirected graph vs. directed graph (digraph)

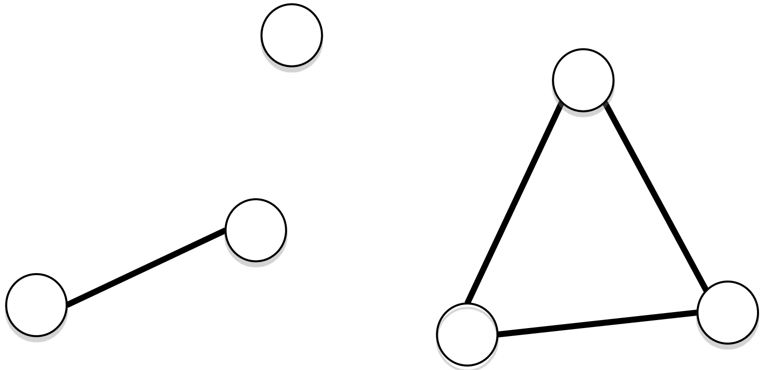




Types of graphs

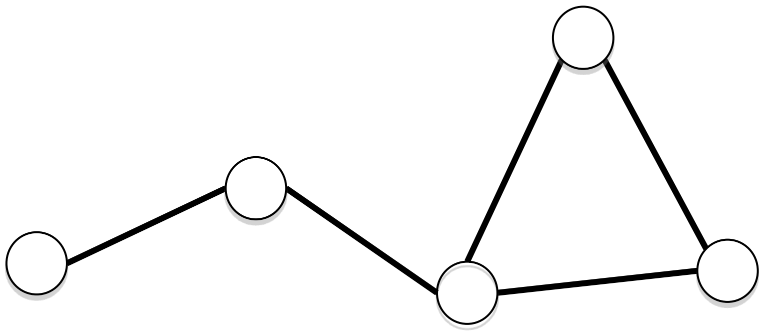
connected graph vs. disconnected graph





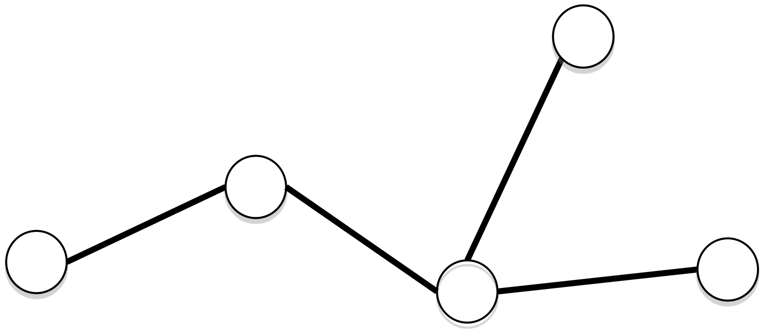
Types of graphs

cyclic graph



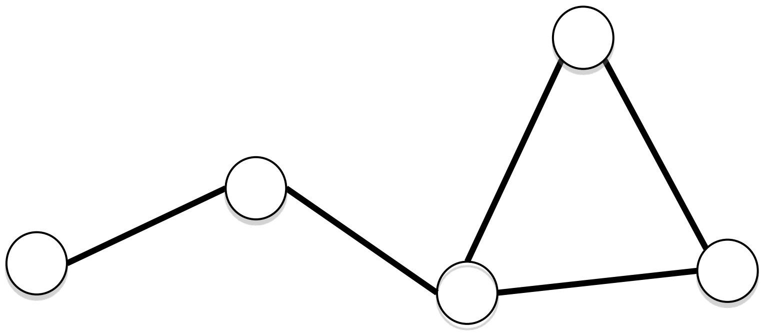
vs.

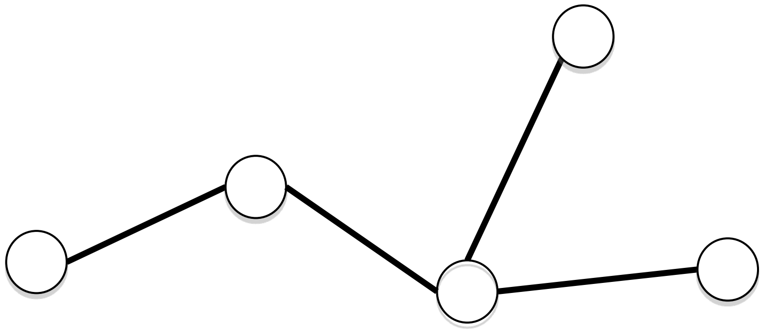
acyclic graph



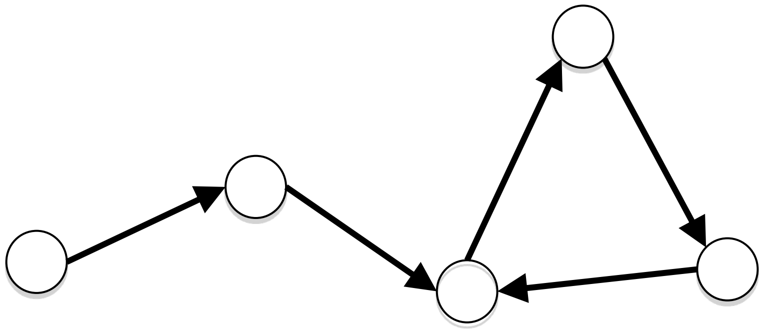
Types of graphs

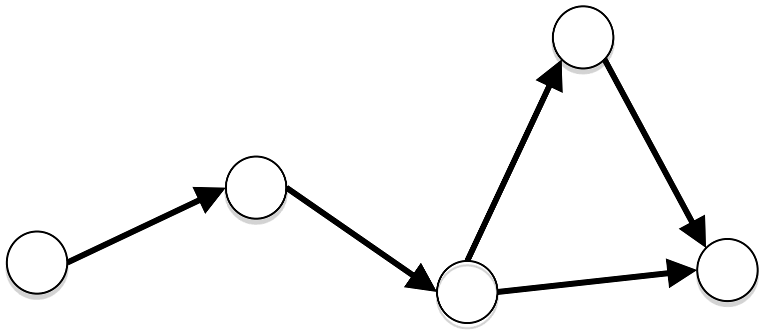
cyclic graph vs. acyclic graph





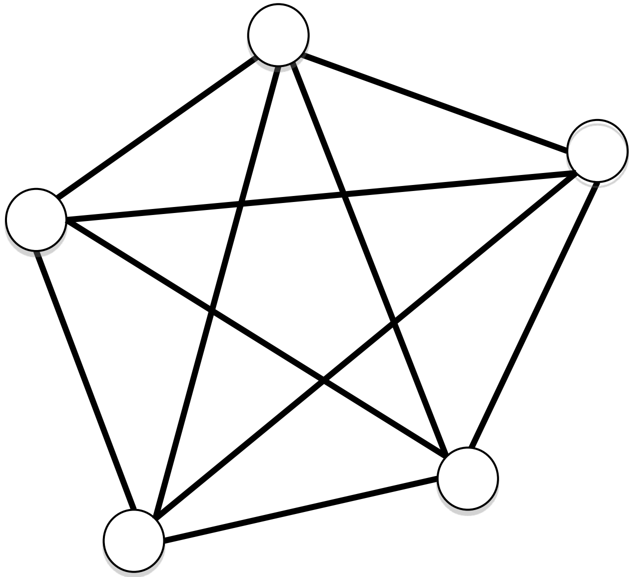
directed cyclic graph vs. directed acyclic graph



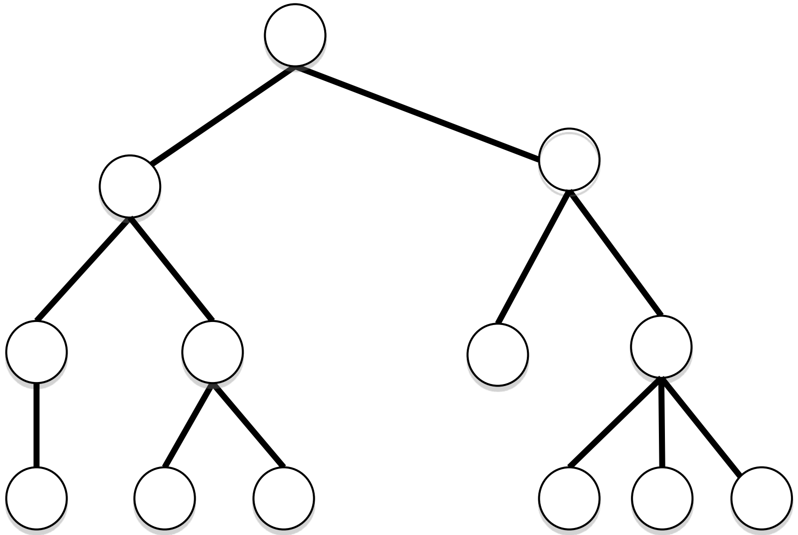


Types of graphs

complete graph



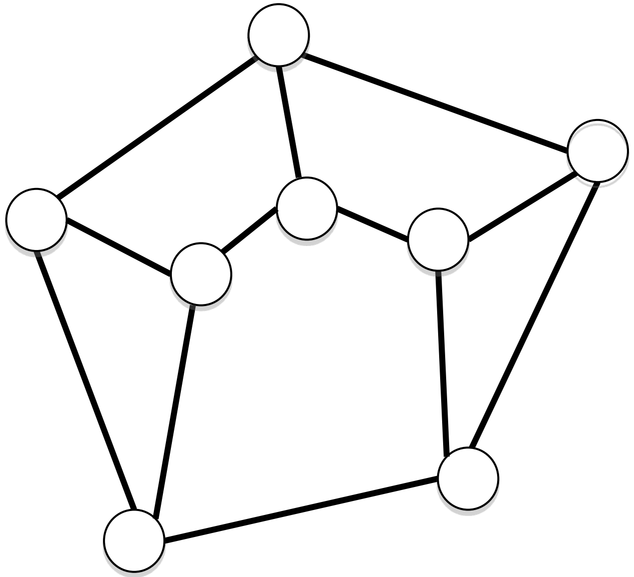
tree (connected acylic graph)



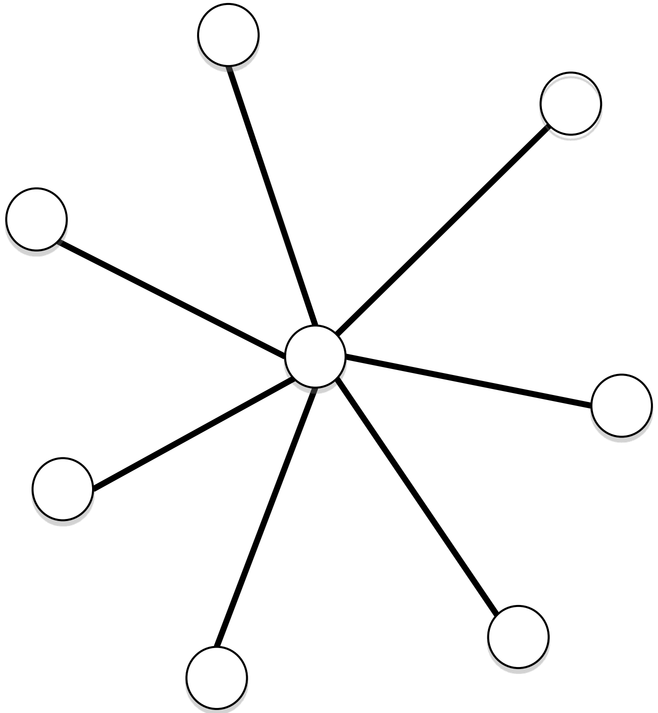
Types of graphs

cubic graph

degree = 3



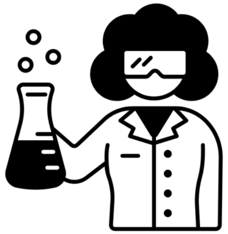
star graph



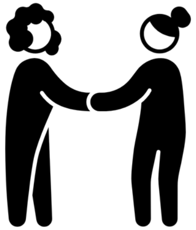
Break

CHILL

WALK



COFFEE OR TEA



MAKE FRIENDS

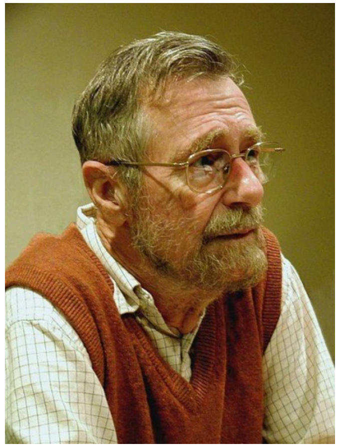
Icons by Iconathon on The Noun Project

Finding the shortest path

General problem

- How do you find the shortest path between two or more points?
- Dijkstra-Algorithm (most common approach)

Edsger Wybe Dijkstra (1930-2002)



© 2016 Wikipedia (H. Richards)

Model

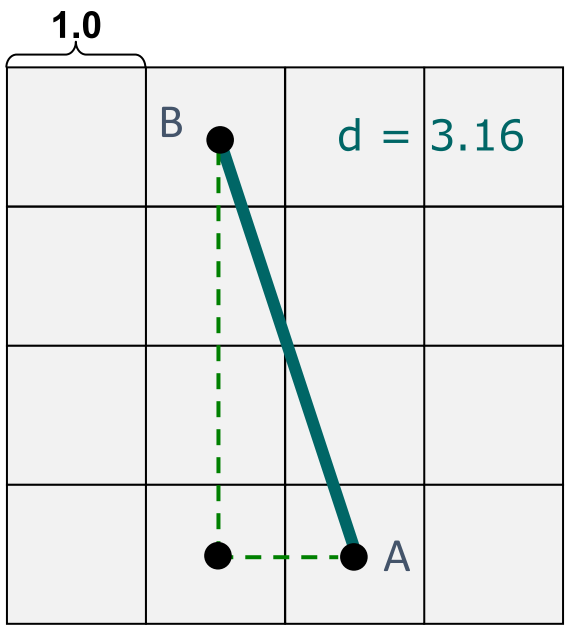
- Road map is a weighted network
  - Nodes = cities/towns/regions/locations on a map
  - Links = road segments/bridges/paths
  - Weights = distance/time/money

Applications

- Navigation Systems, Route-Planning-Software
- Central role in every network analysis tool

Physical Distances

Distance computation via Euclidean Distance



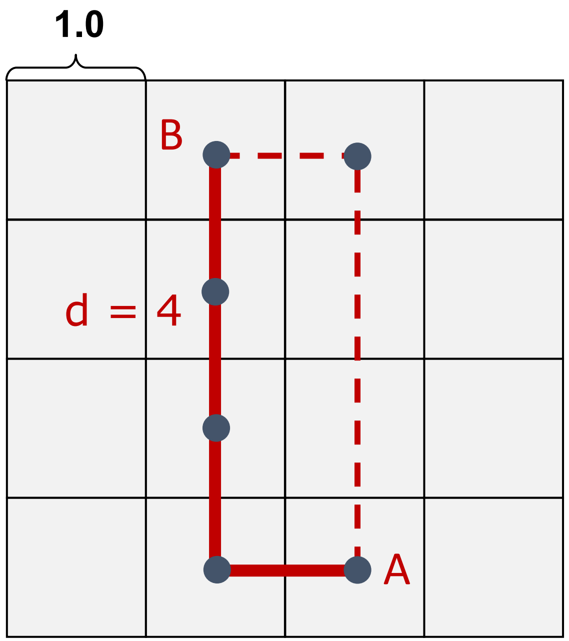
Euclidean Distance

$$d = \sqrt{a^2 + b^2} = \sqrt{1^2 + 3^2}$$
$$= \sqrt{10}$$
$$= 3.16$$

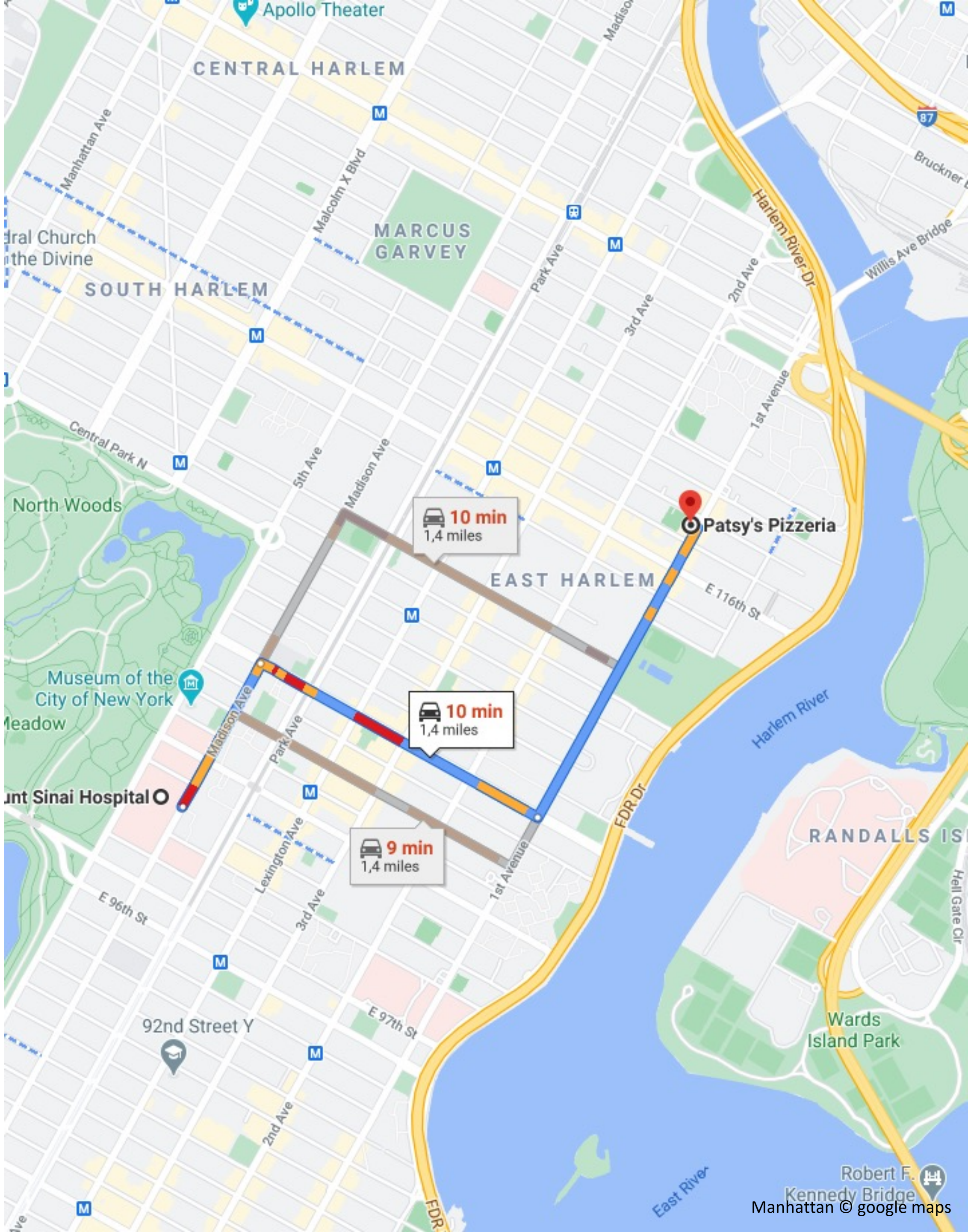
https://en.wikipedia.org/wiki/Pythagoras

Physical Distances

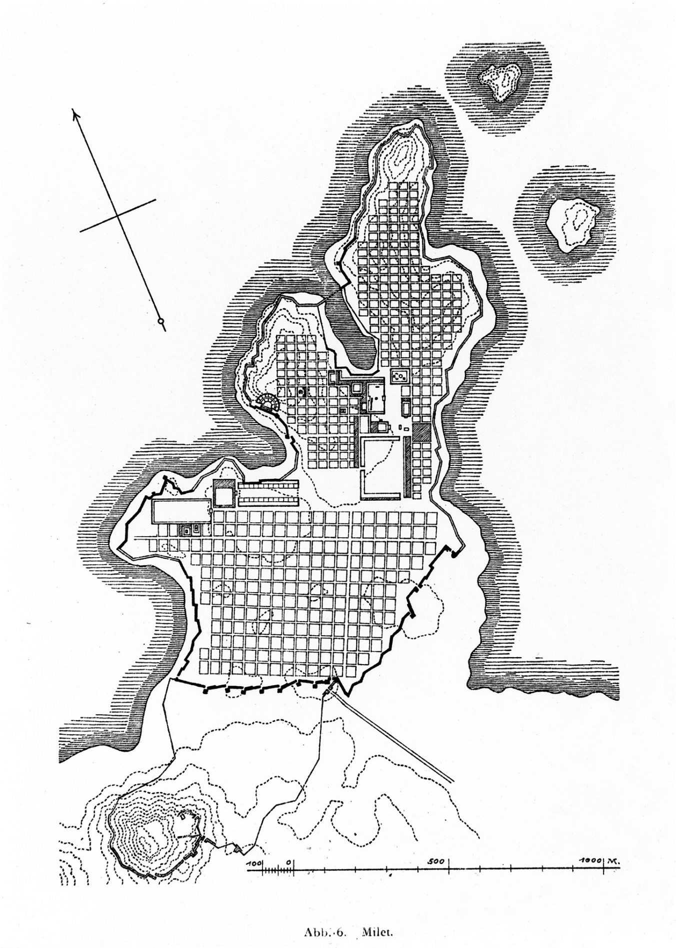
Distance computation via
Manhattan Distance



Manhattan Distance
4 neighbour

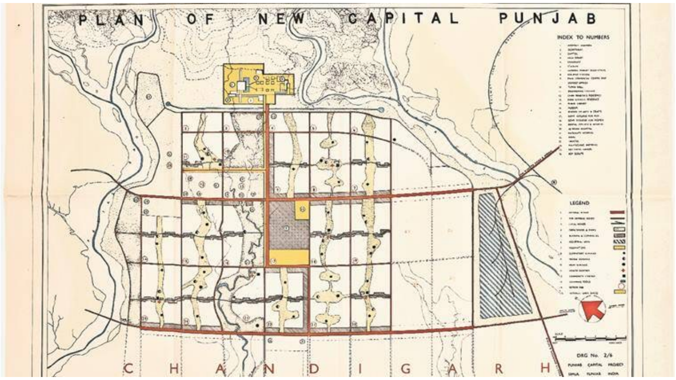


Manhattan © google maps



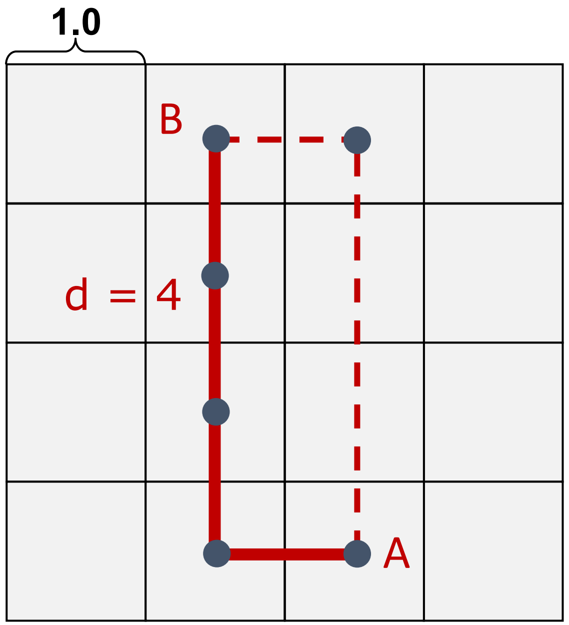
Ancient Greece

Buenos Aires

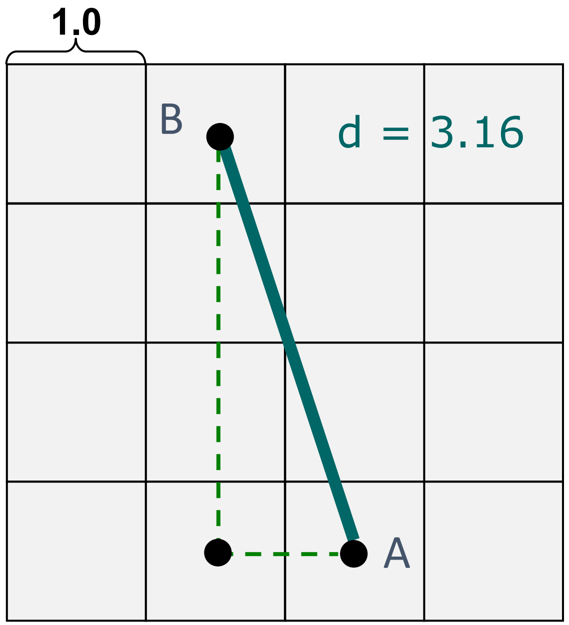


Chandigarh, India

Physical Distances

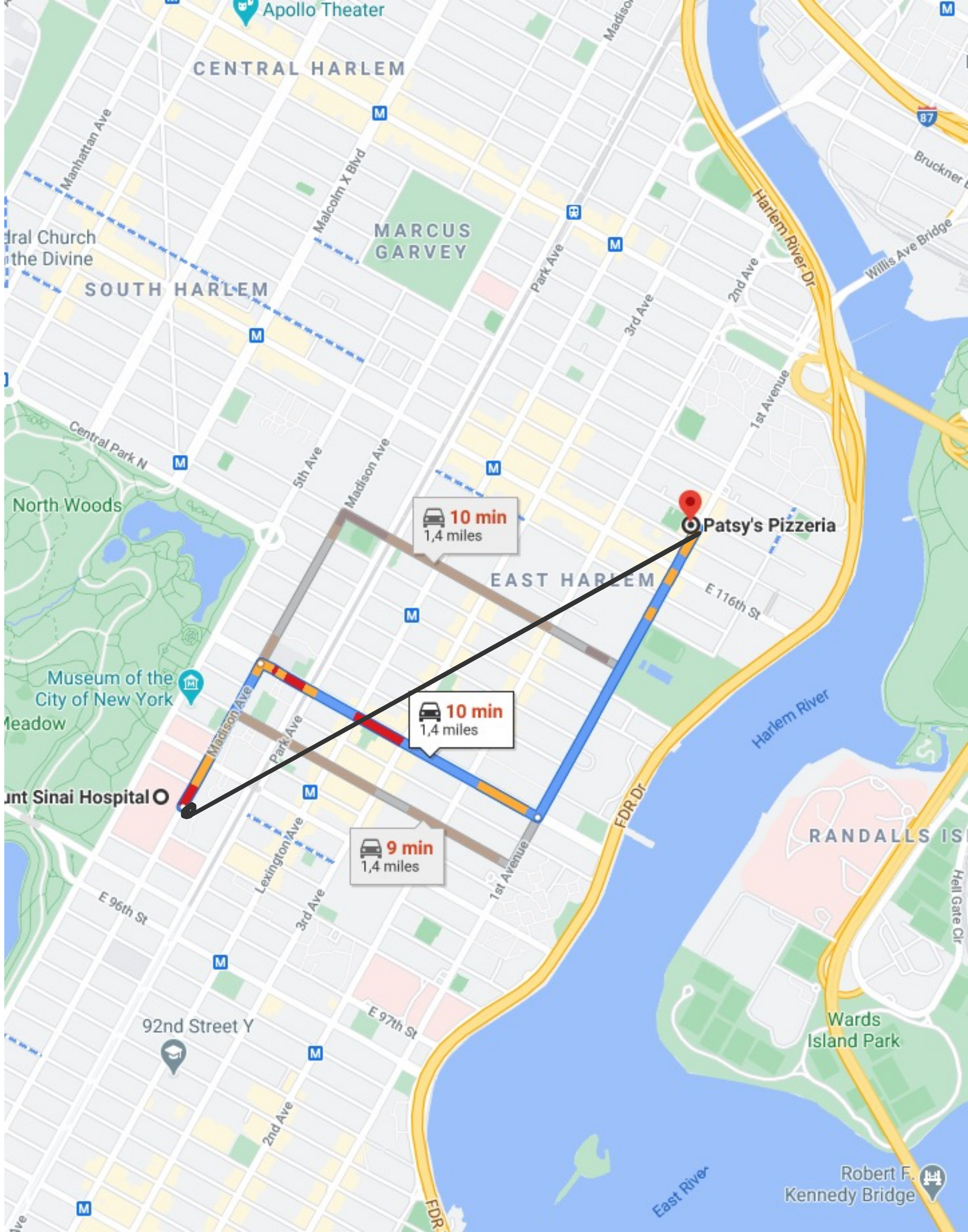


Manhattan Distance
4 neighbour

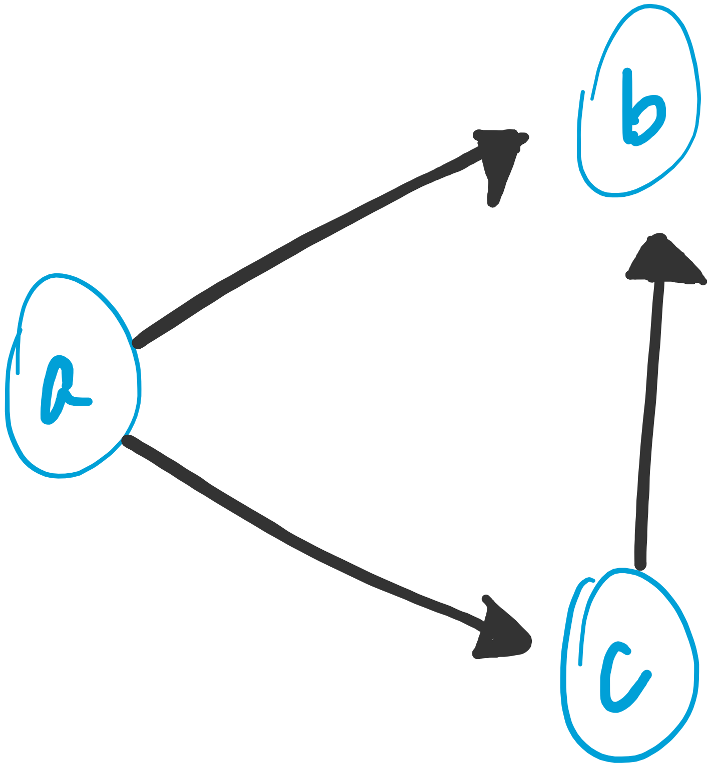


Euclidean Distance

$$d = \sqrt{a^2 + b^2} = \sqrt{1^2 + 3^2}$$
$$= \sqrt{10}$$
$$= 3.16$$

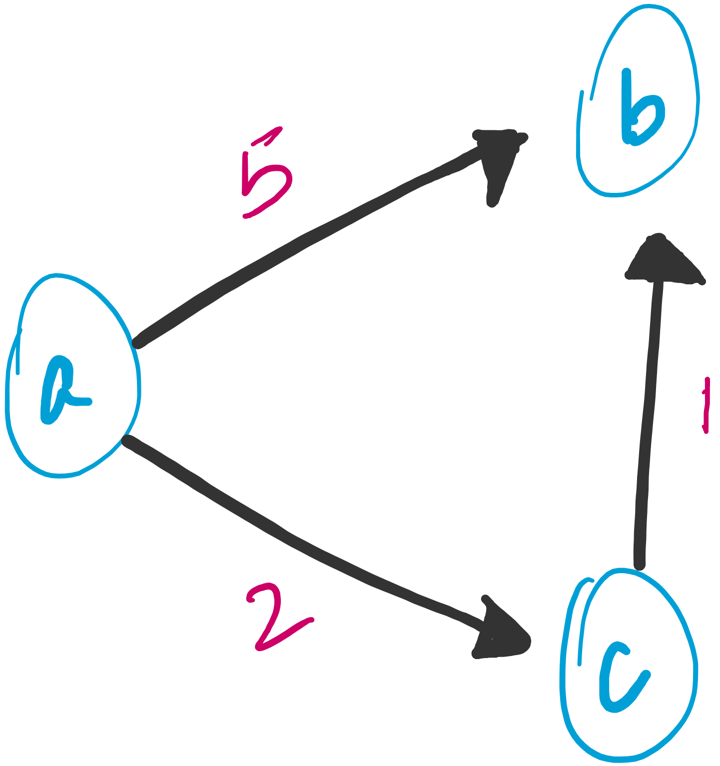


Finding the shortest path



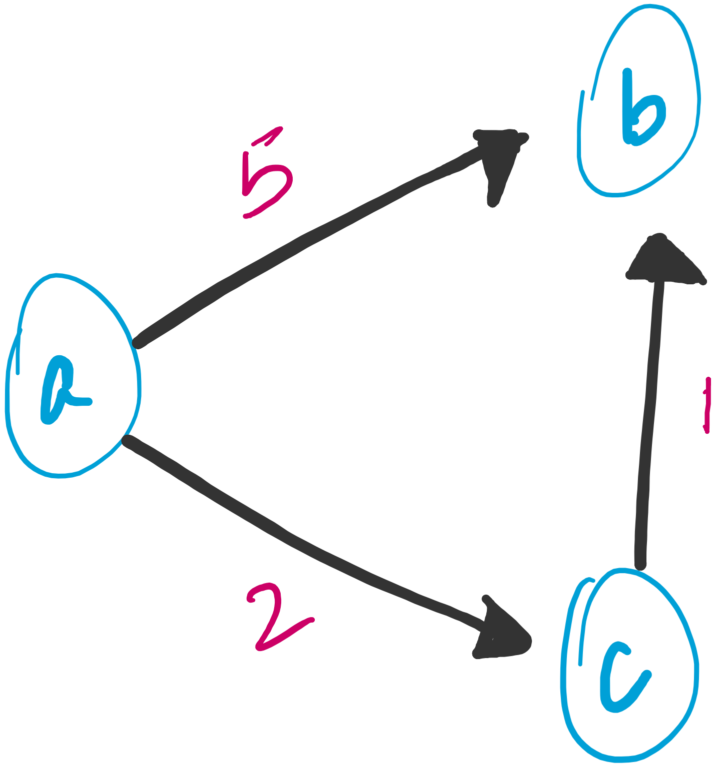
What is the shortest path from (a) to (b)

Finding the shortest path



What is the shortest path from (a) to (b)

Finding the shortest path



What is the shortest path from (a) to (b)

(1) Taking the route from node (a) → (b) will cost us **5**

(2) Taking the route from (a) → (c) → (b) will cost us 2+1 = **3**

Space, formally

For a statistical method to be **explicitly spatial**, it needs to contain some representation of the geography, or **spatial context**

One of the most common ways is through **Spatial Weights Matrices**

- **(Geo)Visualization**: translating numbers into a (visual) language that the human brain "*speaks better*"
- **Spatial Weights Matrices**: translating geography into a (numerical) language that a computer "*speaks better*".

Core element in several spatial analysis techniques:

- Spatial autocorrelation
- Spatial clustering / geodemographics
- Spatial regression

W as a formal representation of Space

W

N x N positive matrix that contains spatial relations between all the observations in the sample

$$w_{ij} = \begin{cases} x > 0, & if\ i\ and\ j\ are\ neighbours \\ 0, & otherwise \end{cases}$$

w_{ii} *= 0 by convention*

…What is a neighbour???

Types of W

A neighbour is "somebody" who is

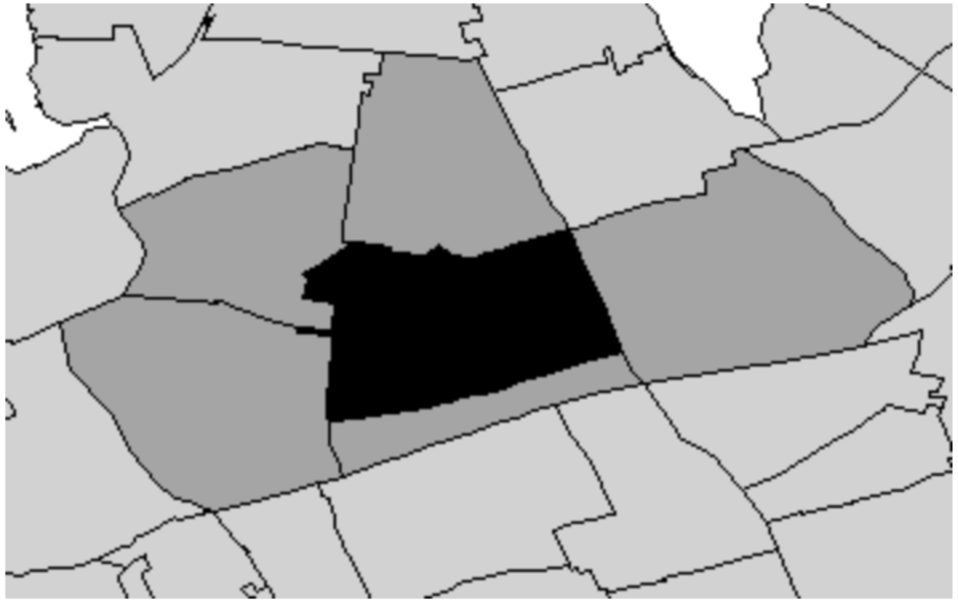
- Next door → **Contiguity**-based Ws
- Close → **Distance**-based Ws
- In the same "place " as us → **Block** weights

Contiguity-based weights

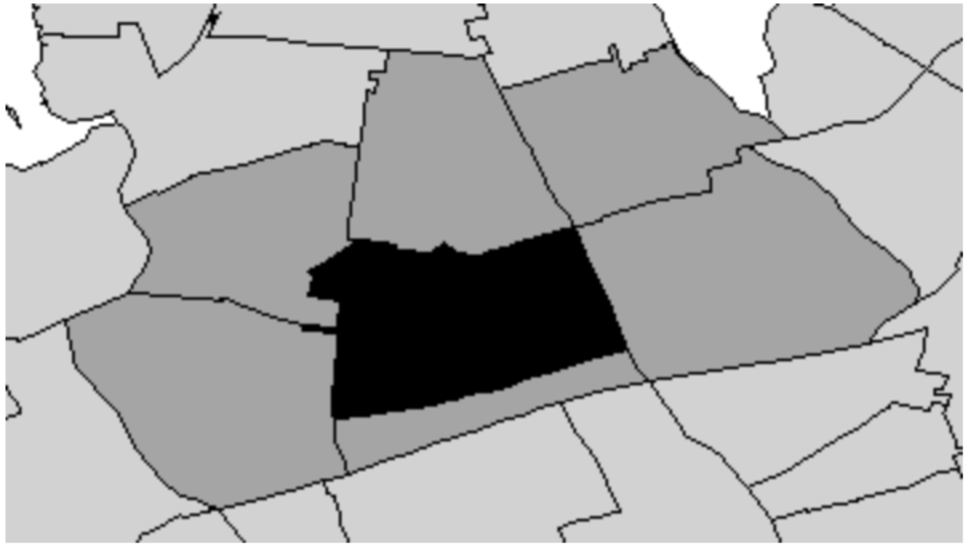
Sharing **boundaries** to any extent

- Rook
- Queen
- …

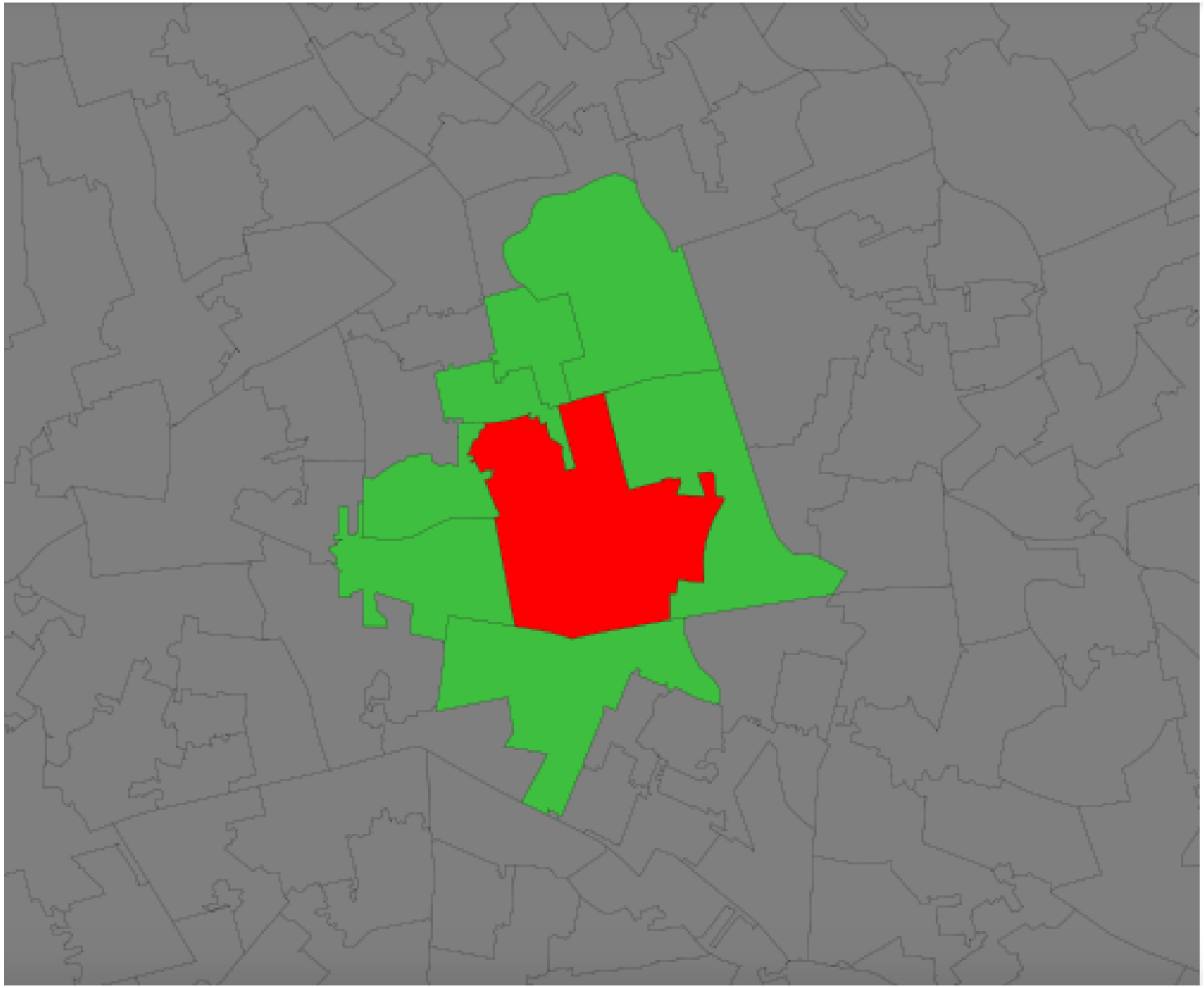
Rook



Queen



Images taken from: Arribas-Bel, D. (2019). A course on geographic data science. *Journal of Open Source Education*, *2*(16), 42.

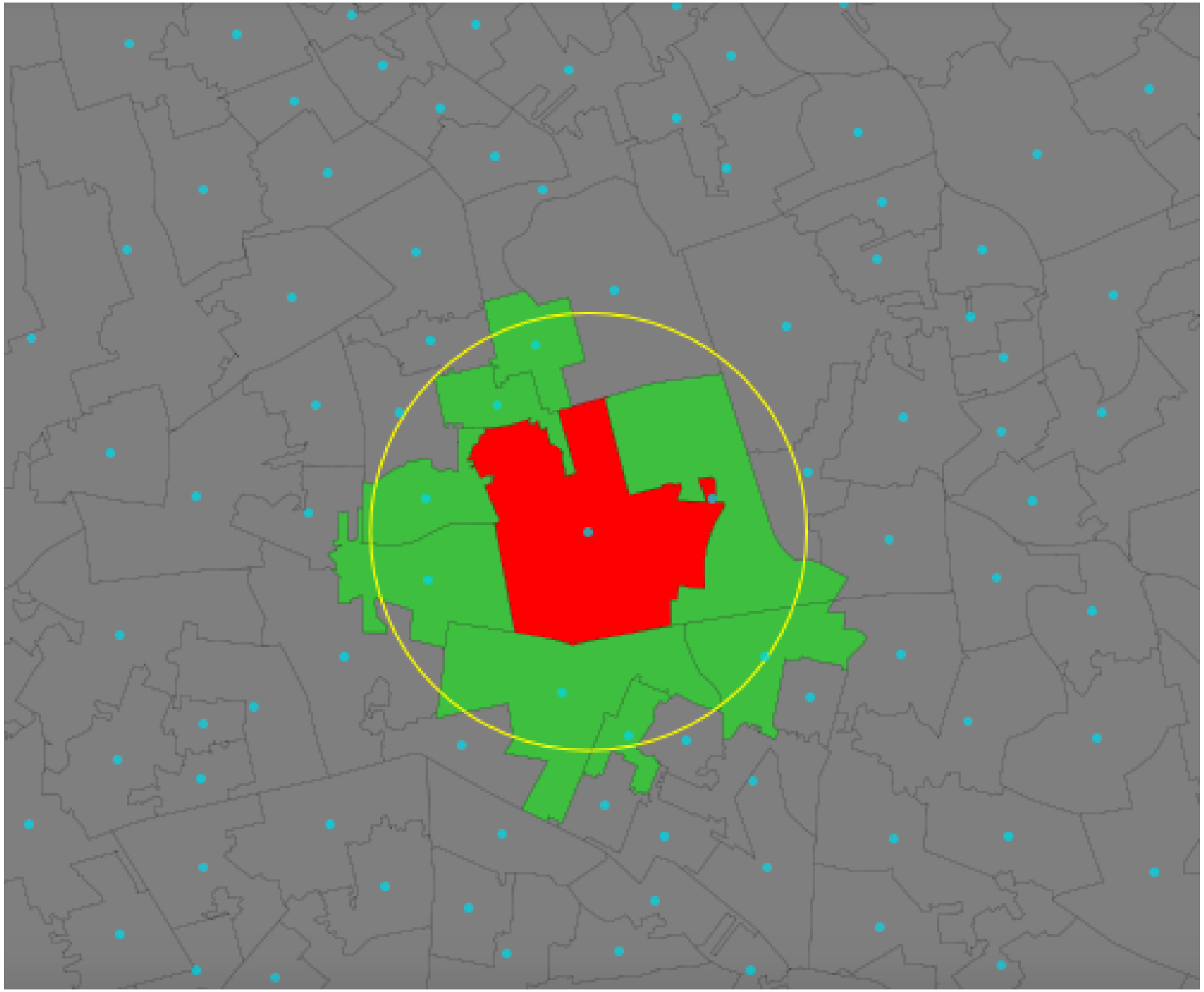


Images taken from: Arribas-Bel, D. (2019). A course on geographic data science. *Journal of Open Source Education*, *2*(16), 42.

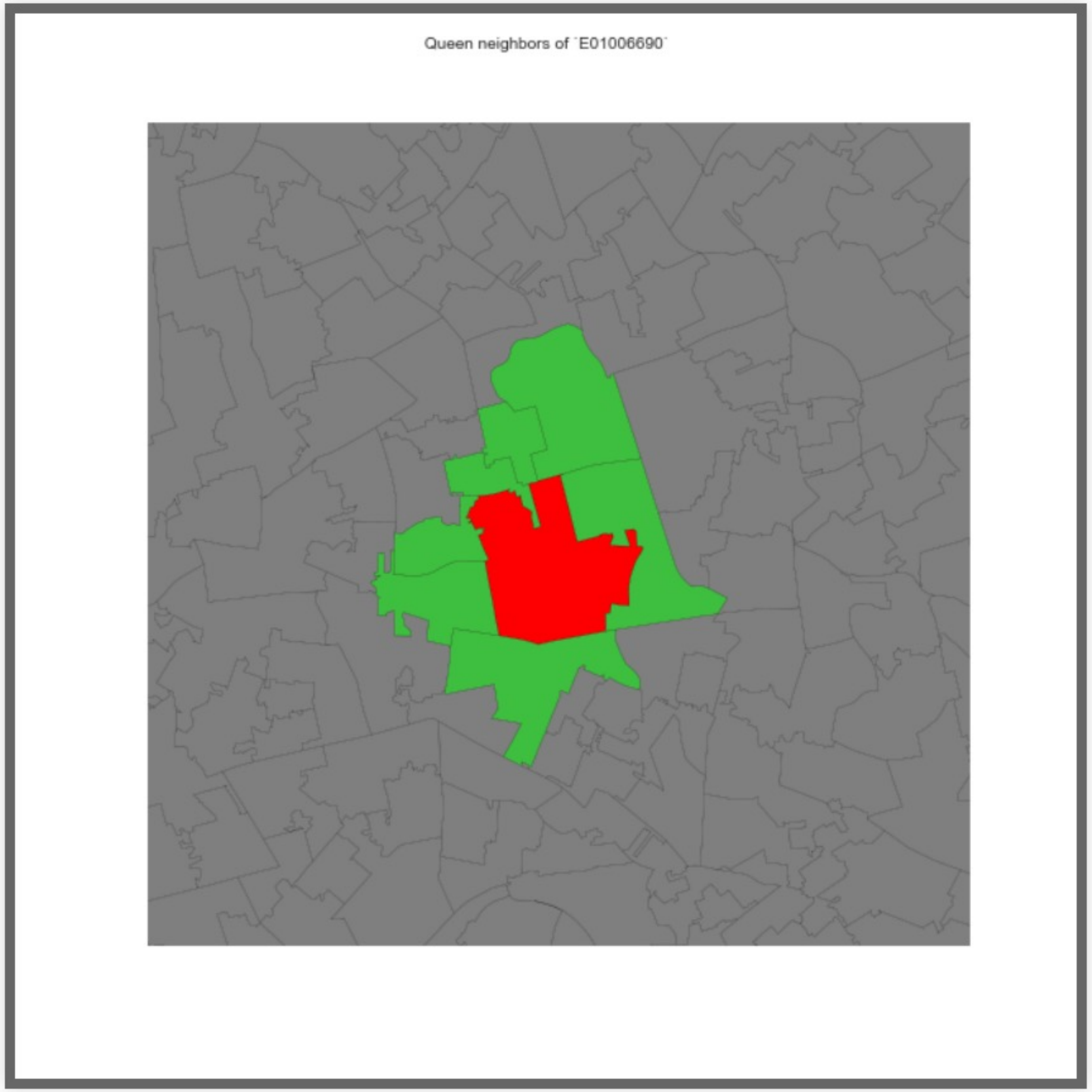
Distance-based weights

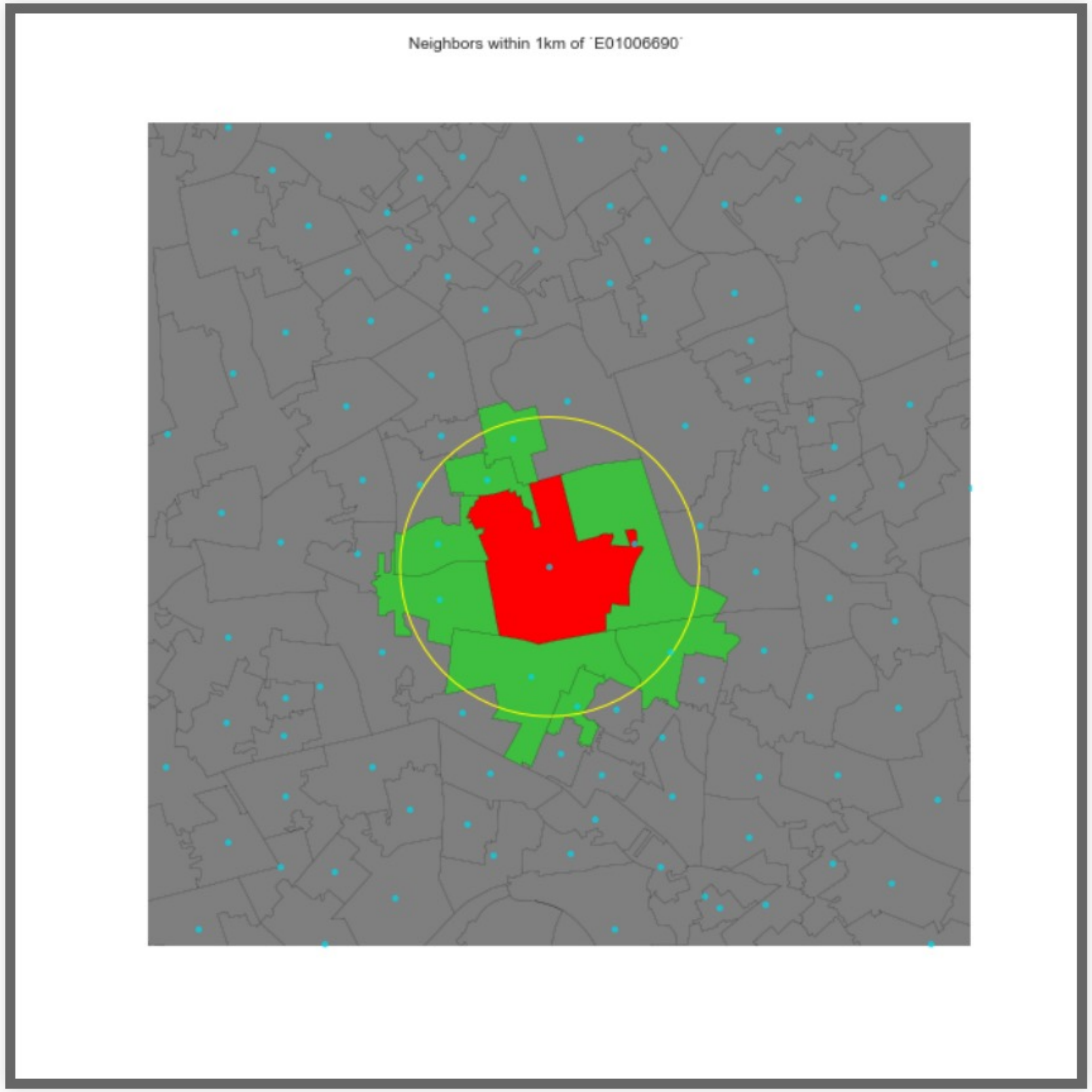
Weight is (inversely) proportional to distance between observations

- Inverse distance (threshold)



Images taken from: Arribas-Bel, D. (2019). A course on geographic data science. *Journal of Open Source Education*, *2*(16), 42.





Images taken from: Arribas-Bel, D. (2019). A course on geographic data science. *Journal of Open Source Education*, *2*(16), 42.

Block weights

Weights are assigned based on discretionary rules loosely related to geography

For example:

- Buurts into Wijks
- Post-codes within city boundaries
- Counties within states
- …

How much of a neighbour?

Not a neighbour? receive zero weight: w_{ij} = 0

Neighbours, it depends, w_{ij} can be:

- One: w_{ij} = 1 → Binary
- Some proportion (0 < w_{ij} < 1, continuous) which can be a function of:
  - Distance
  - Strength of interaction (e.g., commuting flows, trade, etc.)

Choice of W

Should be based on and reflect the **underlying channels of interaction** for the question at hand.

Examples:

- Processes propagated by immediate contact (e.g. disease contagion) → Contiguity weights
- Accessibility → Distance weights
- Effects of county differences in laws → Block weights

Standardisation

In some applications (e.g. spatial autocorrelation) it is common to *standardize* W

The most widely used standardization is row-based: divide every element by the sum of the row:

$$w'_{ij} = \frac{w_{ij}}{w_{i.}}$$

where $\boldsymbol{w_{i.}}$ is the sum of a row

Spatial Lag

Spatial Lag

Weighted average of neighbouring values

- Neighbour definition comes from spatial weights w_{ij}

$$Y_{iL} = w_{i1}Y_1 + w_{i2}Y_2 + w_{i3}Y_3 + \dots w_{in}Y_n$$

Spatial Lag variable has a *smaller* variance than Y because it is a smoother function

Spatial Lag

- Measure that captures the behaviour of a variable in the neighborhood of a given observation i.
- If W is standardized, the spatial lag is the weighted average value of the variable in the neighborhood (good for comparison and scaling)

Spatial Lag

- Common way to introduce space formally in a statistical framework
- Heavily used in both **ESDA** and spatial regression to delineate neighborhoods.
- Examples (covered in next lecture):
  - Moran’s I
  - LISAs
  - Spatial models (lag, error…)

Recapitulation

- Everything is connected and must be considered so
- Spatial Weights matrices: matrix encapsulation of space
- Different types for different cases (contiguous, distance and blocks)
- Useful in many contexts, like the spatial lag and Moran plot, but also many other things!

For next class..

Finish Labs to practice programming

Complete Homework for more practice

Check Assignment contents and due date

See “To do before class” for next lecture (~ 1 hour of self-study)