NIHR | Applied Research Collaboration South West Peninsula

@penCHORD_UoE
@peninsula_ARC

Module 5 : Geographic Modelling & Visualisation

Session 5A : Introduction

Elliott Coyne

"Bring on Wonderful"

HSMA 5

#hsma5isalive

NIHR | Applied Research Collaboration South West Peninsula

@penCHORD_UoE
@peninsula_ARC

Module 5 : Geographic Modelling & Visualisation
Session 5A : Introduction
Elliott Coyne

"Bring on Wonderful"

HSMA 5

#hsma5isalive

Welcome to Module 5!

Geographic Modelling

Over the next four sessions we’ll be looking at….

- **General topic introduction, Travelling Salesman Problem (TSP)**
- Geospatial problems
- Geopandas
- QGIS (Geographic Information System)

What's it all about?

Geographic Modelling

Where?

Where to place services…

Where will patients go…

Where do we find people with this disease…

Where is the demand...

Predictions

…if our demand continued to grow?

…if we close these services?

…and how is it spreading?

…and how will users have equitable access

What are geospatial problems in health?

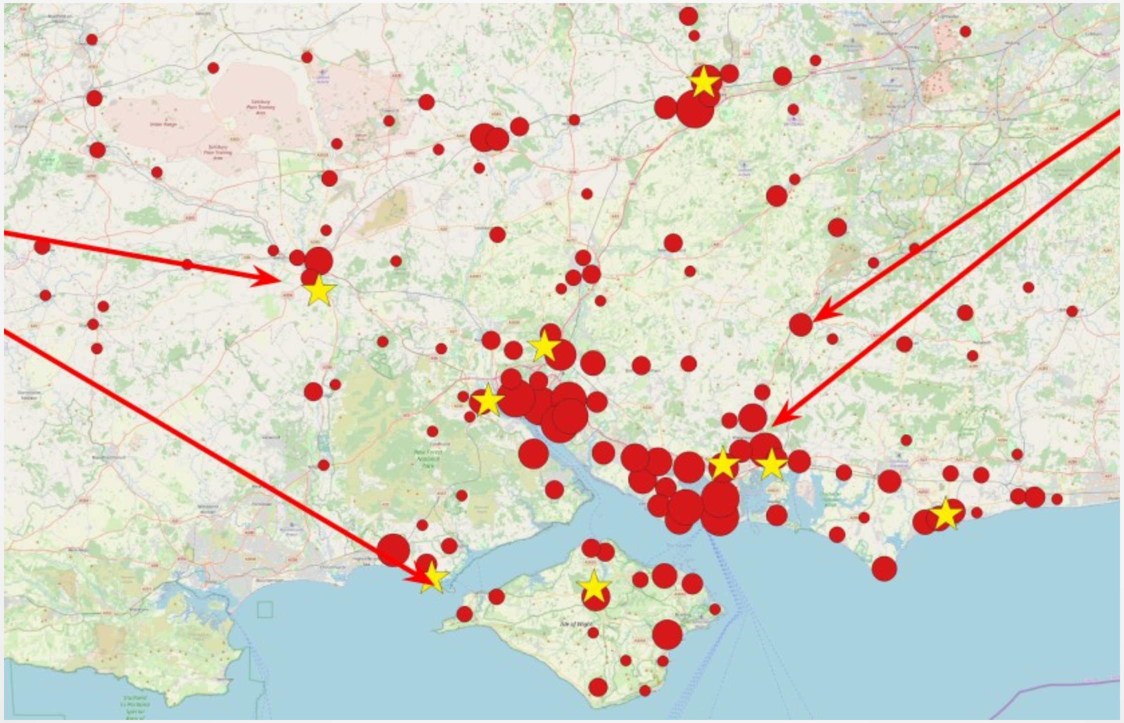
Health care delivery or service configuration problems that are related to, and complicated by geography.

They include...

- **Routing and scheduling problems:** mobile health workforce travelling to service users i.e., community, or district nurses
- **Location-allocation** problems: Location of health service provision facilities i.e., sexual health clinics, or mobile ophthalmology services
- Key issues:
  - availability of services
  - equity of access for service users
  - capacity / throughput
  - ‘travel times’ or ‘distances’

Example Location-Allocation Problem

★ Dialysis treatment centres (i.e., hospitals, or smaller treatment centres)



‘Location’ of patients using dialysis service (size represents density)

Wessex

Considerations:
- demand
- equity/ access
- deprivation
- access to transport...

Location-allocation modelling

The *health service* perspective

Customers of this type of analysis: individual provider, regional (i.e., CCG) or national level of organisation (NHS England).

Examples of different levels...

- A city hospital providing ophthalmology services need support to decide where to locate 5 mobile clinics most equitably for users.
- A public health team (i.e., within local authorities or councils) and sexual health provider organisation need support in reducing the number of locations for outpatient clinics in a region whilst still offering equitable access. *Spoiler alert – there will be winners and losers!*
- NHS England want support in centralising hyperacute (i.e., emergency) stroke care across England, to maximise equity, quality (throughput) and patient outcomes. There may be benchmarks to meet i.e., *"unless you hit 600 stroke admissions per year, we can't call you a 'comprehensive' centre"*

Location-allocation modelling

The *data scientist* perspective

Types of needs to address...

- Where to locate facilities in a new region
- Measurement of the effectiveness of existing/ previous locations i.e., *was it a good decision?*
- Finding set of optimal sites in an existing system configuration
- Generating alternative solutions (i.e., improve existing system)


ELSEVIER

European Journal of Operational Research

Volume 123, Issue 3, 16 June 2000, Pages 437-452

Invited Review

Use of location-allocation models in health service development planning in developing nations

Shams-ur Rahman [a], David K. Smith [b]

Show more

Add to Mendeley Share Cite

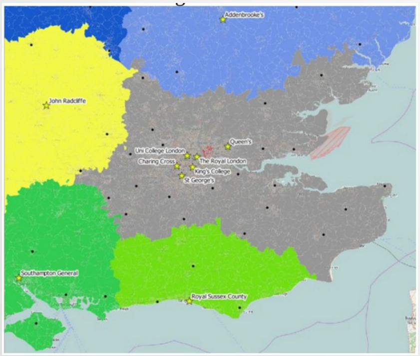
Rahman and Smith (2000), Use of location-allocation models in health service development planning in developing nations, European Journal of Operational Research, Volume 123, Issue 3, 2000, Pages 437-452, https://doi.org/10.1016/S0377-2217(99)00289-1.

Location-allocation modelling

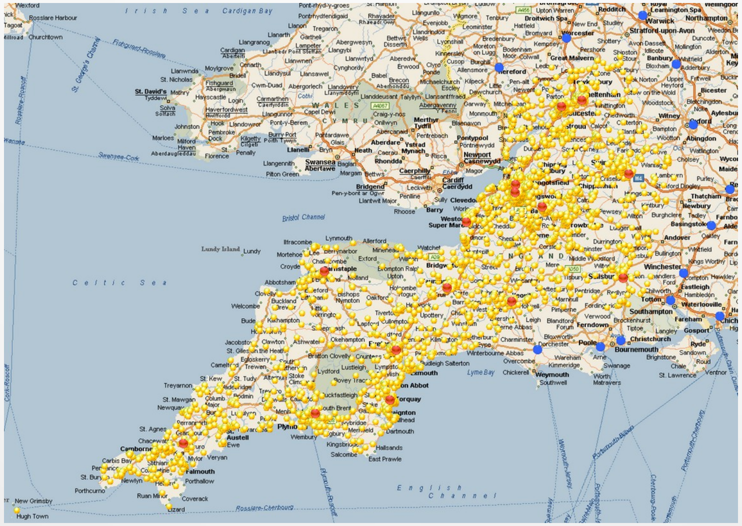
Does everyone know what *heuristic means?

One size doesn't fit all…
Different size projects include….

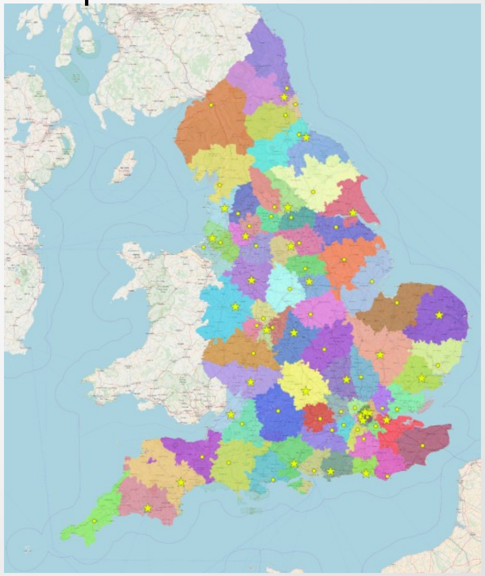
Small scale problems in health - best supported using using a combination of simple mapping and analysis inc. descriptive statistics



Medium scale problems in health
- brute force to assess each combination
- Optimisation can be either single, or multi-objective



Large scale problems in health
- search for good solutions (heuristic algorithms)
- possibly multi-objective in nature
- decision support – not just optimisation



Examples of Types of Location Data

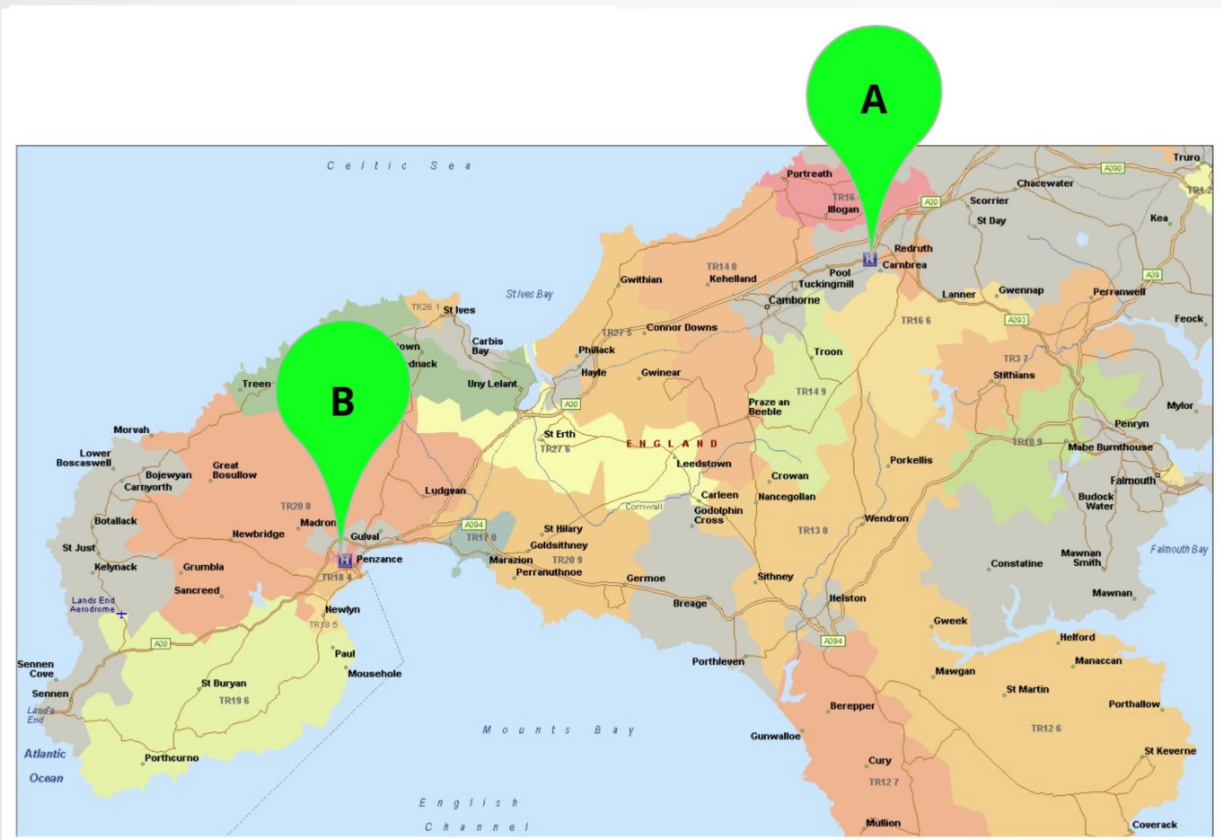
Type	Description	Example
Latitude & Longitude	Exact geographic coordinates	50.85023, -1.06993
Northings and Eastings	Exact coordinates associated with the Universal Transverse Mercator coordinate system (UTM)	5634945.02, 635862.59
Full Postcode	There are about 1.7 million unit postcodes	PO6 3LY
Postcode Sector	On average 180 postcodes	PO6 3
Lower Super Output Area (LSOA)	An average population of 1500 people or 650 households	East Devon 020B
Middle Super Output Area (MSOA)	MSOAs have an average population of 7500 residents or 4000 households	Cornwall 049

Geographic size will vary

Types of Location Data

Facility Locations

Penzance & Cambourne

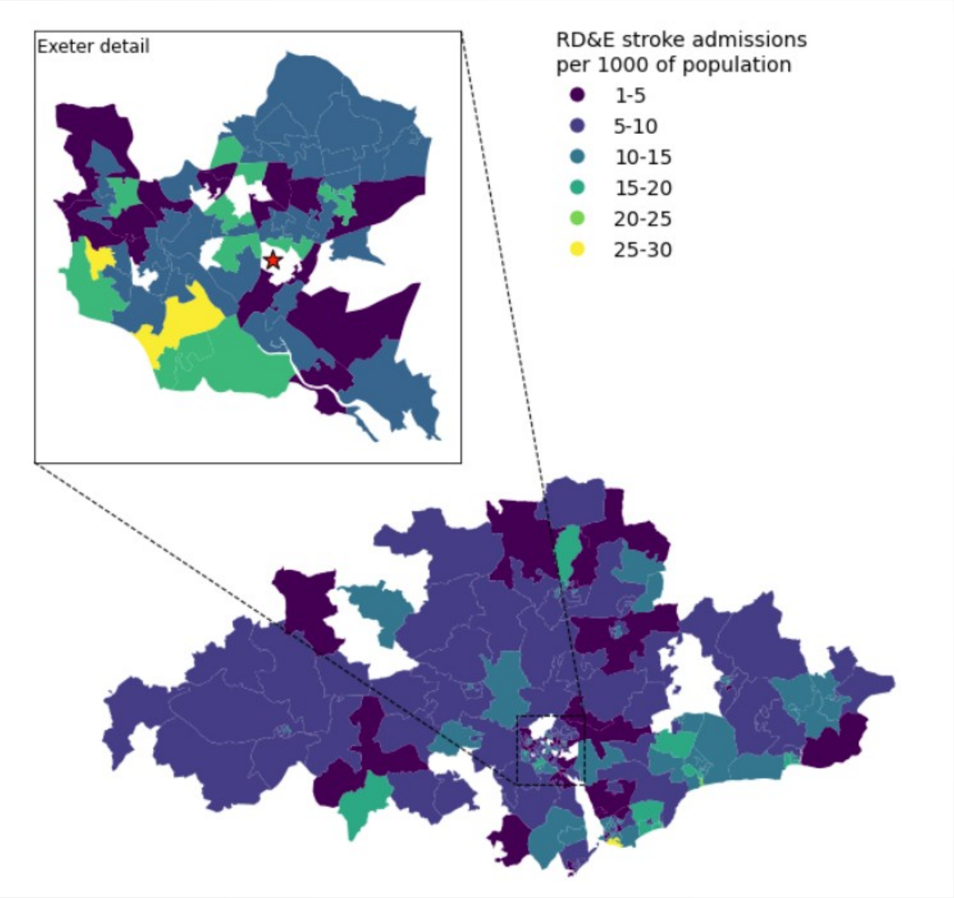


Not sensitive data e.g.

- Full Postcode
- Lat / Long
- Northing Eastings

Types of Location Data

Patient Locations



Sensitive data

Care to be taken if sharing externally to your organisation

Likely you will use:

- Postcode sector
- Lower super output area (or higher)

Further data protection

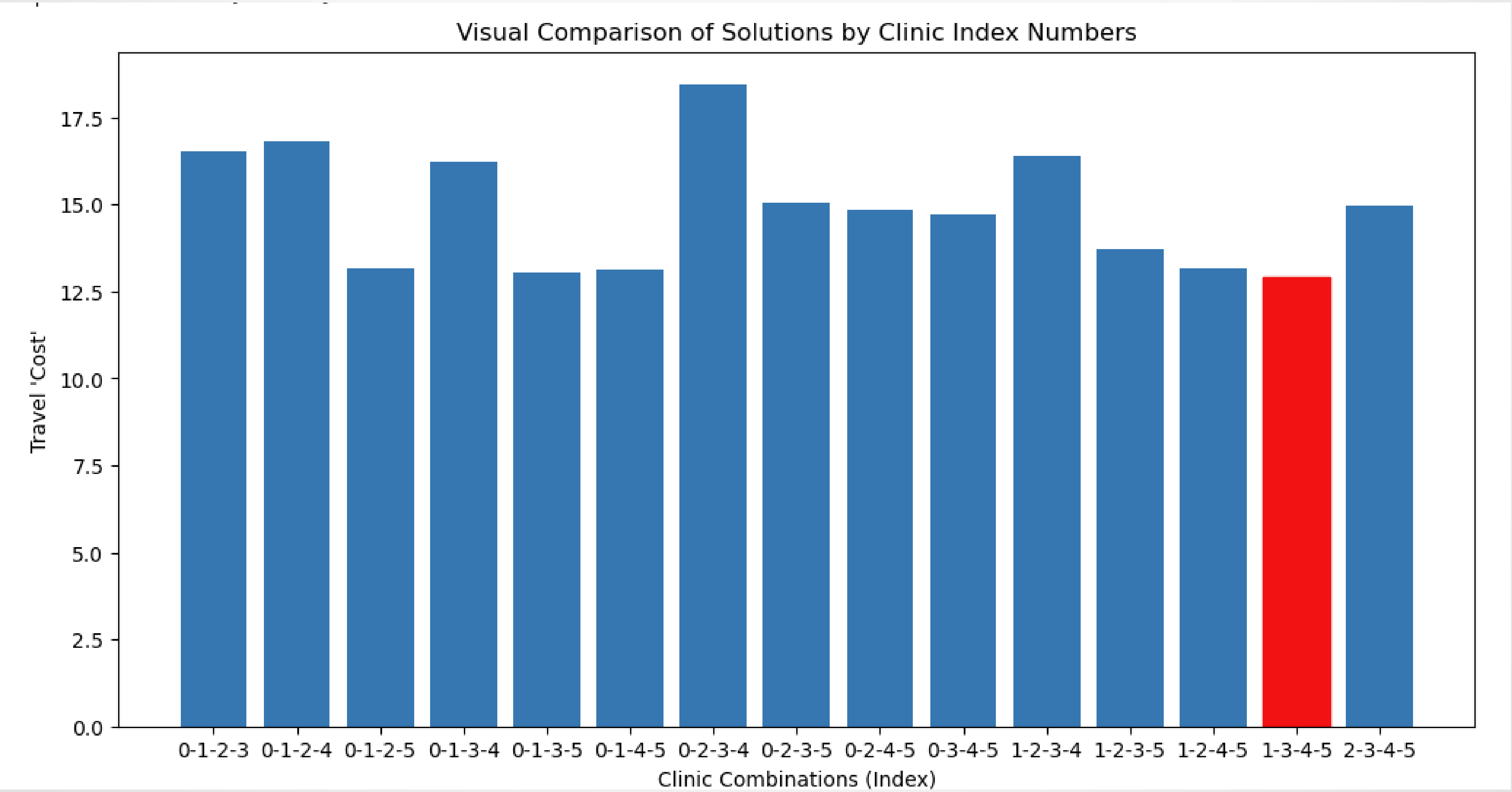
- E.g. postcode sectors with less than 6 patients of interest may be censored

Location-allocation modelling

Use of travel times

- People often have to travel to health facilities to use services
- Generally felt that the closer the service to its community the better
  - May have life critical consequences in emergency health care.
- Travel times usually taken from Geographic Information Software (GIS) software
  - E.g. Routino (Linux), Open Source Routing Machine
- These are typically **car** travel times
  - may mean limitations to the analysis? (Also consider public transport travel times)
  - Local ambulance service may be able to provide historical blue light travel times.
- Public transport times more difficult to access
  - frequency and duration
  - car ownership in study population?

Does the lowest always win?



Formulations of a location-allocation problem

- **p-median problem**
  - Given discrete demands (i.e. LSOAs) locate a number (*p* or less) of health facilities so that the total weighted travel distance or time between facilities and demand centres is minimised (***See next slide***)

- **Location set covering problem**
  - Find the minimum number of facilities (and their locations) such that each and every demand centre is covered by at least one facility within a given maximal service distance (time) *E.g., 'everyone within 30 minutes of a service' – this may require 50 pins → 45 mins may only take 27 pins*

- **Maximal covering location problem**
  - Locate the facilities in such a way that as few people as possible lie outside the desired service distance *i.e., trade off between number of facilities and number of service users that are close to them. Relaxes the constraint of 'location set coverage problem') i.e. able to identify who are the winners and losers.*

‘p-median’ Problem aka Weighted Average

Given discrete demands, locate a number (p or less) of health facilities so that the total weighted travel distance or time between facilities and demand centres is minimised

Average
10 + 30 / 2 = 20 minutes

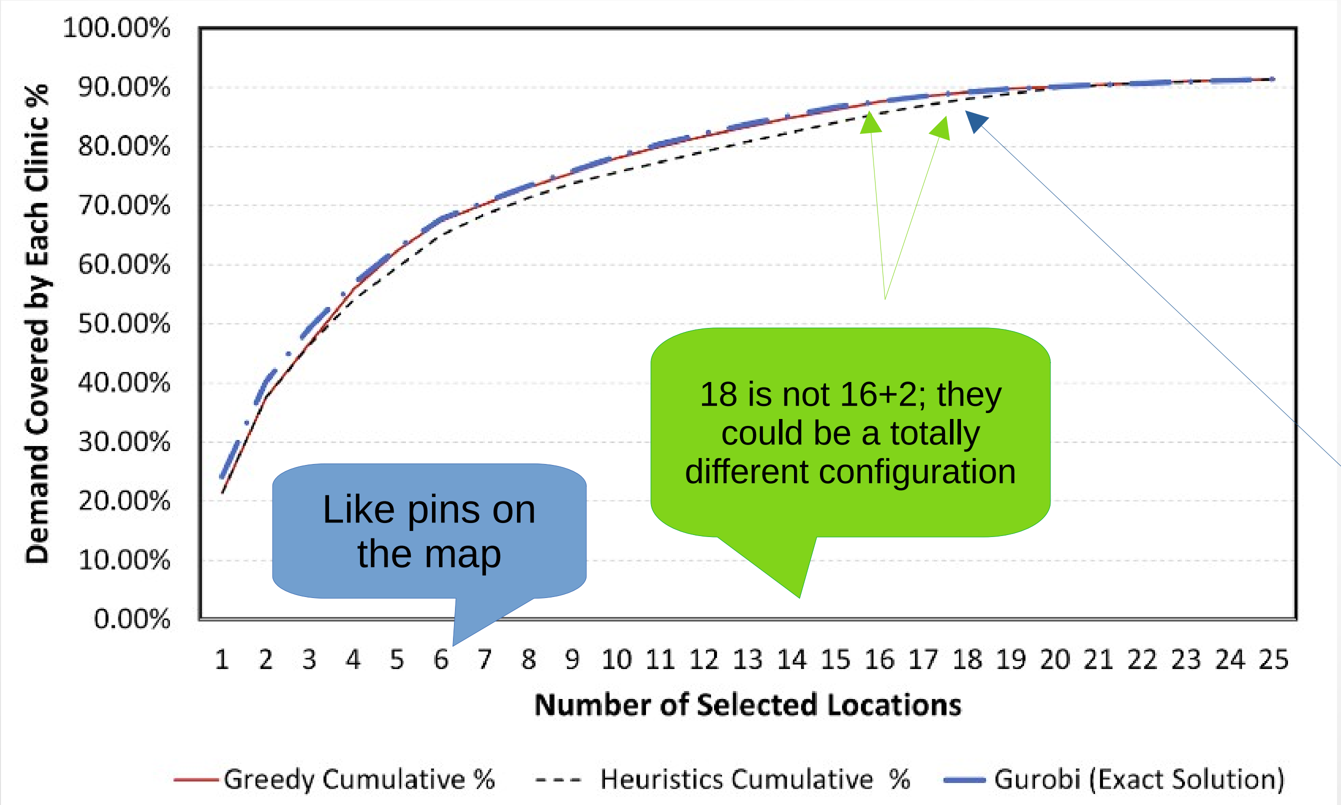
Weighted Average (WA)
A: 10 mins x 5 = 50
B: 30 mins x 1 = 30
80

WA = 80 / 6 trips = 13.3 mins

Problem with this approach? It doesn’t consider max travel time.

Maximal covering location problem

Locate the facilities in such a way that as few people as possible lie outside the desired service distance



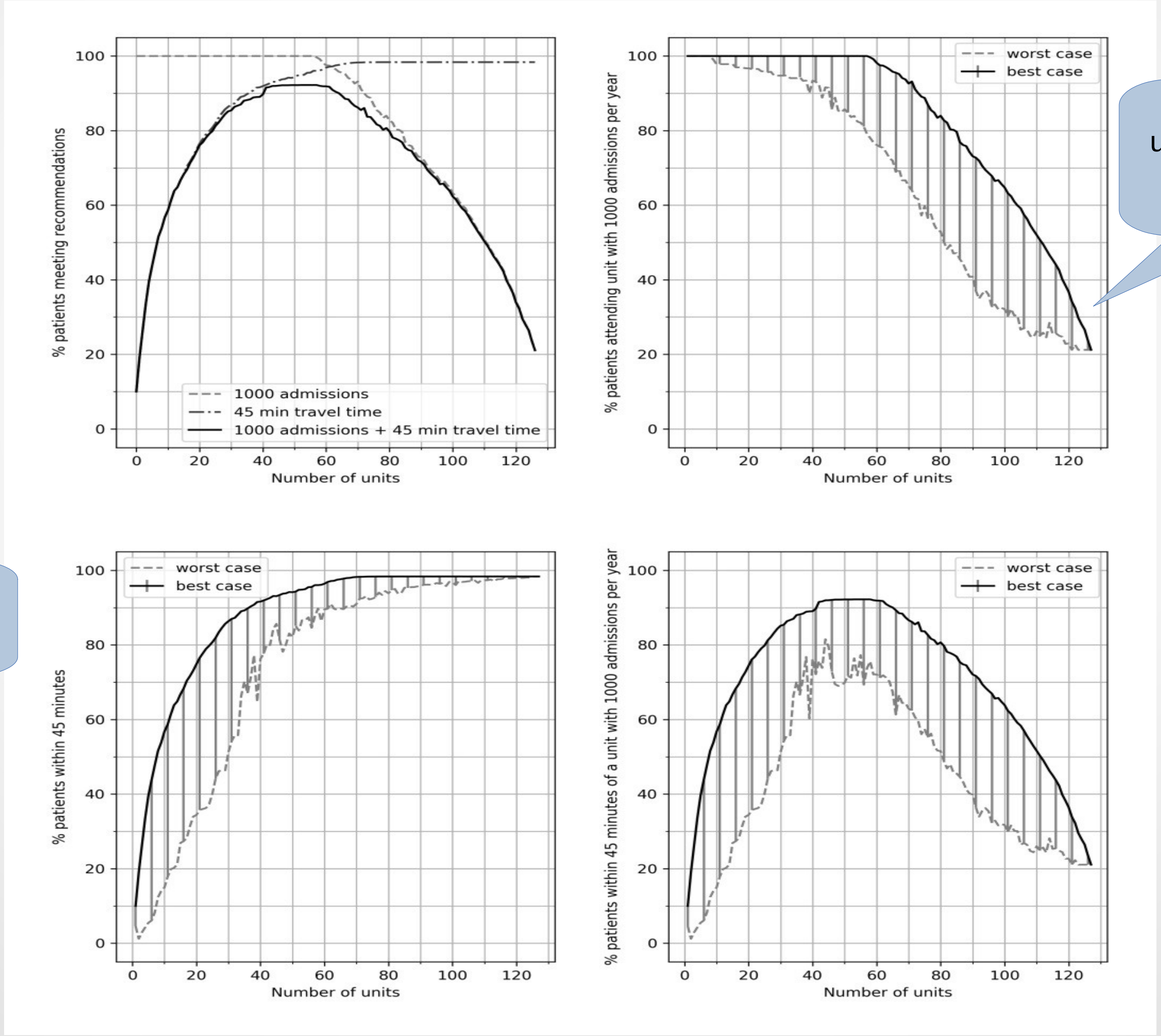
Helps people understand the trade off between number of locations Vs demand covered

… And also that there are diminishing returns

Location-allocation modelling

Multiple-objective problem

% pats meeting recommendations for both considerations when considering factors in isolation and together



% pats attending unit that has >1000 admissions annually

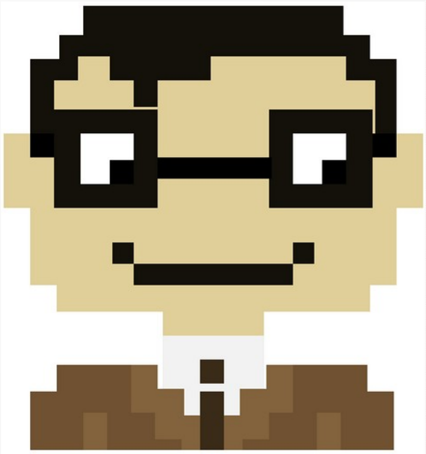
% pats within 45 minutes travel time

Break for 15 mins then…

Practical Issues

Decision support versus optimisation

Great! The NHS has got a location problem. I'll just grab my computer and show them the optimal solution. Whoop Whoop!



Don't be like Alan the Analyst!!!

- You cannot model everything (and don't want to)
- There's lots of qualitative factors in play – some may not even become apparent until the solution is presented to stakeholders
- *"The patients don't like the car parks in Whiteley"*
- Your work is one form of evidence taken into account when making a decision.

Often multiple 'good' solutions will be presented to help stakeholders consider the different options available

Raw vs. standardised geospatial demand

Just because Number of cases in LSOA1 is double that of LSOA2 – the rate is the same. So take care when presenting data!

LSOA	No. Cases
1	100
2	50

Regional borders and edge cases

- If you are analysing an area it is possible demand will drop off as you approach the region's border i.e., pop^{n} from edge of Cornwall, people may be travelling to Devon
  - Data to confirm this may be held by another provider
- You need to assess how influential these 'low demand' border areas are on your analysis and recommendations
  - Not always easy!
  - Sensitivity analysis that excludes low demand areas can help. *Would this change the configuration to be recommended?*
- What happens if you place a new facility near to a border?
  - Is it possible that it may attract new demand not seen before? Known to happen in health care!
  - Can you source extra data? From another NHS provider, or model somehow (i.e. demographic info and rates per age-band within a geographic region)
  - Can you estimate this demand and, would it make a difference to the analysis and recommendations?

Geospatial analysis and complexity

What happens with big problems in optimisation?

- You can’t bruteforce problems when they get big
- Facility locations are an example of combinatorial optimisation problems
- These problems are **NP-Hard** *(i.e., can’t be solved in a reasonable time frame)*
  - For medium to large instances of these problems it is not possible to solve them to optimality.
- But computing power is increasing right?
- Example: Location problem with 2^n combinations
  - Using Summit Supercomputer
  - @200 petaFLOPS (2×10^{17})

n	combinations	Time
10	1024	0 seconds
60	1.15×10^{18}	6 seconds
70	1.18×10^{21}	1.6 hours
120	$1.33 \text{ X } 10^{36}$	210 billion years

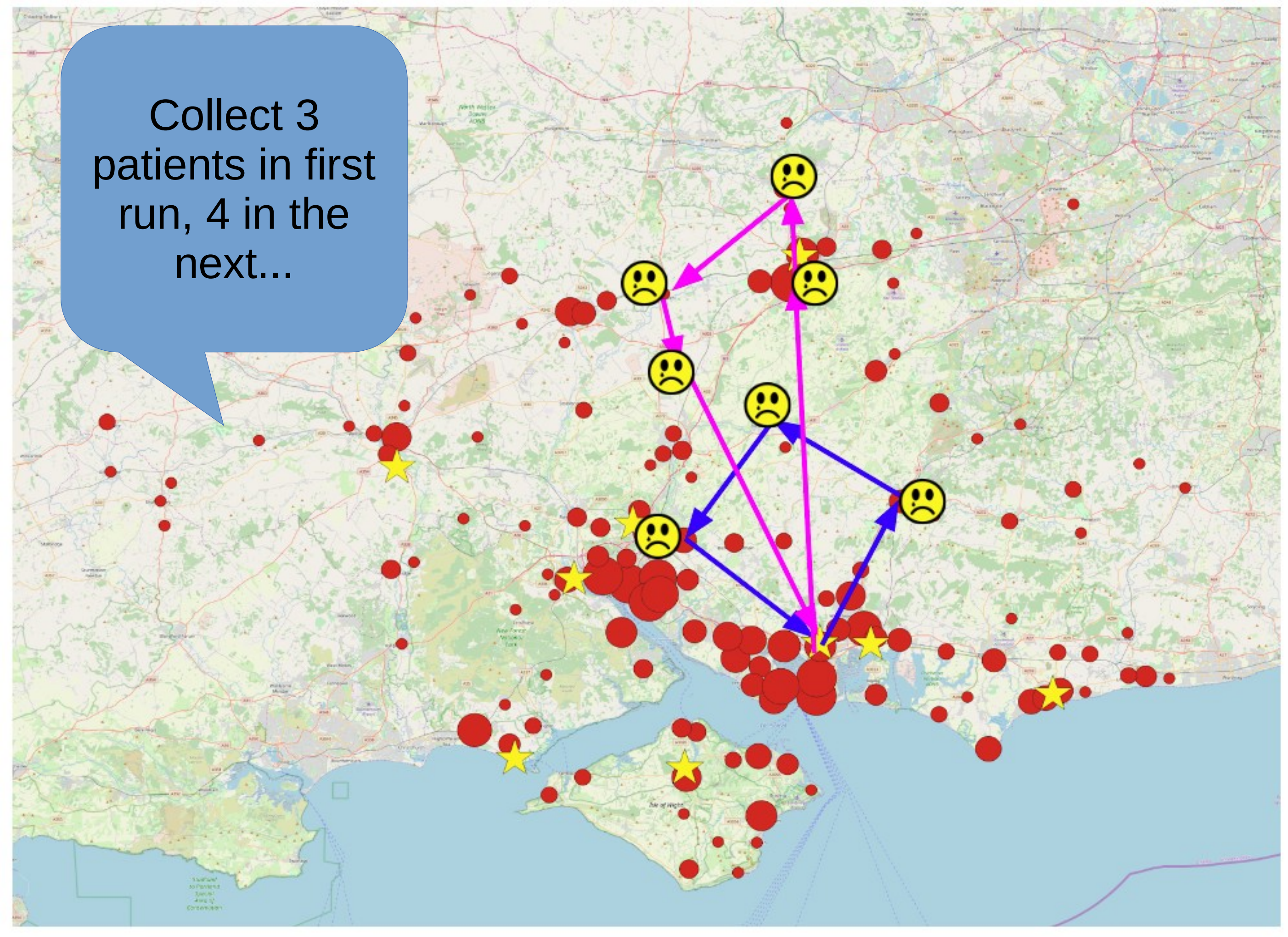
Discussion (30 mins) + Feedback (10 mins)

In your groups discuss any location-allocation problems in your organisation?

- What are the locations
- Who accesses these locations (the patient/community group?)
- What is the current problem
- What are the objectives to define the good solutions?
- What are potential “what-if” scenarios

Introduction to Routing & Scheduling

Routing and scheduling problems



Collect 3 patients in first run, 4 in the next...

Dialysis example again… Ambulance needs to collect patients from their home addresses

Need to consider timings of appointments, etc


If I plan
it right...
·NURSING DISTRICT·
BS = Bedsore
C = Catheter
BED
SORE
REPORT
I can fill a report
between catheters
Lenen.

Coordinating Assets and Needs

List of Services

Service ID	Patient	Address	Time window	Service type	Other considerations
0001	Ann Patient	Road, Postcode	0900-1100	Insulin	Cat
0002	John Smith	Road, Postcode	0800-1700	Other treatment	2 Nurses
0003	John Smith	Road, Postcode	0800-1700	Insulin	At least 30 min after 0002
...	...	...	...	...	...

List of nurses

Nurse ID	Skills	Shift	On-call	Starts from	Preferences
0001	Insulin	Day	Yes	Home (address)	
0002	Blood test	Day	No	Hospital	No cats
0003	Blood test, Insulin	Evening	No	Hospital	
...	...	...	...	...	...

The Vehicle Routing Problem

- Determination of the optimal set of routes to be performed by a fleet of “vehicles” to serve a given set of customers
- It is one of the most important, and studied, combinatorial optimisation problems.
  - Vehicles → Nurses
  - Customers → Patients

Routing and scheduling modelling

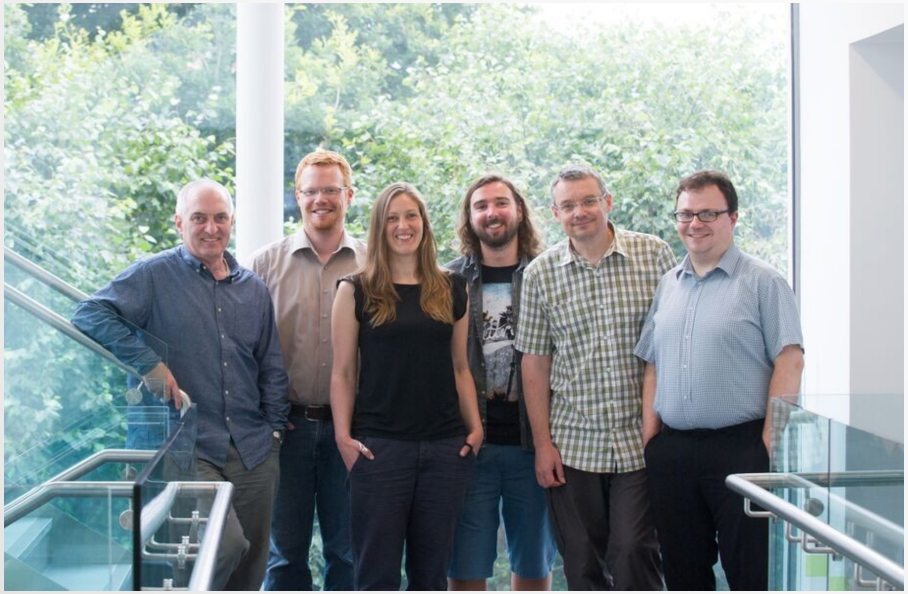
The health service perspective

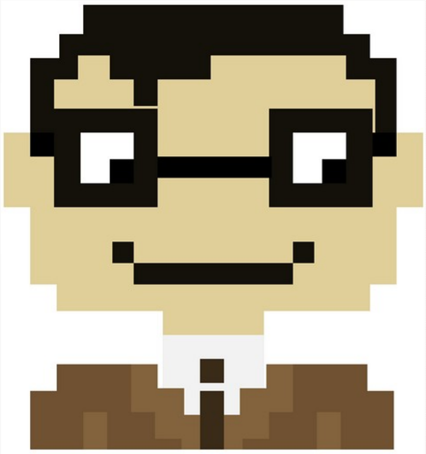
- Customers:
  - A health service provider organisation that must **routinely** deploy assets
- Classic applications:
  - Allocating and routing ‘community’ nurses to patients homes
  - Transporting patients to and from outpatient appointments
- Example complexities
  - Heterogeneous demand e.g. patients with different conditions/needs
  - Heterogeneous assets e.g. nurses with different skill sets
  - Availability of assets e.g. shift work
  - Time windows for pickup and appointments

The Travelling Sales~~man~~person Problem (TSP)

Can you help Alan the Analyst?!

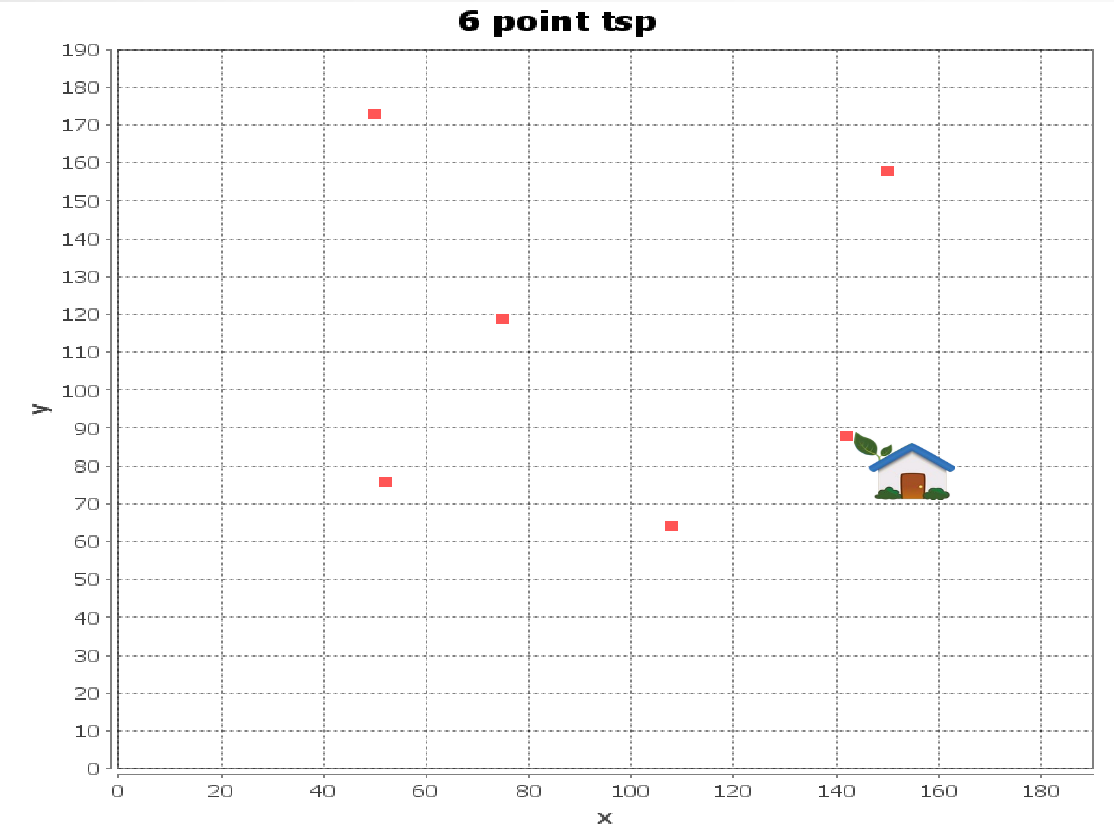
I am off to visit five of my fellow PenCHORD team members, but there is a problem…

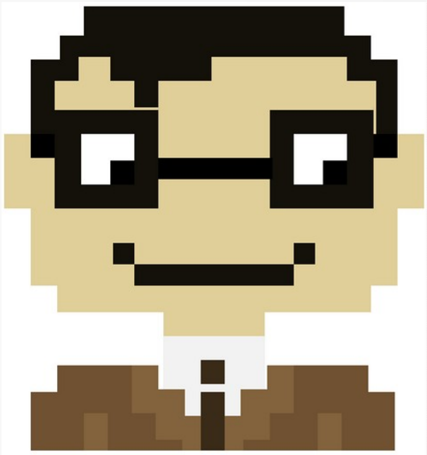




Can you help Alan the Analyst?!

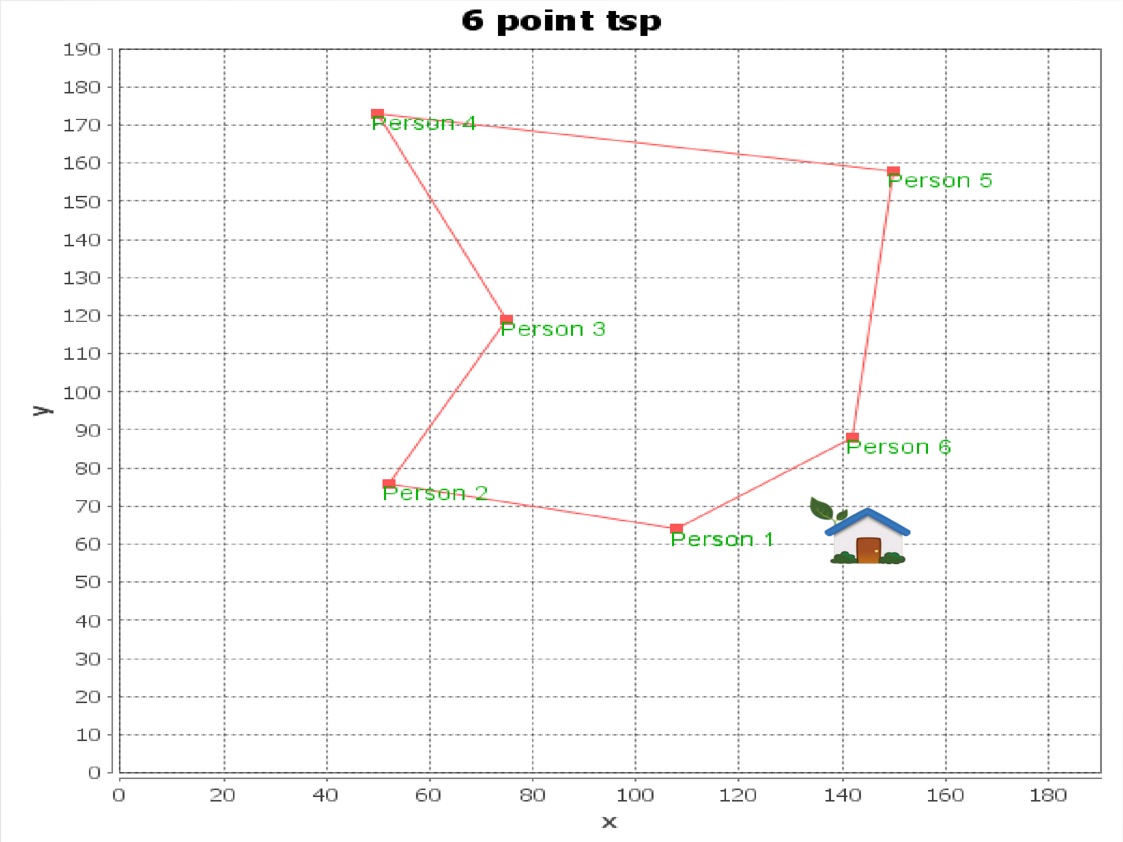
…I can see the locations of the team on my GPS along with my home location. But what is the shortest route to visit them. Please can anyone help me?

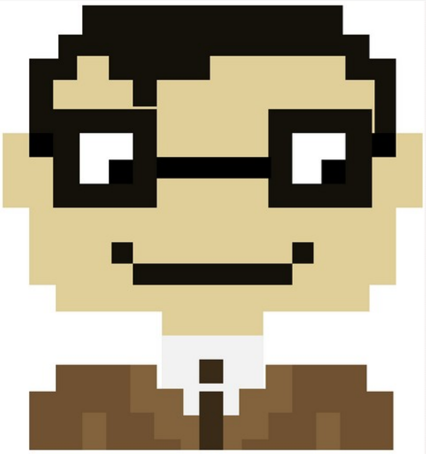




Can you help Alan the Analyst?!

Phew, thanks for that everyone, great work!!



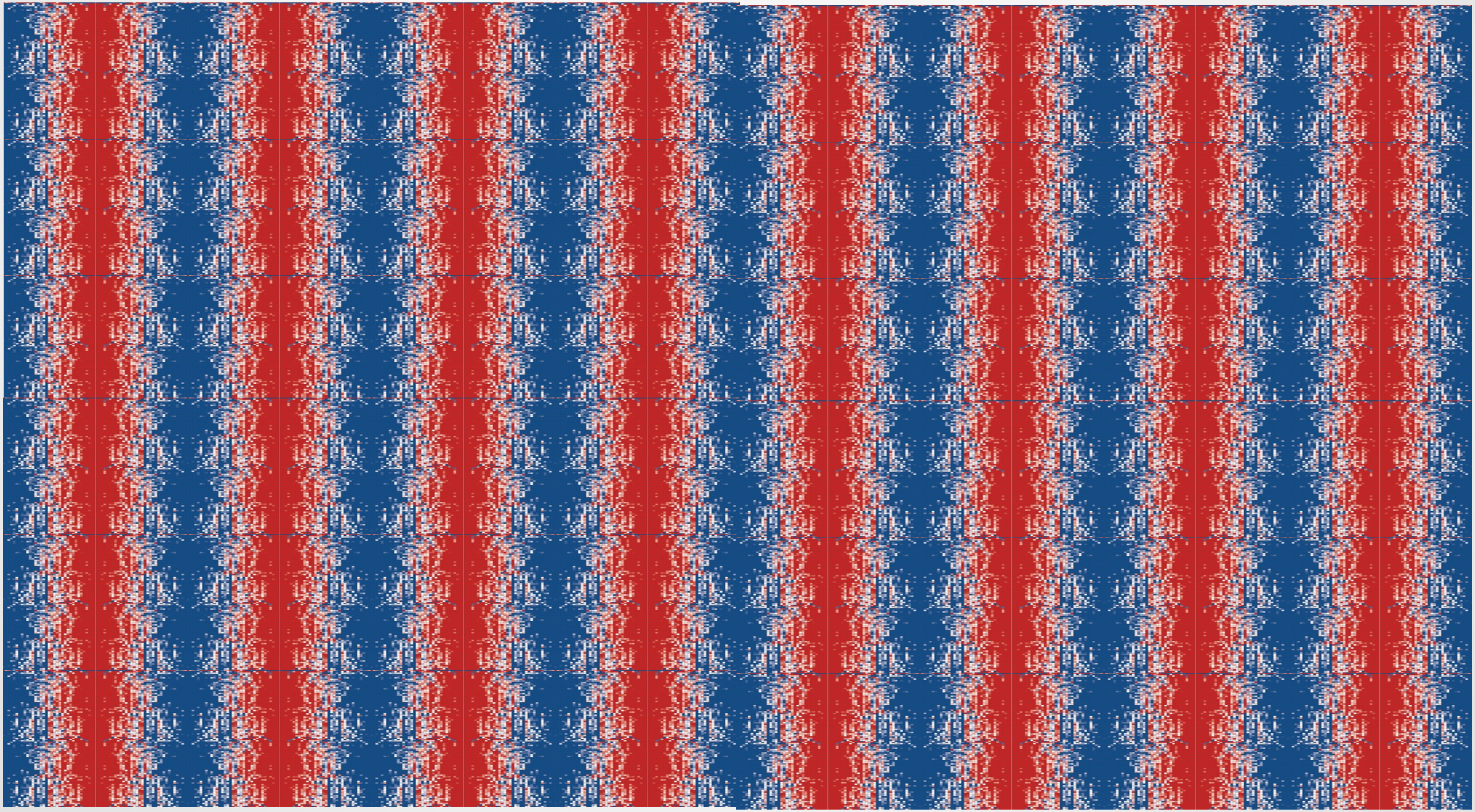


What is a solutions space?

- For simplicity, if we had 4 places to visit, there would be 4! = 24 possibilities for the sequence to visit Locations A, B, C and D (represented below).
- Each route, or combination will have a different distance
- **Low** to **High** (blue to red)

ABCD	**BACD**	**CABD**	**ACBD**	**BCAD**	**CBAD**
DBAC	**ADBC**	**BDAC**	**DABC**	**BADC**	**ACDB**
ABDC	**CADB**	**DACB**	**ADCB**	**CDAB**	**DCAB**
DCBA	**CDBA**	**BDCA**	**DBCA**	**CBDA**	**BCDA**

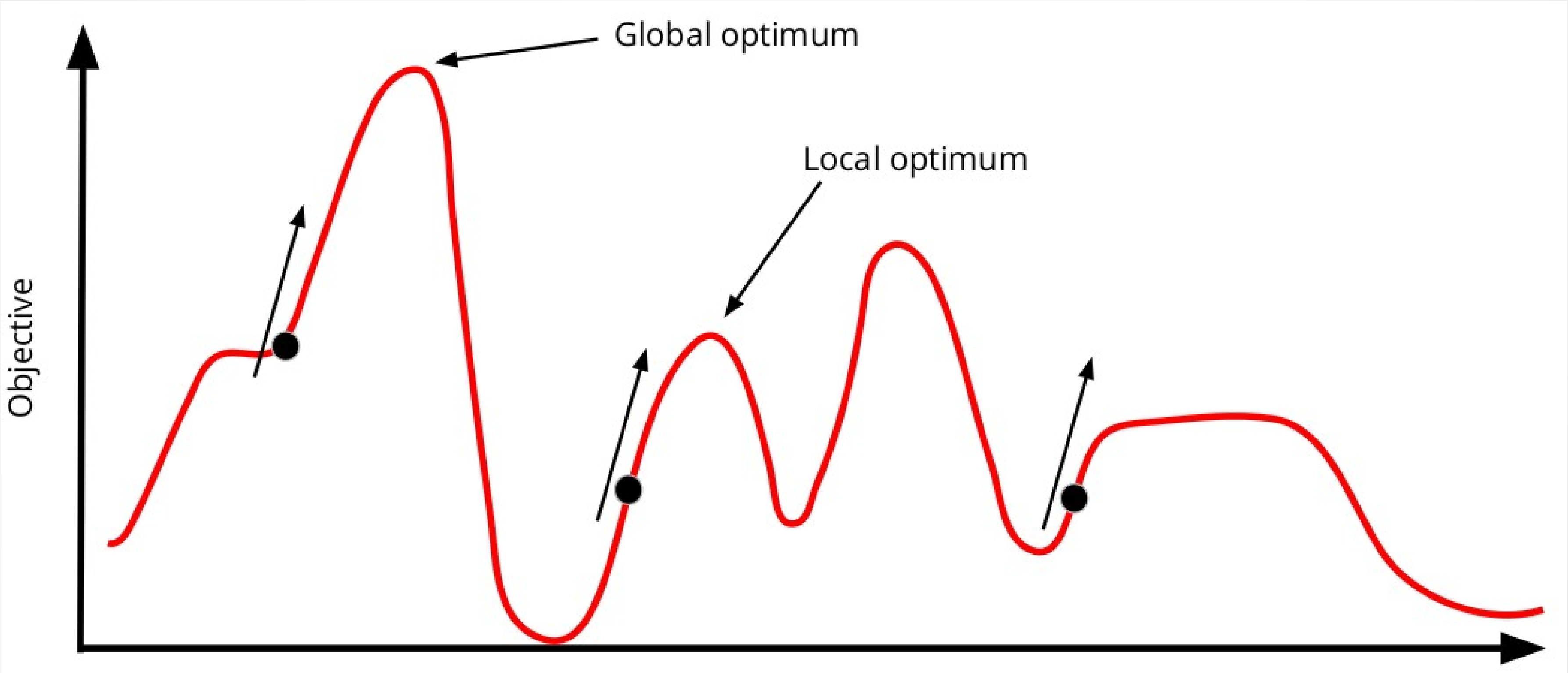
Macro View



What about more complex routes?

- One approach is a ‘hill-climbing’ method based on reversing portions of the route (or a ‘pairwise exchange’ approach).
- Hill-climbing methods perform some search that then leads to them picking the best improvement in that step of the search.
- The search method is then repeated until no further improvement is found.
- Hill-climbing methods may get trapped in local optima in complicated problems.
- The solution found depends on the starting point. To reduce the impact of this the algorithm is repeated multiple times with different start points.

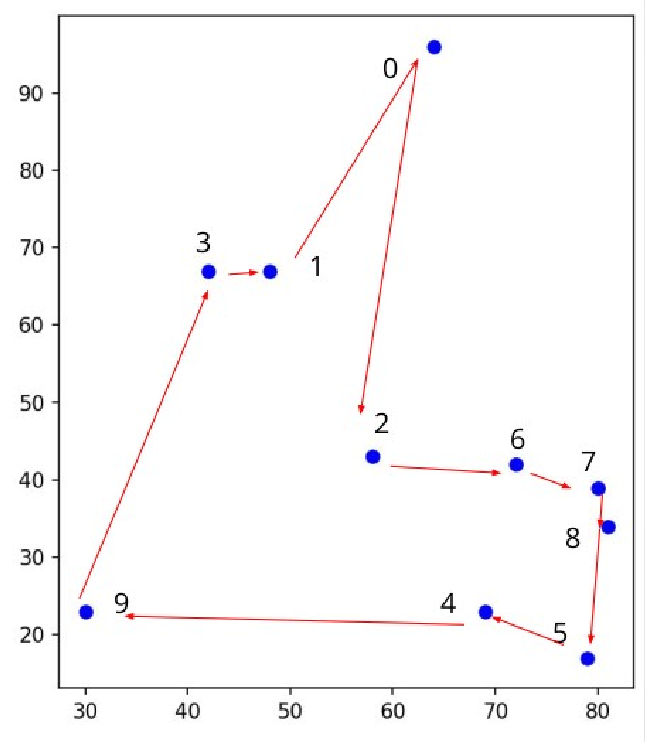
Hill Climbing Algorithms



Hill Climbing Algorithms Illustrated

Representation

0	2	6	7	8	5	4	9	3	1



Symmetric or Asymmetric

1	3	9	4	5	8	7	6	2	0
8	7	6	2	0	1	3	9	4	5

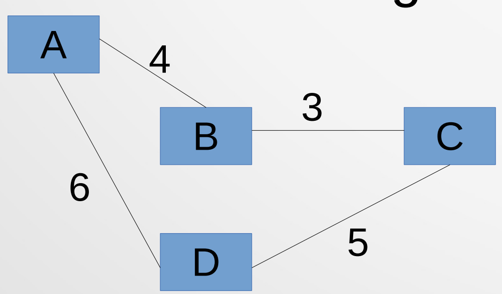
6	3	4	9	1	7	5	2	0	8

Overview of Pairwise Exchange

Aka “Simple Tweak”

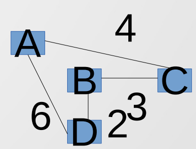
Steps to be performed...

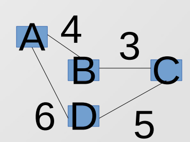
- **Step 1**. Pick a route at random (randomly order points to visit).
- **Step 2**. Examine all possible pairwise exchanges in the route in (e.g. if we have a route A-B-C-D-E-F-G-H, and pairwise exchange **C** and **F** we have a new route A-B-F-E-D-C-G-H). Choose the pairwise exchange that gives the best reduction in measurement metric i.e., distance or travel times).
- **Step 3**. Repeat Step 2. until no further improvement is observed.
- **Step 4**. Repeat from Step 1 for a determined number of repeats (or repeat until a maximum algorithm use time is reached).

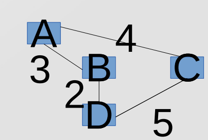


Route: A – B – C – D – A
Distance: 18

Route	Distance
A – C – B – D - A	15
A – D – C – B - A	18
A – B – D – C - A	14



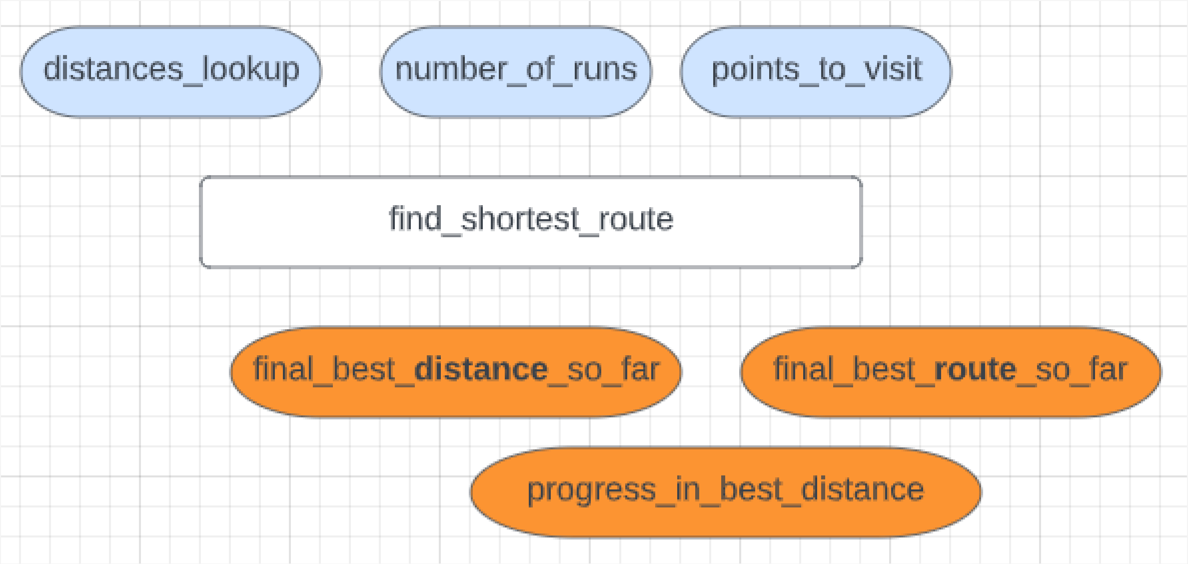




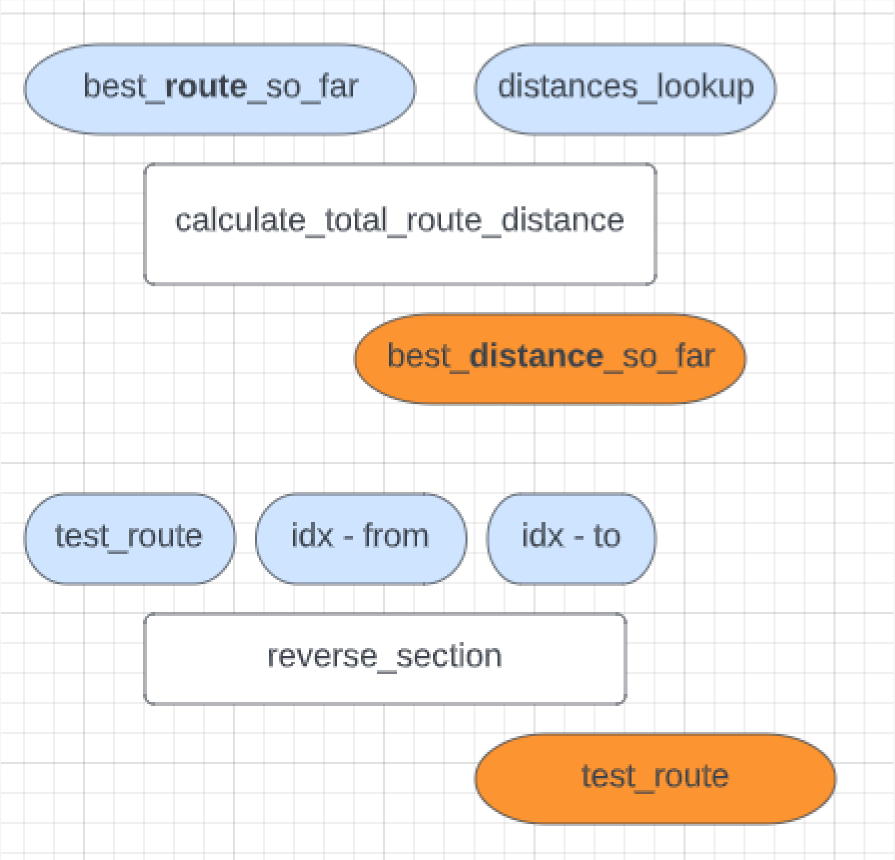
Considerations for Approach

- The distances or travel times between **all** points in the route must be known.
- There are a choice of distances to consider such as Euclidean (i.e., straight line), or actual (which would require appropriate software).
- Travel times will vary i.e., private car Vs. public transport.
- **Remember** this **won’t** look at every possible combination of approaches

Example Illustrated with Code (1)



Example Illustrated with Code (2)



Exercise 1

- The details of 8 of the largest airports in England & Scotland (***airports.csv***)
- In PSGs, open ***5A_tsp_airport.ipynb*** and review the code used for the functions of described in the previous slides and see how the optimal route between the airports is calculated
- The GeoPy library is used to calc distances between different points (lat/ long)
- Remember there are 9! combinations (362,880) however the optimal route is quite obvious.
- We'll regroup in **40** minutes.

Exercise 2

- In PSGs, open ***5A_tsp_delivery_driver.ipynb***
- Take a look at the following functions:
  - *load_points*
  - *create_dictionary_of_distance*
  - *print_results*
  - *plot_route*
  - *main_code*
- We’ll regroup in **20** minutes.

Closing Remarks

- Healthcare geospatial problems are largely split between **routing and scheduling** and **location-allocation**
- Often a lot of benefit for healthcare customers comes from providing straightforward analysis of a problem as well as visualisation via maps (sessions 5C & 5D)
- Your work will often form part of a wider evidence base used to make a decision - particularly in **location-allocation**.