

# Geographic Data Science - Lecture I

## Introduction

Dani Arribas-Bel

# Today

- This course
- The (geo-)data revolution
- (Geo-)Data Science

This course

*More stats than a GIS course, more GIS than a stats course...*

*...but in a fun way!*

# Structure

11 weeks of:

- **Prep. materials:** videos, podcasts, articles... 1h. approx.  
(most recommended!)
- **1h. Lecture:** concepts, methods, examples
- **2h. Computer practical:** hands-on, application of  
concepts, Python (highly *employable*)
- **Further readings:** how to go beyond the minimum

# Philosophy

- (Lots of) **methods** and techniques
  - General overview
  - Intuition
  - Very little math
- Emphasis on the **application**
- Close connection to "real world" applications
- **FUN**

# Assignments

- Mark based on two assignments:
  1. Week 5 (50%)
  2. Week 9 (50%)
- Coursework
- Equivalent to 2,500: report with **code, figures** (e.g. maps), and **text**
- Suggested applications, BUT feel free to pick your own (and talk to me)!

# The (geo-)data revolution

# The (geo-)data revolution

Exciting times to be a:

- Geographer
- Map fan
- Data fan

The world is being "**datafied**"...

# "Datafication"

*Quantification of phenomena through the systematic recording of data*

*“taking all aspects of life and turning them into data” Cukier & (Mayer-Schoenberg)*

Examples: credit transactions, public transit, tweets, facebook likes, spotify songs, etc.

Many implications:

- Opportunities for optimization of systems
- **window** into human **behaviour** (this course)
- issues with **intentionality** and **privacy**
- ...

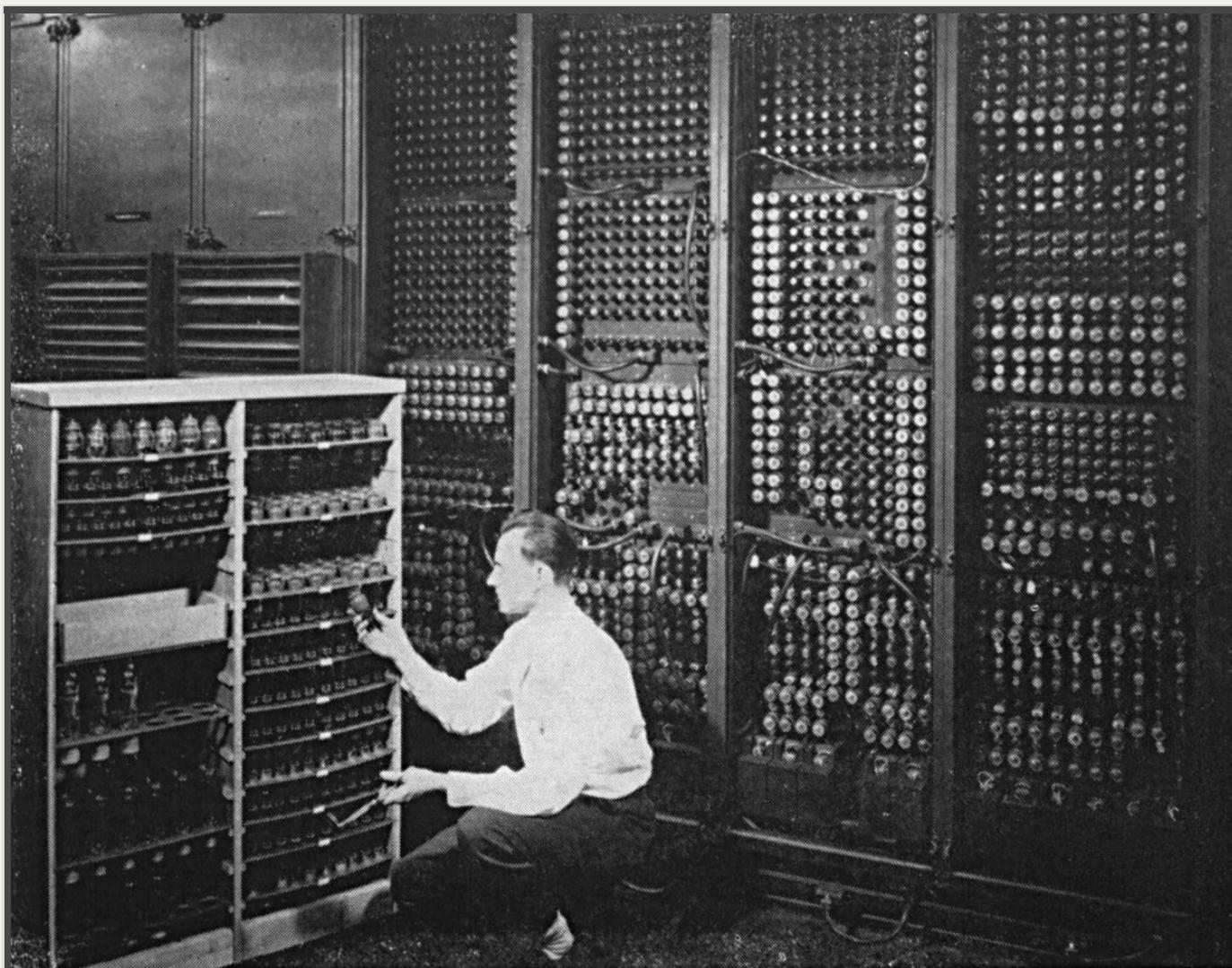
*Why now?*

Advances in:

- Computing power
- Communication
- Geospatial technology

*Why now? --> Computing power*

Source



Replacing a bad tube meant checking among ENIAC's 19,000 possibilities.

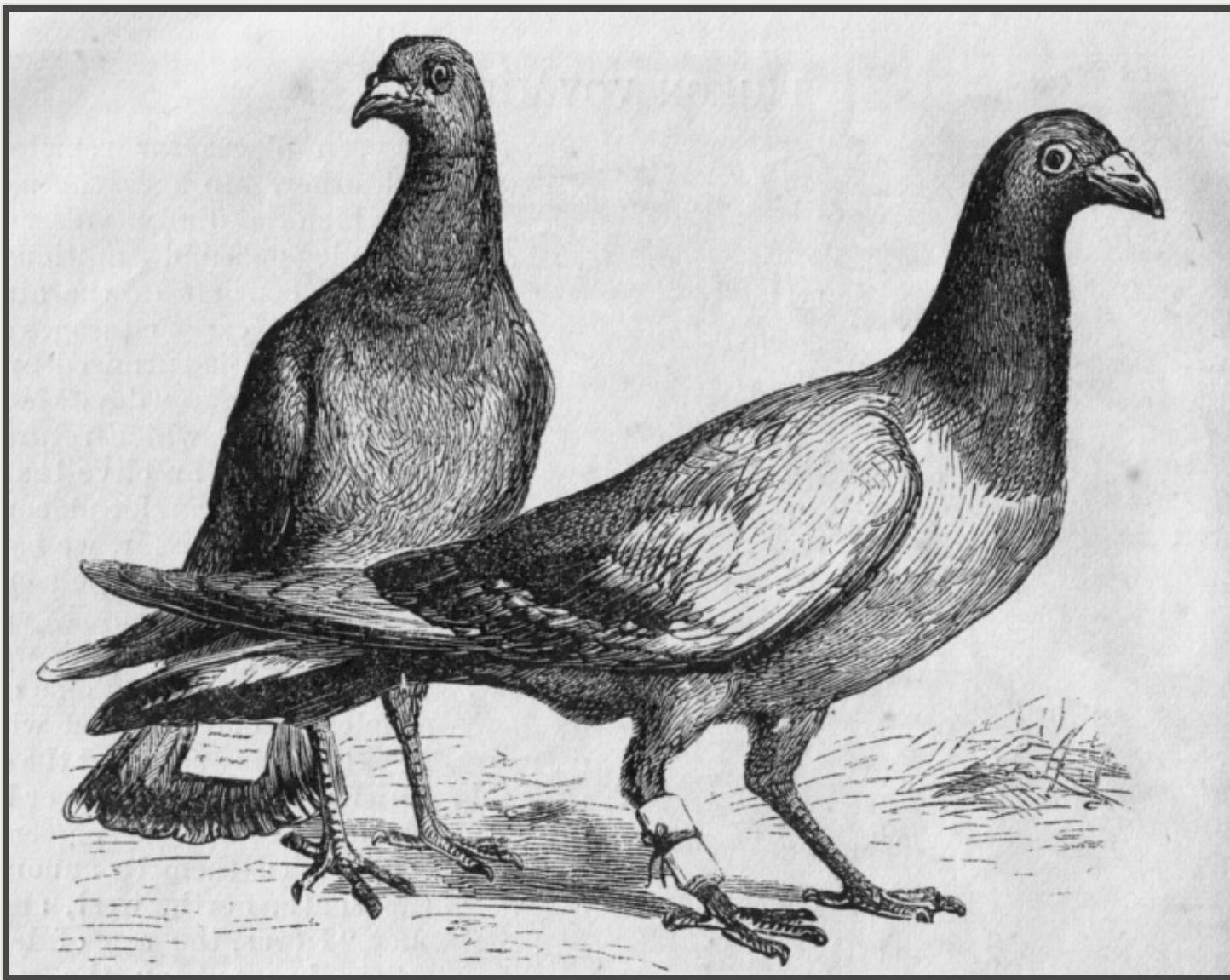
*Why now? --> Computing power*

Source



*Why now? --> Communication*

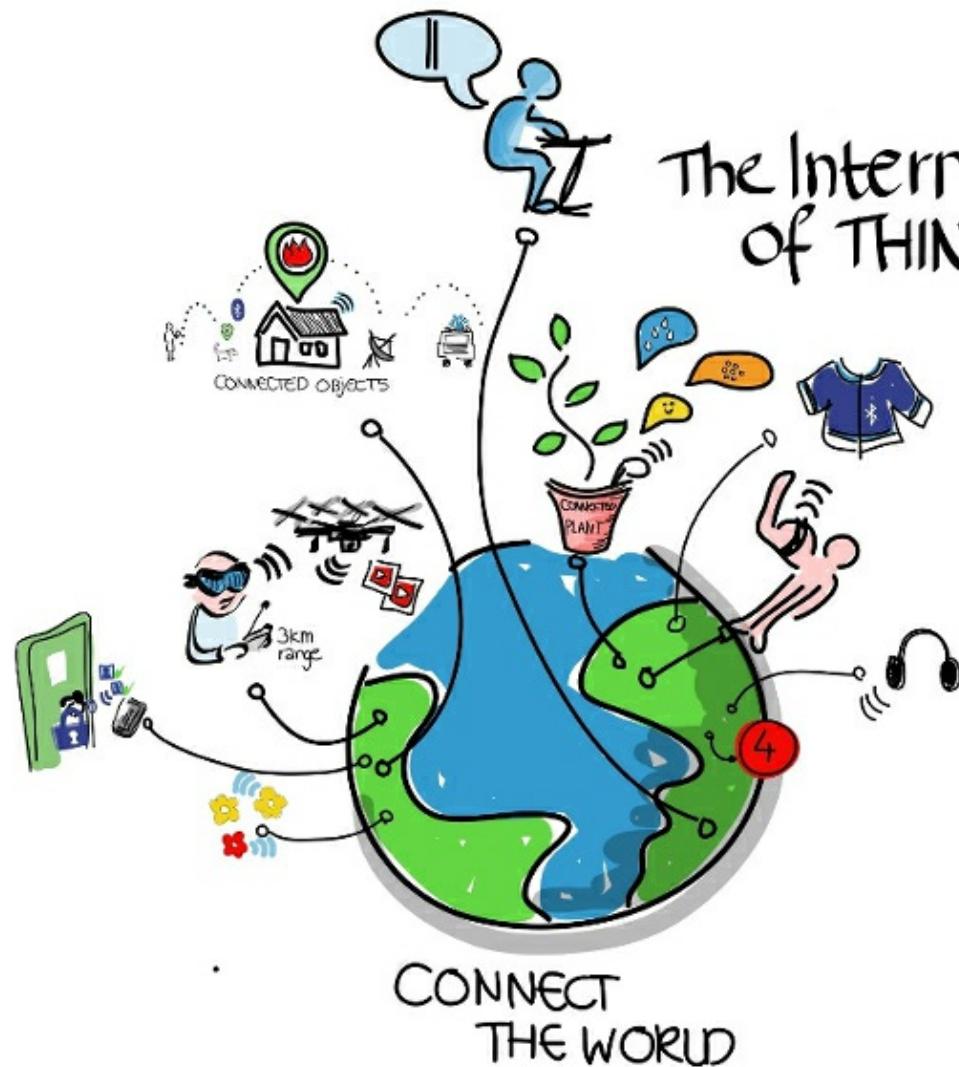
Source



*Why now? --> Communication*

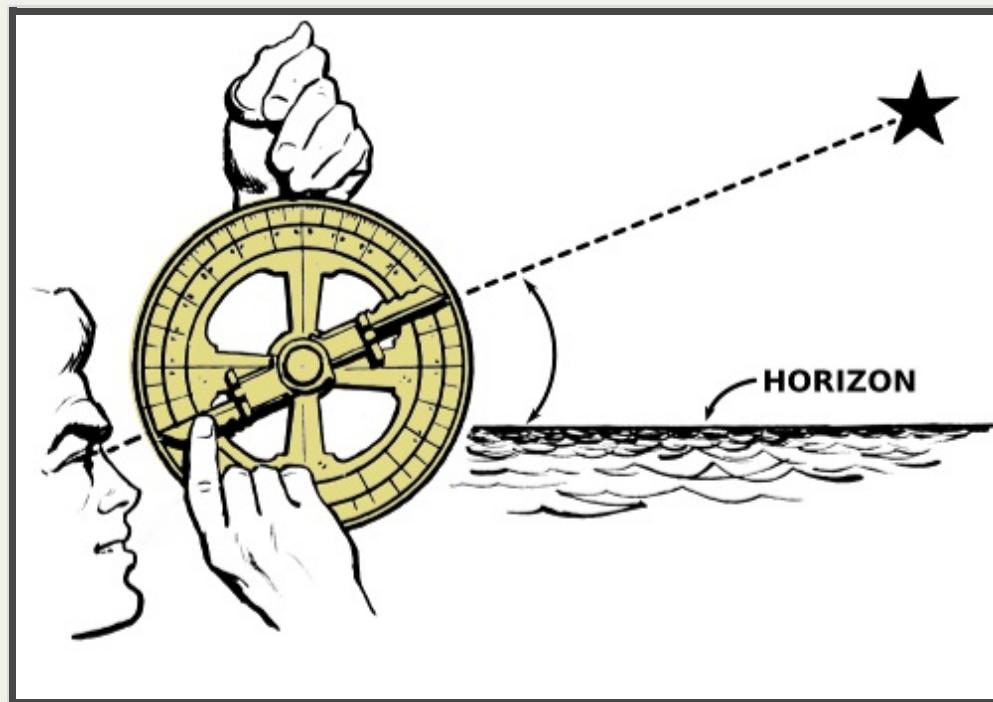
Source

# The Internet of THINGS



*Why now? --> Geospatial technology*

Source





*Why now? --> Geospatial technology*

Source





## The (geo-)data revolution

The confluence of the three (computing, communication and geospatial) is creating large amounts of data.

Now, data in itself is not very valuable:

- Data --> Information --> Knowledge --> Action

# Data Science

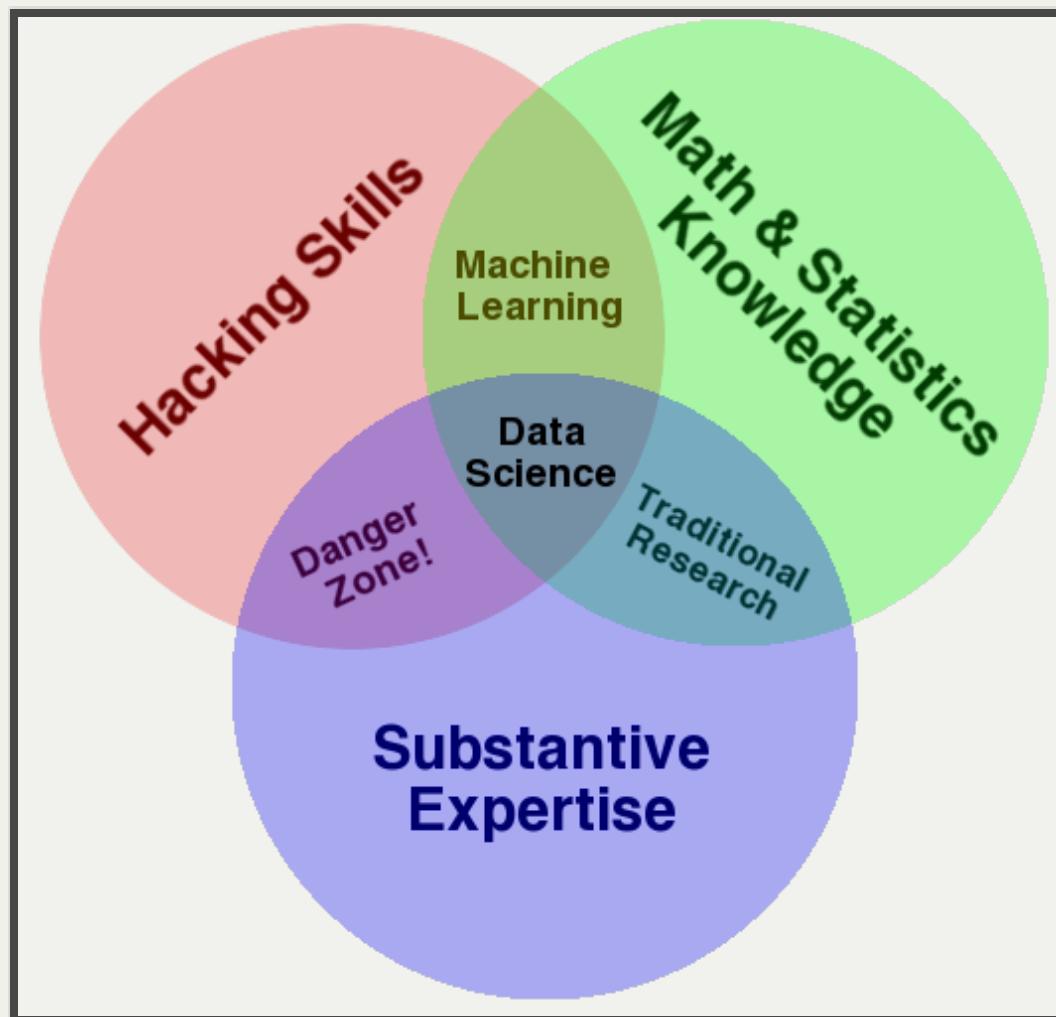
*Methods, tools and techniques to turn data into actionable knowledge*

But wait, isn't statistics just that?

Not only...

# Data Science

Source: Drew Conway





# Data Science

Statistics is a *very* important part of DS...

... but not the only one:

- Computational tools --> Programming (hence this course's tutorials!)
- Communication skills --> "Story telling" (hence this course's assignments)
- Domain expertise --> Theories about why the data are the way they are (hence the rest of your degree)

# Data Science

- Not all new (standing on the shoulders of giants)
- "The data becomes key part in the product"
- Focus on actionability and solving particular problems

Some examples...



Total price: £32.97

Add all three to Basket

*i* These items are dispatched from and sold by different sellers. [Show details](#)

- This item:** Green and Black's Organic Dark Chocolate 85 Percent Cocoa 100 g (Pack of 5) £11.62 (£2.32 / 100 g)
- [Green and Black's Organic Ginger Dark 100 g \(Pack of 5\) £10.40 \(£2.08 / 100 g\)](#)
- [Green and Black's Organic Dark Chocolate Maya Gold 100 g \(Pack of 5\) £10.95 \(£2.19 / 100 g\)](#)

## Amazon

### Customers Who Bought This Item Also Bought



Green and Black's Organic  
Ginger Dark 100 g (Pack of



Green and Black's Organic  
Dark Chocolate Maya Gold



Green and Black's Organic  
Dark Chocolate 100 g



Vivani Organic Dark  
Chocolate with 85% Coco



Join the best free dating site on Earth.

I am a

Straight

Woman

Continue

## Dating sites



Signing up takes two minutes and is totally free.

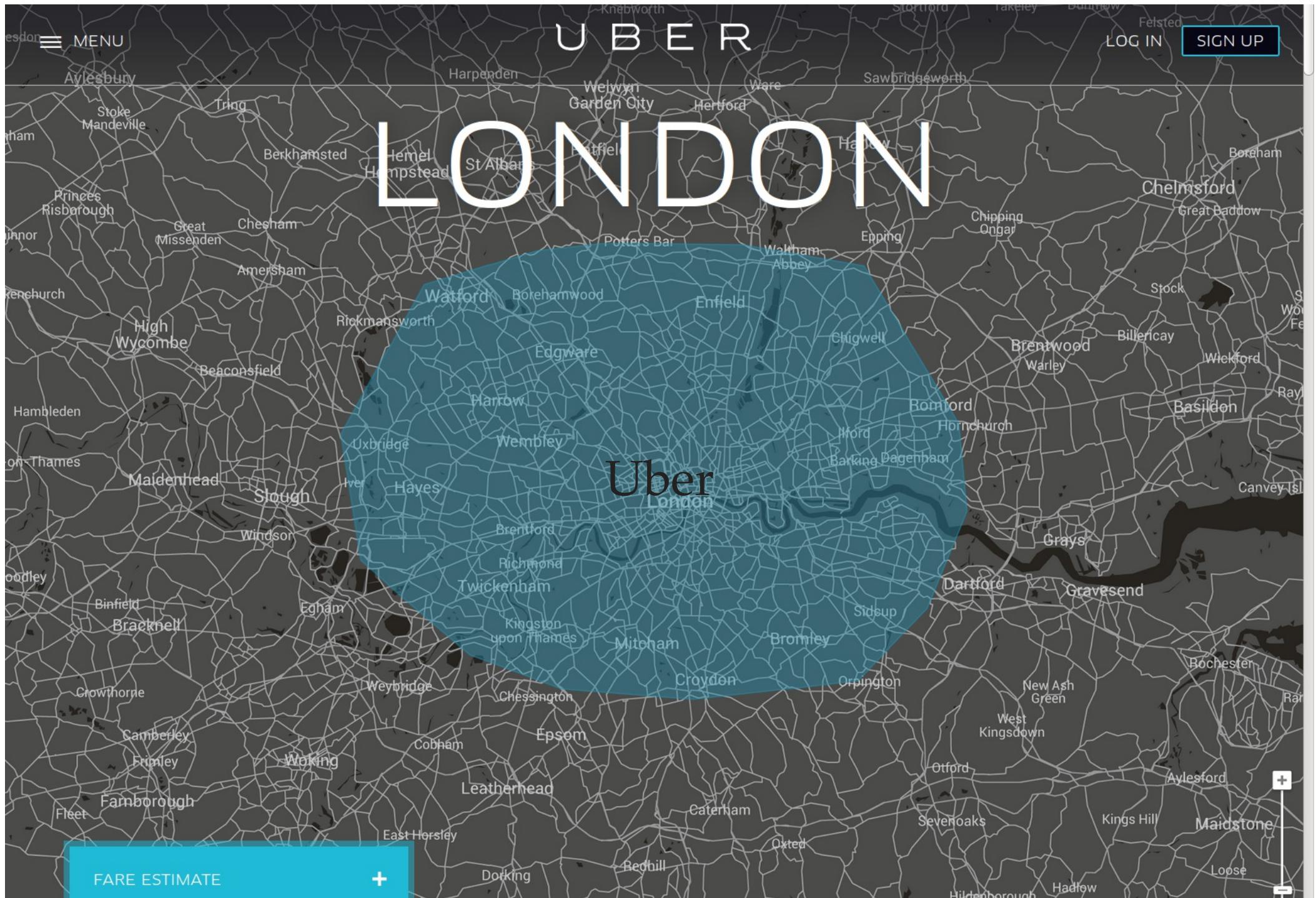


Our matching algorithm helps you find the right people.



iOS or Android?  
You can take us to go.

# U B E R L O N D O N





## Geo-Data Science

- A (very) large portion of all these new data are inherently **geographic** or can be traced back to some location over space.
- Spatial is special.
- Some of the methods require an explicitly spatial treatment --> (Geo-)Data Science

Some examples...



Where are you going?

Browse ▾

Sign Up Log In Help ▾ List Your Space

Cities &gt; London

Saved Neighbourhoods 0

## Find a Neighbourhood in London

What kind of neighbourhood are you looking for?

Dining ×

Artsy 7

Nightlife 9

Shopping 15

Loved by Londoners 11

Touristy 14

Great Transit 13

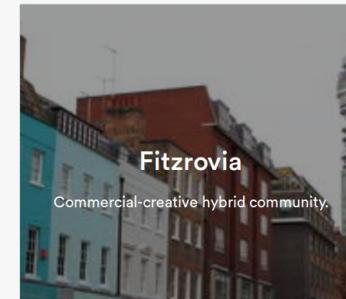
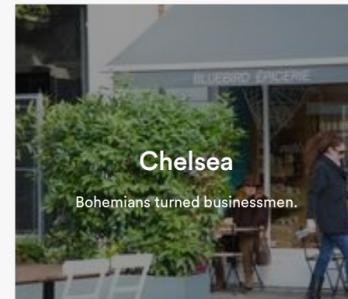
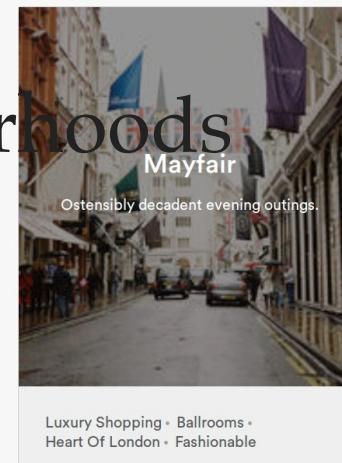
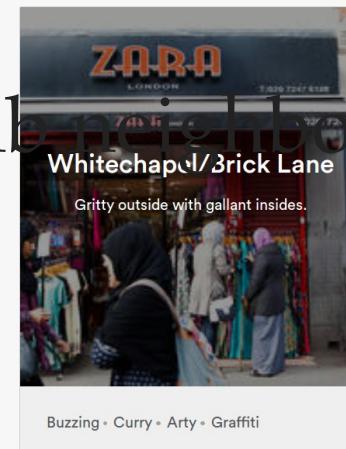
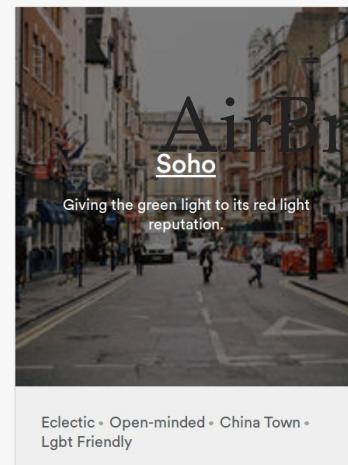
Museums &amp; the Arts 10

Peace &amp; Quiet 7

Along the River Thames 5

West End 7

23 neighbourhoods match Dining. [See all listings](#)



Penny Ln, Live... +

<https://www.google.co.uk/maps/dir/Penny+Ln,+Liverpool+L18+1DE,+UK/University+of+Liverpool,+Liverpool/@53.3974201,-2.9788596,13z>

Walking directions Print X

行人 via A562 49 min  
2.4 miles  
[Show terrain](#)

Use caution - may involve errors or sections not suited for walking

**Penny Ln**  
Liverpool L18 1DE, UK

↑ Head southwest on Penny Ln toward Russell Rd  
79 ft

↗ Turn right onto Russell Rd  
0.2 mi

↖ Turn left onto Smithdown Rd/A562  
1.3 mi

↖ Slight right to stay on Smithdown Rd/A562  
① Continue to follow A562  
364 ft

↗ Turn right toward Smithdown Ln  
325 ft

↑ Continue onto Smithdown Ln  
0.2 mi

↖ Turn left to stay on Smithdown Ln  
0.5 mi

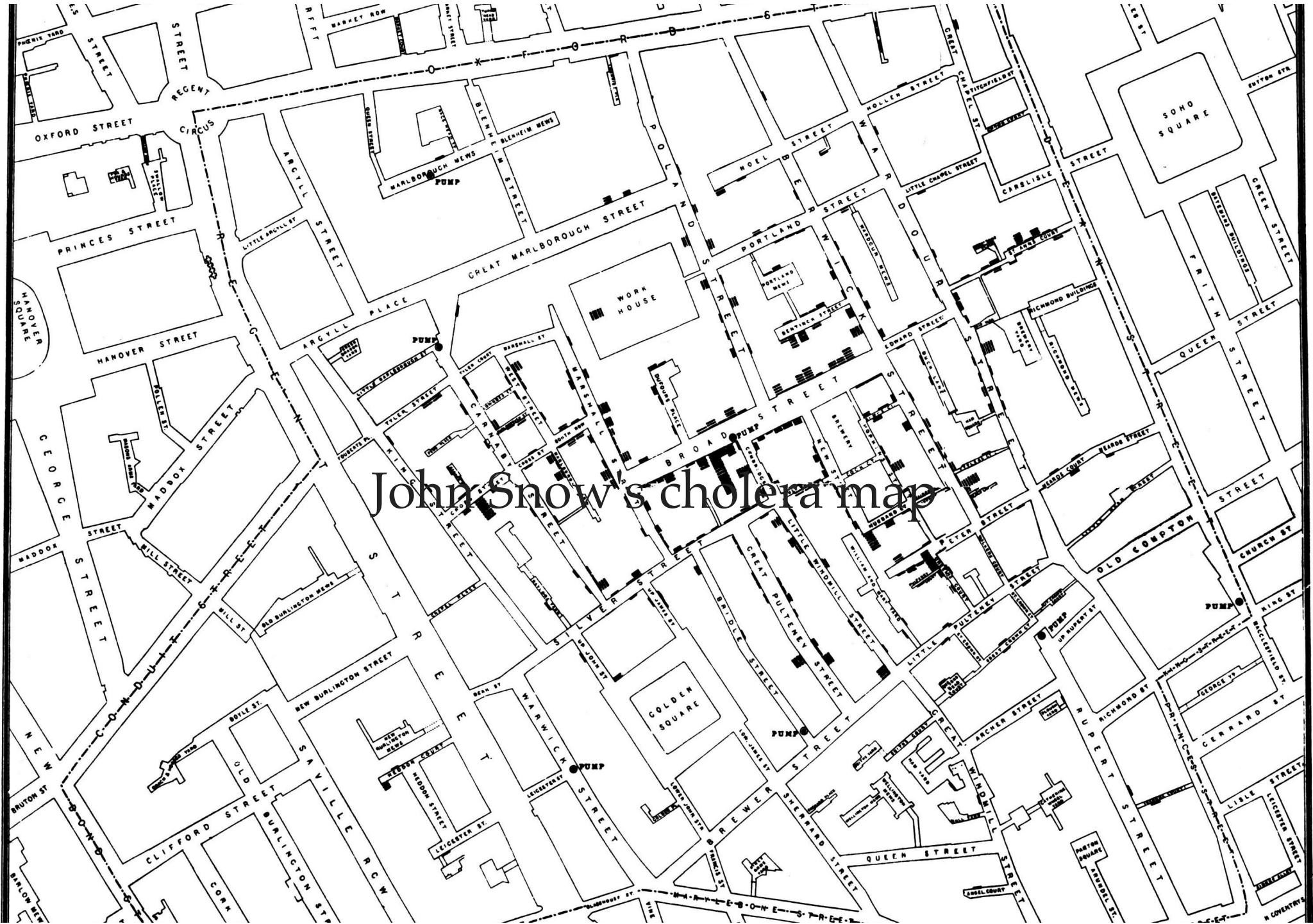
↑ Continue straight onto Brownlow Hill/A5048  
① Continue to follow Brownlow Hill

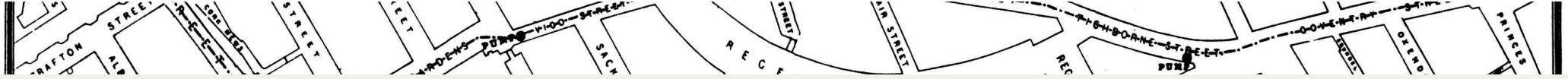
57 min  
2.8 miles

50 min  
2.5 miles

49 min  
2.4 miles

# Google Maps routing





Geographic Data Science'15 - Lecture 1 by Dani Arribas-Bel  
is licensed under a Creative Commons Attribution-  
NonCommercial-ShareAlike 4.0 International License.