

Geographic Data Science - Lecture I Introduction

Dani Arribas-Bel

Today

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

This course

Quiz

- Can you think of a real-world context where data and statistics are being used to make a difference?
- Have you ever heard the term "**Big Data**"?
- Have you ever heard the term "**Data Science**"?
- Have you ever written a line of computer code?

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

...but in a fun way!

Philosophy

- (Lots of) methods and techniques
 - General overview
 - Intuition
 - Very little math
 - Lots of ways to continue on your own
- Emphasis on the application and use
- Close connection to "real world" applications

Logistics – Website

<http://darribas.org/gds17>

GDS17



ENVS363/563

Geographic Data Science

Welcome to Geographic Data Science, a course taught by Dr. Dani Arribas-Bel in the Autumn of 2017 at the University of Liverpool.

The timetable for the course is:

- **Lectures:** *Mondays* 9:00am–10:00am, **SCTH-MR**
- **Computer Labs:** *Mondays* 11:00am–1:00pm, **ENG-HHTC**

Locations

- **SCTH-MR**: South Campus Teaching Hub, Main Room [[Map](#)]
- **ENG-HHTC**: Harrison Hughes Building (Engineering), Computer lab (top floor) [[Map](#)]

Logistics - Format

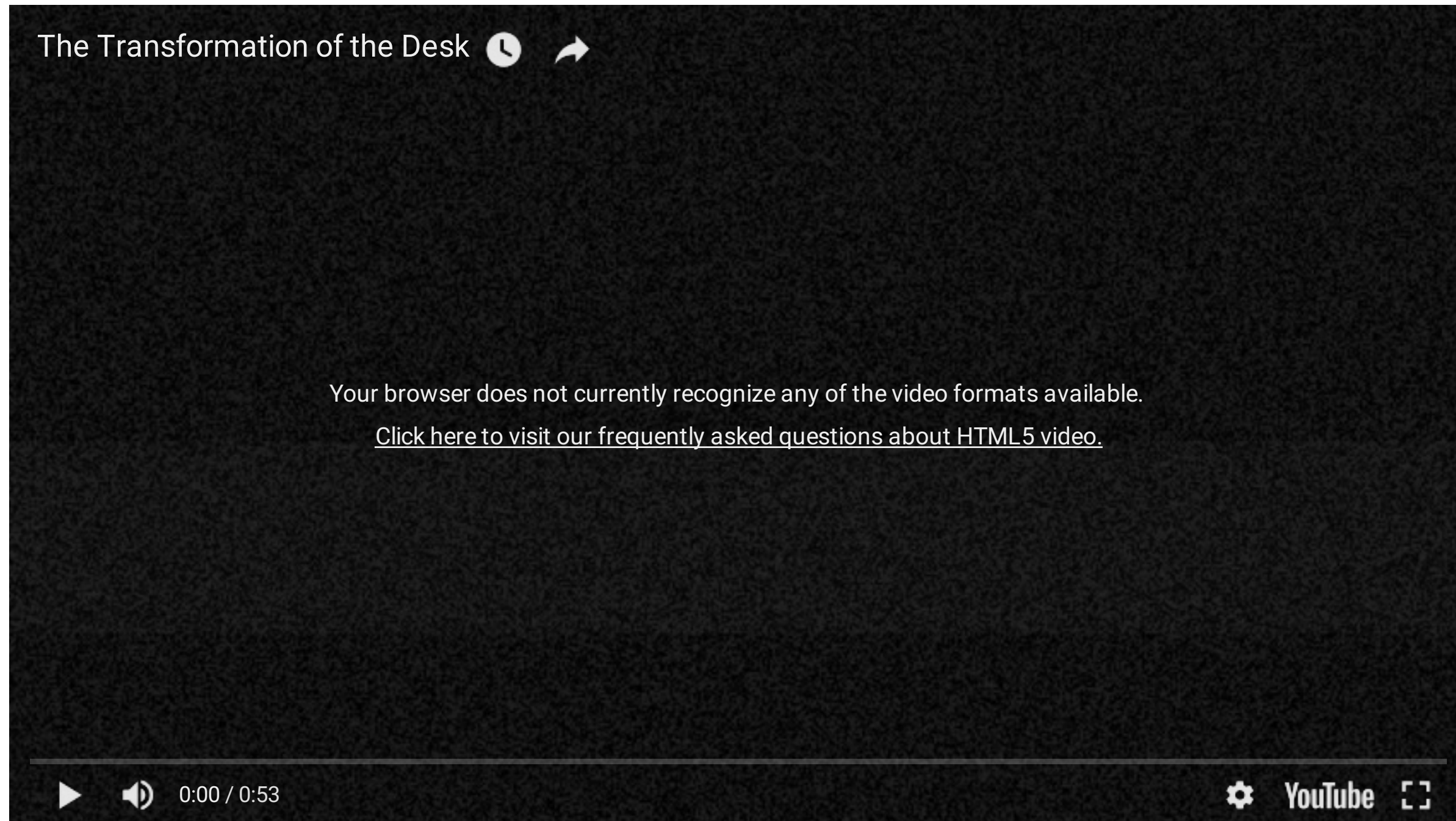
11 weeks of:

- **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

Logistics – Content

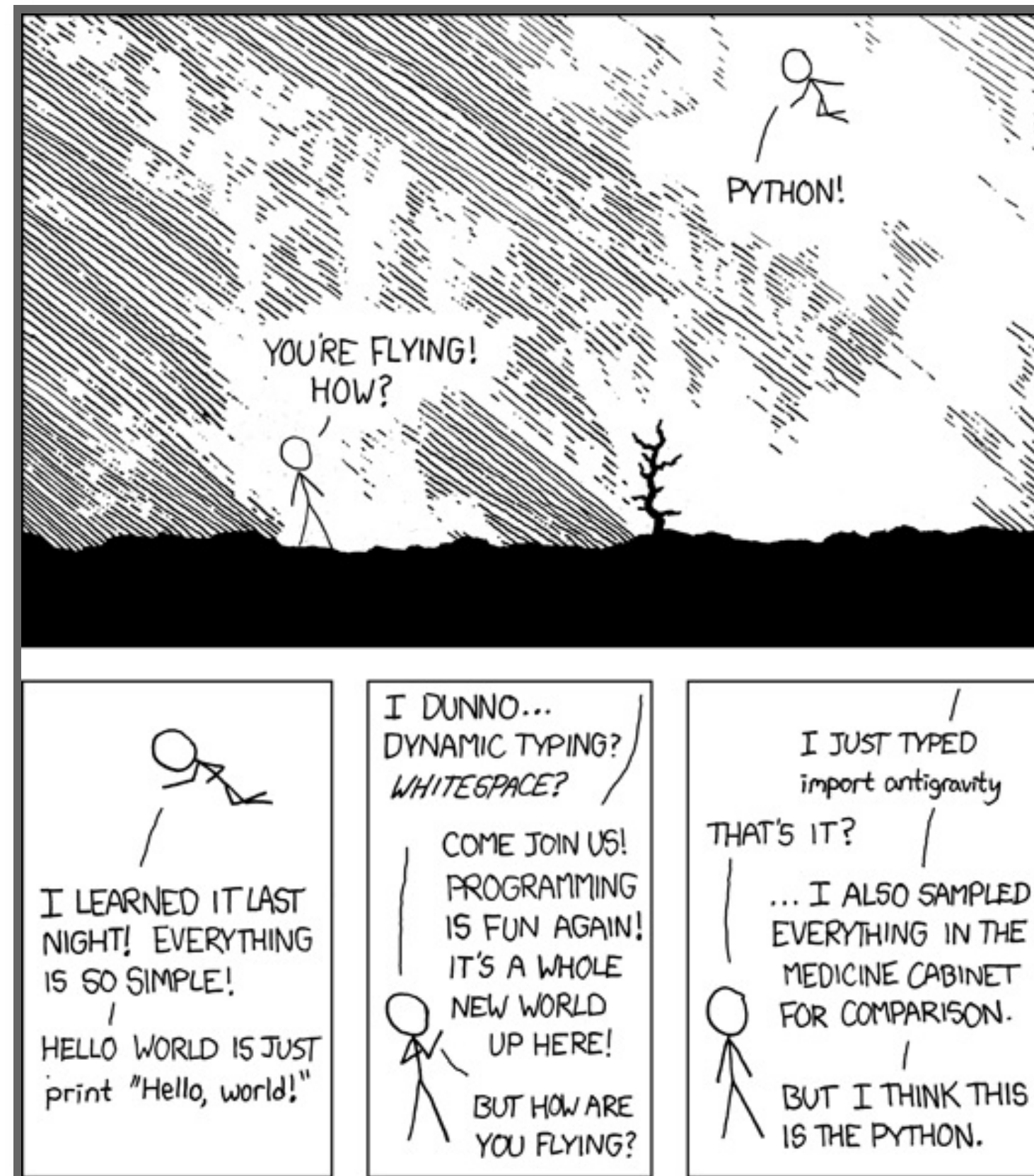
- Weeks 1–3: "big picture" lectures + introduction to computational tools (learning curve)
- Weeks 4–8: "meat" of the course (lots of concepts packed)
- Weeks 9–11: catch up + prepare an awesome Assignment II

Code



"Even if you won't be a poet, you need to know how to write"

Python



Python

- **General purpose programming language**
- Sweet spot between *"proof-of-concept"* and *"production-ready"*
- Industry standard: **GIS** (Esri, QGIS) and **Data Science** (Google, Facebook, Amazon, Netflix, The New York Times, NASA...)

Self-directed learning

Prepare for the labs

- I won't be leading/lecturing at the computer labs
- **Go over the notebooks** before the lecture and the computer lab --> If the first time you see a notebook is at the lab, you won't be able to follow on
- **Bring** questions, comments, feedback, (informed) rants to class/labs
- Use the **forum** (link on VITAL)
- **Collaborate** (it's **NOT** a zero-sum win!!!)

More help!!!

This course is much more about "learning to learn" and **problem solving** rather than acquiring specific programming tricks or stats wizardry

- Learn to **ask** questions (but don't expect exact answers all the time!!!)
- **Help** others as much as you can (the best way to learn is to teach)
- **Search** heavily on Google + Stack Overflow

Assignments

- Mark (mostly) based on two assignments, due:
 - Week 7 (40%), Week 12 (55%)
 - Coursework
 - Equivalent to 2,500: report (*notebook*) with code, figures (e.g. maps), and text
- Discussion board (5%)

NOTE: recommendation letters only for great students (>70)

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.

"Datafication"

Many implications:

- **Window into human behaviour** (this course)
- **Opportunities for optimization of systems**
(Industrial IoT, planning systems...)
- **Issues with intentionality and privacy**
- ...

Why now?

Advances in:

- Computing power and storage
- Connectivity
- Geospatial technology

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:

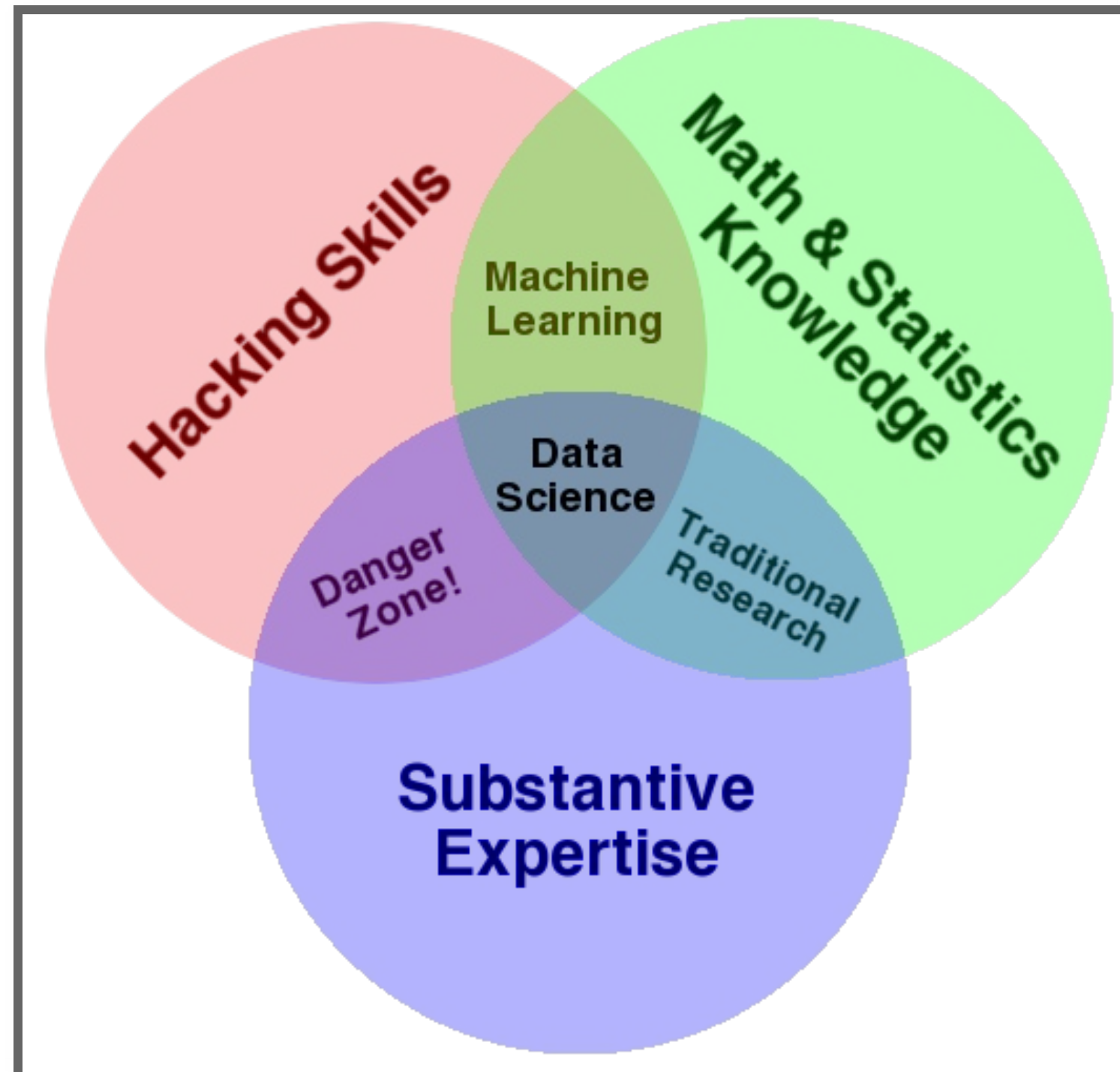
- Data --> Information --> Knowledge --> Action

Data Science

Methods, tools and techniques to turn data into actionable knowledge

Data Science

Source: Drew Conway



Data Science

Statistics + ...

- **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)

Some examples...

Frequently Bought Together



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)

Customers Who Bought This Item Also Bought



Green and Black's Organic Ginger Dark 100 g (Pack of 5)

★★★★★ 15

£10.40



Green and Black's Organic Dark Chocolate Maya Gold 100 g (Pack of 5)

★★★★★ 5

£10.95 **Prime**



Green and Black's Organic Dark Chocolate 100 g (Pack of 5)

★★★★★ 22

£8.20 **Prime**



Vivani Organic Dark Chocolate with 85% Cacao 100 g (Pack of 5)

★★★★★ 25

£11.95 **Prime**

Have an account? [Sign in](#)



Join the best free dating site on Earth.

I am a



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.

"The Google of online dating" - *The Boston Globe*

[About OkCupid](#) | [Legal](#) | © Humor Rainbow, Inc.

Geo-Data Science

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

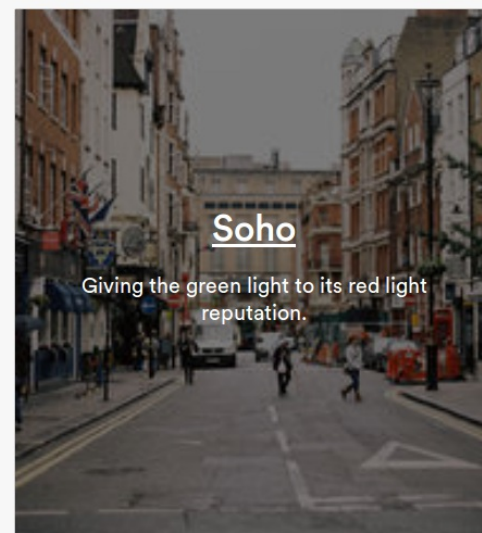
Some examples...

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 & the Arts 10
- Peace & Quiet 7
- Along the River Thames 5
- West End 7

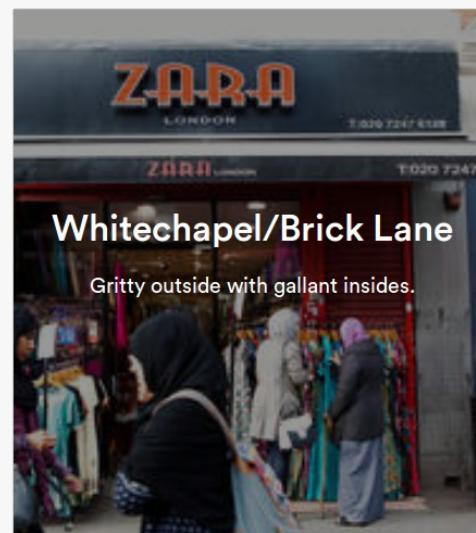
23 neighbourhoods match Dining. [See all listings](#)



Soho

Giving the green light to its red light reputation.

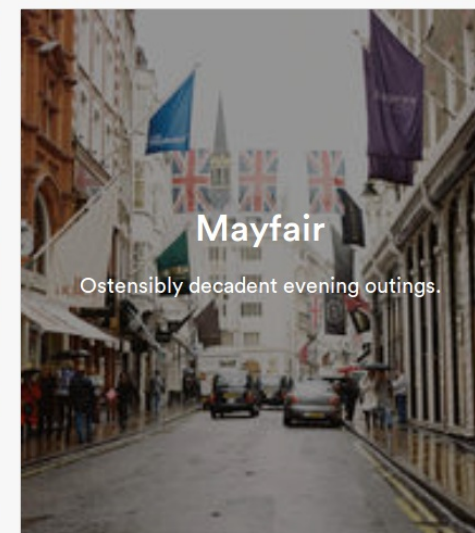
Eclectic • Open-minded • China Town • Lgbt Friendly



Whitechapel/Brick Lane

Gritty outside with gallant insides.

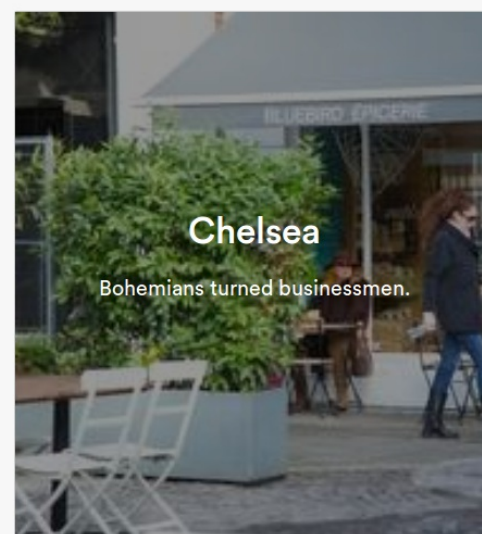
Buzzing • Curry • Arty • Graffiti



Mayfair

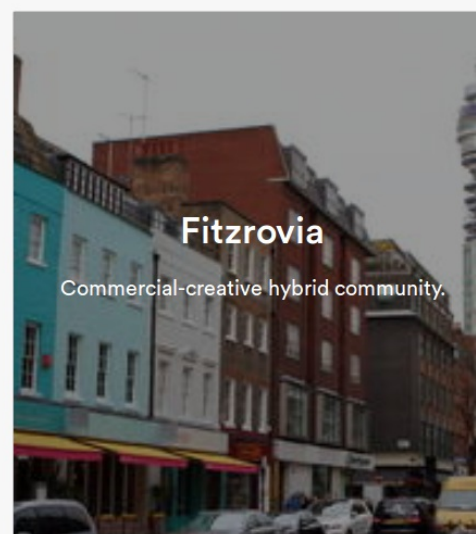
Ostensibly decadent evening outings.

Luxury Shopping • Ballrooms • Heart Of London • Fashionable



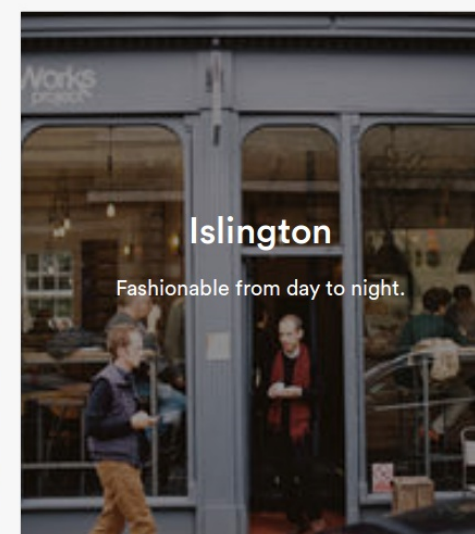
Chelsea

Bohemians turned businessmen.



Fitzrovia

Commercial-creative hybrid community.



Islington

Fashionable from day to night.



WANT TO DRIVE WITH UBER?

BECOME A DRIVER



MENU

LOG IN SIGN UP

UBER LONDON

FARE ESTIMATE +

LONDON X

Google Map data ©2015 Google TERMS OF USE

Walking directions

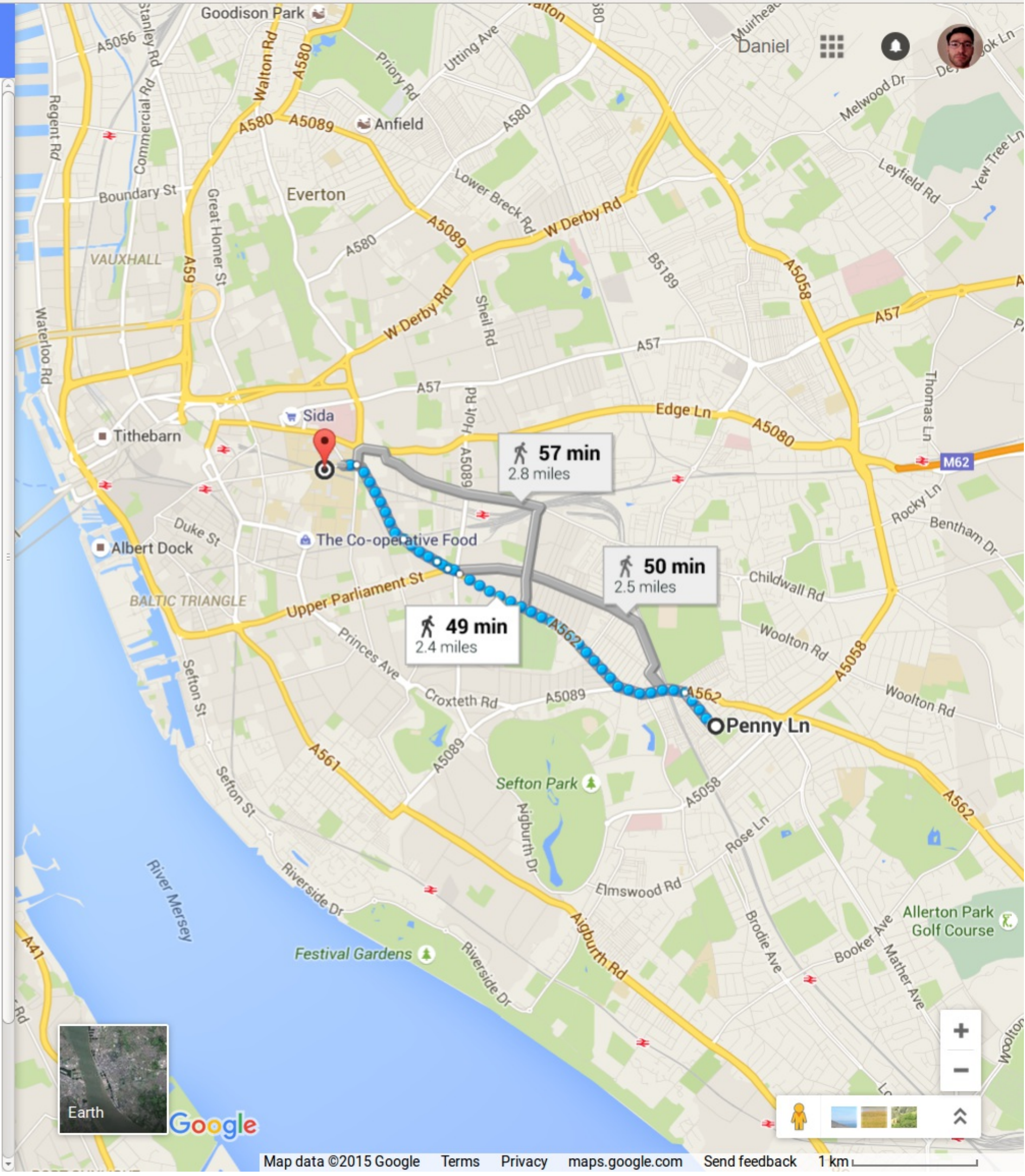
via A562 49 min 2.4 miles

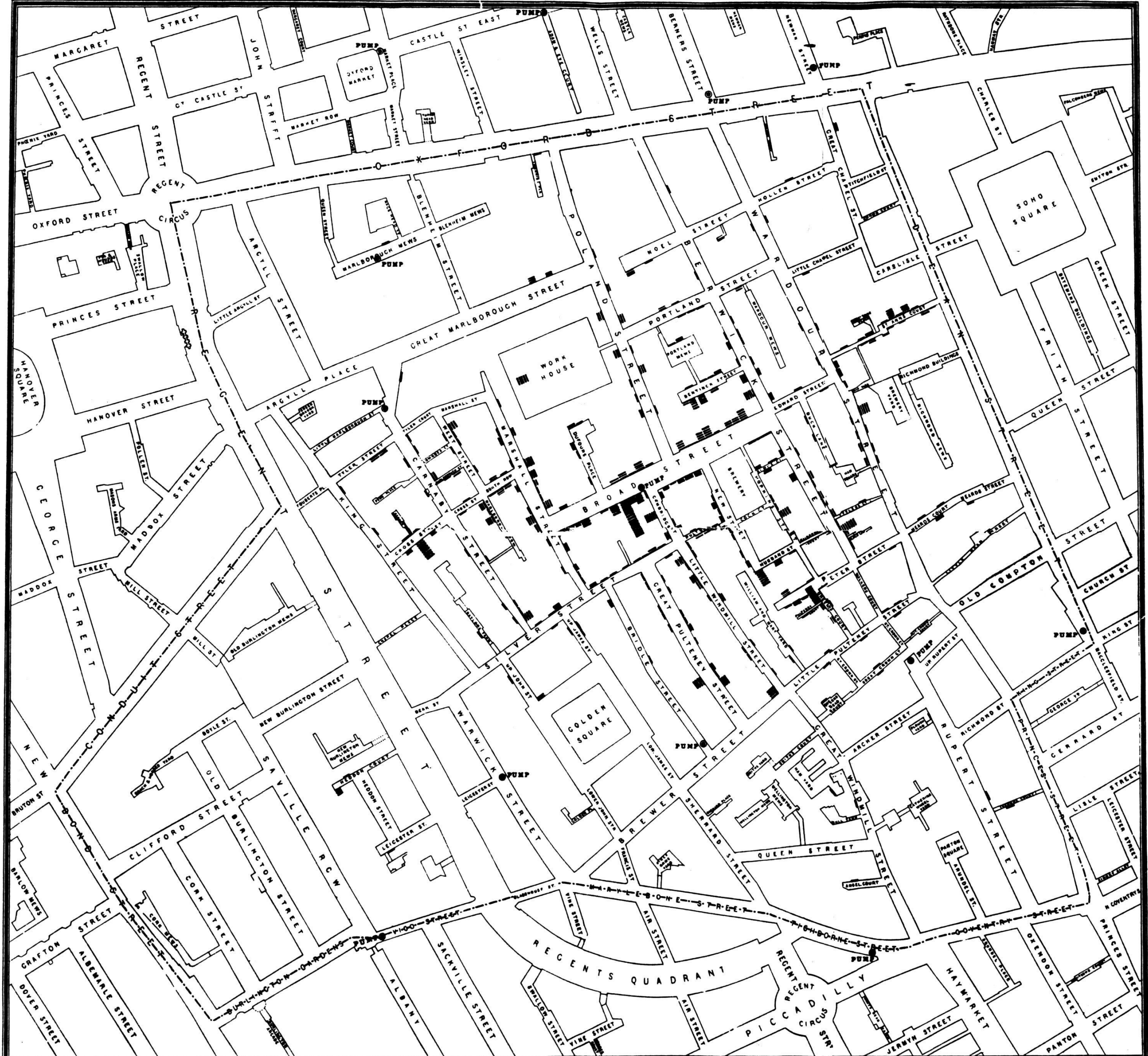
Show terrain Use caution - may involve errors or sections not suited for walking

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 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 Continue to follow Brownlow Hill 325 ft

University of Liverpool Liverpool, Merseyside L69 3BX, United Kingdom







Geographic Data Science'17 – Lecture 1 by Dani Arribas-Bel is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.