Geographic Data Science - Lecture I

Introduction

Dani Arribas-Bel

Today

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

This course

Quiz

- Have you ever used **data** to make decisions in your life?
- 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

With a few twists!

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

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

Content

- **Weeks 1-3**: "big picture" lectures + introduction to computational tools (learning curve)
- **Weeks 4-9**: "meat" of the course (lots of concepts packed) + Week 7 **break**
- **Weeks 10-11**: catch up + prepare an awesome Computational Esssay

Logistics - Website

http://darribas.org/gds19

GDS19

ENVS363/563

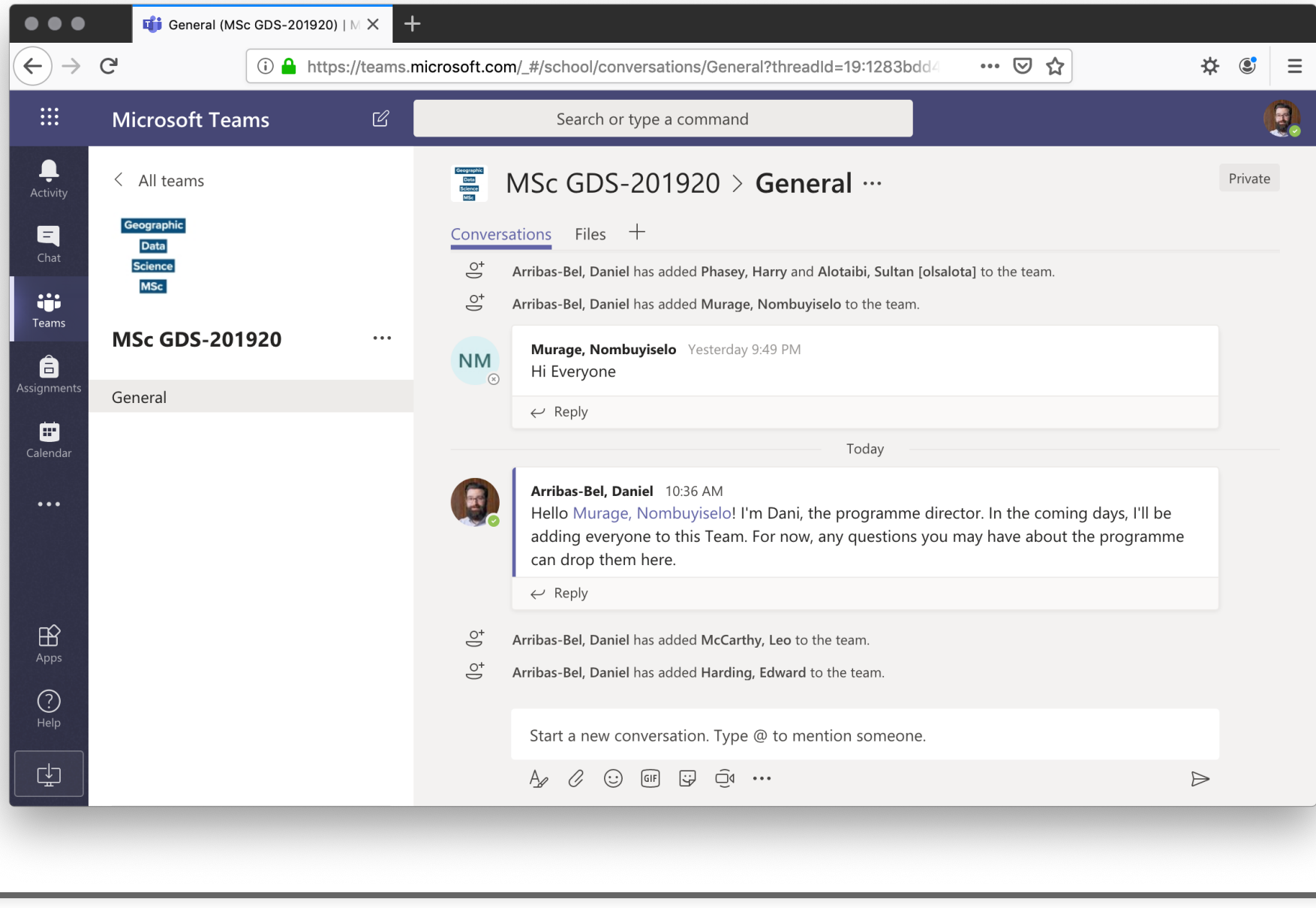
Geographic Data Science

Welcome to Geographic Data Science, a course taught by Dr. Dani Arribas-Bel in the Autumn of 2019 at the University of Liverpool.

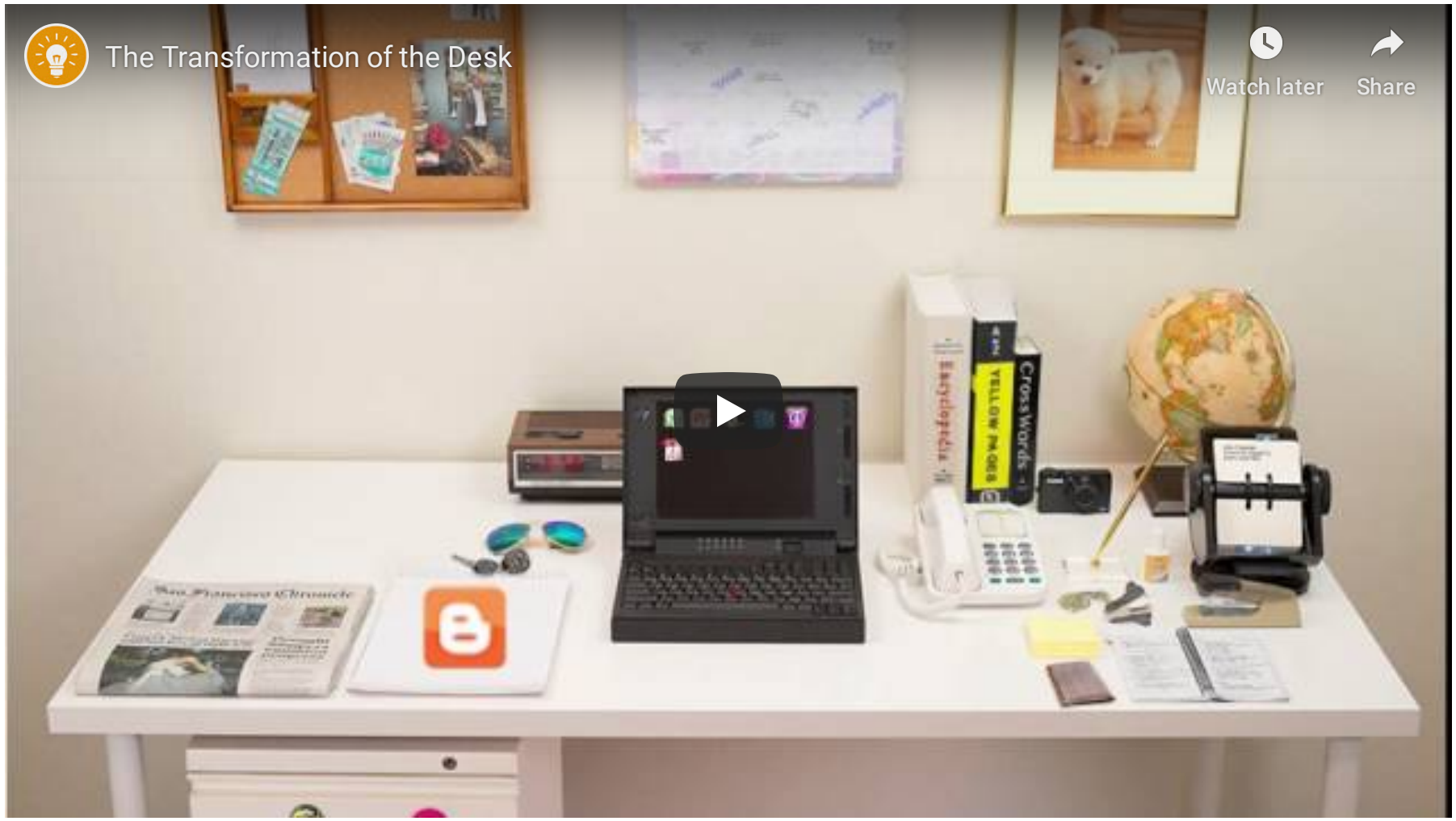
The timetable for the course is:

- **Lectures**: *Monday* 12:00pm-1:00pm, `ERB-ERT`
- **Computer Labs**: *Thursdays* 1:00pm-3:00pm, `CTL-6-PCTC`

Logistics - Teams [URL]

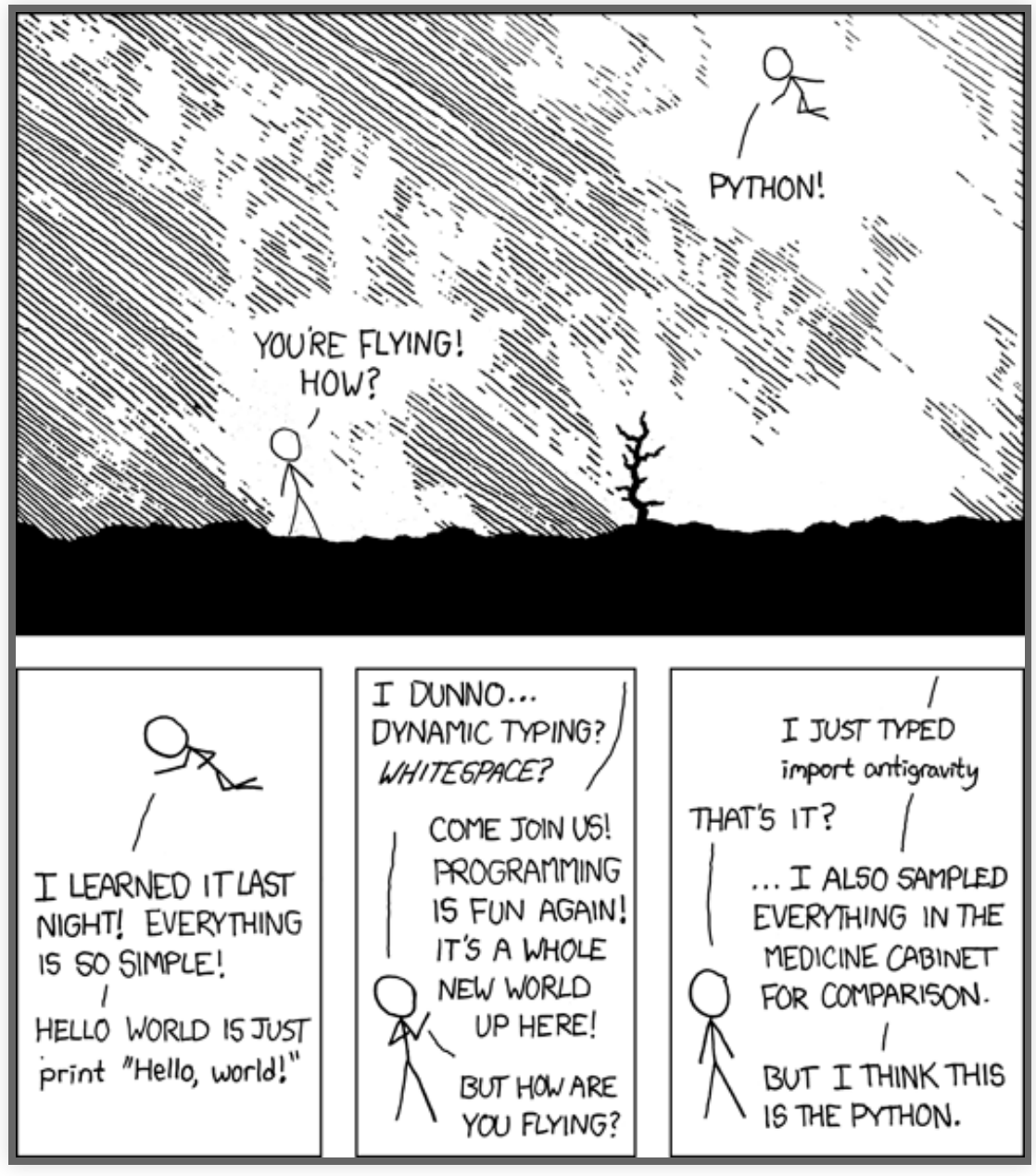


Code



Driving Vs automobile engineering

Python


PYTHON!
YOU'RE FLYING!
HOW?
I LEARNED IT LAST NIGHT! EVERYTHING IS SO SIMPLE!
HELLO WORLD IS JUST print "Hello, world!"
I DUNNO... DYNAMIC TYPING? WHITESPACE?
COME JOIN US! PROGRAMMING IS FUN AGAIN! IT'S A WHOLE NEW WORLD UP HERE!
BUT HOW ARE YOU FLYING?
I JUST TYPED import antigravity
THAT'S IT?
... I ALSO SAMPLED EVERYTHING IN THE MEDICINE CABINET FOR COMPARISON.
BUT I THINK THIS IS THE 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 **Team**
- **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

- In-lab computer tests: W.5 (25%) and W.10 (25%)
- Computational essay (W.12, 50%)
  - Equivalent to 2,500 word
  - 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

Statistics + …

- **Computational tools** –> Programming (hence this course's **tutorials**!)
- **Comunication 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

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% Coco 100 g (Pack of 5)

25

£11.95 Prime

Free Online Dating | OkCupid - Mozilla Firefox (Private Browsing)

Free Online D...

https://www.okcupid.com

Have an account? Sign in

okcupid

Join the best free dating site on Earth.

I am a Straight Woman Continue

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*

Geo-Data Science


It’s called GEOGRAPHIC 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…

Uber - London - Mozilla Firefox (Private Browsing)

Where is Uber... | Uber - London

https://www.uber.com/cities/london

WANT TO DRIVE WITH UBER? BECOME A DRIVER

MENU

UBER

LOG IN SIGN UP

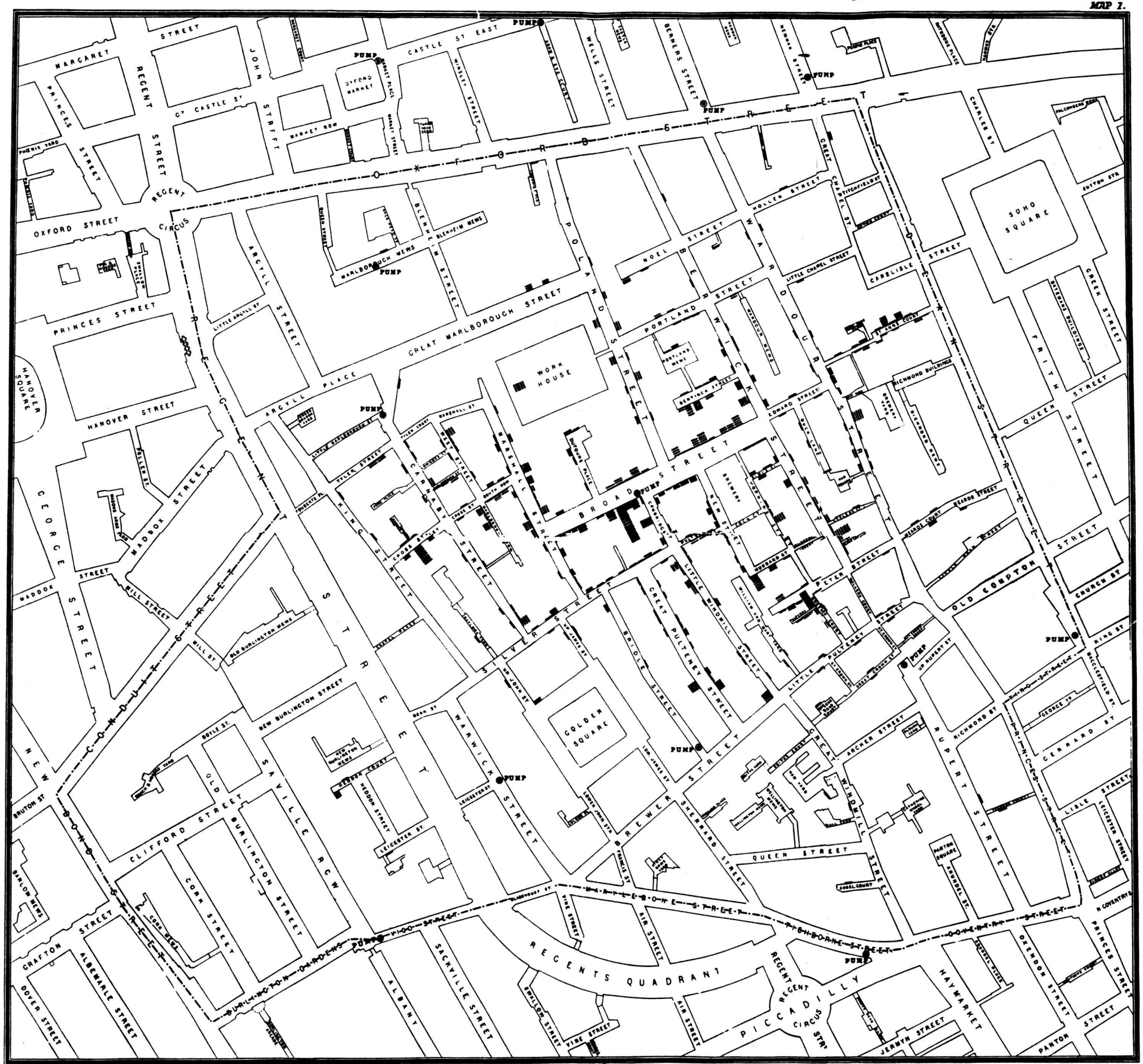
LONDON

FARE ESTIMATE +

LONDON

RIDE WITH UBER

MAP 1.



C. F. Cheffins, Lith. Southampton Buildings London.

SCALE 30 INCHES TO A MILE.

For next week…

For next week…

- Join Teams
- After the lab on Thursday, drop questions about JupyterLab on the `JupyterLab` channel created on Teams
- I'll respond them in class