Web Mapping and Analysis

Elisabetta Pietrostefani and Dani Arribas-Bel

1/23/23

Table of contents

Welcome

This is the website for "Web Mapping and Analysis" for the module **ENVS456** at the University of Liverpool. This is course designed and delivered by Dr. Elisabetta Pietrostefani and Professor Dani Arribas-Bel from the Geographic Data Science Lab at the University of Liverpool, United Kingdom. The module has two main aims. First, it seeks to to provide hands-on experience and training in the design and generation of web-based mapping and geographical information tools. Second, it seeks to provide hands-on experience and training in the use of software to access, analyse and visualize web-based geographical information.

 A compilation of this web course is hosted as a GitHub repository that you can access:

- As an html website.
- As a pdf document
- As a GitHub repository.

Contact

Dani Arribas-Bel - darribas [at] liverpool.ac.uk Professor in Geographic Data Science Office 6xx, Roxby Building, University of Liverpool - 74 Bedford St S, Liverpool, L69 7ZT, United Kingdom.

Elisabetta Pietrostefani - e.pietrostefani [at] liverpool.ac.uk Lecturer in Geographic Data Science Office 6xx, Roxby Building, University of Liverpool - 74 Bedford St S, Liverpool, L69 7ZT, United Kingdom.

Course Structure

Course Structure

Introduction

Block 1

- Introduction Lecture: Introduction to the module
- Lab: Powerful examples & Discussion about interactive map

Data Backends

Block 2 - Data Backends I

- Lecture: The Web's architecture and Economy
- Lab: Non-spatial APIs - What do APIs actually do?

Block 3 - Data Backends II

- Lecture: Data architechtures & formats
- Lab: Creating, manipulating, and integrating web geospatial data

Block 4 - Data Backends III

- Lecture: APIs
- Lab: Acquiring data from the web

Assignment I

Block 5 - Assignment I

- Lecture: Q&A
- Lab: Clinic

Assignment I: *Combining (geo-)data in an interactive map*

Frontend Topics

Block 6 - Frontend Topics I

- Lecture: Map design
- Lab: Designing maps with Mapbox Studio

Block 7 - Frontend Topics II

- Lecture: Interactivity
- Lab: Designing for interactivity (Kepler)

Block 8 - Frontend Topics III

- Lecture: Statistical visualization
- Lab: Choropleths in Kepler

Block 9 - Dashboards

- Lecture: Dashboards: bringing analysis to the web
- Lab: Building Dashboards (Shiny)

Assignment II

Block 10 - Assignment I

- Lecture: Technology gallery
- Lab: clinic

Assignment II: *A dashboard of IMD*

Overview

Aims

This module aims to:

- Provide hands-on experience and training in the design and generation of web-based mapping and geographical information tools.
- Provide hands-on experience and training in the use of software to access, analyse and visualize web-based geographical information.

Learning Outcomes

By the end of the module, students should be able to:

(1) Experience using tile rendering tools to generate content for map-based web sites.

(2) Knowledge of web based mapping infrastructure

(3) Web-based data collection techniques (accessing Twitter, Facebook, Google and Open-Streetmap information)

(4) Network analysis

(5) Programming skills to enable basic online data manipulation and web mapping

Feedback

Formal assessment. Two pieces of coursework (50%/50%). Equivalent to 2,500 words each

Verbal face-to-face feedback. Immediate face-to-face feedback will be provided during computer, discussion and clinic sessions in interaction with staff. This will take place in all live sessions during the semester.

Teams Forum. Asynchronous written feedback will be provided via Teams. Students are encouraged to contribute by asking and answering questions relating to the module content. Staff will monitor the forum Monday to Friday 9am-5pm, but it will be open to students to

make contributions at all times. Response time will vary depending on the complexity of the question and staff availability.

Computational Environment

This course can be followed by anyone with access to a bit of technical infrastructure. This section details the set of local and online requirements you will need to be able to follow along, as well as instructions or pointers to get set up on your own. This is a centralized section that lists *everything* you will require, but keep in mind that different blocks do not always require everything all the time.

Software

To run the analysis and reproduce the code, you need the following software:

- QGIS- the stable version (3.22 LTR at the time of writing) is OK, any more recent version will also work.
- R-4.2.2
- RStudio 2022.12.0-353
- Quarto 1.2.280
- the list of libraries in the next section

To install and update:

- QGIS, download the appropriate version from QGIS.org
- R, download the appropriate version from The Comprehensive R Archive Network (CRAN)
- RStudio, download the appropriate version from Posit
- Quarto, download the appropriate version from the Quarto website

To check your version of:

- R and libraries run `sessionInfo()`
- RStudio click `help` on the menu bar and then `About`
- Quarto check the `version` file in the quarto folder on your computer.

R List of libraries

The list of libraries used in this book is provided below:

- `sf`
- `geojsonsf`
- `mapview`
- `tidyverse`
- `viridis`
- `viridisLite`

You need to ensure you have installed the list of libraries used in this book, running the following code:

list.of.packages.cran <- c(“tidyverse”, “viridis”, “viridisLite”, “ggthemes”, “patchwork”, “showtext”, “RColorBrewer”, “lubridate”, “tmap”, “sjPlot”, “sf”, “sp”, “kableExtra”)

new.packages.cran <- list.of.packages.cran[!(list.of.packages.cran %in% installed.packages()[,“Package”])] if(length(new.packages.cran)) install.packages(new.packages.cran)

for(i in 1:length(list.of.packages.cran)) { library(list.of.packages.cran[i], character.only = T) }

Online accounts

- CDRC Data: we will use some of the data provided by the CDRC, so a (free) account with them will be necessary.
- Mapbox: Mapbox is one of the industry leaders in web mapping. Their free tier is rather generous so will more than suffice for what we will do within the course. You can sign up for a new (free) account here.
- Kepler:

Assessments

Assignment I

- **Title:** *Combining (geo-)data in an interactive map*
- **Type:** Coursework
- Due date: **Thursday March 2nd, Week 5**
- 50% of the final mark
- Chance to be reassessed
- Electronic submission only

In this assessment, you will have the opportunity to explore different sources and combine them in a single tileset that can be explored interactively through a web browser. **The assignment aims to evaluate your knowledge and aptitude in the following areas**:

- Understanding of core "backend" concepts in web mapping such as tilesets, client-server architecture, or APIs.
- Ability to use the web as a resource for original data.
- Design skills to present effectively a diverse set of geospatial data in a web map.

Design, data and assemblage.

This assignment requires you to source data from the web in different formats, assemble them into a tileset, and document the process. To be successful, you will need to demonstrate your understanding not only of the technical aspects involved in the process, but also of the conceptual notions that underpin them. Below are described the required components for your submission.

First, the design. Start by designing a map for an area you are interested in. There are no clear restrictions but, to ensure you are on the right path, check on your ideas with the module leader, who will be able to assess whether potential problems may arise from your choices. This stage should draw some inspiration from the first weeks of the course, where we looked for examples of web maps and spent time discussing what made them good and why.

Second, the data. Draft a list of potential data that would be ideal to use for your map, and try to find out whether they exist and are available. This will be a good guide for which data

you will actually end up using. Do not worry about spending a significant amount of time on this aspect; identifying good data takes time and is at the core of this task. Make sure you include both data you can access from direct downloads (e.g. CDRC) and those you download through an API. Once you know which datasets you need, go ahead and do the work required to download them for the map you want to build.

Third, the assemblage. With all data you have at your disposal from the previous stage, create a tileset that allows to embed the map in an HTML file and explore it through the browser. Pay attention to the design aspects involved in this step too. For example, what is the extent of your map (not necessarily the extent of each of your data)? What are the zoom levels your map will allow? Do you have the same "map" for every zoom level? These are questions you will have to ask (and answer!) yourself to complete this stage successfully.

Presentation of your work

Once you have created your map, you will need to present it. An important aspect of this stage is that it is not really the map you need to present, but the *process* of creation you have followed and the design choices you have made that should go into the text. Additionally, you will need to provide evidence that you understand the concepts behind some of the technologies you have used. Write up to 1,000 words and include the following:

- Map brief
  - About 250 words introducing the map. This should cover what it tries to represent (what is it about?) and the choices you have made along the way to take that idea into fruition.
  - About 250 words discussing and motivating the sources of data you have used. Here you should engage not only with what data you are using but why and what they bring to the map. Everything should be in the map for a reason, make sure to spell it out clearly.
- Conceptual background
  - About 250 words with your description of what an API is, how it works and how it has made your map possible.
  - About 250 words with your description of how tile-based maps work.

Submit

Once completed, you will need to submit the following:

1. A static HTML version of an R-markdown that includes two parts:
    1. All your narrative about the map brief and conceptual background.
    2. A second section with any code you may have used to complete the assignment, **documented** in detail. **NOTE**: this section will not contribute towards the word count.
2. A compressed .zip file containing you tileset and a HTML file that allows a user to browse through the tileset.

The assignment will be evaluated based on four main pillars, on which you will have to be successful to achieve a good mark:

1. **Map design abilities.** This includes ideas that were discussed in the course in Blocks 1 and 2.
2. **Technical skills**. This includes your ability to master technologies that allow you to create a compelling map, but also to access interesting and sophiticated data sources.
3. **Overall narrative**. This assesses your aptitude to introduce, motivate and justify your map, as well as you ability to bring each component of the assignment into a coherent whole that "fits together".
4. **Conceptual understanding of key technologies** presented in the course, in particular of APIs and tile-based mapping.

How is this assignment useful?

This assessment includes several elements that will help you improve critical aspects of your web mapping skills:

Design: this is not about making maps, this is about making good maps. And behind every good map there is a set of conscious choices that you will have to think throug to be successful (what map? what data? how to present the data? etc.). Technology: at the end of the day, building good web maps requires solid understanding of current technology that goes beyond what the average person can be expected to know. In this assignment, you will need to demonstrate you are proficient in a series of tasks manipulating geospatial data in a web environment. Presentation: in many real-world contexts, your work is as good as it can come across to the audience it is intended to. This means that it is vital to be able to communicate not only what you are doing but why and on what building blocks it is based on.

Assignment II

- **Title:** *A dashboard of IMD*
- **Type:** Coursework
- Due date**: Thursday April 27th, Week 10**
- 50% of the final mark
- Chance to be reassessed
- Electronic submission only

This assignment requires you to build a dashboard for the Index of Multiple Deprivation. To be successful, you will need to demonstrate your understanding not only of technical elements, but of the design process required to create a product that can communicate complex ideas effectively. There are three core building blocks you will have to assemble to build your dashboard: basemap, main map(s), and widgets. Let us explore each of them more in detail.

First, the basemap. Design your own basemap using Mapbox. Think about the data in the background, which colors, the zoom levels that will be allowed, and how it all comes together to create a backdrop for your main message that is conducent to the experience you want to create. The basemap is like a good side dish: it's there to make you like the main course even more.

Second, the main map(s). One you have your own basemap from Mapbox… move to R. This is where the core of your dashboard should come to shine. What you want to show, how, which interactive elements you will allow the user to access and how they will let them modify the experience of your dashboard. The main course of the meal, make it count!

Third, additional widgets. One of the advantages of dashboards in comparison to standard web maps is that they allow to bring elements of analysis to a more finished product. Think about what you want your users to be able to analyse, why, and how that will modify the main map. This is the icing on the cake!

Submit

Once completed, you will submit a report through Turnitin that includes the following:

- A link to the published dashboard, which needs to be reachable online
- About 250 words for the overall idea of the dashboard. What do you want to communicate? What is the story you want to tell?
- About 250 words for the data used. Which datasets are you using? Why? What new information do they bring and how they complement each other?
- About 250 words to describe your design choices in the basemap and other layers presented (e.g. choropleths).

- About 250 words to describe your design choices around interactivity, including both cartographic elements (e.g. zooming, panning) as well as additional interactivity built around components such as widgets.

The assignment will be evaluated based on:

1. *Overall design of the experience.* It is very important you think through every step of preparing this assignment as if it was part of something bigger towards which it contributes. Because that is exactly what it is. Everything should have a reason to be there, and every aspect of the dashboard should be connected to each other following a common thread. And, of course, make this connection and holistic approach come alive in your report.
2. *(Base)map design.* Critically introduce every aspect you have thought about when designing the maps, and explicitly connect it to the overal aim of the dashboard. Be clear in your descriptions and critical in how you engage every design choice.
3. *Interactivity design.* Your dashboard should use interactivity when necessary to deliver a more compelling and fuller experience that better gets your message across. Be sure to clearly lay out in your report which elements are used and why.
4. *Narrative around the description of the process.* Finally, the final mark will also take into account not only how good your dashboard is, but how well you are able to introduce it. Start with the key goals, and then unpack every element in an integrated and compelling way.

How is this assignment useful?

This assignment combines several elements that will help you improve critical aspects of web mapping:

- *Design*: this is not about making maps, this is about making good maps. And behind every good map there is a set of conscious choices that you will have to think throug to be successful (what map? what data? how to present the data? etc.).
- *Technology*: at the end of the day, building good web maps requires familiarity with the state-of-the-art in terms of web mapping tools. In this assignment, you will need to demonstrate your mastering of some of the key tools that are leading both industry and academia.
- *Presentation*: in many real-world contexts, your work is as good as it can come across to the audience it is intended to. This means that it is vital to be able to communicate not only what you are doing but why and on what building blocks it is based on.

Marking Criteria

This course follows the standard marking criteria (the general ones and those relating to GIS assignments in particular) set by the School of Environmental Sciences. Please make sure to check the student handbook and familiarise with them. In addition to these generic criteria, the following specific criteria will be used in cases where computer code is part of the work being assessed:

- 0-15: the code does not run and there is no documentation to follow it.
- 16-39: the code does not run, or runs but it does not produce the expected outcome. There is some documentation explaining its logic.
- 40-49: the code runs and produces the expected output. There is some documentation explaining its logic.
- 50-59: the code runs and produces the expected output. There is extensive documentation explaining its logic.
- 60-69: the code runs and produces the expected output. There is extensive documentation, properly formatted, explaining its logic.
- 70-79: all as above, plus the code design includes clear evidence of skills presented in advanced sections of the course (e.g. custom methods, list comprehensions, etc.).
- 80-100: all as above, plus the code contains novel contributions that extend/improve the functionality the student was provided with (e.g. algorithm optimizations, novel methods to perform the task, etc.).

1 Introduction

Dani Arribas-Bel

Lecture: Introduction to the module

Lab: Powerful examples & Discussion about interactive map

1.1 Lecture

Slides can be downloaded here

1.2 Lab: Powerful examples

This lab has two main components:

1. The first one will require you to find a partner and work together with her/him
2. And the second one will involve group discussion.

1.2.1 Paired activity

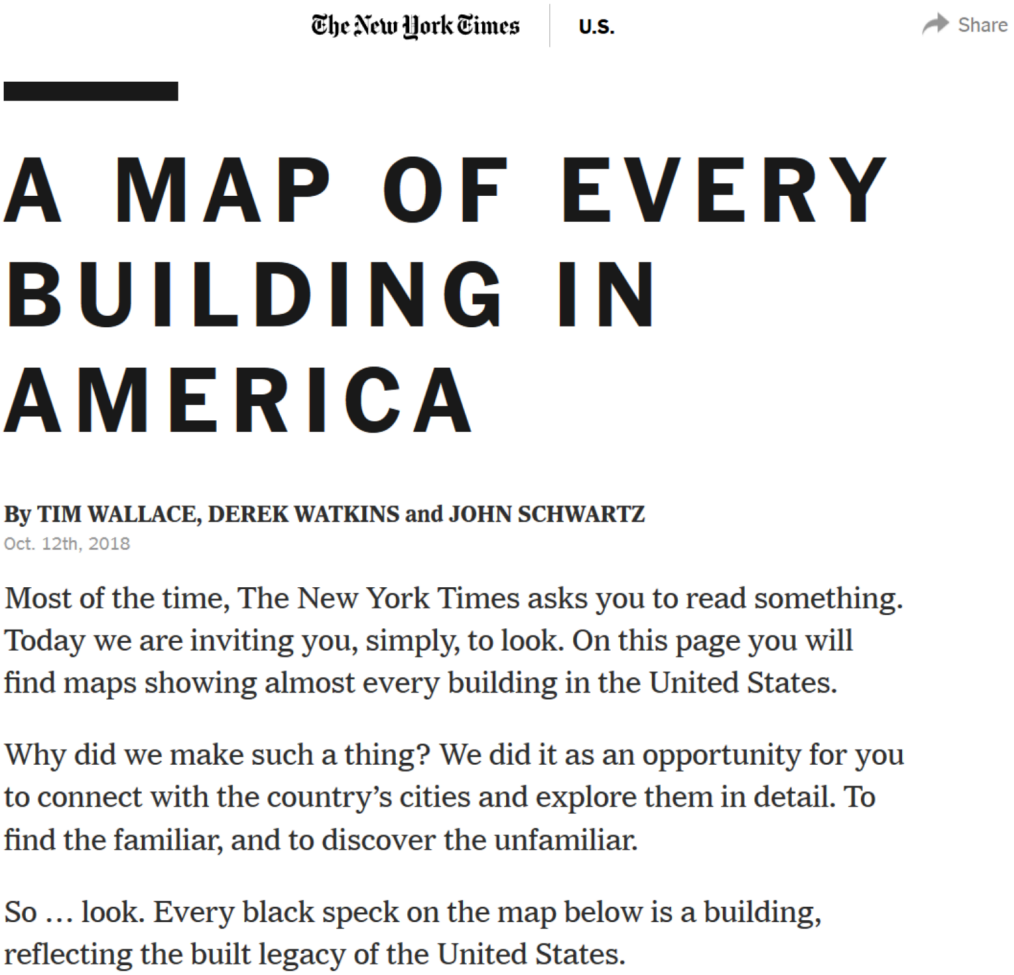
In pairs, find three examples where web maps are used to communicate an idea. Complete the following sheet for each example:

- **Substantive**
  - `Title`: Title of the map/project
  - `Author`: Who is behind the project?
  - `Big idea`: a "one-liner" on what the project tries to accomplish –
  - `Message`: what does the map try to get accross
- **Technical**
  - `URL`:
  - `Interactivity`: does the map let you interact with it in any way? Yes/No

- `Zoomable`: can you explore the map at different scales? Yes/No
- `Tooltips`:
- `Basemap`: Is there an underlying map providing geographical context? Yes/No. If so, who is it provided by?
- `Technology`: can you guess what technology does this map rely on?

Post each sheet as a separate item on the Teams channel for Lab No.1

As an example, below is the sheet for the project "A map of every building in America"



- **Substantive**
  - `Title`: A map of every building in America
  - `Author`: The New York Times

  - `Big idea`: Show patterns in the built environment through building footprints
  - `Message`: The project displays a map of the US where only building footprints are shown. This element is used to show a) how much buildings tell about places and the build environment, and b) some interesting specific examples that illustrate large trends in society (e.g. urban sprawl)

- **Technical**
  - `URL`: `https://www.nytimes.com/interactive/2018/10/12/us/map-of-every-building-in-the-u`
  - `Interactivity`: Yes
  - `Zoomable`: Yes
  - `Tooltips`: No
  - `Basemap`: Yes, Mapbox/OpenStreetMap
  - `Technology`: MapboxGL

1.2.2 Class discussion

We will select a few examples posted and collectively discuss (some of) the following questions:

1. What makes them powerful, what "speaks" to us?
2. What could be improved, what is counter-intuitive?
3. What design elements do they rely on?
4. What technology do they use?

1.3 References

- For an excellent coverage of "visualisation literacy", Chapter 11 of Andy Kirk's "Data Visualisation" is a great start. Lab: Getting up to speed for web mapping
- A comprehensive overview of computational notebooks and how they relate to modern scientific work is available on Ch.1 of the GDS book.
- A recent overview of notebooks in Geography is available in Boeing & Arribas-Bel (2021)

2 Web architecture

Dani Arribas-Bel

Lecture: The Web's architecture and Economy

Lab: What do APIs actually do? (non-spatial APIs)

2.1 Lecture

Slides can be downloaded here

2.2 Lab: What do APIs actually do?

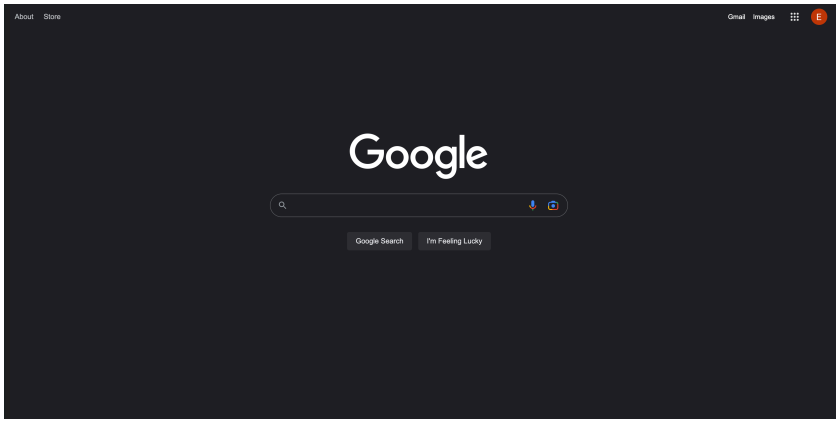
In this lab, we will unpack how Application Programming Interfaces ("APIs") work and we cover the basics of accessing an API using R. Instead of downloading a data set, APIs allow programmers, statisticians (or students) to request data directly from a server to a local machine. When you work with web APIs, two different computers — a client and server — will interact with each other to request and provide data, respectively.

Web APIs

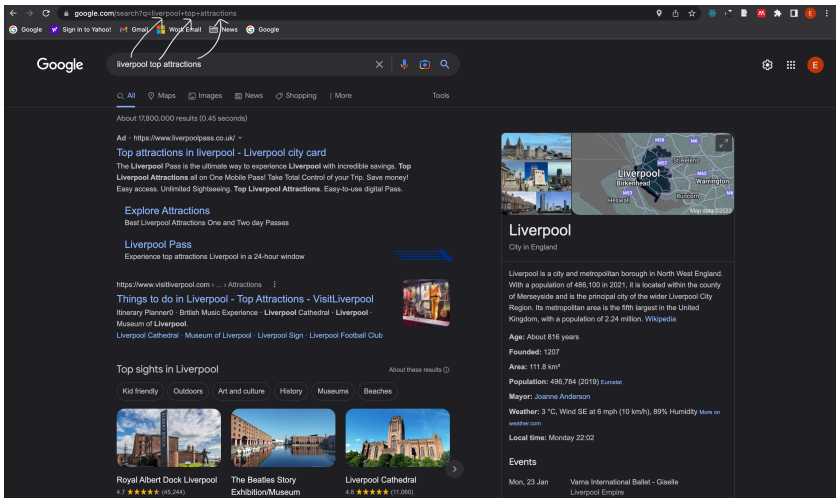
- Allow you query a remote database over the internet
- Take on a variety of formats
- Adhere to a particular style known as Representation State Transfer or REST (in most cases)
- RESTful APIs are convenient because we use them to query database using URLs

2.2.1 RESTful Web APIs are all around you.

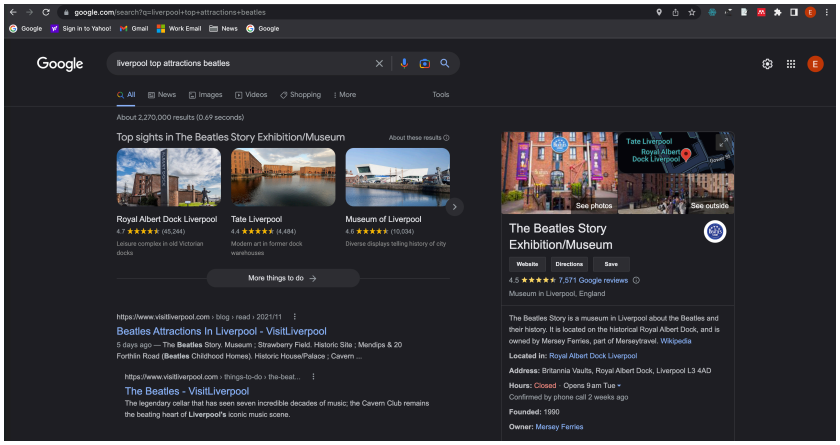
Consider a simple Google search:



Ever wonder what all that extra stuff in the address bar was all about? In this case, the full address is Google's way of sending a query to its databases requesting information related to the search term *liverpool top attractions*.



In fact, it looks like Google makes its query by taking the search terms, separating each of them with a +, and appending them to the link https://www.google.com/#q=. Therefore, we should be able to actually change our Google search by adding some terms to the URL:



Learning how to use RESTful APIs is all about learning how to format these URLs so that you can get the response you want.

2.2.2 Paired activity

Using your friend the internet, look up some answers to the following questions:

1. What is a URL and how can it help us query data?
2. What is a "GET request"?

2.2.3 Class discussion

TO DO Non-spatial APIs - tidycensus - Unpacking what that function is doing. - What the API is actually doing? - Wanted to understand - abstracts everything away. Unpack. - REST type of API. - Function web address.

2.2.4 Paired activity answers

1. Uniform Resource Location (`URL`) is a string of characters that, when interpreted via the Hypertext Transfer Protocol (`HTTP`). URLs point to a data resource, notably files written in Hypertext Markup Language (`HTML`) or a subset of a database
2. `GET` requests a representation of a data resource corresponding to a particular `URL`. The process of executing the `GET` method is often referred to as a `GET request` and is the main method used for querying RESTful databases. `HEAD`, `POST`, `PUT`, `DELETE`: other common methods, though mostly never used for database querying.

2.3 References

- Brief History of the Internet, by the Internet Society, is a handy (and free!) introduction to how it all came to be.
- Haklay, M., Singleton, A., Parker, C. "Web Mapping 2.0: The Neogeography of the GeoWeb". Geography Compass, 2(6):2011–2039
- A blog post from Joe Morrison commenting on the recent change of licensing for some of the core software from Mapbox