

02.526

Interactive

Data Viz

Hello

My name is Chi-Loong (chiloong@vslashr.com) and I'll be your adjunct lecturer for this course.

I run my own frontend UX and visualization studio at [V/R](#).

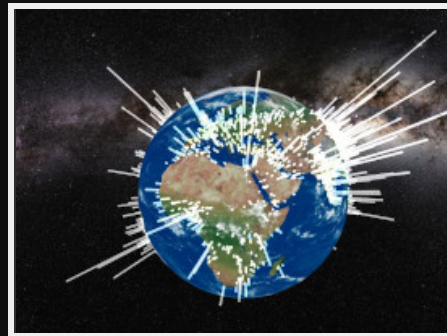
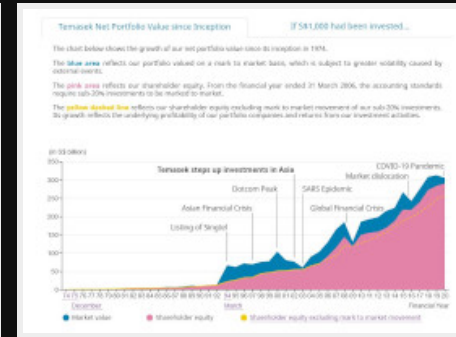
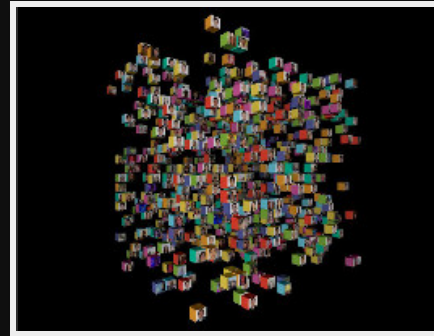
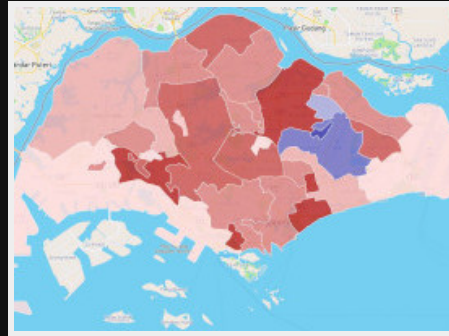
Past lives

I spent half my career in storytelling. I was in journalism and public relations for a decade.

The other half is in technology — I have background in computer science — also approaching a decade.

In the last 8 years I have been running a consulting-based visualization / frontend UX studio.

Samples



I also do visualization prototypes and dashboards for software products. Clients include A*Star (for MOT), PSA (for MPA), Attilatech (cyber security), etc.

Some other credentials

Beyond work experience, I also have some experience teaching and giving workshops at universities (e.g. NUS, SIT), code schools (e.g. General Assembly , SGCodeCampus), mainly on UX and visualization.

Lastly, I believe in ground-up community grassroots. I run and am involved in the local Singapore [Data Viz SG](#) and [HHSO](#) meetup groups.

Course vision

As I am an adjunct lecturer from the industry, the way I envision this course will be different from academic lecturers.

For masters students, I feel that the greatest value is in discussing case studies and networking experience.

Hence I am hoping to give more of a flavour of that from guest lectures by people in the industry.

Example: [Coronavirus, one year on in Singapore and Asia](#), Straits Times

Example: [Abandoned at sea: Desperate journeys of Rohingya refugees](#), Kontinentalist

Design and code

This course is split up into two components — **design** and **code**.
In the **design** portion, we will go through the academia / research / case studies in what makes for good data visualization design.
In the **code** portion, we will be building interactive data visualizations for the web.

On technology

Visualizations can be built with all kinds of software tool / cloud-based platforms.

However, for an interactive visualization course I will going through how to do it natively for the web.

This namely means using **D3 (Data Driven Documents)**, the defacto data visualization library for the web.

Why D3

- If you're interested in analytics, Excel / Tableau / PowerBI are good non-coding choices.
- If you're familiar with data science languages, Python or R has graphing libraries (ggplot2, plotly, matplotlib, etc.)
- D3 is used because you want to **build beautiful interactive visualizations** for the web. You can go really deep to get the interactions you want.
- It also has comprehensive built-in **design patterns for visualizations**, like geo projections, force simulations, hierarchies, etc.
- Lastly, it is open source and free.

Why web

- For web visualization there are many libraries / tools / platforms beyond D3.
- If you are pressed for time for a standard chart type, it might be easier to use other tools / libraries built on top of D3. For example, **NVD3** or **C3.js**.
- You can also use other libraries (leaflet, highcharts, etc.) to build the components you need.
- There are even full platforms - Mapbox, Cesium, etc. that will allow you to host and build on top of those platforms.

This is my bias. You are not going to get glassy visualizations for the web without coding it natively for the web.

Course expectations

Basic web development is a technical requirement for this course. It is expected that you know basic **HTML / CSS / JS**.

We'll go through some recap refreshers, but it will be quick.

If you absolutely have no experience in this area you will have to do your own reading and self-revision.

There are lots of tutorials on the web for HTML / CSS / JS.

It will be great if you are more technical and know more — CSS /JS frameworks, web architecture frontend / backend, etc. — but it is not required for this course.

Course expectations 2

On the JS side, you should also know how to pull data from APIs (AJAX), data formats like JSON (GeoJSON, TopoJSON).

Beyond HTML / CSS / JS, we will also be using **Git**, a modern version control tool.

Students will be publishing their assignments and work on **GitHub** and **GitPages**.

If you do not have a GitHub account, please do sign up for one.

I will do a survey on Thursday's practical lab, and also go through with the class on self-introduction and understand your expectations of the course.

Course expectations 3

However, if you prefer to use other web-based tools for the final project, you are free to do so.

The only requirement is that you need to be able to publish your work on a GitHub repository (through GitPages).

For coordination for this module, we will be using communication tool **Slack**.

Again, we'll discuss on Thursday on details.

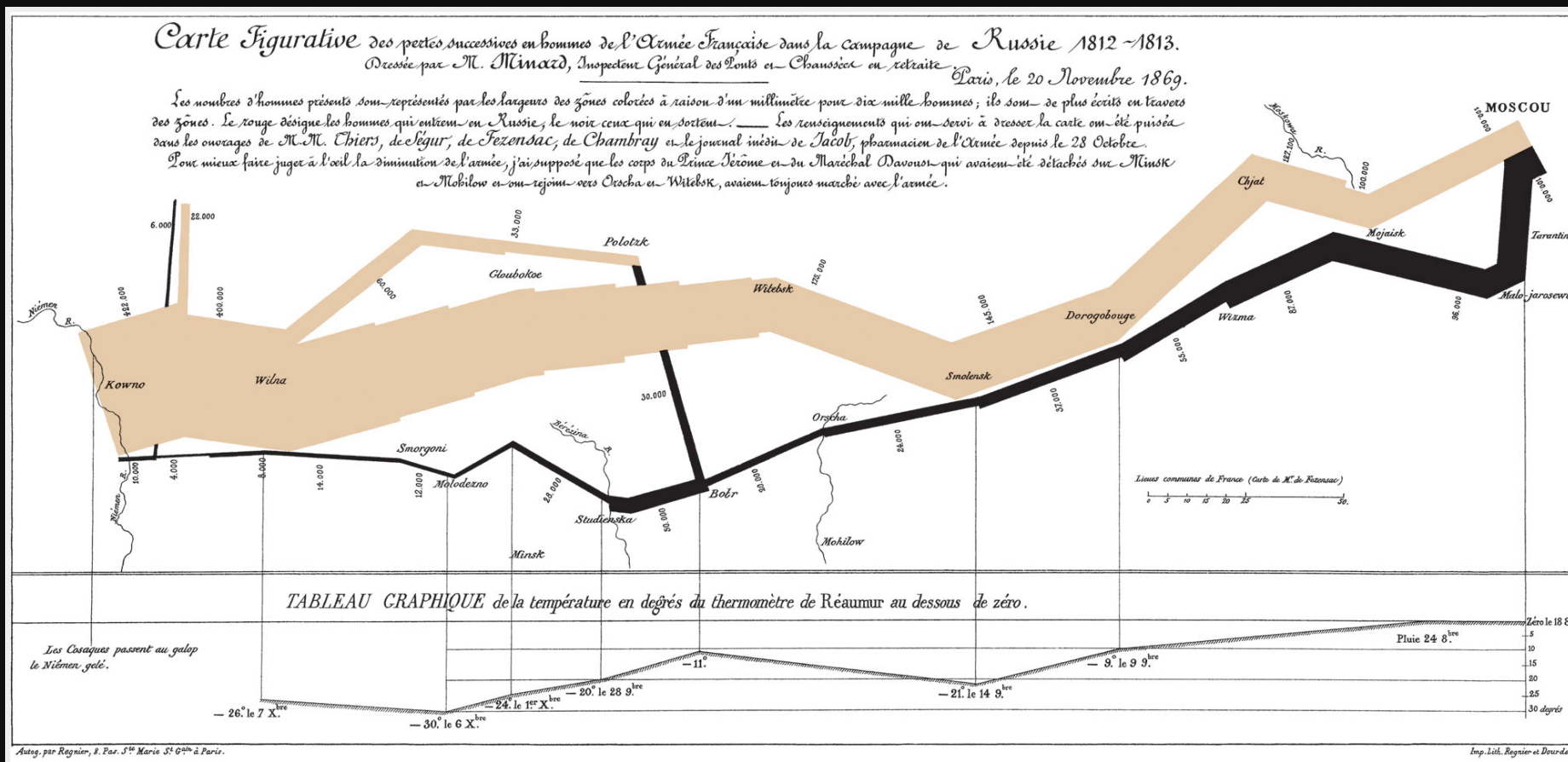
1a: Introduction to Data Viz

Why data visualization?

Wiki definition

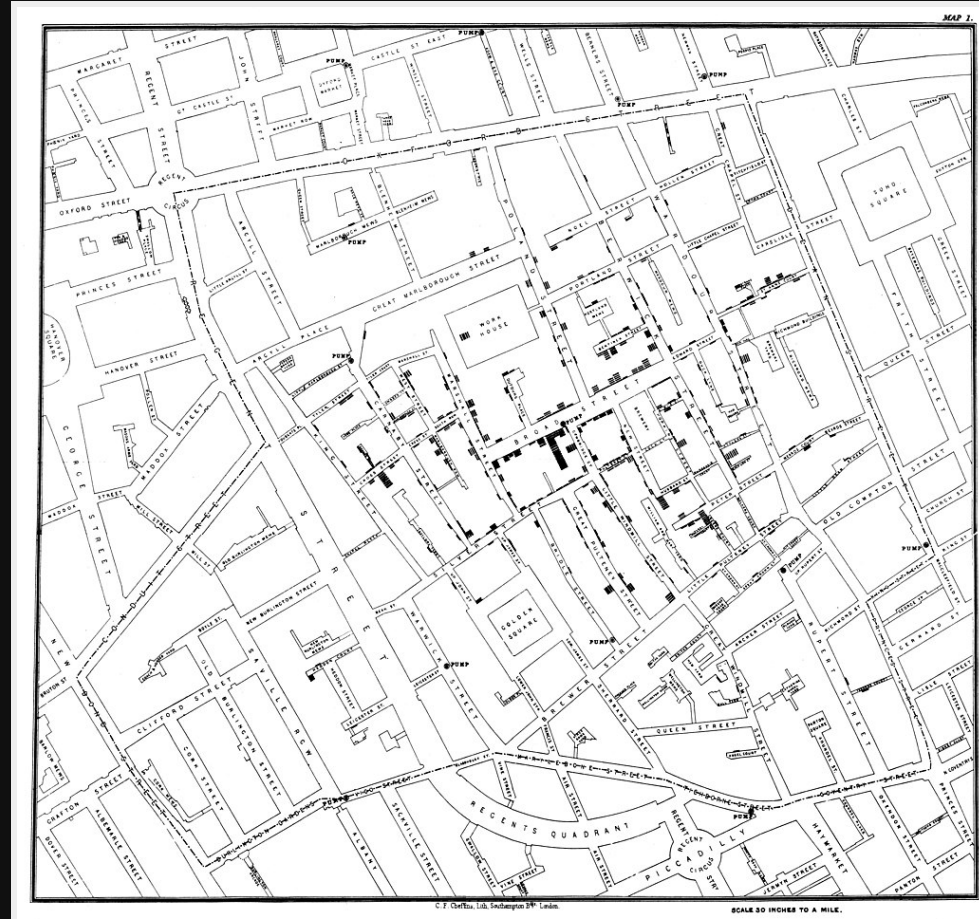
To communicate data / information clearly and effectively by encoding it as graphical visual objects (e.g. points, lines).

Historical viz



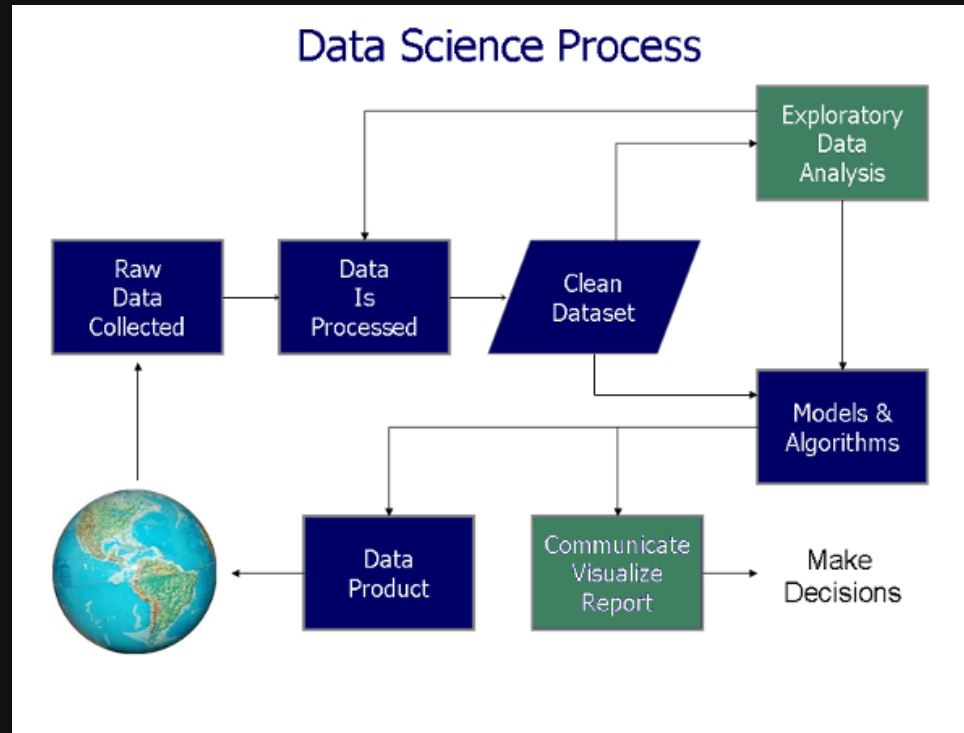
Charles Joseph Minard's 1869 diagram of Napoleonic France's invasion of Russia. It shows the number of men in Napoleon's 1812 Russian campaign army, their movements, as well as the temperature they encountered on the return path.

Historical viz 2



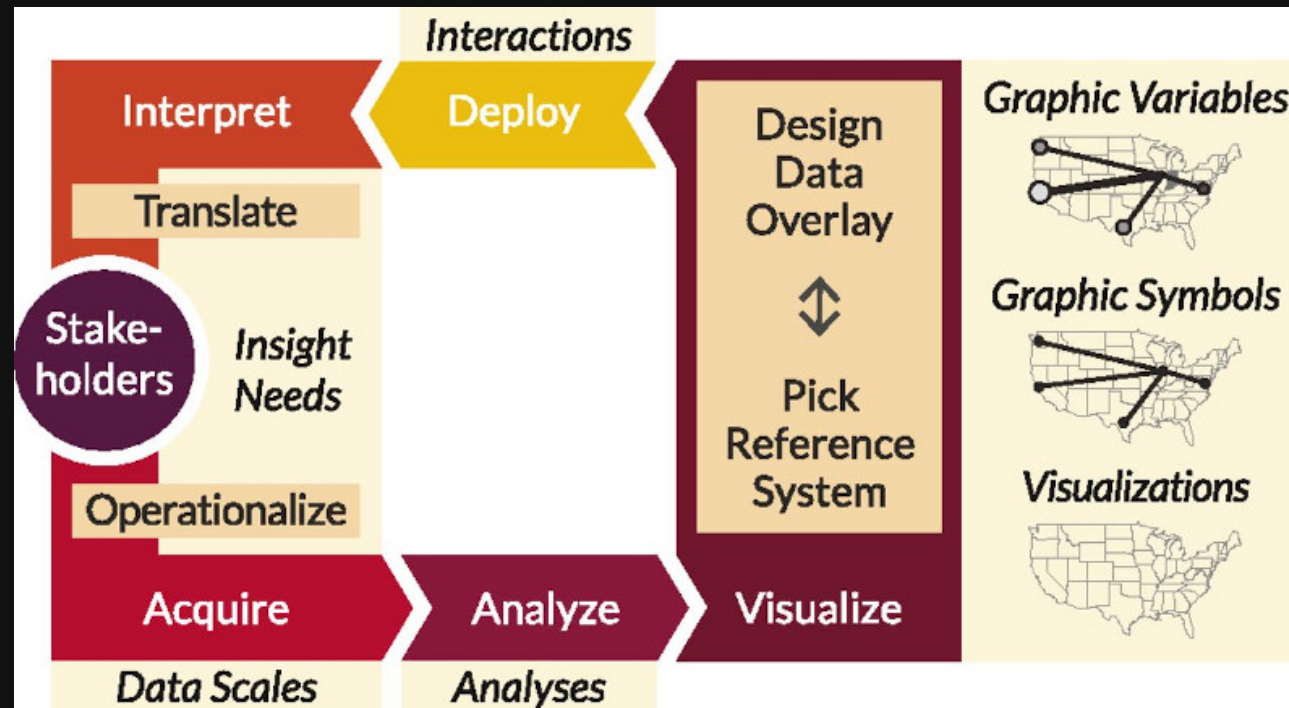
John Snow's map of the 1854 cholera outbreak in Soho, which helped debunk the then-dominant miasma theory.

Exploratory or explanatory analysis?



A 3rd category: Engineering a data visualization product / tool to automate exploratory / explanatory analysis.

Visualization analysis



- **What:** What data does the user see?
 - **Why:** Why does the user want to use a viz tool?
 - **How:** How is the visual encoding and design choices chosen?
- These three questions form the **data — task — idiom** trio.

A broader definition

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

Reading: [Data visualization literacy framework](#)

Reading: [Chapter 1, Visualization Analysis and Design \(Tamara Munzner\)](#)

Big picture questions 1

- Why have a human in the decision making loop?
- Why have a computer in the loop?
- Why use an external representation?
- Why depend on vision?
- Why show data in detail?
- Why use interactivity?

Big picture questions 2

- Why is the viz idiom design space huge?
- Why focus on tasks?
- Why focus on effectiveness?
- Why are most designs ineffective?
- Why is validation difficult?
- Why are there resource limitations?

Questions?



Chi-Loong | V/R