



# Assignment 3: Addressing Complexity

Worth 40% of the Mark for the Module

Due: 07/12/2024  
@23:59 hrs

# Assignment 3 : Addressing Complexity

[Worth 40% of module. Design, implementation, and presentation]

## Motivations :

- ✦ apply concepts discussed in lectures to visualize a data set of some complexity
- ✦ design and implement a visualization comprised of more than a single static idiom (or chart type) OR that requires a something more than a well-established chart type

## Objectives [see details in following slides]

1. **Choose** an **interesting dataset**, that necessitates visualization with more than a single established chart
2. **Design and implement** a **novel** visualization for this data set
3. **Choose** any **tool** (you may even use a combination of several)
4. Write a **report**, discussing your design in the context of the issues discussed in the course
5. Create a 2-minute **presentation video**

# 1. Choice of Data

You are free to choose any interesting data set. Some suggested data sources are provided on blackboard. It is up to you to choose the data set (or subset thereof) that you will visualize.

There is no explicit bottom-limit to how complex the data-set should be, but you will have to explain why you feel the data warrants visualisation (essentially, you should propose an analysis question worth exploring through visualisation)

- ♦ Higher complexity/effort is likely to lead to a higher contribution but you should try to balance this with feasibility (for a time-budget of about 40 hours of work), and avoid contributing effort on superfluous unnecessary detail.
- ♦ As a rough approximation, consider the examples in Assignment 2 (Minard, Gapminder) as a baseline. Try to find something at least as complex as this in some way.

**Complexity can be based on various criteria, including (but not limited to)...**

- ♦ *Volume*: Large items (high-dimensional, many attributes), many items, overall size
- ♦ *Variety*: Multi-variate/varied attribute types, heterogenous, non-uniformly distributed
- ♦ *Structure*: task is related to complex relationships between elements
- ♦ *Visual complexity* needed for the task: e.g. prone to over-plotting, many/complex graphical elements required

# 2. Implementation: A Novel Visualization

**Requirement #0: Novelty.** This merely means ‘individual’ and that you are NOT simply replicating an existing visualization of your chosen data. i.e. either the dataset has not yet been visualized OR it has not been previously visualised in the way that you are going to visualize it.

**Requirement #1 :** at least one of the following should apply to your result

- ✦ The overall visualization **uses more than one idiom**, combined in an individual way (you’re not merely replicating an existing visualisation), OR
- ✦ The visualization **uses an idiom that is unique in some way** e.g., combining several encoding channels or a completely new idiom

**Requirement #2:** your visualization must feature at least one of the following technical elements:

- ✦ **Multiple facets** (see Lecture 10) e.g. small multiples, juxtaposition/superimposition/embedding of different idioms (see Lecture 10), OR
- ✦ **Animated/dynamic view** (animation may be linear and passive i.e. without any input from user), OR
- ✦ **Interactive** manipulation of the view (some form of real-time changes can be made by the user to what is show on screen)

Animation &  
Interaction will be  
discussed in Week 8

**Some additional guidelines:**

- ✦ The full visualization should be visible on one page/screen (legible at approx. A4 size) without scrolling or turning the page (separate tabs should be avoided). You may allow the user to change the view interactively (e.g. modify some part of the view by clicking), but you should not submit several unconnected visualizations.
- ✦ The submitted visualization should be a ‘finished product’; the user should be able view/use it without looking at your code or having to change your code/script (e.g. submitting just your raw jupyter notebook with instructions for the reader to change code in ‘line X’ is not acceptable)
- ✦ The choice of (and number of) idioms used should be justified and not merely to ‘show off’ (“function over form”, see Lecture 11).

# 3. Choice of Tool

You may choose any generalized\* visualization tool.

\* Essentially, the only thing to AVOID is using a third-party black-box solution that has been specifically created to visualize your dataset format with little/no design input from you

e.g. using an off-the-shelf “MRI Visualization tool” for an MRI scan OR merely plugging new data into a dashboard someone else created without any changes would score poorly in terms of novelty/complexity.

If you are using an extremely high-level tool that does a lot of the visualization work for you, you should try to be more ambitious in terms of data & task complexity; you should explain what efforts you have added to visualise the data [note that marks are going for visualisation designing and development merely mining the data is not sufficiently relevant to this module]

This is not an exhaustive list but examples of **tools you may certainly use include**: processing, d3.js, Tableau, qlikview, PowerBI, seaborn, HighCharts, GoogleCharts / Google Data Studio, plotly, three.js, p5.js, jupyter notebook, VtK, OpenGL, R, ggplot, matplotlib, shiny, matlab, visme, datawrapper, rawgraphs, bokeh, altair, Mathematica, ...

# 4. The Report

Write a report about your implementation of max 3 pages (not including references) + 1 full page image (or screenshot) of your visualization

- ✦ Template provided on blackboard. You may simply replicate this as best as you can according to the sample PDF that also describes the specification.

## The report should discuss

- ✦ The **tools/technologies used** and (in brief) any steps taken to pre-process the data; comment briefly on the reason why you chose the tool
- ✦ The **dataset** : general description, more detailed analysis of dataset/data types, attribute types, derived attributes, why it is particularly complex that it necessitates visualization
- ✦ The **tasks** that you have intended to support in your visualization (should be more than one)
- ✦ A discussion about the **novelty** of the visualization; if/why the implementation is particularly complex that it warrants visualisation
- ✦ **Encoding channels and idioms** (including interactive operations, if any) AND a **justification** for why particular encodings are used
- ✦ A **critical analysis** of the strengths and weaknesses of the visualization

You *may* support this discussion with additional images, but these must be included within the 3-page limit

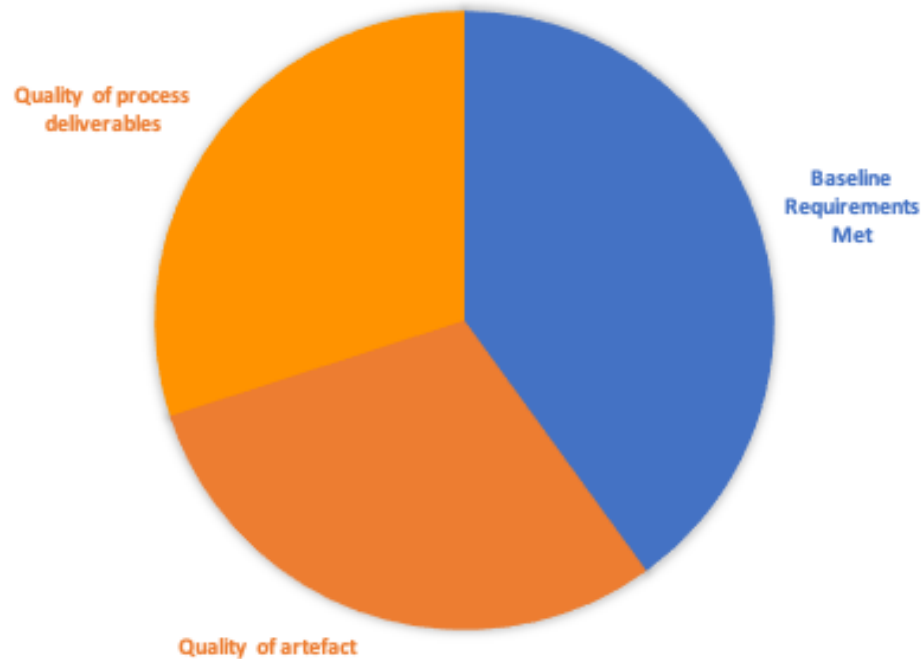
Do not assume the reader has seen the design you discussed in Assignment 2 (if you are re-using this design, you should summarise it again)

# 5. The Presentation

A presentation video of max 2 minutes outlining the technical contributions of the paper

- ✦ This should essentially summarise the main points of the technical paper supported by visuals of your visualization result
  - ✦ Use captions or voice over to present a summary of both the artefact and the paper
- 
- ✦ Please note that these may be made privately available to students in the class (and future classes) to view as an example. If you wish to remain anonymous, do not include your name in the video (meta-data on the page will be removed in any case).

# Marking Scheme



This project is worth 40% of the module, broken down as follows:

- ✦ 0 to 16/40 for meeting the baseline requirements: all deliverables submitted on time, artefact successfully visualizes the data, the process paper and video together reasonably describe the data, task, idioms and encodings channels employed.
- ✦ + 0 to 12/40 marks for quality of the artefact\* beyond the baseline including but not limited to: Novelty, Complexity and How well it achieves the intended data visualization task(s) using the idioms employed
- ✦ + 0 to 12/40 marks for the quality of process deliverables beyond the baseline: Paper is well written (clear and concise with minimal errors and suitable references); video is well put together; together they summarize the paper and artefact to someone not previously familiar with it, express clearly the complexity/novelty of the visualization, explains/justifies the design choices made, makes a good attempt at discussing strengths and weaknesses; reasonably consistent with the discussion in the module notes and related reading or equivalent sources

\* “*artefact*” = your implemented solution (the actual visualisation you created, separate from the report or presentation video; could be an image, animation, interactive visualisation)



# High-level Rubric

	Excellent (70-100%)	Good (40-69%)	Poor (0-39%)
<b>Baseline Requirements</b> [Weighting = 0.4]	All deliverables submitted on time and in line with the specification; artefact successfully visualizes the data in some way; the process paper and video together describe the data, task, idioms and encodings channels employed sufficiently well to make an assessment of the project.	All deliverables submitted with some shortcomings in compliance with specifications, the report or video is partially lacking details to fully assess the project.	Missing significant elements of deliverables. Insufficient detail in the report and/or video to adequately assess the project.
<b>Quality of the artefact</b> [Weighting = 0.3]	The artefact submitted demonstrates a high degree of novelty, complexity and is highly effective in achieving the intended data visualization task(s) using the idioms employed. The visualization is easily understandable by a third party and doesn't distort the presented data or information. The visualisation avoids excessive superfluity or clutter.	The artefact submitted demonstrates some degree of novelty and complexity beyond average, and is reasonably effective in achieving the intended data visualization task(s). The visualisation has some flaws potentially including: superfluity, clutter, data distortion or it is not fully self-explanatory.	The artefact submitted demonstrates little or no novelty, has minimal complexity, and/or has significant flaws with regards to effectiveness at its intended data visualization task(s), superfluity, clutter, data distortion or is not self-explanatory.
<b>Quality of presentation</b> [Weighting = 0.3]	Video and report are well put together, clear and concise with minimal errors and suitable references. The discussion summarizes the artefact very well; clearly expresses the complexity/novelty of the visualization; explains/justifies the design choices made; provides a meaningful analysis/critique of the developed artefact.	Video or report have some flaws but are reasonably put together and generally clear. The discussion summarizes the artefact reasonably well; expresses the complexity/novelty of the visualization to some degree; provides partial explanation/justification the design choices made; and provides a meaningful analysis/critique of the developed artefact.	Video or report have significant shortcomings. The discussion largely fails to clearly present how the develop artefact address the requirements of the project; to present the novelty of the visualisation; to justify the need for the visualisation; and/or fails to justify the reason for the specific design choices made. The critical analysis is unfounded or is weakly supported by evidence or examples from the presented solution.

# Examples of Short Presentations

Spring Break vs. COVID19: The Real Impact of Ignoring Social Distancing, short video presentation by Tectonix GEO [ [LINK](#) ]

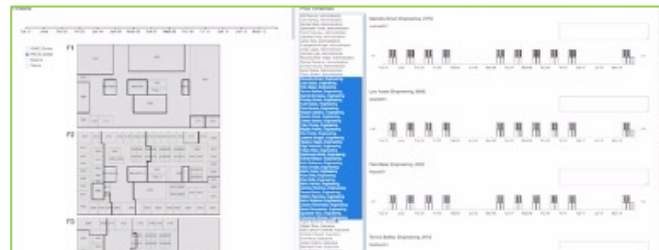
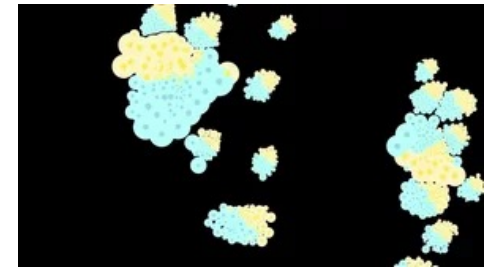
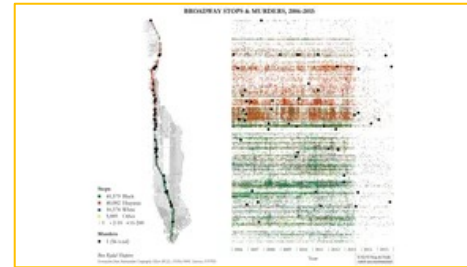
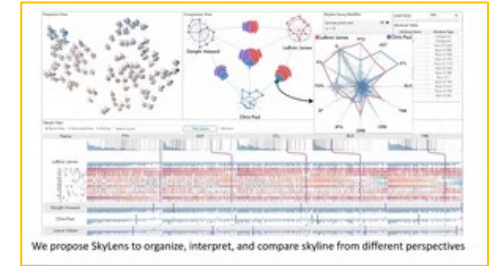
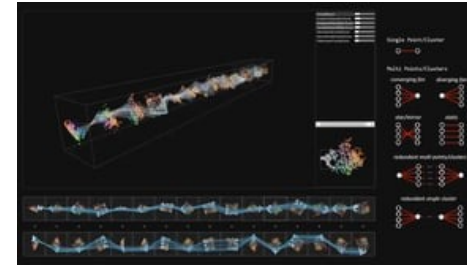
<https://www.youtube.com/watch?v=cq2zuE3ISYU>

Vis 2017 : 30 Second Videos [ [LINK](#) ]

- ✦ Shorter than required but shows how much detail can be expressed in a short video
- ✦ These are technical papers : you are not really expected to provide scientific novelty as they are providing in these conference videos

VAST Challenge Videos [ [LINK](#) ]

- ✦ Closer to the brief of this project i.e., novel implementation of visualization of some specific data (but probably more dev time)
- ✦ Mostly longer than 2 mins and there is a lot detail of also the user process of analyzing the data (not really expected in this assignment)



# Tips on creating video

[These are JUST SUGGESTIONS. ALL Optional. You can find many other alternative ways to create a video]

One of the simplest options for this projects is to create a PowerPoint presentation and save this as a video.

- ✦ Here's one of many tutorials on the web that show you how to do this (ignore the second half transitions are really not needed)
  - ✧ <https://helpdeskgeek.com/office-tips/how-to-turn-a-powerpoint-presentation-into-a-video/>
- ✦ You should add a voice-over or captions / subtitles to describe the video

Some free screen capture options:

- ✦ Microsoft Stream [through browser, any platform] – free with TCD id at <https://www.microsoft365.com/launch/stream>
- ✦ Quicktime [Mac – bundled by default] : <https://libguides.rowan.edu/c.php?g=248114&p=4711659>
- ✦ Xbox Game bar [free on Windows] : <https://www.microsoft.com/en-us/windows/learning-center/how-to-record-screen-windows-11>
- ✦ The following are other suggested 3<sup>rd</sup> party software that can be used free (usually with watermark, this is fine for the project)
  - ✧ Camtasia: <https://www.techsmith.com/download/camtasia/> [mac/windows] – free version with watermark
  - ✧ Bandicam: <https://www.bandicam.com/> [windows] – free version with watermark
  - ✧ Fraps: <http://www.fraps.com/download.php> [windows] – free version with watermark

For more advanced editing, there are many free tools out there e.g. iMovie[Mac], Clipchamp [Windows], movavi, filmora, BUT you should just keep it functional. I'm not really expecting a fancy video

# IMPORTANT: Permissions to Shared Videos/Files

Numerous students in the past have submitted links to shared files that are protected and inadvertently can't be accessed by the examiners. This affects marking progress. Also, note that I am not the only one that needs to see these. Therefore, you should ensure the following:

## 1. You MUST either:

- a) Make the file viewable by anyone with the link OR
- b) Make the file viewable to [cs7ds4.staff@gmail.com](mailto:cs7ds4.staff@gmail.com) OR
- c) Password protect the file and submit the password and link in the text of your report

## 2. You MUST not change the file after submission; the file MUST remain online (unchanged) until August 30<sup>th</sup>, 2026

REMINDER: you may not submit a link to the report, nor insert it in a zip file. This must be in PDF format uploaded to blackboard

# Deliverables

## 1. Technical Report

- ✦ This should be in PDF format of upto 3-pages (+ references/declarations)
- ✦ For marking reasons, you should not add the report inside a zip file; and should not submit it as a link.

## 2. A LINK to a presentation video of max 2 minutes outlining the technical contributions of the paper

- ✦ DO NOT SUBMIT THE ACTUAL VIDEO ON BLACKBOARD (ONLY A LINK)

## 3. Visualization “artefact” i.e. the actual visual output or programme that you have created:

- ✦ This could be ...
  - ✦ a link to your interactive visualization if it is online;
  - ✦ a locally executable binary (you must provide instructions on how to run it);
  - ✦ a video if the visualization is animated but non-interactive;
  - ✦ or a high resolution image/PDF if it is static
- ✦ You MUST submit a link to download the project source files and data set
- ✦ DO NOT SUBMIT ANY FILE LARGER THAN 10 MB ON BLACKBOARD (ONLY A LINK)

**NB. If your artefact itself is a video, you may submit two videos: an artefact video and a presentation video (there is no limit on the length of the artefact video but if larger than 10Mb, submit a link)**

# Required format / information for Report

You should submit a single PDF of MAX 4 pages (+1 for any references) assuming text of roughly 10-12pt standard font (an optional template is provided on blackboard)

- ◆ Your submission MUST include the following information

- ◆ CS7DS4 / CSU44065 Data Visualization 2025-26

- ◆ Assignment 3

- ◆ Student Name : *<your name>*

- ◆ Student No: *<your student number>*

- ◆ Declaration:

"I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at <http://www.tcd.ie/calendar>.

I have also completed the Online Tutorial on avoiding plagiarism 'Ready Steady Write', located at [http://tcd-  
ie.libguides.com/plagiarism/ready-steady-write](http://tcd-<br/>ie.libguides.com/plagiarism/ready-steady-write)."

To expediate processing, include your name in the title e.g., A4\_SurnameFirstname.pdf

# Assessment Policy

If contacting me by Email: [John.Dingliana@tcd.ie](mailto:John.Dingliana@tcd.ie); Please include the module code 'cs7ds4' or 'csu44056' in the subject header to ensure it is seen

Do NOT submit Assignments by Email

## Late submission:

- ♦ In this module, for all assignments, unless stated otherwise in writing, a **10% penalty** applies for each day late up to a maximum of 6 days. After this, the mark will be capped at 40%.
- ♦ No further submissions will be accepted after 31/12/2025 – the grade will be returned as 0 after this date.
- ♦ If you submit after the deadline, you should Email me to let me know that a submission has been made late.
- ♦ There will be a grace period of 2 hours after the deadline for potential accidental/technical issues. Thus, there is no need to email me if you are just a few minutes late

**Extensions (for exceptional circumstances):** Submit a request for extension by Email **before the deadline** providing reasons and any relevant documentary evidence of exceptional circumstances

**Querying grades/feedback:** Email me so that I can refer to your submission and respond in due course. It's a large class, thus I generally won't be able to discuss specific feedback in the classroom.