Visual Analytics: An Introduction

Daniel Archambault
Pronounced: Arshambo-

Self Introduction

Self Introduction

- *Me:* Hi, my name is Dan.

Self Introduction

- *Me:* Hi, my name is Dan.
- *Class:* Hi Dan.

Self Introduction

- *Me:* Hi, my name is Dan.
- *Class:* Hi Dan.
- And I have a visual analytics problem

Self Introduction

- *Me:* Hi, my name is Dan.
- *Class:* Hi Dan.
- And I have a visual analytics problem
  - Many of them actually

Self Introduction

- *Me:* Hi, my name is Dan.
- *Class:* Hi Dan.
- And I have a visual analytics problem
  - Many of them actually
  - Fortunately, most of them are for you
  - **Bwhahahahahah!**

Self Introduction

- *Me:* Hi, my name is Dan.
- *Class:* Hi Dan.
- And I have a visual analytics problem
  - Many of them actually
  - Fortunately, most of them are for you
  - **Bwhahahahahah!**
- I am teaching this class for the for the third time

Self Introduction

- *Me:* Hi, my name is Dan.
- *Class:* Hi Dan.
- And I have a visual analytics problem
  - Many of them actually
  - Fortunately, most of them are for you
  - **Bwhahahahahah!**
- I am teaching this class for the for the third time
  - I am truly fortunate to be here

Self Introduction

- *Me:* Hi, my name is Dan.
- *Class:* Hi Dan.
- And I have a visual analytics problem
  - Many of them actually
  - Fortunately, most of them are for you
  - **Bwhahahahahah!**
- I am teaching this class for the for the third time
  - I am truly fortunate to be here
  - And still very nervous

More About Me..

- I'm Canadian, eh...
- Grew up in the capital of the country
  - that's Ottawa, Ontario
- I am fortunate to have studied/lived in many countries
  - Queen's University, Kingston, Canada
  - Univ. of British Columbia, Vancouver, Canada
  - Univ. of Victoria, Victoria, Canada
  - INRIA Bordeaux Sud-Ouest, Bordeaux, France
  - Univ. College Dublin, Dublin, Ireland
  - Swansea University, Swansea, Wales
- CS has taken me to some pretty cool places

The Teaching that I'm doing with Everyone

- Human-Centred Visual Analytics CSCM27/327
  - Information Visualisation
  - Data Analytics
  - How they work together
  - Graph Mining and Visualisation
- Assessment
  - One Assignment: M27 30%, 327 35%
  - Project presentation M27 15%, 327 25%
  - Project report 40%
  - Paper summary: M27 Only 15%
- Class
  - Lecture 1: Tues. 10am-11am, 1hr, School of Management 011
  - Lecture 2: Tues. 4pm-6pm, 2hrs, School of Management 011

More on the Project

- The project will be done in teams of 4-5
  - select a publicly available data set
  - create tools to analyse and visualise it
  - present your findings
- More information will be available in future

Programming in This Module

- You are expected to code in this module
- You are expected to learn libraries and programming languages **on your own**
  - Python, Jupyter Notebooks, Altair
- You are expected to practice programming on your own
- If you show better learning, you earn higher marks

Installing Altair and Jupyter

- You should install Jupyter and Altair
- You will likely need conda (miniconda) for this
- You are responsible for getting this working on your machine
- Instructions here:

  `https://altair-viz.github.io/getting_started/installation.html`

- I would suggest using jupyter notebooks and not lab.

Office Hours and Contact

- My Office: CoFo 334 (eventually)
- Office Hours: **Tues. 2pm-4pm (should not conflict)**
- Personal questions `d.w.archambault@swansea.ac.uk`

- I'm much better at explaining things in person
- Please make use of my office hours!

How to Study for this Course

- No required textbooks. Suggested ones:
- Colin Ware, **Information Visualization: Perception by Design**, 2012.
- Tamara Munzner, **Visualization Analysis and Design**, 2014.

1. Be prepared to take notes
   - I tend to organise a course around lectures
   - and use texts as support and alternate explanations

Class Ground Rules

1. Treat everyone in the class with respect
   - that includes your fellow students
   - that includes me (you'd be surprised)
2. There is no such thing as a stupid question
   - never be afraid to ask a question
   - but don't try to negotiate marks with me
3. Give your best effort for all parts of the assignment
   - the worst thing you can do, in my mind, is not submit
   - or demonstrate very little effort
4. Be positive and have fun!

Cheating

Cheating

- It sucks.

Cheating

- It sucks.
- Don’t do it.

Cheating

- It sucks.
- Don’t do it.
- Not even unintentionally
  - For individual assignments,
    - Do collaborate on white board or laptop
    - Before end of session, erase all recorded data from the session

Cheating

- It sucks.
- Don’t do it.
- Not even unintentionally
  - For individual assignments,
    - Do collaborate on white board or laptop
    - Before end of session, erase all recorded data from the session
- Don’t plagiarise off of other groups

Cheating

- It sucks.
- Don’t do it.
- Not even unintentionally
  - For individual assignments,
    - Do collaborate on white board or laptop
    - Before end of session, erase all recorded data from the session
- Don’t plagiarise off of other groups
- Don’t exchange code or copy text from each other

Cheating

- It sucks.
- Don’t do it.
- Not even unintentionally
  - For individual assignments,
    - Do collaborate on white board or laptop
    - Before end of session, erase all recorded data from the session
- Don’t plagiarise off of other groups
- Don’t exchange code or copy text from each other
- Obviously, no copying of code from the Internet...

Keys to Success

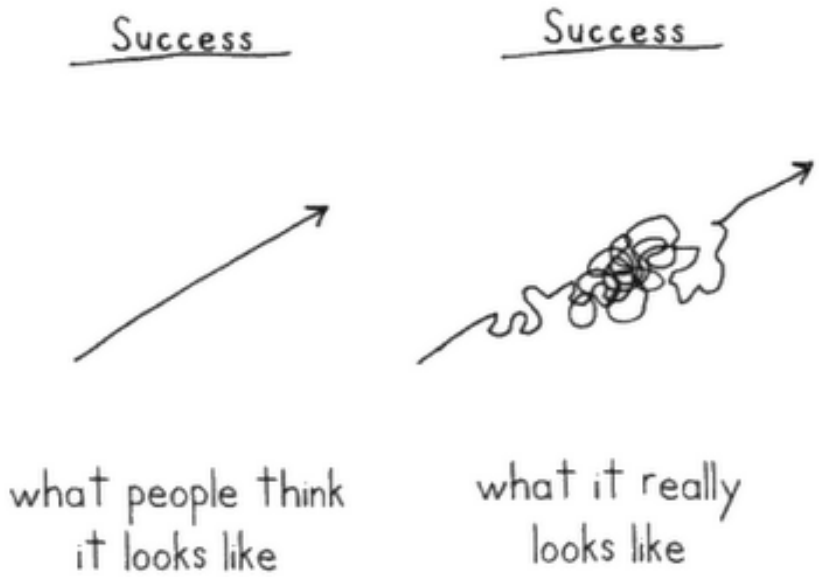
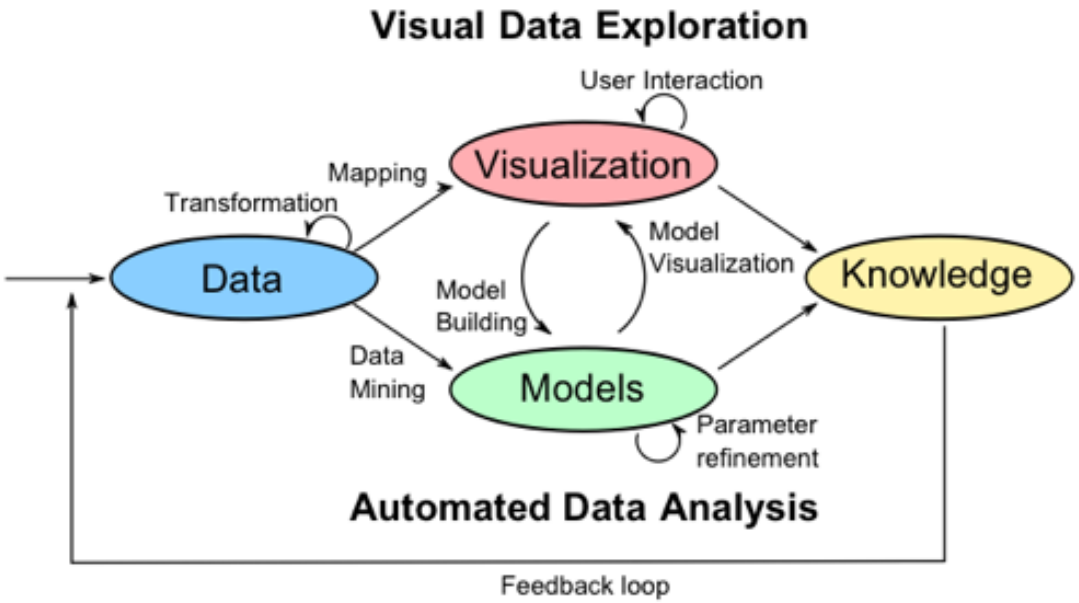


Image by Bernard Goldbach

1. Accept that you still have many things to learn.
2. Accept that everyone is unsure of themselves
   - corollary: some people good at looking confident (i.e. Trump)
3. Work hard and be prepared to make lots of mistakes
   - and don't beat yourself up over them.

Questions?

- Any Questions?



http://www.visual-analytics.eu/faq/

- **What is visual analytics?**

What is the pattern in this data?

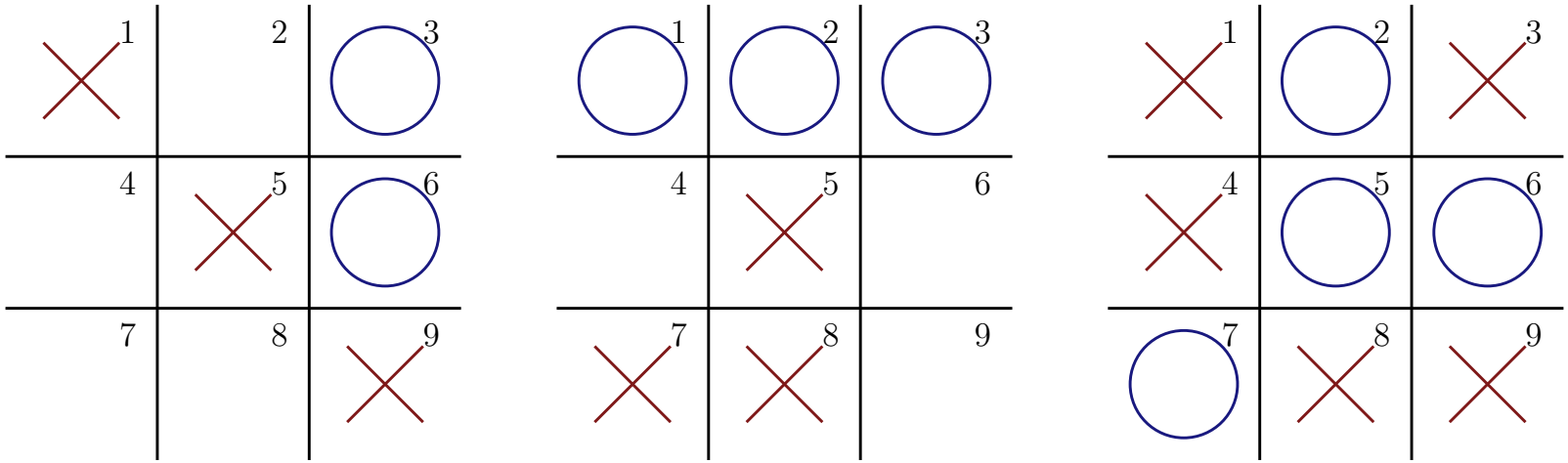
- Accepted strings of numbers:
  1. 5, 6, 1, 3, 9
  2. 5, 1, 8, 2, 7, 3
  3. 1, 5, 4, 7, 3, 2, 8, 9, 6
- Rejected strings of numbers:
  1. 1, 10, 4, 7, 3, 2, 8, 6, 9
  2. 0, 1, 2,
  3. 1, 5, 4, 4

Based on example by Pat Hanrahan

- What is the pattern?

The Answer

- The right visual representation helps us understand



- Visualisation helps us understand
- Visualisation helps us explain

Now It's Time to Scale Up...

- The tick-tac-toe board board is 10^6 squares
  - Can we visualise everything now?
  - How do we begin to look for patterns?
- Patterns are now complex and usually unknown

Real Data is Large and Complex

- Can we see the lines of noughts here?
- Visual analytics sought to answer these questions

History of Visual Analytics

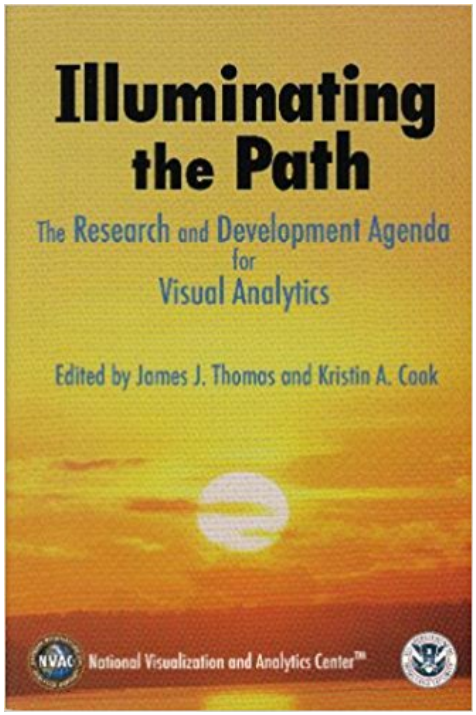
- Visual analytics has been around for a while before it was given a name
  - early years existed as part of visualisation in various forms

History of Visual Analytics

- Visual analytics has been around for a while before it was given a name
  - early years existed as part of visualisation in various forms
- Then, 9-11 happened ...
- Governments wanted to prevent it from happening again
  - large data sets
  - automatic processing for scale
  - visual tools that leverage human creativity
  - support for analytical reasoning and presentation
- Computers and data processing were viewed as a possible solution

Illuminating the Path and Jim Thomas

- Jim Thomas often credited with formally defining visual analytics



- First IEEE Visual Analytics Conference in 2006
- Your lecturer was there!
- New area to help find solutions to these problems

My View...

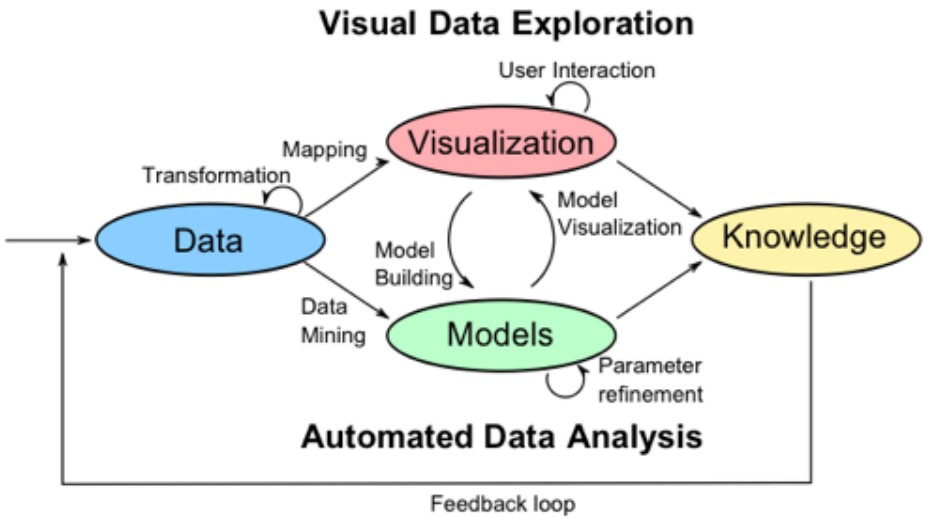
- Political beliefs aside...
- Visual analytics defines tools to support analytical reasoning about data
- It involves two main ingredients
  - data analytics techniques and data mining
  - visualisation techniques for exploration and explanation
- Many diverse applications exist for this technology

Modern Definitions of Visual Analytics

- New areas spend time defining what they are and what they are not
- VAST was part of the visualisation community for years
- It has broadened beyond its initial intention
  - security
  - biology and bioinformatics
  - social media analysis
  - smart cities and intelligent infrastructure
  - machine learning + vis
    - 3 ML paper sessions + 1 panel at this year's conference
- Intent: scalably support human analytical reasoning process

Definition: Daniel A. Keim

- A way of processing information in a way that is transparent for analytic discourse
  - visualisation communicates to user
  - machine learning automates data process
  - leverage strengths of human and machine



Keim D., Andrienko G., Fekete JD., GÃ¶rg C., Kohlhammer J., MelanÃ§on G. (2008) Visual Analytics: Definition, Process, and Challenges. In: Kerren A., Stasko J.T., Fekete JD., North C. (eds) Information Visualization. Lecture Notes in Computer Science, vol 4950. Springer, Berlin, Heidelberg

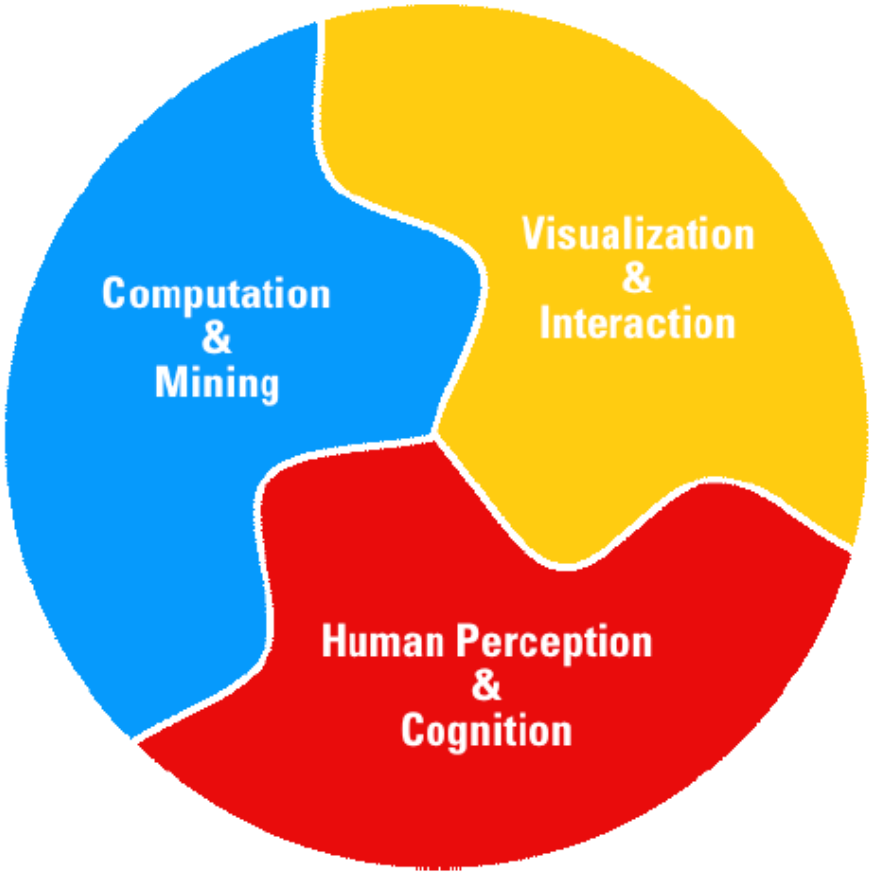
Shneiderman Mantra for Visualisation, 1996

- overview first, zoom and filter, then details on demand
- overview first, zoom and filter, then details on demand
- overview first, zoom and filter, then details on demand
- overview first, zoom and filter, then details on demand
- overview first, zoom and filter, then details on demand
- overview first, zoom and filter, then details on demand
- ...

Keim Mantra for Visual Analytics, 2005 & 2008

- analyze first, show the important, zoom filter & analyze, then details on demand
- analyze first, show the important, zoom filter & analyze, then details on demand
- analyze first, show the important, zoom filter & analyze, then details on demand
- analyze first, show the important, zoom filter & analyze, then details on demand
- analyze first, show the important, zoom filter & analyze, then details on demand
- analyze first, show the important, zoom filter & analyze, then details on demand
- ...

Visual Analytics: Silvia Miksch



- Interplay between areas of visualisation, mining, and perception

Conclusion

- Visual analytics leverages the advantages of visualisation and data mining
- Tackles problems that cannot be solved using one or the other exclusively
- In this module, we design systems to visualise and automatically process data sets to support analytical reasoning