

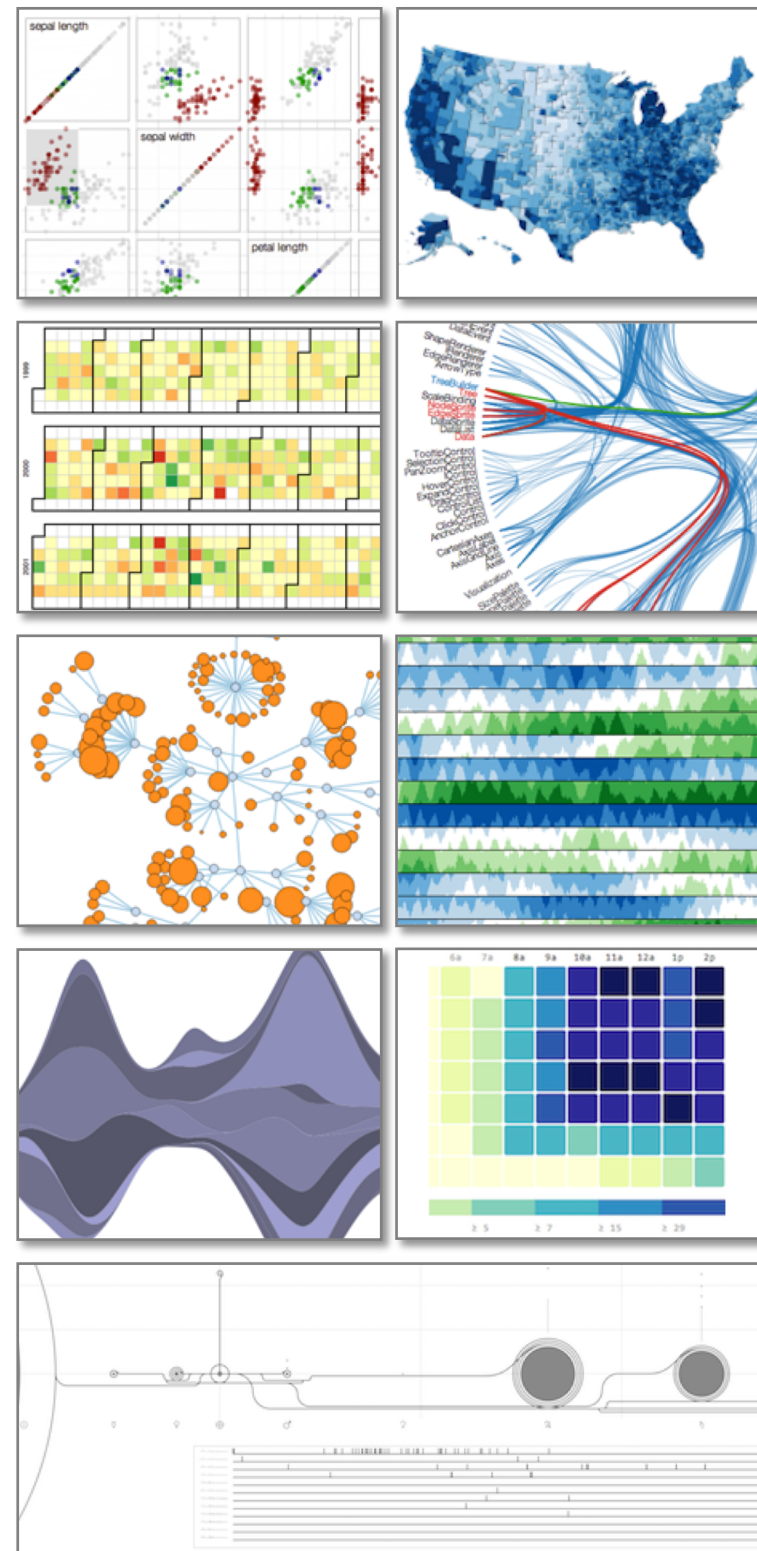
# Data Visualization (INF552)

Fundamental principles of visual perception  
for the design of interactive visualizations

Visualization techniques for:

- Multi-variate data
- Time-series
- Hierarchical structures and networks
- Geographical information systems and multi-scale spaces

*The course addresses both theoretical and practical aspects,  
with hands-on development of visualizations  
for all of the above types of data,  
using D3 (Web-based development).*



# Data Visualization

Emmanuel Pietriga

## Course Organization

9 sessions, each split into:

- 2 hours of lecture,
- 2 hours of exercises (PC)

## Course Description

The visual representation of data takes full advantage of the human visual system in terms of perception and cognition. Elaborate patterns, interesting data points and outliers can easily be identified, individual data points and sets can efficiently be compared and contrasted, provided that the data is properly represented. Visualization enables users to explore their data in an interactive manner, to get overviews and drill down to detailed views, following processes that yield insights that would be difficult to obtain using fully automated data analysis techniques from the fields such as, e.g., data mining or machine learning. The two serve different purposes, but can complement one another very effectively. Visualization can for instance help formulate hypotheses, that can then be tested using statistical tests or other elaborate data analysis techniques. Beyond these exploratory aspects, data visualization can also support decision making, and plays a central role in the communication of findings to a wide range of audiences. This course first gives an overview of the field of data visualization. It then discusses fundamental principles of human visual perception, focusing on how they help inform the design of visualisations. The following sessions focus on visualization techniques for specific data structures and discuss them in depth from both design and implementation perspectives, including: multi-variate data, hierarchical structures, networks, time-series, and geographical data. All exercises are based on the D3 software library (Data-Driven Documents), that enables developers to create a wide range of interactive, Web-based visualizations that run on a wide variety of platforms, ranging from desktop workstations to mobile devices.

## Books

- Tamara Munzner. *Visualization Analysis and Design*. A K Peters/CRC Press, 2014.
- Edward Tufte. *The Visual Display of Quantitative Information (2nd edition)*. Graphics Press, 2001.
- Colin Ware. *Information Visualization - Perception for Design (2nd edition)*. Morgan Kaufmann/Elsevier, 2004.

# Session details

---

## Session 1 - Introduction to Data Visualization

### Course:

- Human-Computer Interaction, from researchers to practitioners
- The benefits of visualising data
- Data visualization and Scientific Visualization
- A brief history of the field, covering seminal examples and state-of-the art techniques
- Visualization on the Web

### Exercises:

- The exercises of this session are designed to make sure all students have basic knowledge about the languages and paradigms that will be used throughout the following exercise sessions: HTML/SVG DOM manipulation in Javascript, basic CSS rules, event-based programming with Javascript.

---

## Session 2 - Fundamental Principles of Visual Perception for Visualization

### Course:

- Marks and visual variables
- Pre-attentive channels
- Gestalt laws
- Color spaces and color usage
- Representation and presentation of data
- Visual mappings, characteristics of individual visual variables

### Exercises:

- The exercises of this session will introduce the D3 library using simple examples that illustrate some of the principles discussed during the course, focusing on how to achieve mappings between data and visual variables, and how to update visualizations dynamically.

---

## Session 3 - Multi-variate Data Visualisation

### Course:

- Multi-variate data formats
- Nominal, ordinal and quantitative data
- Typical visualisation pipelines
- Multi-variate data visualization techniques
- The pitfalls of 3D data visualization
- Foundational concepts, including coordinated visualizations, brushing & linking, small multiples, dynamic queries.

### Exercises:

- The exercises of this session will introduce D3's asynchronous data resource loading mechanism and show how to create highly interactive visualizations that combine multiple variables using reusable components: bar charts, scatterplots, etc.

---

## Sessions 4 & 5 - Visualization of Hierarchical Structures and Graph Structures

### Course:

- Trees and networks
- Node-link diagrams, adjacency matrices, hybrid representations
- Interesting motifs and graph measures
- Layout algorithms
- An overview of tree and network visualization techniques
- Visualization of static and dynamic networks

### Exercises:

- The exercises of this session will focus on reusing existing D3 source code samples and adapt tree and graph datasets for their visualization using different layout and aggregation techniques, including force-directed graphs, hierarchical edge bundling, reorder-able adjacency matrices.

---

## Session 6 - Visualization of Temporal Data (time-series)

### Course:

- Visualization techniques for time-series: line charts, horizon graphs, colorfield, linear vs. circular and spiral layout, animations
- Multi-scale visualization of time-series (different temporal granularities)
- Exploratory visualization of time-series (pattern identification, signal processing)
- Time-series querying

### Exercises:

- The exercises of this session will focus on how to create interactive time-series visualizations for both static and real-time (monitoring) data, investigating alternative ways to combine multiple time-series.

---

## Session 7 - Animated Transitions and Multi-scale Navigation

### Course:

- Benefits and pitfalls of animated transitions in interactive visualization
- Animation pacing functions, perceptual continuity and cognitive load
- Multi-scale navigation interface schemes: panning & zooming, focus+context techniques, overview+detail techniques
- Space-scale diagrams
- Semantic zooming
- Quadtrees and related structures to encode multi-scale representations

### Exercises:

- The exercises of this session will focus on D3's animated transition API, ranging from very simple, purely declarative examples to more elaborate ones that involve staging, tween transitions or generic timers that give more control on the computation of transition trajectories.

---

## Session 8 - Geovisualization & Geographical Information Systems

### Course:

- Introduction to GIS UI front-end core concepts
- Data structures and formats, data feeds, map tiling
- Map projections
- Geovisualization techniques for static and dynamic data
- Spatial and temporal multiplexing

### Exercises:

- The exercises of this session will focus on D3's geo module: handling geojson data, projecting and drawing maps, overlaying data from additional sources, adding support for interactive panning & zooming.

---

## Session 9 - Collaborative visualization in multi-surface environments

### Course:

- Beyond the desktop: platforms for the interactive visual and collaborative exploration of large datasets, including cluster-driven ultra-high-resolution wall displays, collaborative virtual environments (CAVEs), tabletops, handheld tablets, and multi-surface environments that combine those devices
- Software architectures and libraries for such environments
- Some example contexts of use: emergency response & crisis management, real-time monitoring of complex systems

### Exercises:

- The exercises of this final session show how web sockets can be used to remotely control the parameters of a D3 visualization from another device such as, *e.g.*, a smartphone, a tablet, or even another computer, showing how a visualization system can be deployed across devices in a multi-surface environment.