

Lecture 5: Fundamental Concepts, Infrastructure, and Case Study

Ways To Deal With Large Datasets

- * Temporal Partitioning
 - * Navigation: Pan, Rotate, Zoom
 - * Geometric vs. Semantic Zooming
- * Spatial Partitioning
 - * Multiple Coordinated Views
 - * Overview + Detail
- * Aggregation & Filtering

Temporal Partitioning

- * Look various aspects of the data, but not at the same time
- * Partition the data or partition the aspects over time
- * Navigate: **Panning**
 - * Have large dataset and is supported by many pieces of software
 - * Scroll down in Microsoft Word
 - * Have access to the whole data but do not see at the same time
- * Navigate: **Rotation**
 - * For 3D spatial representation
 - * Rotate the data
- * Navigate: **zooming**
 - * **Geometric Zooming**: make think bigger
 - * **Semantic Zooming**: adding information when zooming in
 - * Google Map is a combination of **Geometric** and **Semantic** zooming

Spatial Partitioning: Multiple Coordinated Views (MCV)

- * Chop up the dataset and put it into different spatially parts on your screen or on the area you have for visualising
- * Linked views: those different spatial parts are linked to some extent, and could be linked on different levels
 - * **Link the data**
 - * Three types: all data shared, overview + details, small multiples
 - * Linking: coordination between views
 - * Brushing: select groups of data points
 - * **Link the visual encoding**:
 - * use the different encodings or visualisations and share them in multiple visualisation
 - * **Link the navigation**:
 - * you zooming into one visualisation, it also zooms automatically into another visualisation
- * Major design choices
 - * Share Encoding
 - * Share Data
 - * Share Navigation

Shard Data:

- * All data Shared: same data, different encodings
- * Overview + Detail: one view shows entire dataset, other one user selected subset
- * Small multiples: different partition of dataset in each view, same encoding
 - * Matrix consists of multiple small scatterplots and each scatterplot shows a different partition of the data

Visual Information Seeking Mantra by Ben Shneiderman (1996)

- * How people interact with large datasets
 - * Overview First
 - * Zoom and Filter
 - * Details on Demand

Aggregation:

- * By time: if you have a network that changes over time, you could just aggregate everything that was changed in the last week
- * By space: people on Twitter and how they interact: city -> country -> continent

Collapsible Force Layout: what can you do with the aggregation hierarchy:

- * Elaboration: Focus on details
- * Abstraction: Reducing details

Selection/Filtering:

- if the user selects data in one visualisation or one of the small multiples, those are highlighted and it changes the visual encoding in other parts

General Purpose Tool vs. Narrow Tool for Specific Purpose

- * General purpose tool
 - * Can flexible handle a wide range of data
 - * Can often not be used to solve complex domain specific problems
- * Narrow tool
 - * Designed for specific context and datasets
 - * Vis designer has made lots of choices
 - * User cannot override them