Interactive graphics. Text Visualization

Lecture 5

732A98

Interactive graphics

- Key tool for visual analytics
- Much more efficient than static graphics

Examples:

- Navigation (panning, rotation, zooming)
- Selection (highlighting)
- Connecting (linked views)
- Filtering (sample)
- Reconfiguring (change aesthetics)
- …

.导航(平摇 旋转 聚焦)
.选择(高亮)
.链接(链接视图)
.过滤(样品)
.变换(变换美感)

2/37

Theory of interaction

- Current visualization systems contain certain interaction
  - Limited Features
- Why do we need a theory?
  - To understand what can be interacted and how
  - To see limitations of existing visualization software
  - To be able to propose new relevant interaction tools missing

3/37

Interaction operators

交互操作

Navigation operator:

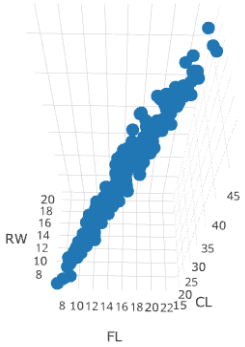
导航操作
.相机位置
.观测方向
.细节程度（等级描述）

- Camera location
- Viewing direction
- Level of details (e.g. hierarchical representations)

4/37

Interaction operators

Example:
Australian crabs



Selection operators

选择操作

- User isolates a subset of objects
  - Highlighting
  - Masking
  - Focusing
- How to implement?
  - Click
  - Click+hold
  - Bounding box, lasso

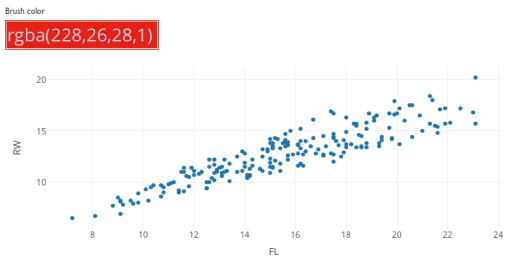
用户选择数据一部分
.高亮
.遮盖
.聚焦
如何操作
.点
.点加保持
.方状选择，套索选择

Selection operator

Example:
Australian crabs

Brush color

rgba(228,26,28,1)



Connection operators

- Related observations are linked in corresponding views
- Selection operator+Connection operator = Brushing
  - Persistent and transient

链接操作

有关的数据在相关的视图内链接
选择操作+链接操作=涂刷
永久的和临时的

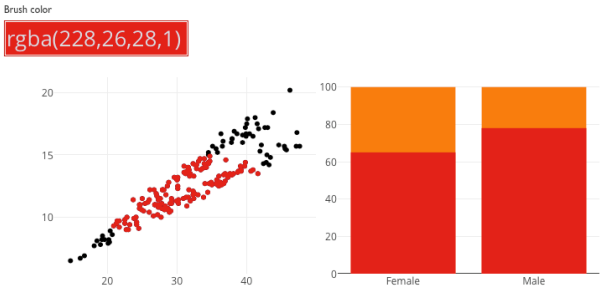
Connection operators

- Challenges
  - How to link data with various mappings
    - Ex: Histogram and scatter plot
    - Ex2: Contour plot and bar chart
    - Ex3: link in which direction? (Hierarchically connected)
  - How to define the corresponding link (key)?
  - Allow for disconecting views?

单向连接

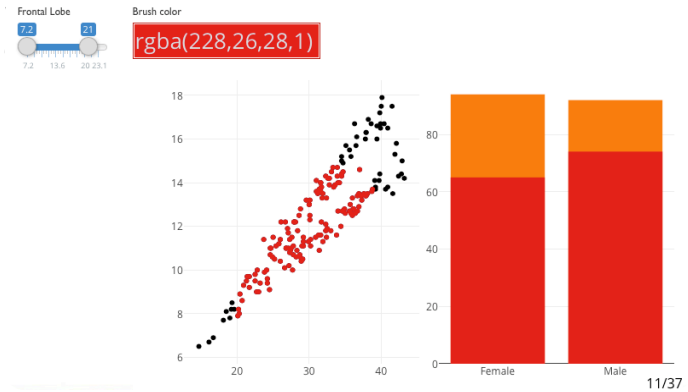
Connection operators

- Australian crabs:
- Which sex do the upper-cluster-crabs have?



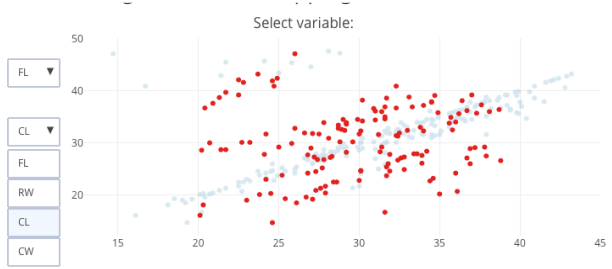
Filtering operators

- Reducing data acc. to specifications



Reconfiguring operators

- Transforming data
  - Sorting rows, reorder columns, MDS
  - Change aesthetics mapping

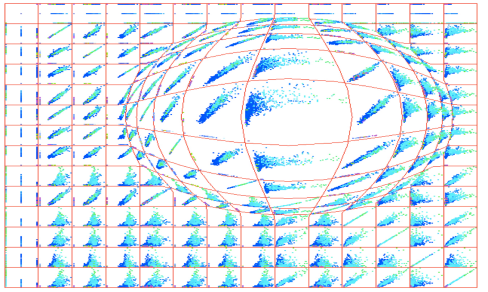


Encoding operators

- Changing the visualization type
- Changing aesthetics
- Another color map
- Change shapes
- …

Abstraction operators

- Distoring objects locally or globally



Interaction operands

- What can the operators be applied to?
- What actions does it imply?

Screen space:

- Navigation: pixel-wise actions
- Selection: sets of pixels (boxes, polygons,…)
- Abstraction: distortion of image (fisheye)
- Filtering: removing some pixels

Interaction operands

Data value space

- Operate observations instead of pixels
- Navigating: translate the axis range
- Zooming: increase/decrease axis range
- Filtering: sample the data, sample dimensions
- Reconfiguring: sorting observations, dimensions, nonlinear transformations

Interaction operands

Data structure space

- Different data structures exist: matrix, list, graph,…
- Navigation operator: how navigate the view in the long tree?
- Selection operator: node in the tree is selected -> subbranches must be selected
- Filtering operator: Social network: click on node - nodes that are X links away disappear
- Abstraction/Elaboration: histogram with zooming - recompute bars?

Attribute space

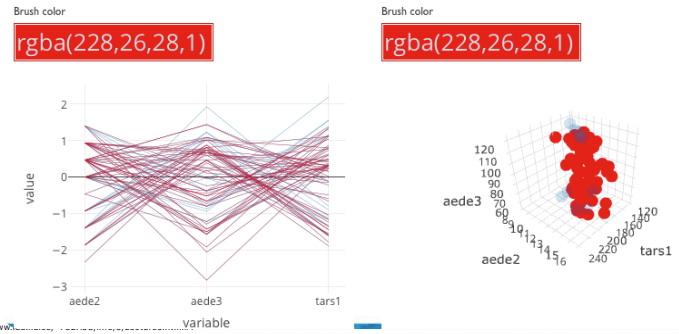
- Working with aesthetics
- Navigation: change range of aesthetics to certain interval (show certain range of colors)
- Encoding operator: change shapes of symbols, non-linear color mapping
- Selection: highlight certain ranges of aesthetics (highlight stars)

Visualization structure space

- Dashboards, scatter plot matrices
  - How to navigate user in these?
  - Which components of dashboard can user hide, move, rearrange?
  - Distortion of certain elements

Example: flea data

- Measurements of fleas and their types
- Which operators are available here?
- Which clusters do you see?



Interactive visualization in Plotly

Without Shiny: key ingridients

- Create cross-talk objects from your dataframes by `d<-SharedData$new(data, key, group)`
  - key parameter should point to same observations in dataframes
    - Not specified: row id used
  - Group - a unique name related to the same data

Interactive visualization in Plotly

Without Shiny: key ingridients

1. `highlight`-function:
   - applied to Plotly object
   - parameters:
     - on: 'plotly_click', 'plotly_selected',…
     - persistent: TRUE/FALSE 持久
     - dynamic:T/F - enables color selector
     - selectize: T/F text field for selection

Interactive visualization in Plotly

```
d <- SharedData$new(crabs)
scatterCrab <- plot_ly(d, x = ~CL, y = ~RW) %>%
  add_markers(color = I("black"))

barCrab <-plot_ly(d, x=~sex)%>%add_histogram()%>%layout(barmode="overlay")

subplot(scatterCrab,barCrab)%>%
  highlight(on="plotly_select", dynamic=T,
            persistent = T, opacityDim = I(1))%>%
  hide_legend()
```

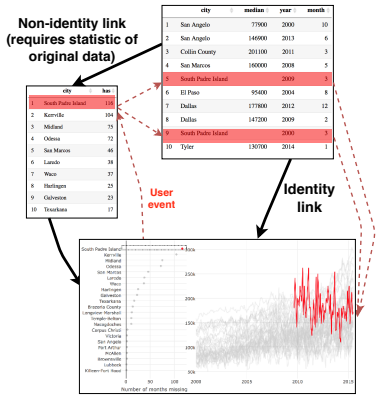

Interactive visualization in Plotly

```
plotly_data(barCrab)
```

```
## # A tibble: 200 x 9
##    species sex   index    FL    RW    CL    CW    BD .crossTalkKey
##  * <fct>   <fct> <int> <dbl> <dbl> <dbl> <dbl> <dbl> <chr>
##  1 Blue    Male      1   8.1   6.7  16.1  19     7   1
##  2 Blue    Male      2   8.8   7.7  18.1  20.8   7.4 2
##  3 Blue    Male      3   9.2   7.8  19    22.4   7.7 3
##  4 Blue    Male      4   9.6   7.9  20.1  23.1   8.2 4
##  5 Blue    Male      5   9.8   8    20.3  23     8.2 5
##  6 Blue    Male      6  10.8   9    23    26.5   9.8 6
##  7 Blue    Male      7  11.1   9.9  23.8  27.1   9.8 7
##  8 Blue    Male      8  11.6   9.1  24.5  28.4  10.4 8
##  9 Blue    Male      9  11.8   9.6  24.2  27.8   9.7 9
## 10 Blue    Male     10  11.8  10.5  25.2  29.3  10.3 10
## # ... with 190 more rows
```

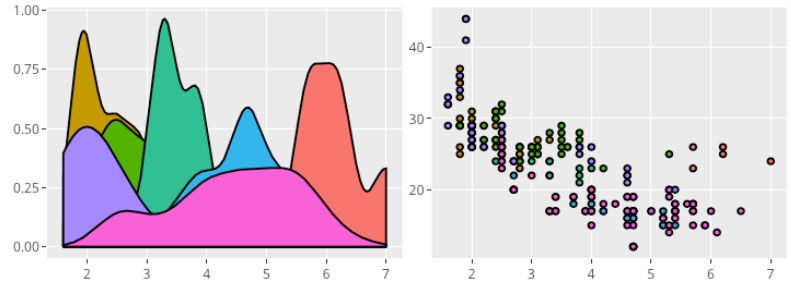
Interactive visualization

- Link can be one-to-many or even dynamic (scatter/barchart)



Interactive visualization

```
m <- SharedData$new(mpg)
p1 <- ggplot(m, aes(displ, fill = class)) + geom_density()
p2 <- ggplot(m, aes(displ, hwy, fill = class)) + geom_point()
subplot(p1, p2) %>% highlight("plotly_click") %>% hide_legend()
```



Text visualization

Applications:

- Articles, books
- Emails, blogs, websites
- Program Logs
- Collections (corpus) of books, blogs,..

Analysis:

- Understanding structure/context of text
- Group similar documents

Text processing

- Vector Space Model
  - Count occurance of each word in a document
  - Columns: frequencies of words
  - **Term/document matrix** (for collection of documents): columns=words, rows=documents
- TF-IDF model
  - Measures relative importance of word in document
  - $Tf(w)$=frequency of word, $df(w)$=frequency of documents, N=#documents

Word cloud

- Words are placed in 2D
- Layout decided algorithmically
- $Tf(w) \rightarrow$ size of words
- Another aesthetics: color

Word cloud

Issues:

- Stopwords need to be removed
- Words sharing the same stem aggregated
- Synonyms
- "Satisfied"/"not satisfied" example
- Incorrect spelling?
- Hyphens and apostrophes 连字号 省略号
- Size mapping inaccurate (long words)
- Does not show the structure

WordTree

- Root= word of interest
- Branches = contexts the word used
  - Type a combination of words and observe the tree, check its branches

https://www.jasondavies.com/wordtree/
(https://www.jasondavies.com/wordtree/)

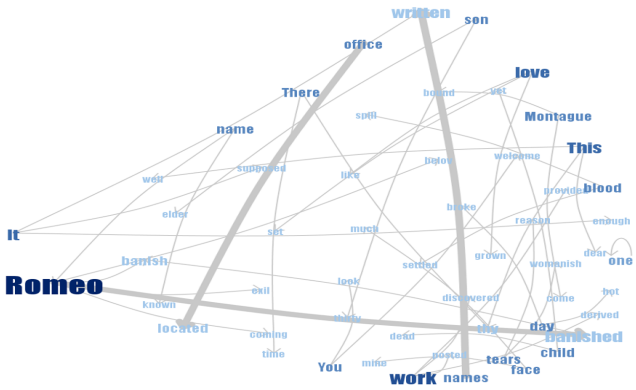
Phrase net

- Graph that analyses co-occurances
  - 'A and B', 'A of B', 'A's B', 'A B', 'A is B',…
- Understanding of context without reading

1. Go
   https://www.cg.tuwien.ac.at/courses/InfoVis/HallOfFame/2011/Gruppe08/H
   (https://www.cg.tuwien.ac.at/courses/InfoVis/HallOfFame/2011/Gruppe08/H
2. Download soft (need Java installed)
3. Upload File
4. Define connector word (is, are, [space], of, and,…)

Phrase net

- Example "Romeo and Julia" (is, are)



Phrase net

- Size of the word = word freq
- Thickness of the connection =co-occurance freq
- Color: dark colors -> word often to the left
- Close on map -> often close in the document

Analysis:

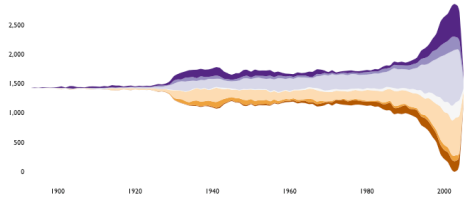
- Analyse most frequent
- Analyse connections and paths
- Analyse strength of paths and intensity of colors
- Click a specific word and look at the respective paths

Steam graphs

- Analyses thematic changes in document collections over time

https://hrbrmstr.github.io/streamgraph/ (https://hrbrmstr.github.io/streamgraph/)

- Example: #movies by genre,different years
- Perception problem: angles in areas



Text visualization

- Other tools for document collection:
- Place similar documents close in 2D (MDS, clustering, SOM)
- Graph/network visualization can be used for text
- A plenty of other visualizations for text:

http://textvis.lnu.se/ (http://textvis.lnu.se/)