36-315: Statistical Graphics and Visualization

Handout 26
Date: April 28, 2003

Iconic displays for hypervariate data

How can we improve the parallel-coordinate plot? Increase prominence, reduce clutter, without giving up connection.

Principles:

- Data-ink, prominence
- Visual encoding of magnitudes
- Visual connection between values

Star plot—Represents each profile with a circular shape. Variables are in a fixed position around the circle, with wedge area proportional to the value (as in Florence Nightingale's original wedge diagram). Variables are scaled to the range $[0, 1]$ and ordered by their correlation. The stars (cases) are themselves ordered according to a PCA projection onto one dimension. The variables are meant to be viewed as one shape, and cases can be compared by mentally overlaying and discriminating the shapes.

These star icons can be used in a map, or a time spiral.

Data image—Represents the data matrix with shaded squares. Variables are scaled and ordered as in the star plot.

Faces—Represents each profile with a cartoon face. Each parameter of the face (head size, mouth shape, eyebrow slant, etc.) is controlled by a variable. The variables are not represented symmetrically. Also, the parameters influence each other in complex ways, e.g. eye slant and eye shape are easier to see when the eyes are large.

Compared to parallel-coordinate plot, star plot and data image can show more cases (about 100). They are less cluttered because they use alignment and proximity instead of connecting lines. But star plot has a weaker encoding—hard to compare values across cases. Data image can compare across cases and across variables. But it requires a good ordering of cases, which is only possible when the variables are highly correlated.

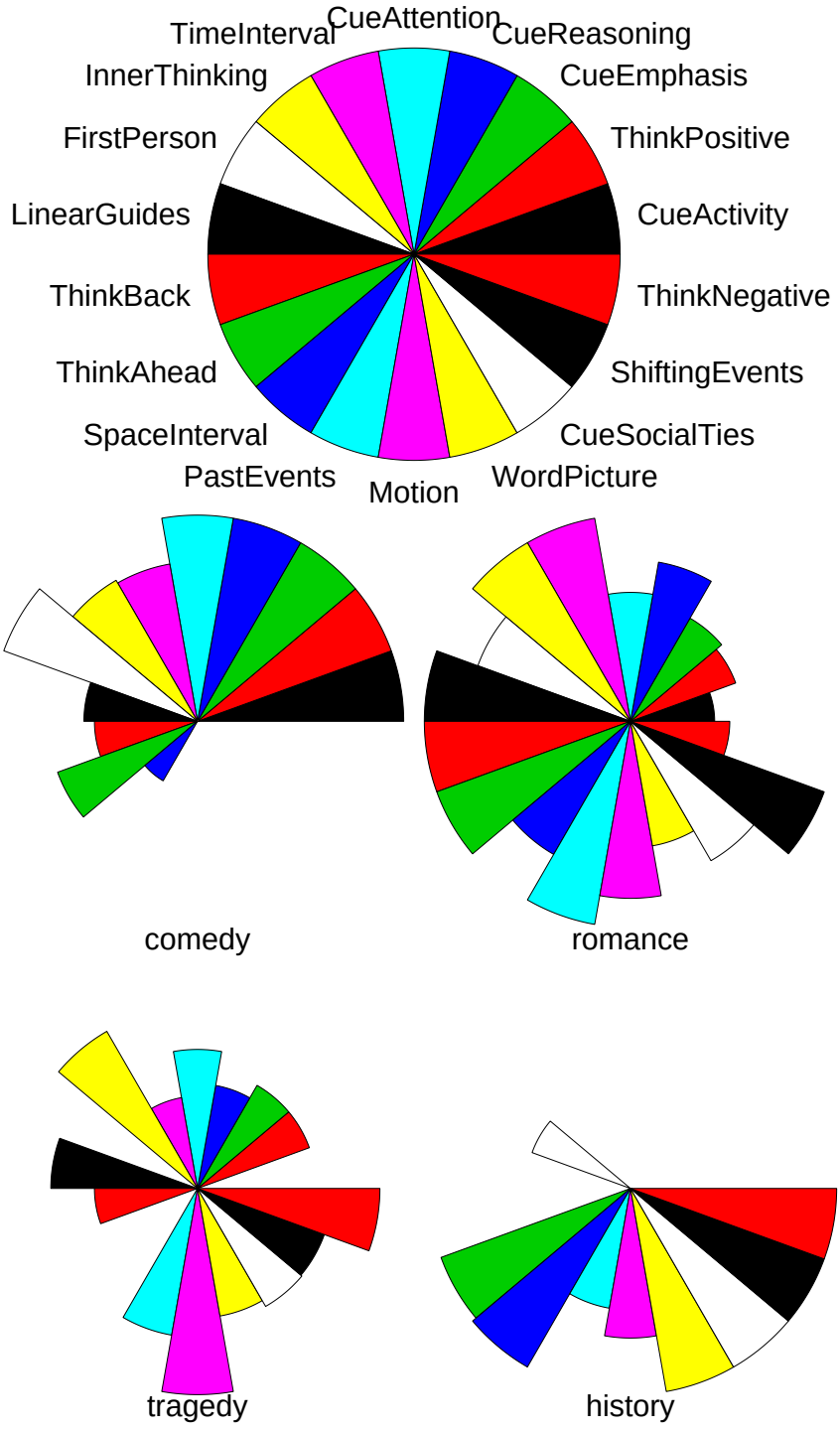
List of figures:

1. Stars on a map
2. Star plots of Shakespeare's plays
3. Data-images of Shakespeare's plays
4. Vote 2000 in Florida counties (image and stars)
5. Faces of Shakespeare's plays

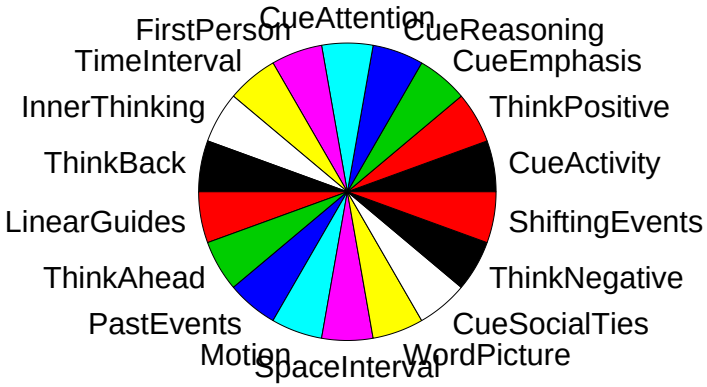
References

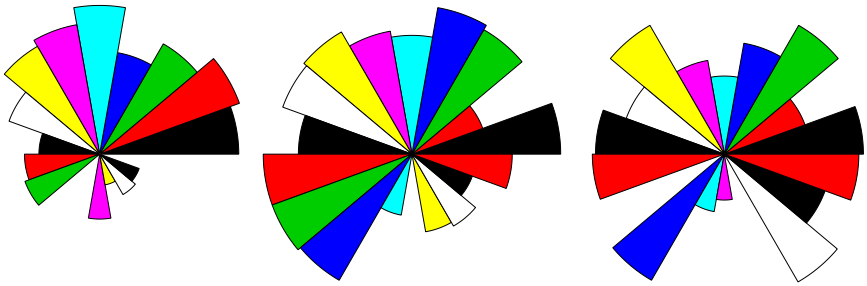
[1] H. Chernoff. "The Use of Faces to Represent Points in K-Dimensional Space Graphically", Journal of the Americal Statistical Association 68:361-368, 1973

[2] M. C. Minnotte and R. W. West. "The Data Image: A Tool For Exploring High Dimensional Data Sets", Proceedings of the ASA Section on Statistical Graphics, 1998. `http://math.usu.edu/~minnotte/research/pubs.html`

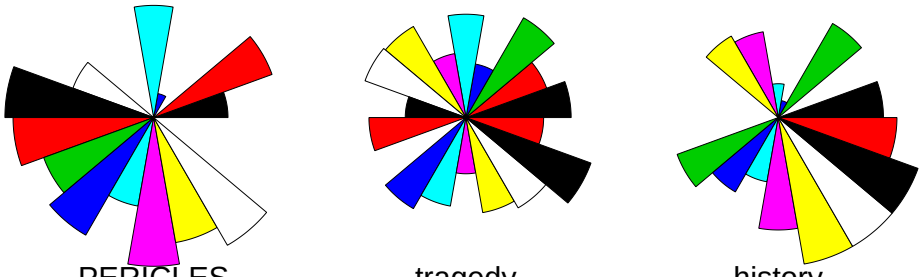


Stars of Shakespeare

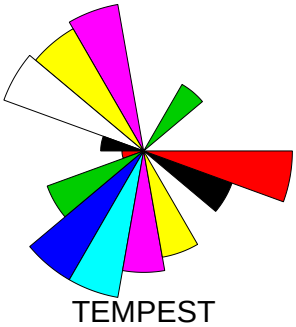




comedy WINTERS TALE CYMBELINE

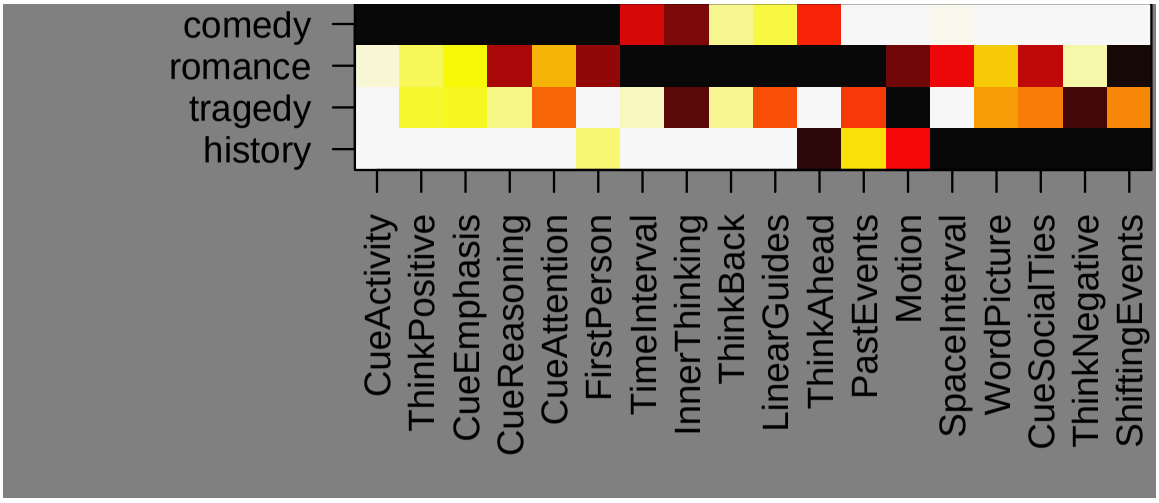


PERICLES tragedy history

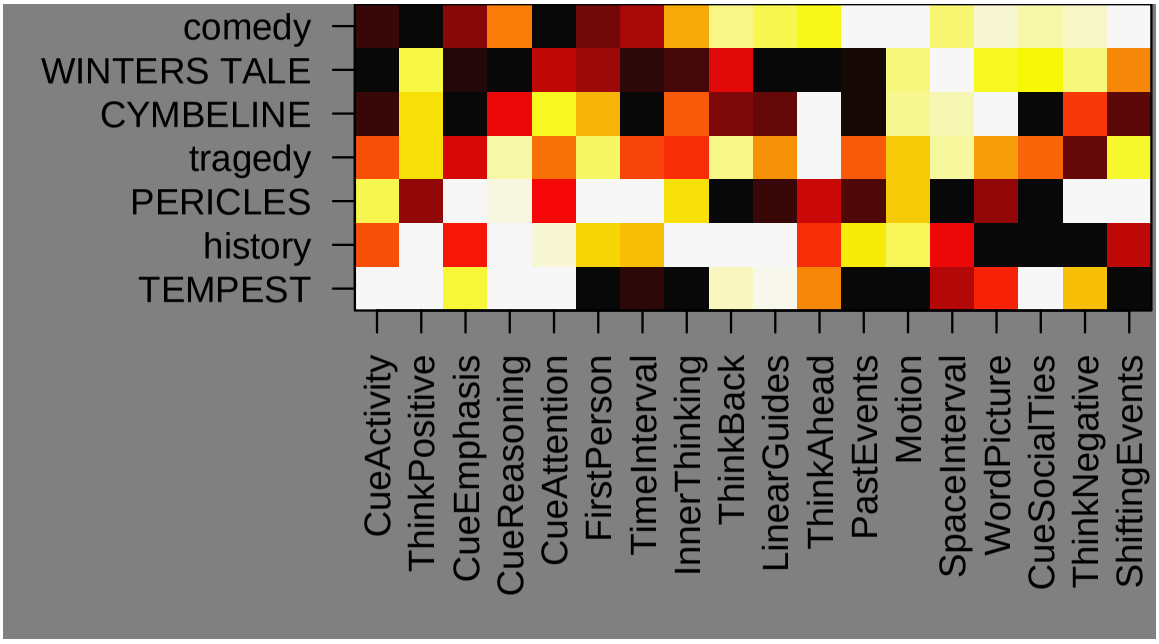


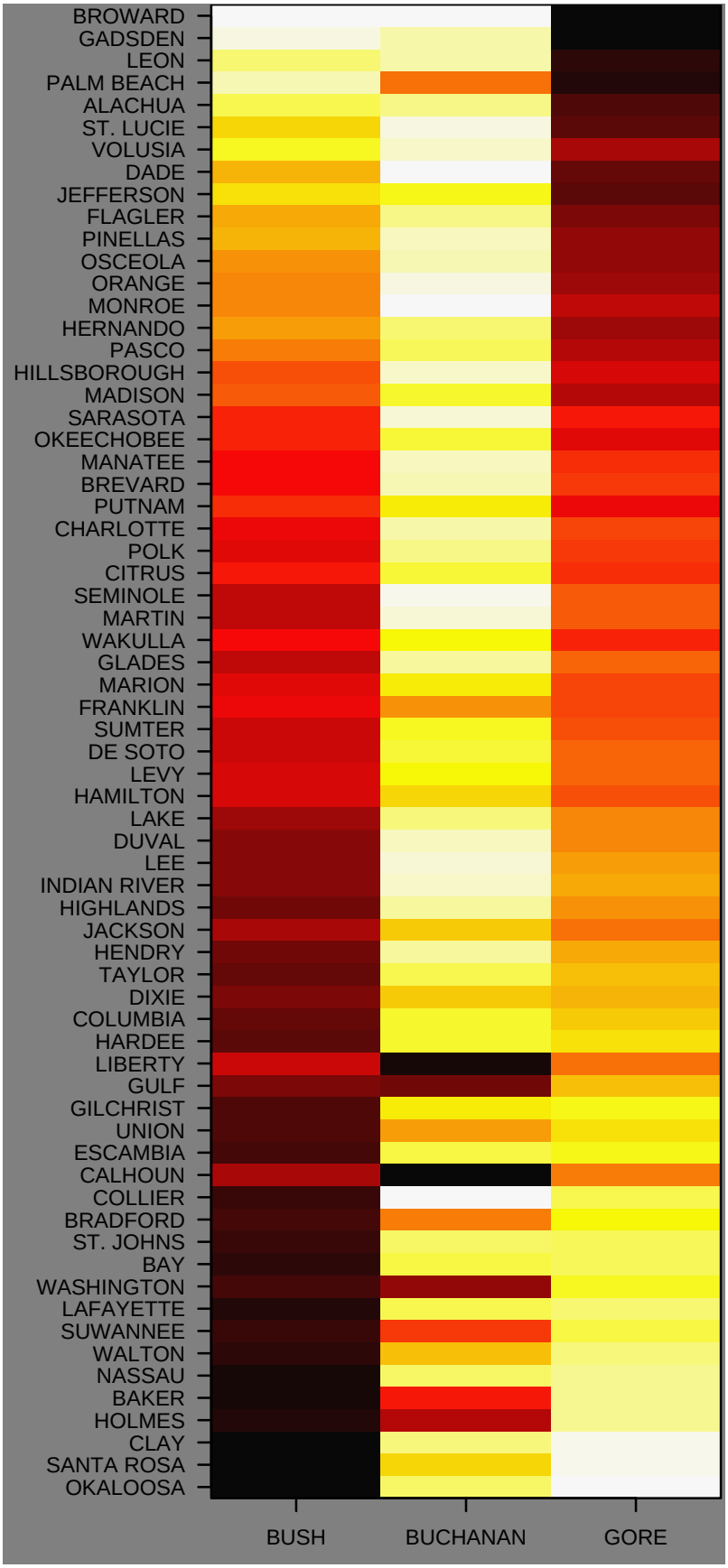
TEMPEST

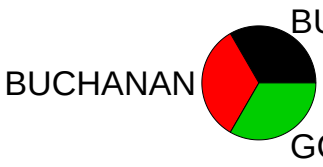
Can you group the romances?



Shakespeare's data-image




BROWARD
GADSDEN
LEON
PALM BEACH
ALACHUA
ST. LUCIE
VOLUSIA
DADE
JEFFERSON
FLAGLER
PINELLAS
OSCEOLA
ORANGE
MONROE
HERNANDO
PASCO
HILLSBOROUGH
MADISON
SARASOTA
OKEECHOBEE
MANATEE
BREVARD
PUTNAM
CHARLOTTE
POLK
CITRUS
SEMINOLE
MARTIN
WAKULLA
GLADES
MARION
FRANKLIN
SUMTER
DE SOTO
LEVY
HAMILTON
LAKE
DUVAL
LEE
INDIAN RIVER
HIGHLANDS
JACKSON
HENDRY
TAYLOR
DIXIE
COLUMBIA
HARDEE
LIBERTY
GULF
GILCHRIST
UNION
ESCAMBIA
CALHOUN
COLLIER
BRADFORD
ST. JOHNS
BAY
WASHINGTON
LAFAYETTE
SUWANNEE
WALTON
NASSAU
BAKER
HOLMES
CLAY
SANTA ROSA
OKALOOSA
BUSH
BUCHANAN
GORE

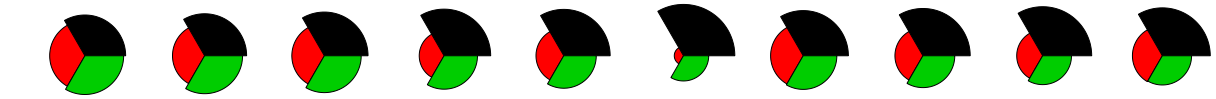


BROWARD LEON PALM BEACH ST. LUCIE MONROE
GADSDEN DADE ALACHUA VOLUSIA ORANGE

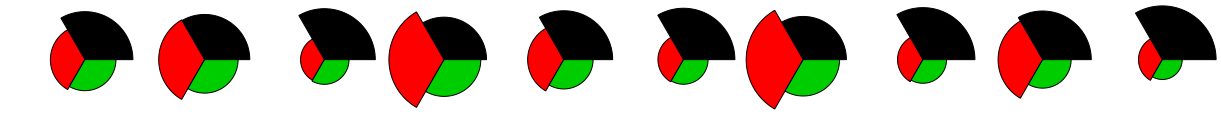
PINELLAS FLAGLER HERNANDO SARASOTA SEMINOLE
OSCEOLA JEFFERSON HILLSBOROUGH PASCO MADISON

MANATEE BREVARD OKEECHOBEE GLADES PUTNAM
MARTIN CHARLOTTE POLK LEE CITRUS

WAKULLA DUVAL HIGHLANDS MARION HENDRY
INDIAN RIVER LAKE DE SOTO SUMTER LEVY



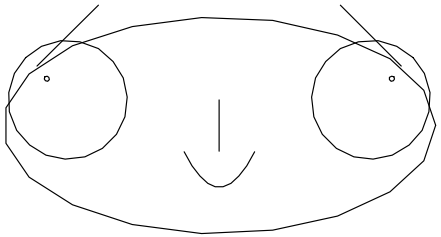
FRANKLIN JACKSON COLUMBIA DIXIE ESCAMBIA
HAMILTON TAYLOR COLLIER HARDEE GILCHRIST



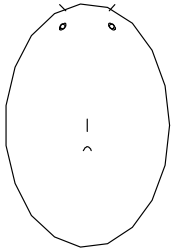
UNION ST. JOHNS BRADFORD CALHOUN WASHINGTON
GULF LIBERTY BAY LAFAYETTE NASSAU

SUWANNEE BAKER HOLMES OKALOOSA
WALTON CLAY SANTA ROSA

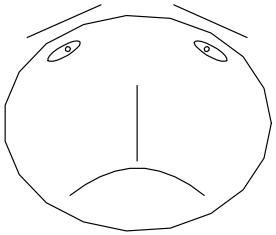
The Faces of Shakespeare



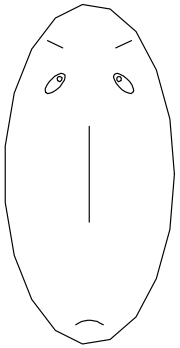
comedy



history

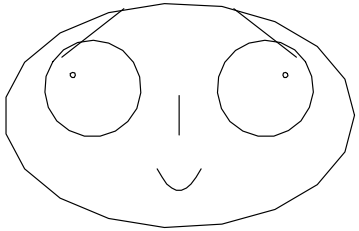


romance

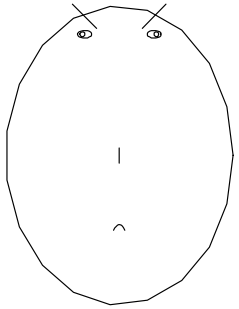


tragedy

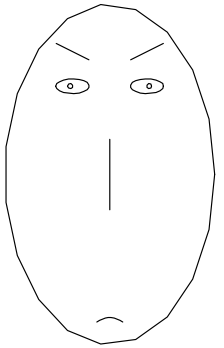
The Faces of Shakespeare



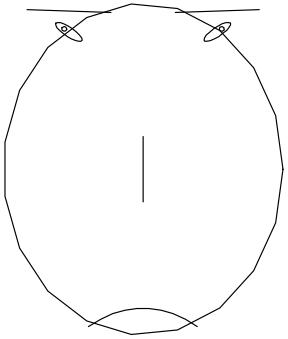
comedy



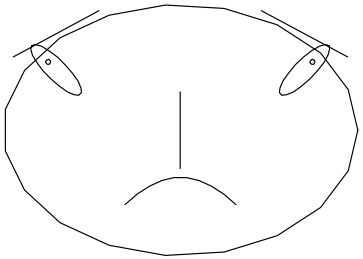
history



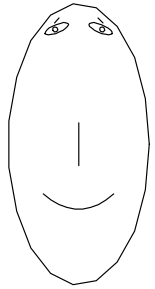
tragedy



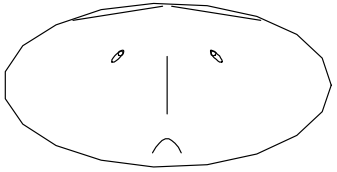
CYMBELINE



WINTERS TALE



PERICLES



TEMPEST

```
       area of face = CueEmphasis
      shape of face = FirstPerson
     length of nose = InnerThinking
  location of mouth = ThinkAhead
     curve of smile = ThinkPositive
     width of mouth = ThinkBack
   location of eyes = CueSocialTies
 separation of eyes = CueReasoning
      angle of eyes = CueActivity
      shape of eyes = PastEvents
      width of eyes = CueAttention
  location of pupil = Motion
location of eyebrow = ShiftingEvents
   angle of eyebrow = ThinkNegative
   width of eyebrow = TimeInterval
```