Interactive Statistical Visualization of Medical Data

Jürgen Symanzik

Utah State University, Logan, UT, USA

***e-mail: symanzik@math.usu.edu**

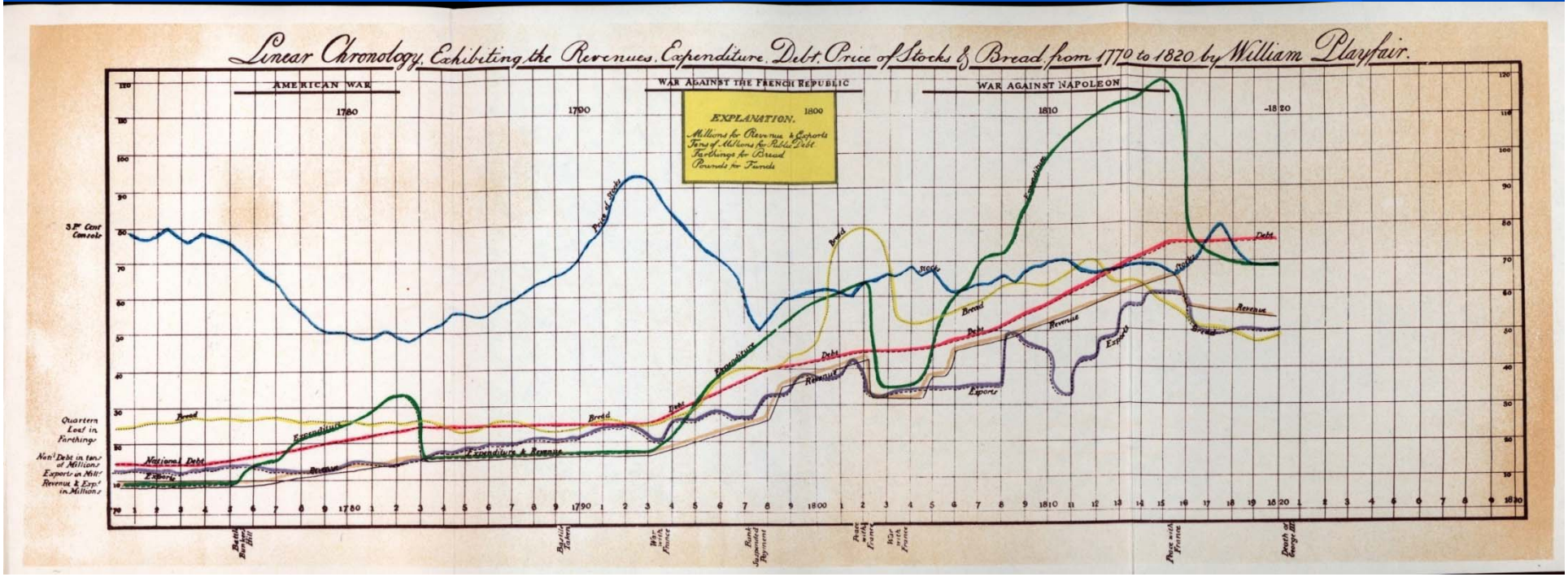
WWW: http://www.math.usu.edu/~symanzik

Contents

- Why Graphics for “Outsiders”?
- Exploratory Graphics for Actigraphy Data
- Carpal Tunnel Syndrome Data
- Micromaps for West Nile Virus and Oral Clefts
- Conclusion

Historical 250th Birthday

- William Playfair: Born on September 22, 1759
- "Father" and "Inventor" of Statistical Graphics
- Modern plots still look very similar:

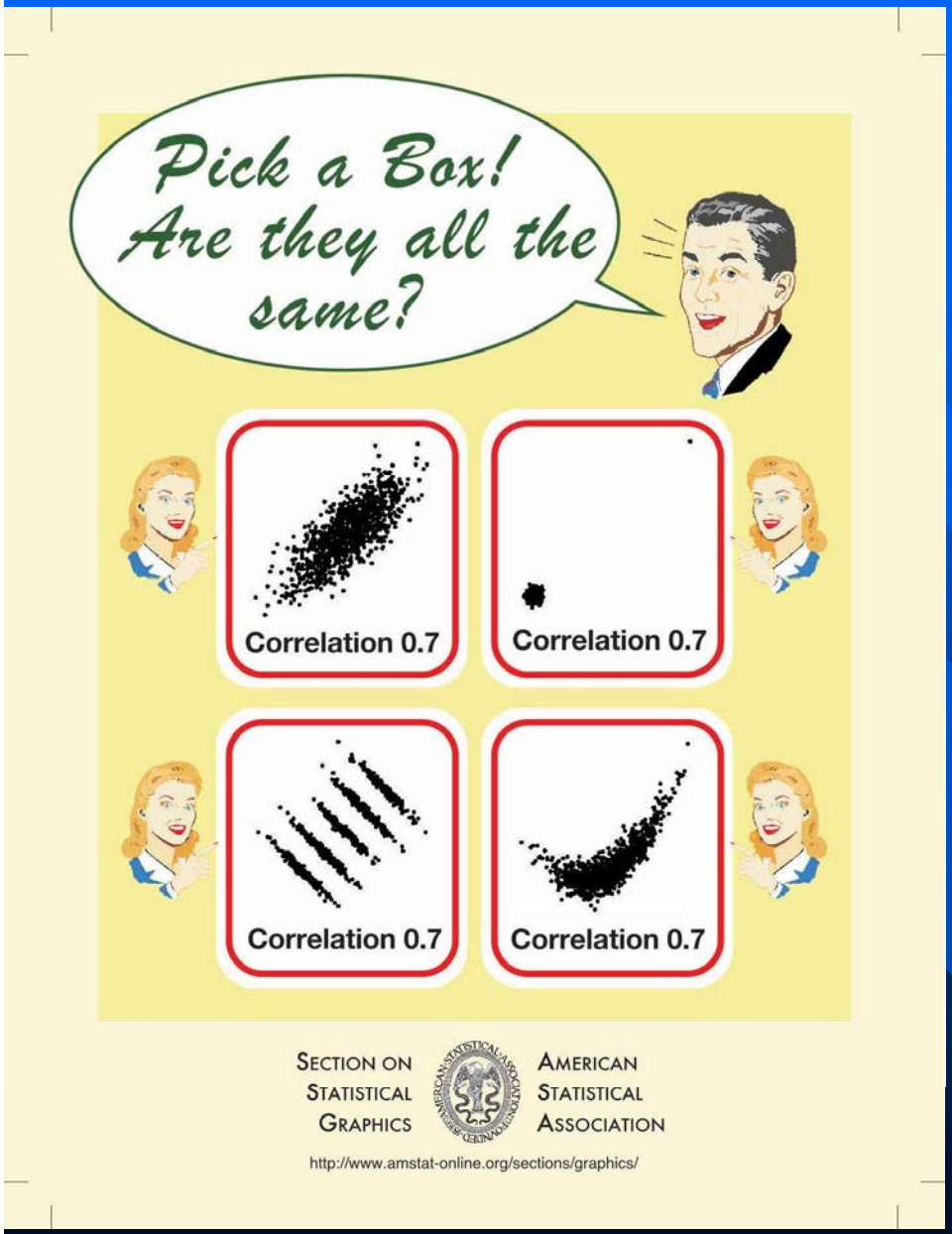


Interview with Andreas Buja*

- " … when I think back about what really may have had the most impact in what I did in the various labs that I worked, it's graphics! You know whenever I made a striking picture, people actually went "aahh," "wow," "that's great!", "Why don't we do more of this?" Pictures really, really speak. …"

- *Computational Statistics (2008) 23:177–184

Inspiration for Statistical Graphics



American Statistical Association
Statistical Graphics Section
Poster Series (~2004)

http://www.public.iastate.edu/~dicook/Stat.Graphics/posters.html

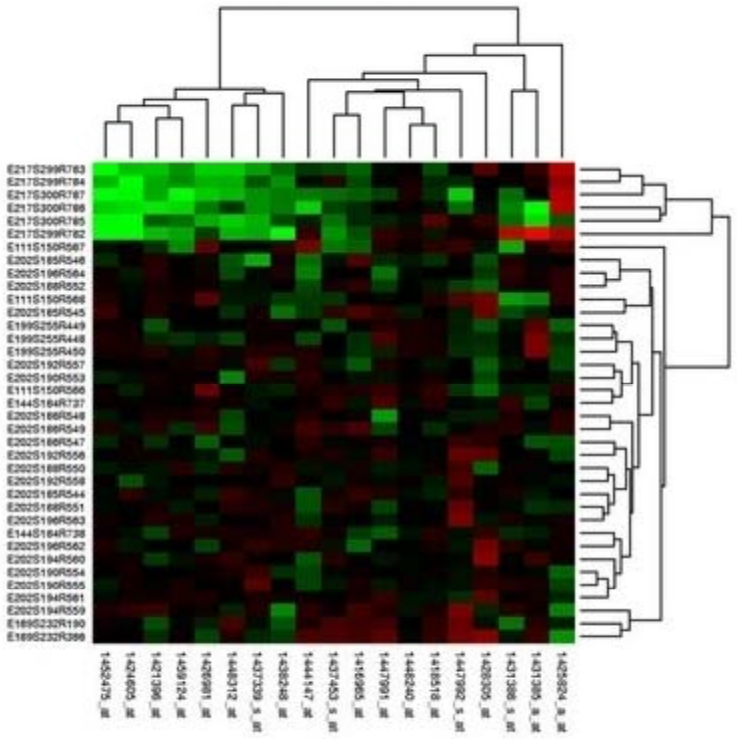
Side Note: Heatmaps for People with Color Vision Deficiencies

- From http://www.vischeck.com/vischeck/vischeckImage.php

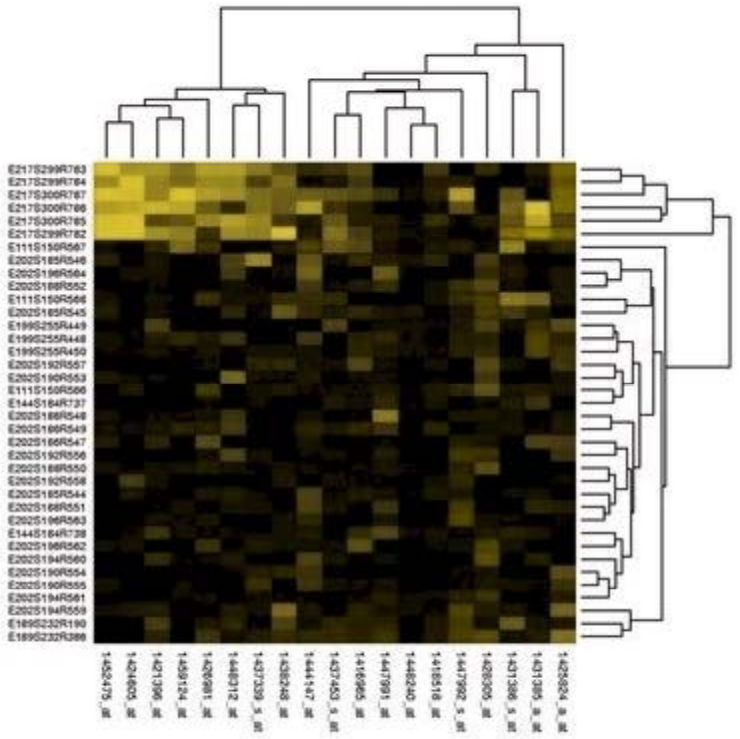
Try Vischeck on Your Image Files

Your Results:

Original Image



Deuteranope Simulation



- **Example:**
  Exploratory Graphics for Functional Actigraphy Data in Sleep Medicine

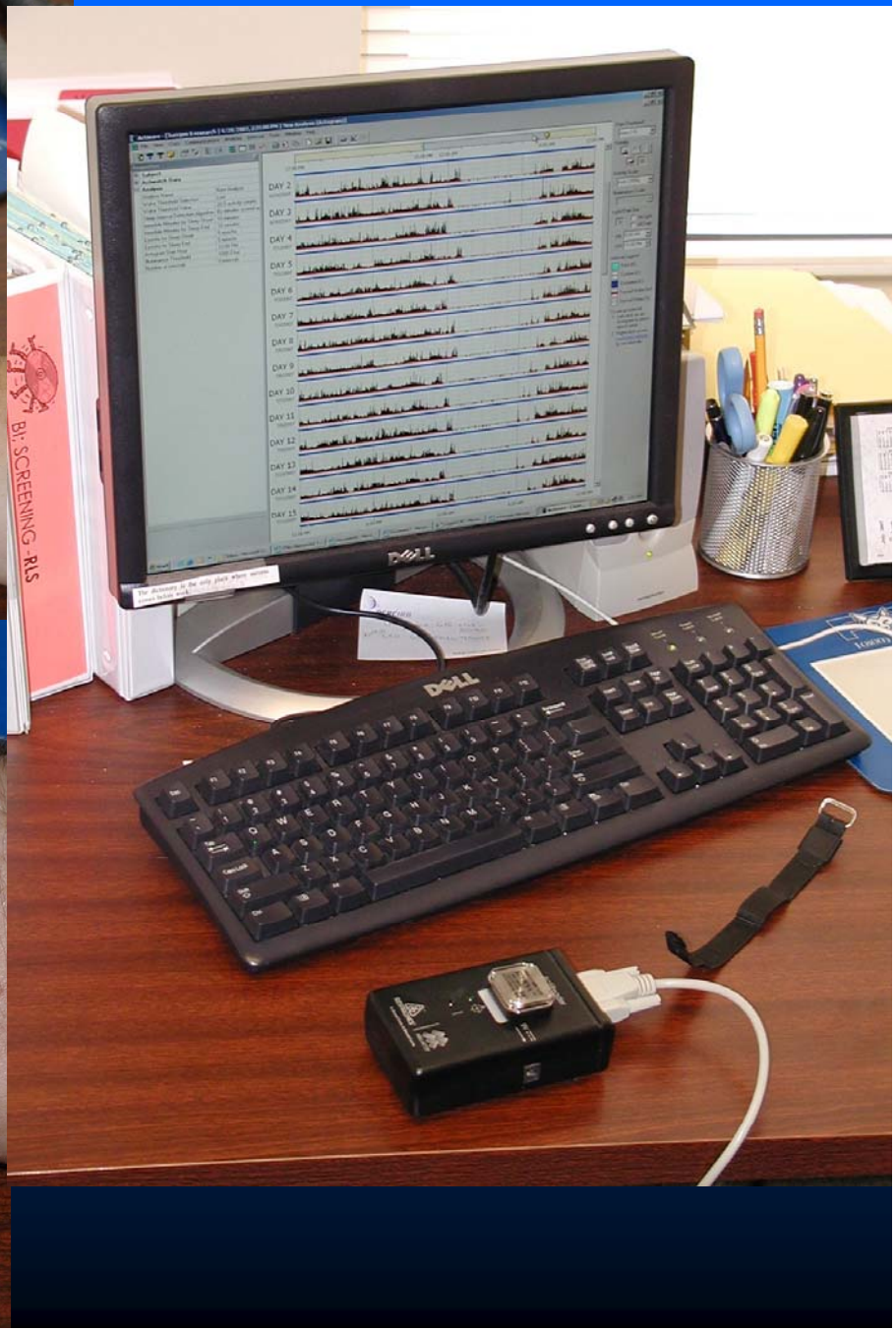
- **Reference:**
  Symanzik, J., Shannon, W. (2008): Exploratory Graphics for Functional Actigraphy Data, American Statistical Association, CD.

Background (1)

- Actigraphy: emerging technology for measuring a patient's activity level continuously over time
- Actigraph: watch-like device (attached to the wrist or a leg) that uses an accelerometer to measures (human) movements (every minute or more often)


• Respiratory
• Accessibility
• Rehab
• Bronchodilators
• DME


Ctrl
Actiwatch

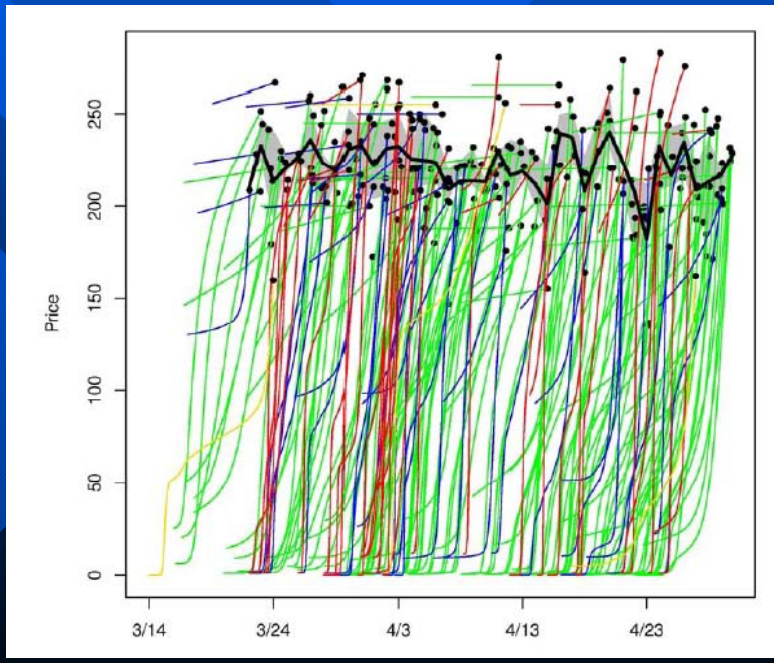

BI: SCREENING - RLS
DELL
DELL

Background (2)

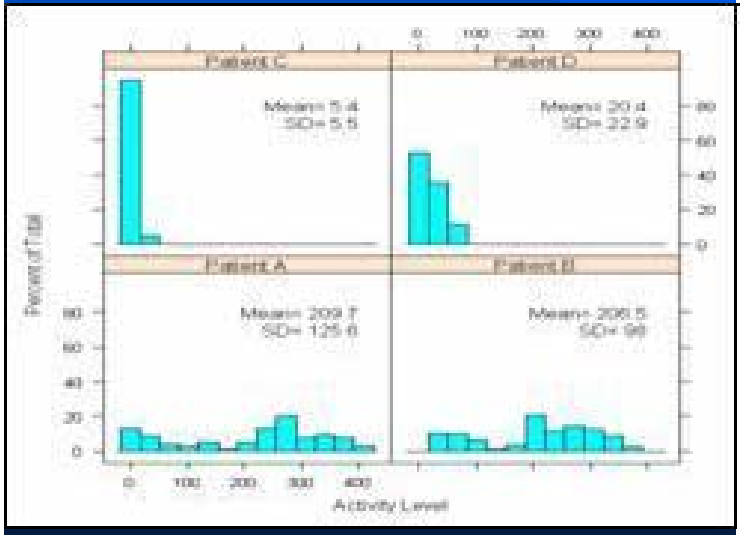
- Analysis of Human Actigraphy Data: Useful for detecting sleep, for assessing insomnia and restless leg syndrome, for tracking recovery after heart attacks, and as an assessment tool for overall status of HIV patients
- Actigraphy Data can be best described as functional data

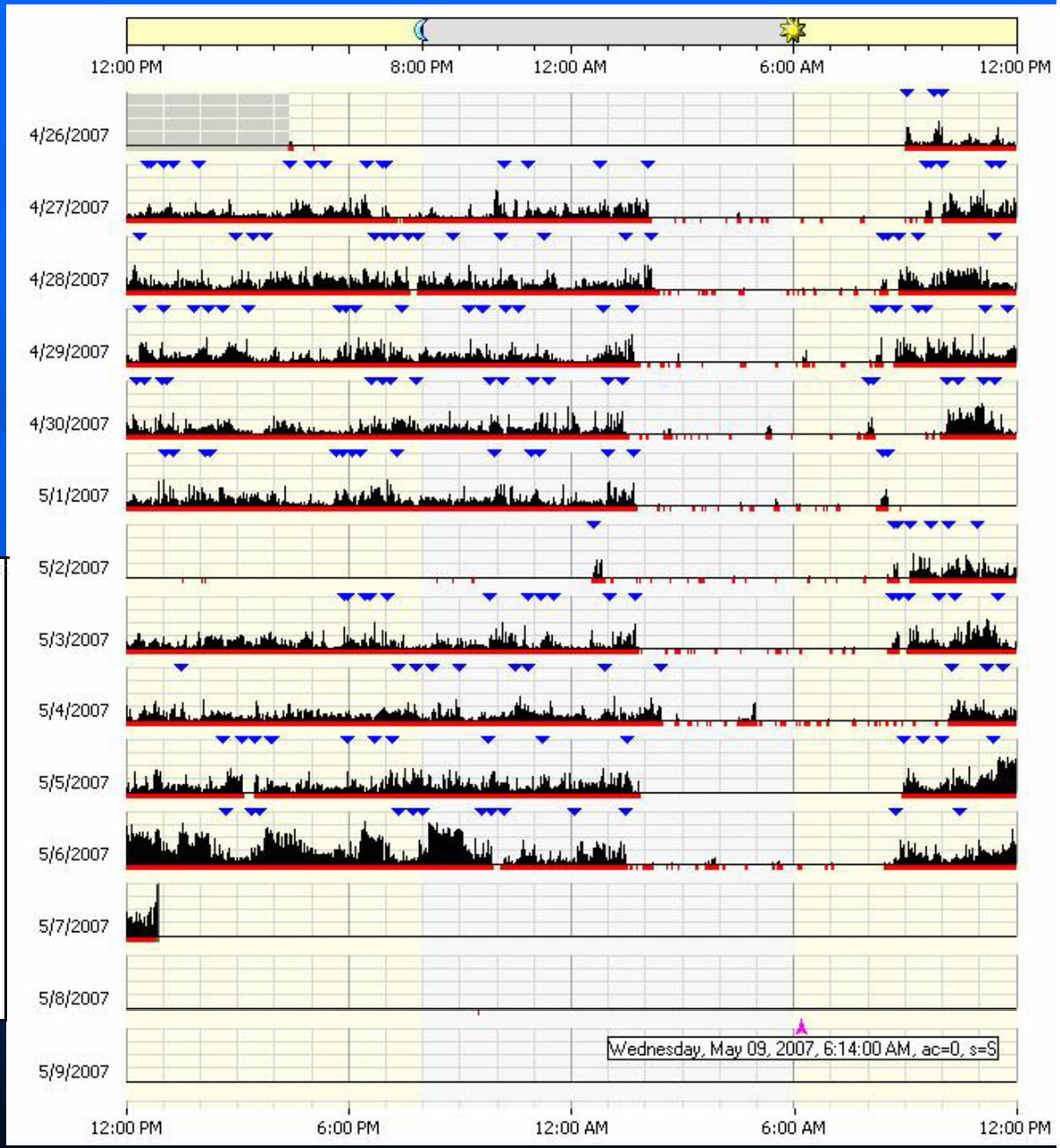
Visualization of Functional Data

- Very limited ! A rare example is:
  Jank, W., Shmueli, G., Plaisant, C., Shneiderman, B. (2008): Visualizing Functional Data with an Application to eBay's Online Auctions, In: Chen, C., Härdle, W., Unwin, A. (Eds.), Handbook of Data Visualization, Springer, Berlin/Heidelberg, 873-898.
- Figure from
  http://www.smith.umd.edu/faculty/wjank/DIV-Berlin2006.pdf
  (page 30).



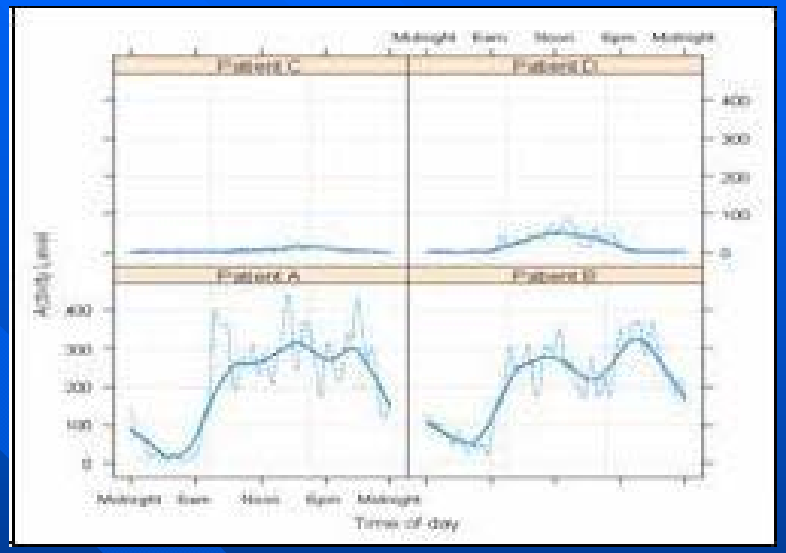
Current Visualization of Actigraphy Data





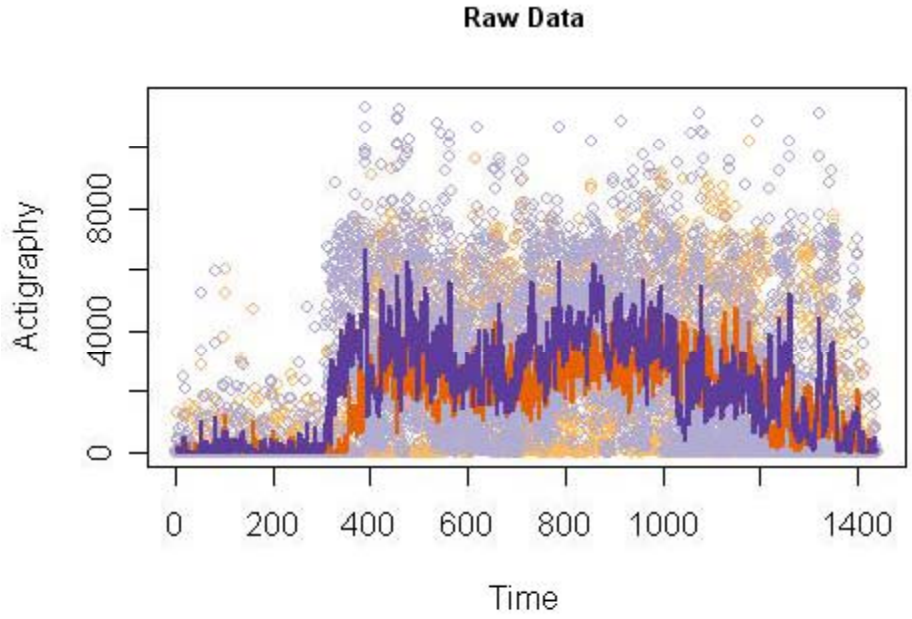
Suggested Future Visualization of Actigraphy Data (1)

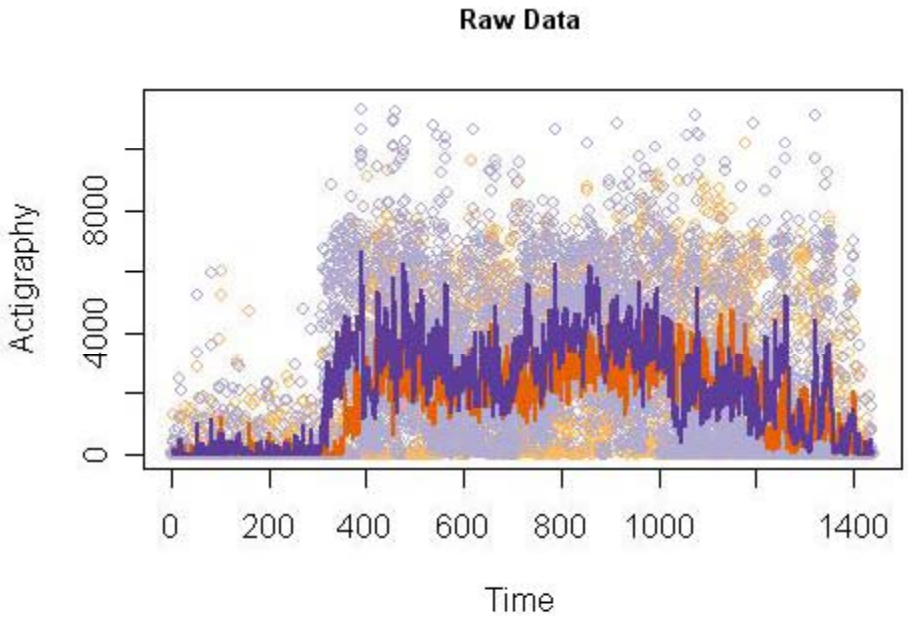
- Displays for
  - Raw data
  - Smoothed data
  - Averages etc.
  - Velocity (First Derivative)
  - Acceleration (Second Derivative)
  - Brushing & Linking
  - Cumulative Sums
- Example: 1 Subject
  - Orange: 5 Days at Baseline
  - Purple: 5 Days after 6 Months

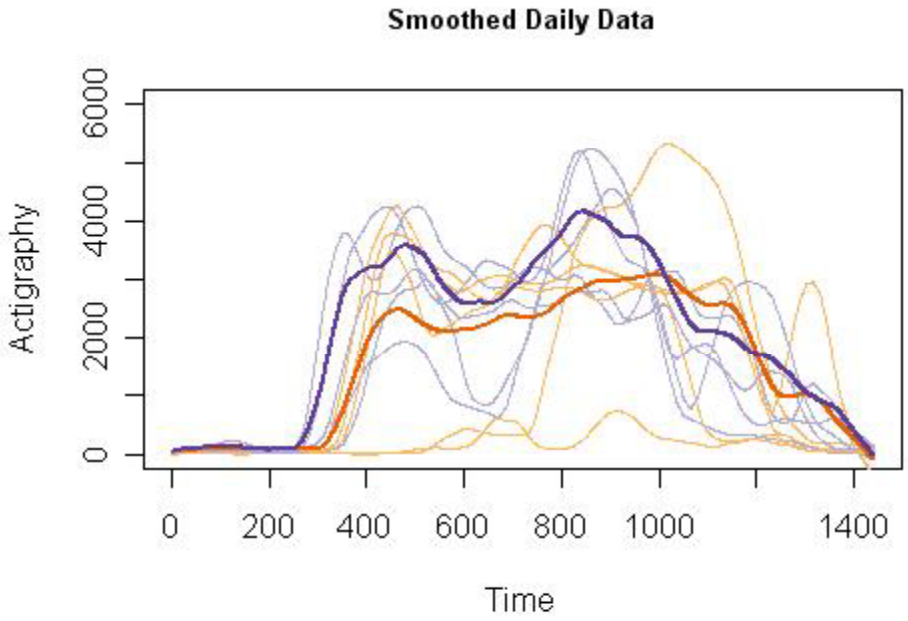


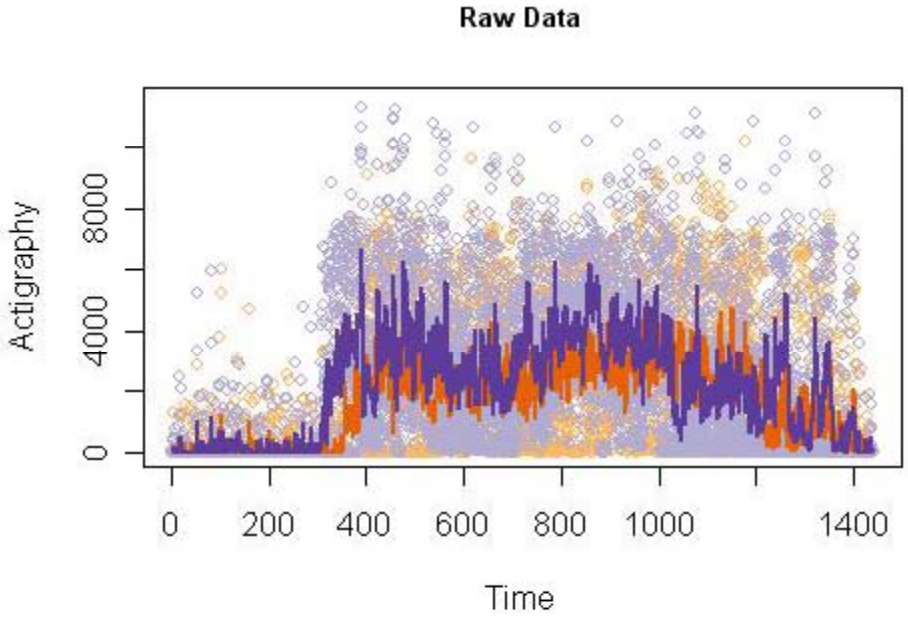
Live Demo

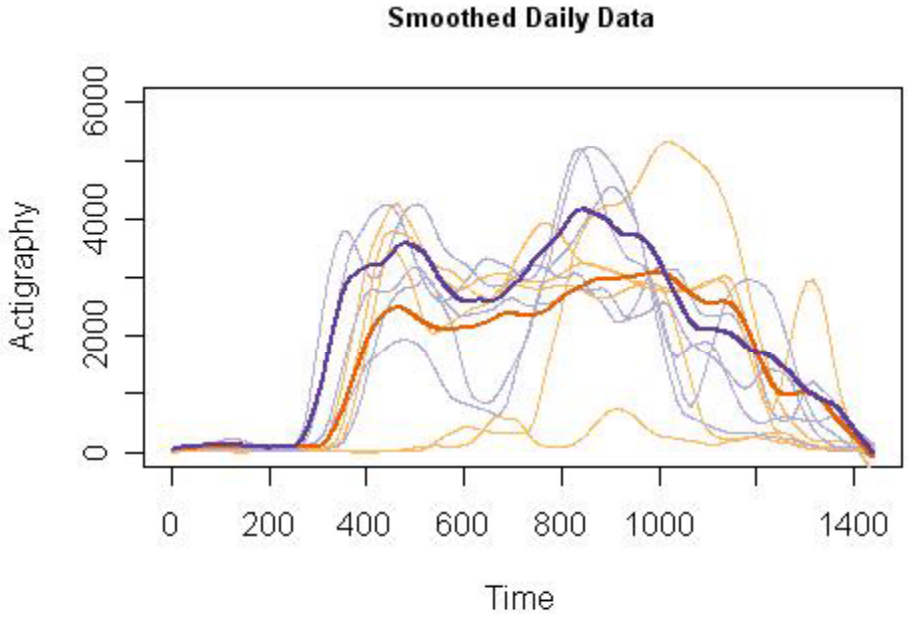
-  AVAD Software (Web Interface to R)
- Prototype Developer: Abbass Sharif

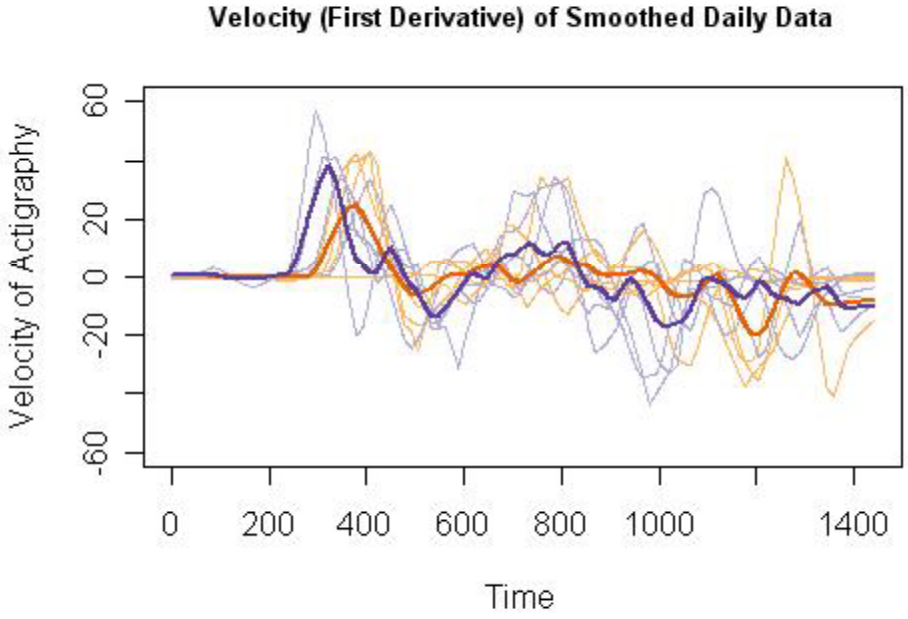

Raw Data
Actigraphy
0
4000
8000
0
200
400
600
800
1000
1400
Time

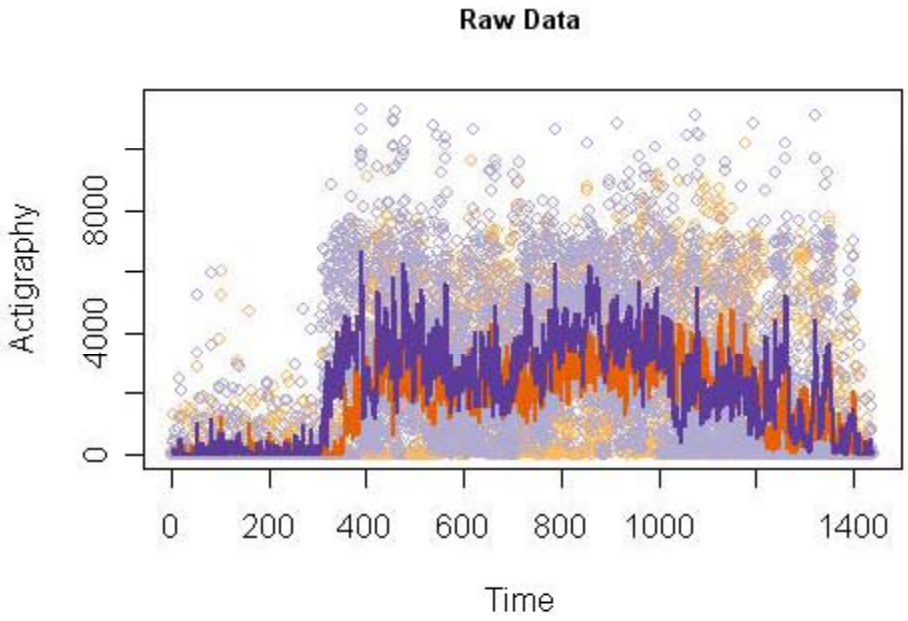

Raw Data
Actigraphy
0
4000
8000
0
200
400
600
800
1000
1400
Time

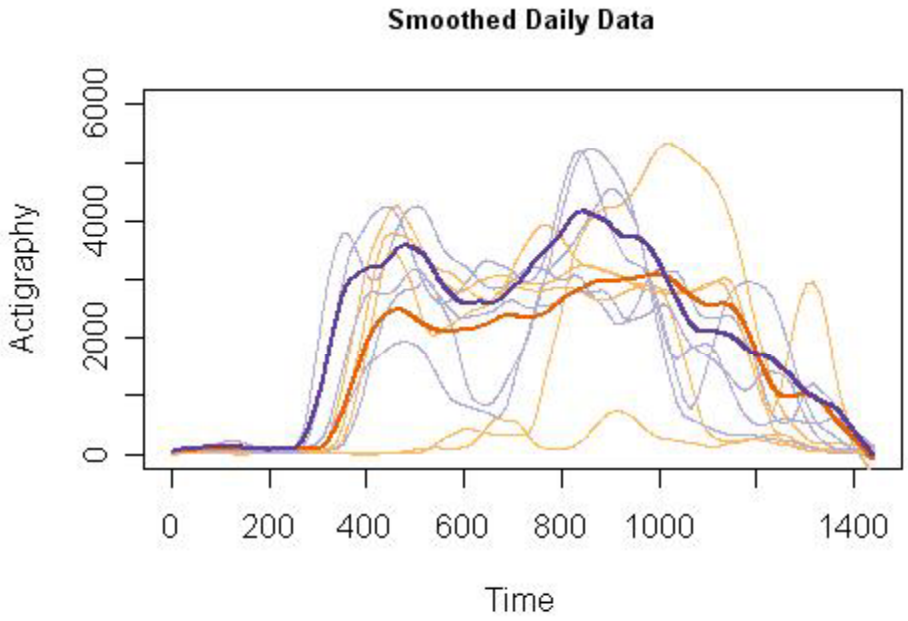

Smoothed Daily Data
Actigraphy
0
2000
4000
6000
0
200
400
600
800
1000
1400
Time

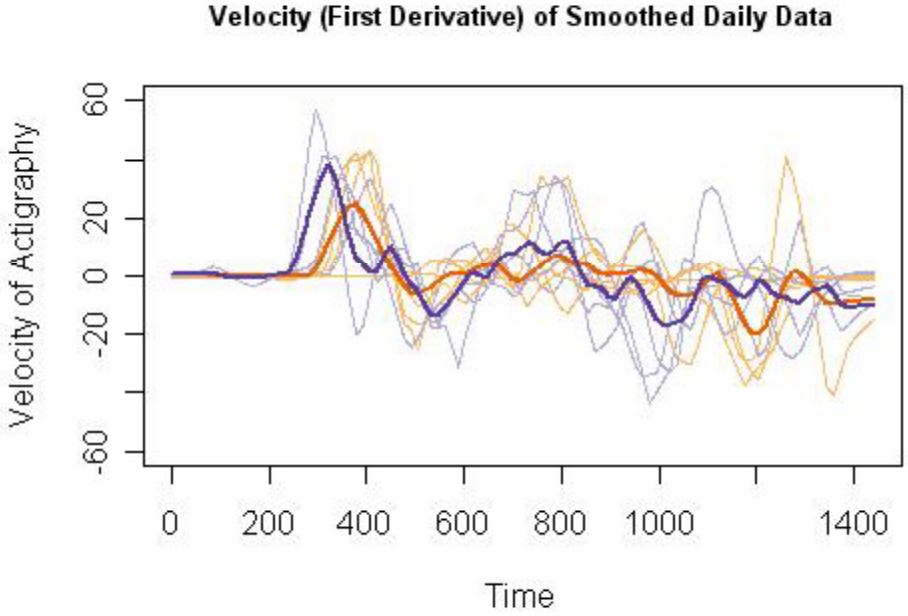

Raw Data
Actigraphy
0
4000
8000
0
200
400
600
800
1000
1400
Time

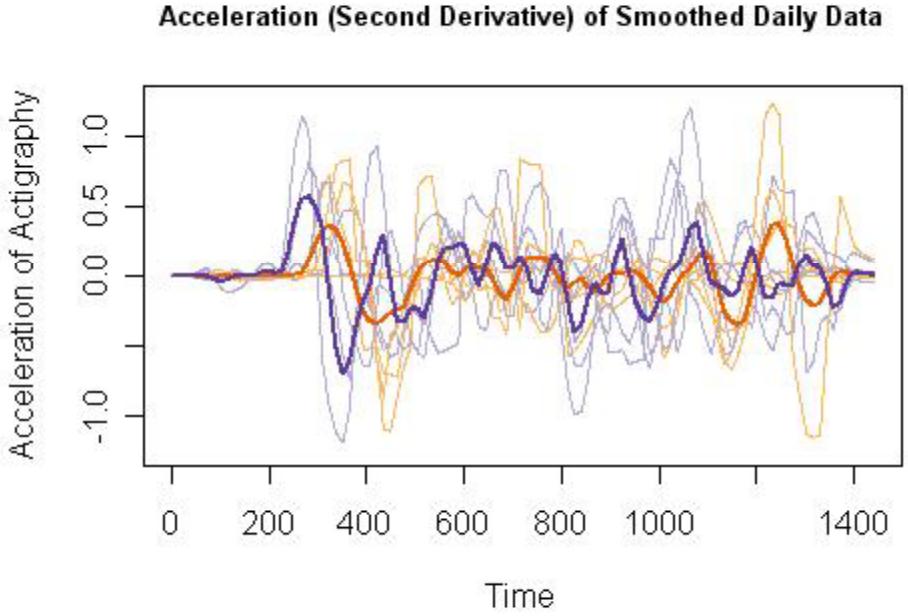

Smoothed Daily Data
Actigraphy
0
2000
4000
6000
0
200
400
600
800
1000
1400
Time

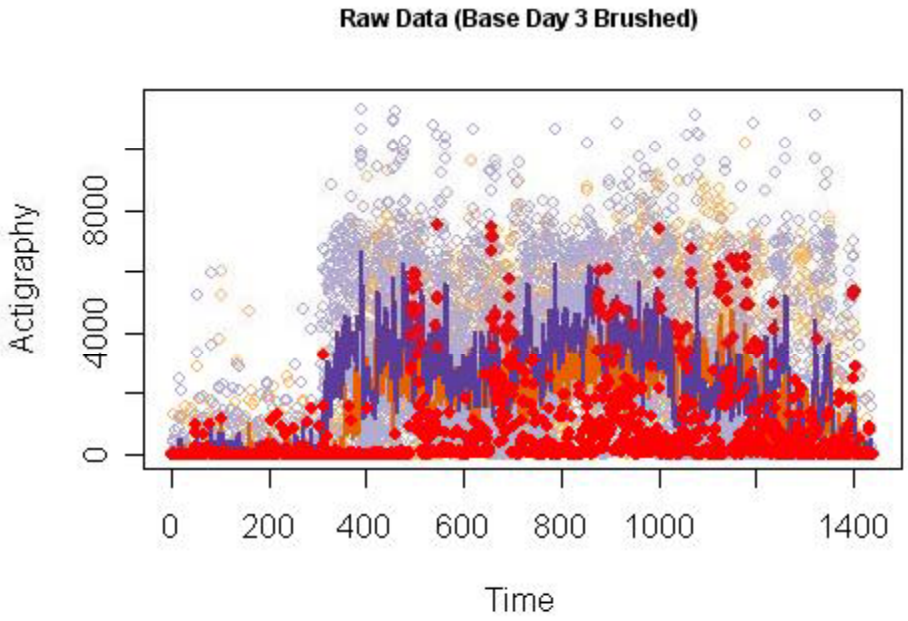

Velocity (First Derivative) of Smoothed Daily Data
Velocity of Actigraphy
-60
-20
0
20
60
0
200
400
600
800
1000
1400
Time

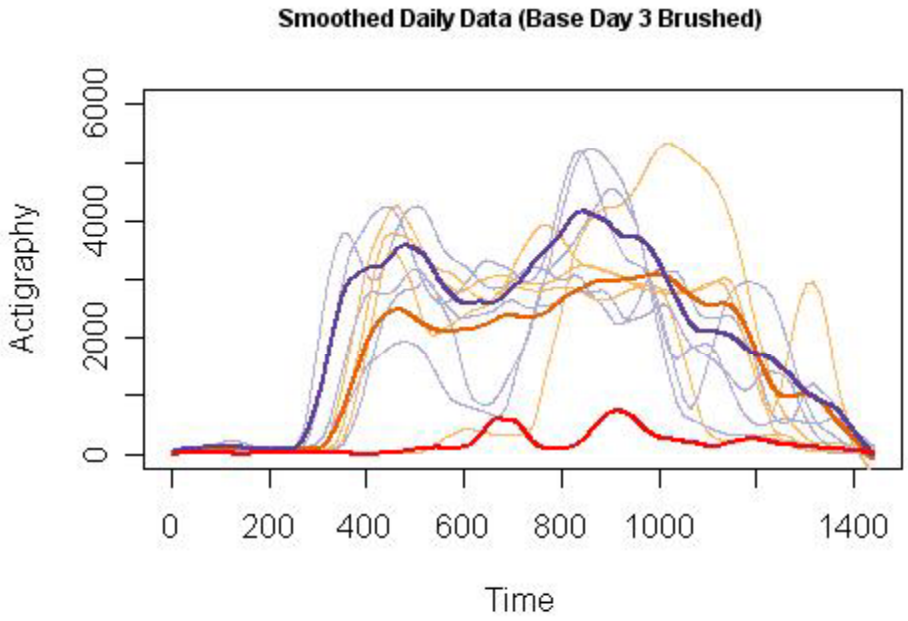

Raw Data
Actigraphy
0
4000
8000
0
200
400
600
800
1000
1400
Time

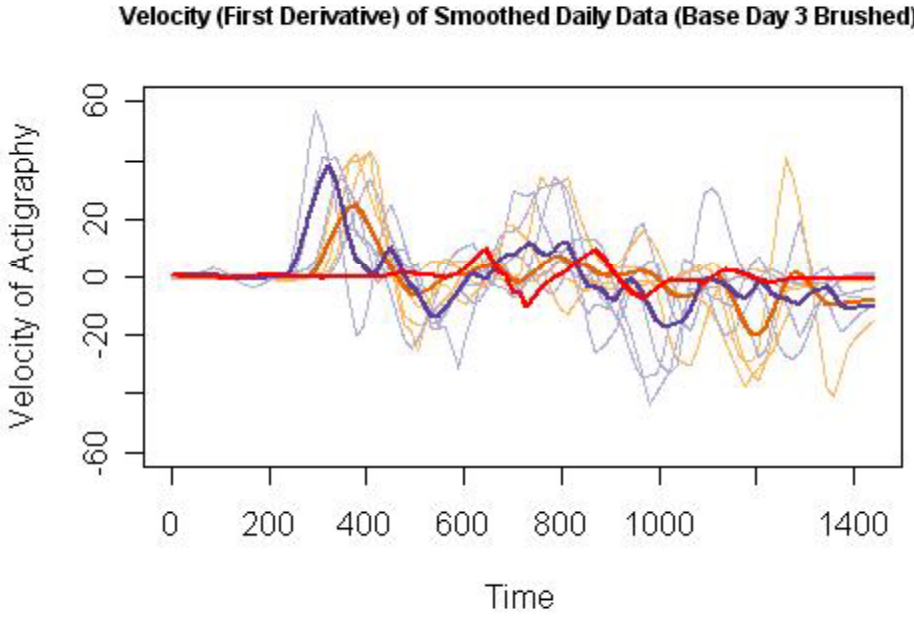

Smoothed Daily Data
Actigraphy
0
2000
4000
6000
0
200
400
600
800
1000
1400
Time

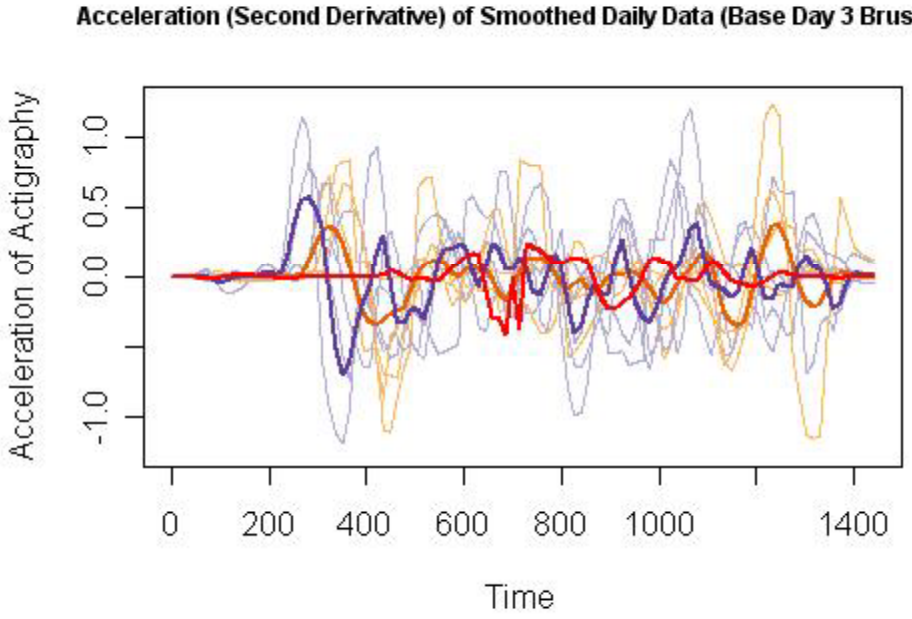

Velocity (First Derivative) of Smoothed Daily Data
Velocity of Actigraphy
-60
-20
0
20
60
0
200
400
600
800
1000
1400
Time

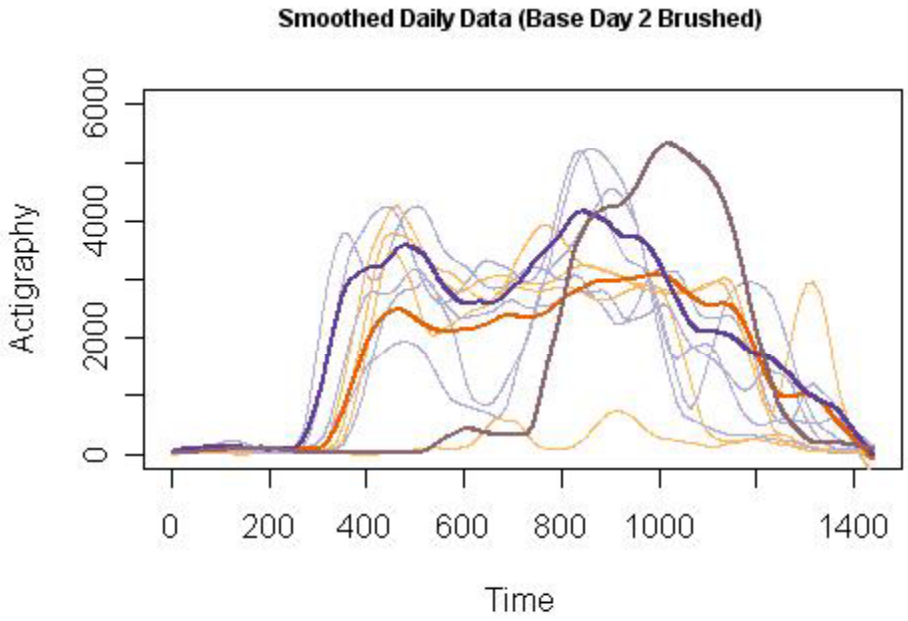

Acceleration (Second Derivative) of Smoothed Daily Data
Acceleration of Actigraphy
-1.0
0.0
0.5
1.0
0
200
400
600
800
1000
1400
Time

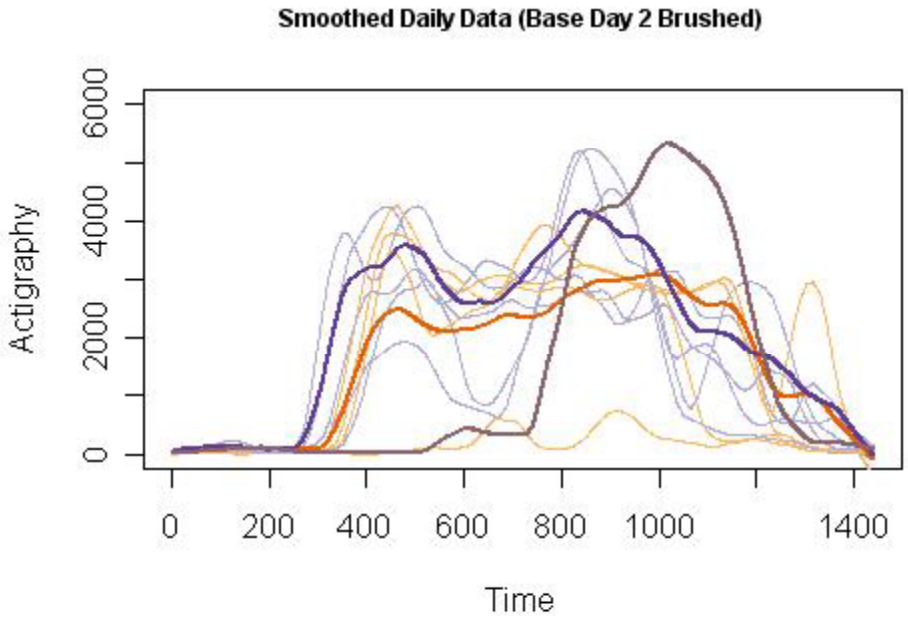

Raw Data (Base Day 3 Brushed)
Actigraphy
0
4000
8000
0
200
400
600
800
1000
1400
Time

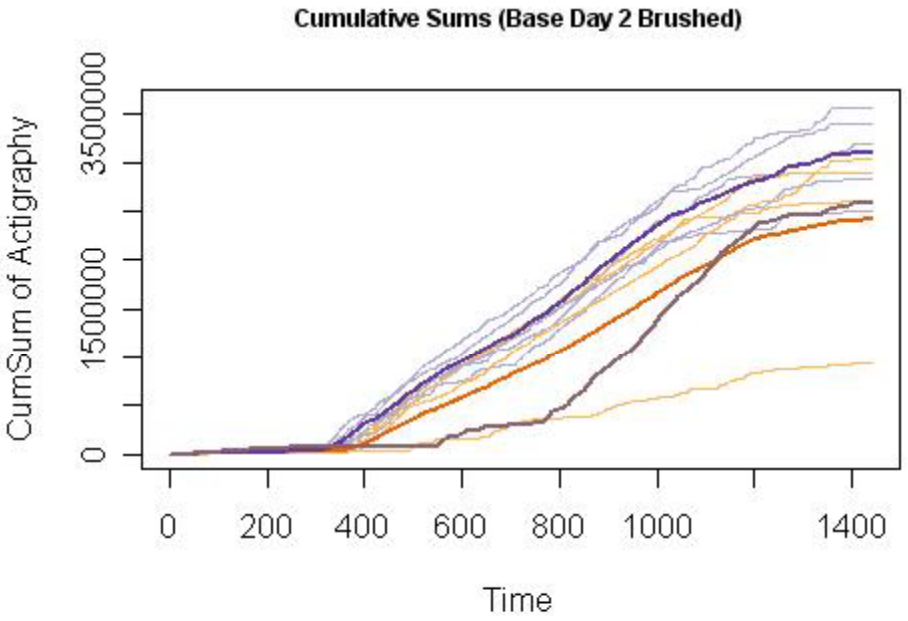

Smoothed Daily Data (Base Day 3 Brushed)
Actigraphy
0
2000
4000
6000
0
200
400
600
800
1000
1400
Time

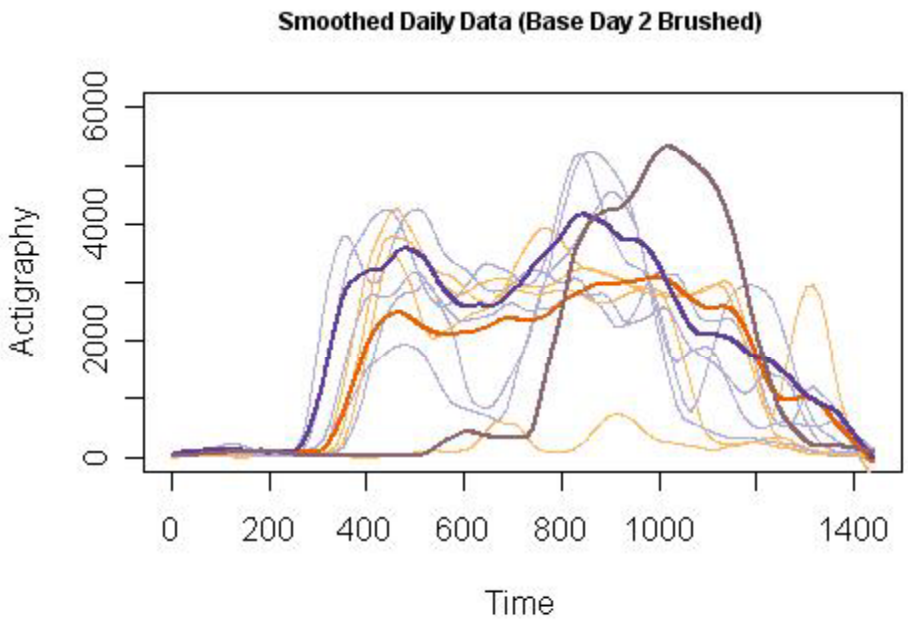

Velocity (First Derivative) of Smoothed Daily Data (Base Day 3 Brushed)
Velocity of Actigraphy
-60
-20
0
20
60
0
200
400
600
800
1000
1400
Time

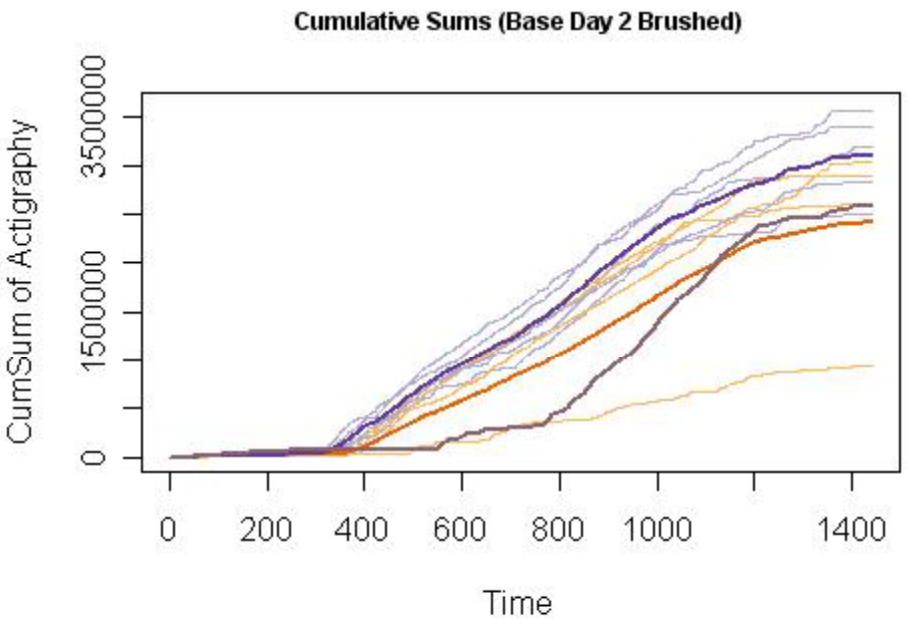

Acceleration (Second Derivative) of Smoothed Daily Data (Base Day 3 Brushed)
Acceleration of Actigraphy
-1.0
0.0
0.5
1.0
0
200
400
600
800
1000
1400
Time

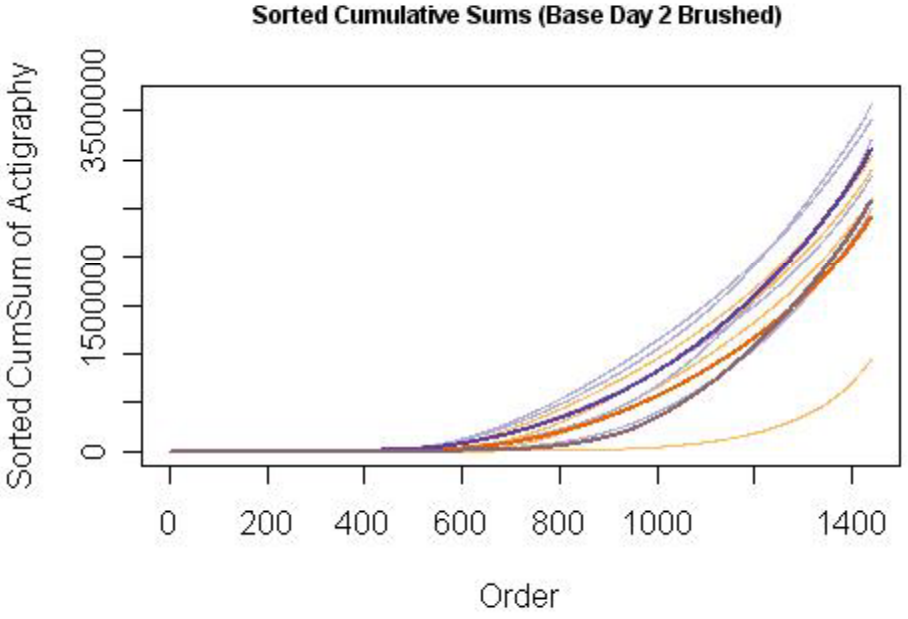

Smoothed Daily Data (Base Day 2 Brushed)
Actigraphy
0
2000
4000
6000
0
200
400
600
800
1000
1400
Time


Smoothed Daily Data (Base Day 2 Brushed)
Actigraphy
0
2000
4000
6000
0
200
400
600
800
1000
1400
Time


Cumulative Sums (Base Day 2 Brushed)
CumSum of Actigraphy
0
1500000
3500000
0
200
400
600
800
1000
1400
Time


Smoothed Daily Data (Base Day 2 Brushed)
Actigraphy
0
2000
4000
6000
0
200
400
600
800
1000
1400
Time


Cumulative Sums (Base Day 2 Brushed)
CumSum of Actigraphy
0
1500000
3500000
0
200
400
600
800
1000
1400
Time


Sorted Cumulative Sums (Base Day 2 Brushed)
Sorted CumSum of Actigraphy
0
1500000
3500000
0
200
400
600
800
1000
1400
Order

Conclusions

Visualization of Actigraphy Data provides

- Potential for application in various medical fields
- Additional insights into actigraphy data
- Ease to compare baseline and past-treatment data
  - » of a single patient
  - » of multiple patients
  - » to identify outliers
  - » to compare averages

- **Example:**
  Agreement in Carpal Tunnel Syndrome (CTS) Studies

References:

Dale, A.M., Strickland, J., Symanzik, J., Franzblau, A., Evanoff, B. (2008): Reliability of Hand Diagrams for the Epidemiologic Case Definition of Carpal Tunnel Syndrome, Journal of Occupational Rehabilitation, 18(3):233-248.

Symanzik, J. (2010): Interactive and Dynamic Statistical Graphics, In: Lovric, M (Ed.), International Encyclopedia of Statistical Science, Springer, Forthcoming.

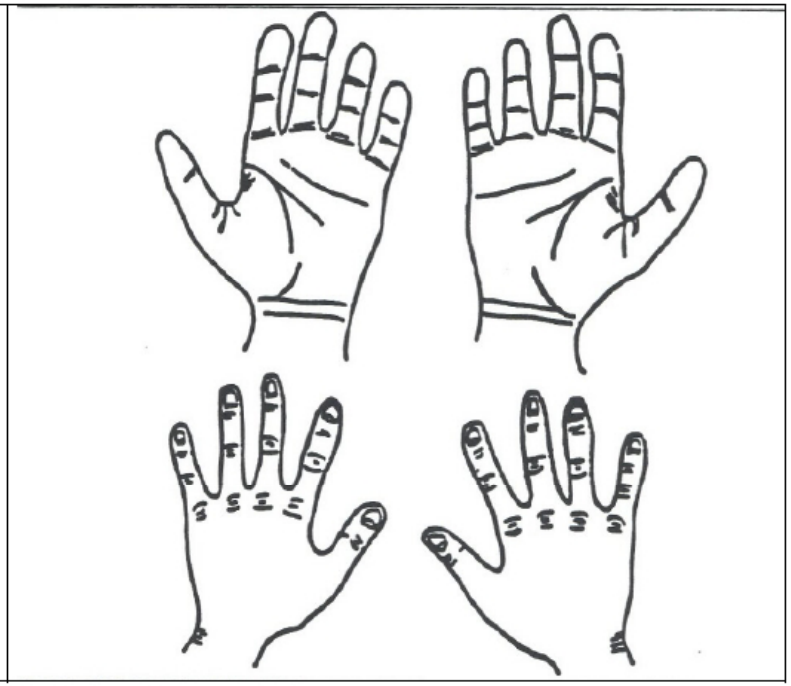
Study Background

- Subjects have to indicate where they experience what kind of "pain" on their fingers/hands
- Background:
  - 393 Subjects
  - 494 Hand Diagrams
- Expert raters assign values from 0 (unlikely) to 3 (classic), describing severity of CTS
- Interested in rating agreement among 3 expert raters

	RIGHT WRIST	LEFT WRIST	RIGHT HAND	LEFT HAND	RIGHT FINGERS	LEFT FINGERS
Burning/ Pain	O	O	O	O	O	O
Tightness/ Stiffness	O	O	O	O	O	O
Soreness/ Cramping/ Aching	O	O	O	O	O	O
Numbness/ Tingling	O	O	O	O	O	O

Please show on the diagram to the right where you have experienced numbness, tingling, burning, or pain by **shading in the problem area.**

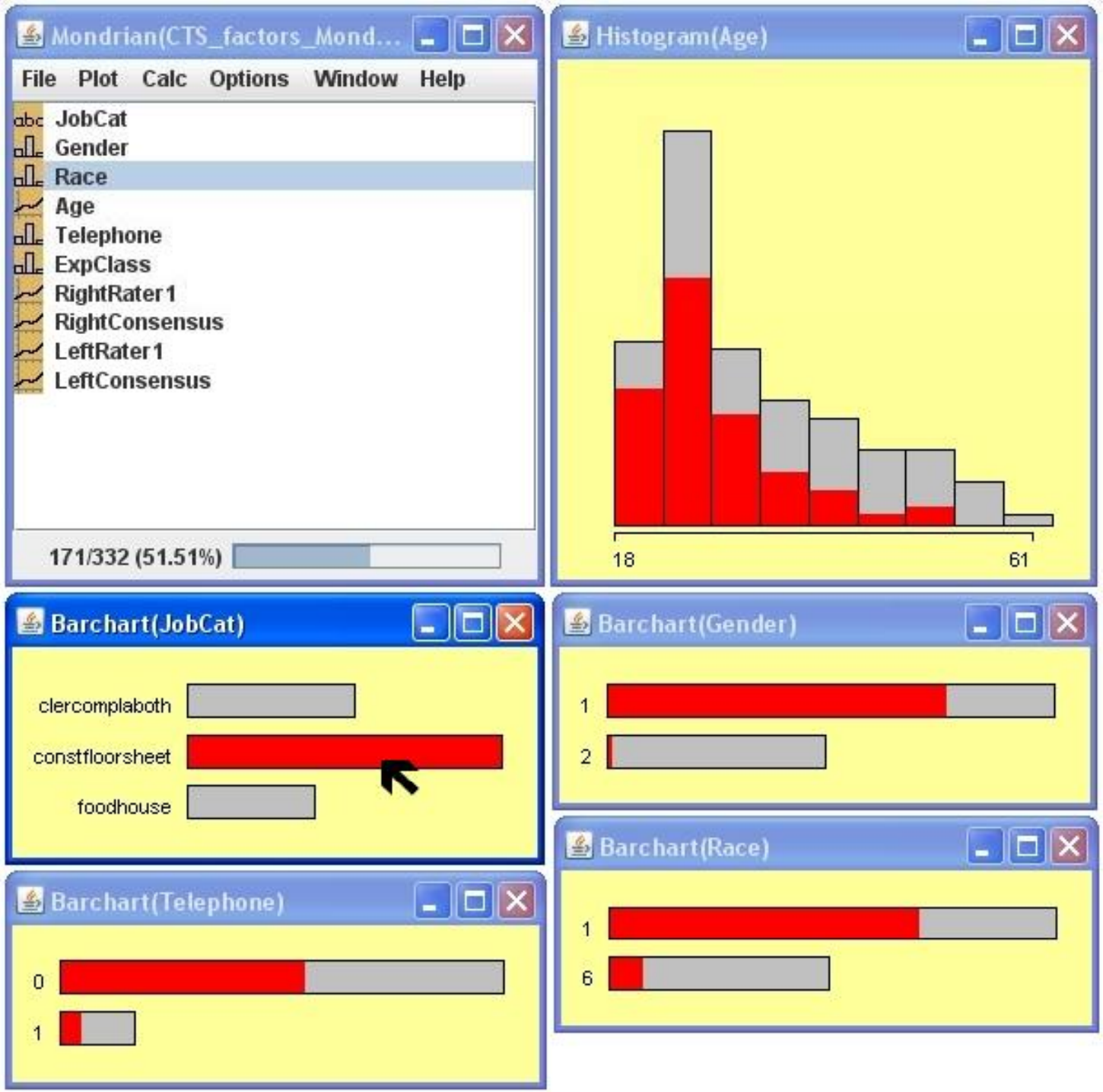
If you have not experienced these symptoms, please skip to the next question.

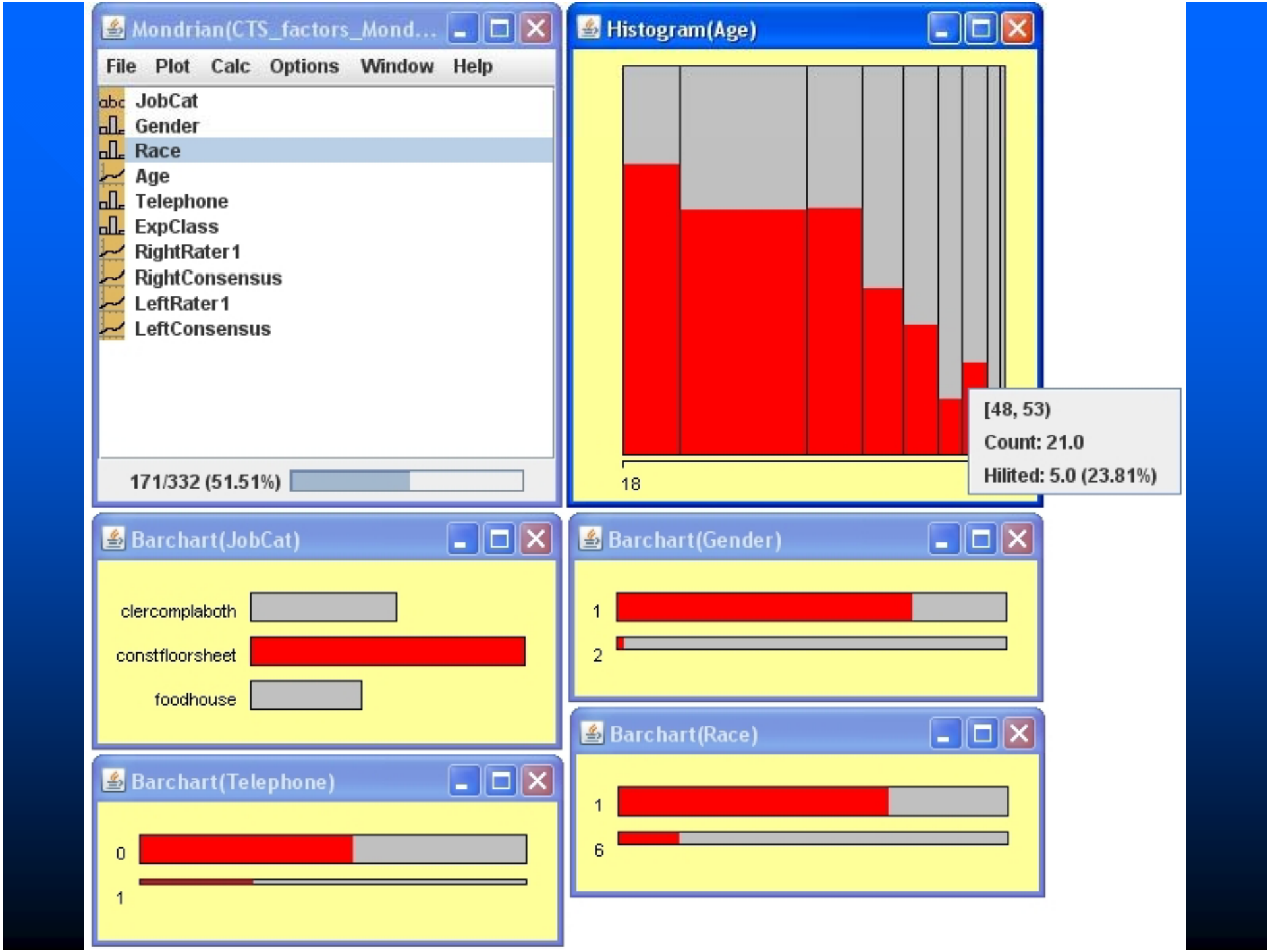


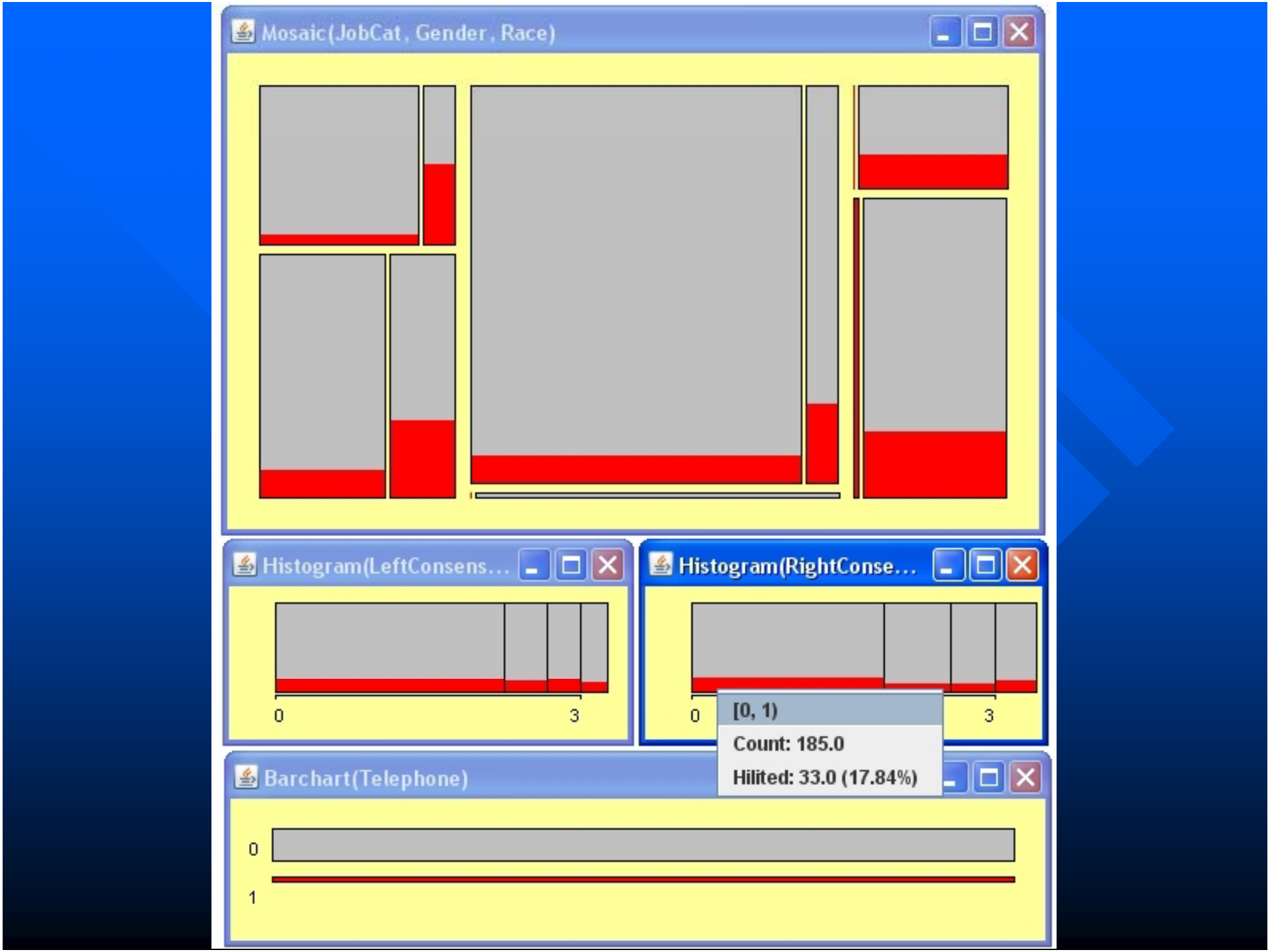
Live Demo

-  Mondrian

- *http://www.rosuda.org/Mondrian/*
- Visualization of categorical and geographic data
- Main Developer: Martin Theus


Mondrian(CTS_factors_Mond...
File Plot Calc Options Window Help
JobCat
Gender
Race
Age
Telephone
ExpClass
RightRater1
RightConsensus
LeftRater1
LeftConsensus
171/332 (51.51%)
Histogram(Age)
18
61
Barchart(JobCat)
clercomplaboth
constfloorsheet
foodhouse
Barchart(Gender)
1
2
Barchart(Telephone)
0
1
Barchart(Race)
1
6


Mondrian(CTS_factors_Mond...
File Plot Calc Options Window Help
JobCat
Gender
Race
Age
Telephone
ExpClass
RightRater1
RightConsensus
LeftRater1
LeftConsensus
171/332 (51.51%)
Histogram(Age)
18
[48, 53)
Count: 21.0
Hilited: 5.0 (23.81%)
Barchart(JobCat)
clercomplaboth
constfloorsheet
foodhouse
Barchart(Gender)
1
2
Barchart(Race)
1
6
Barchart(Telephone)
0
1

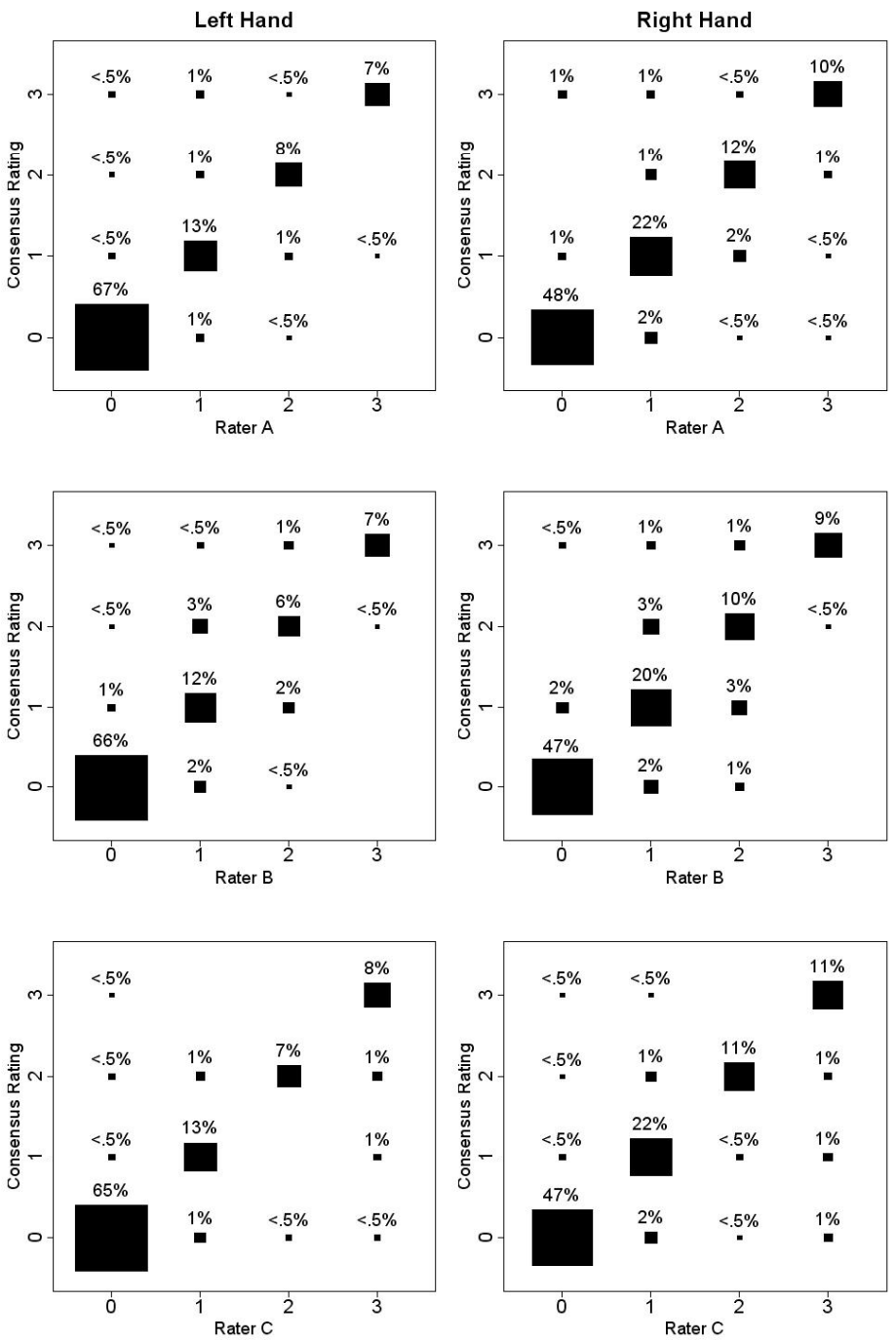

Mosaic(JobCat, Gender, Race)
Histogram(LeftConsens...
0
3
Histogram(RightConse...
0
3
[0, 1)
Count: 185.0
Hilited: 33.0 (17.84%)
Barchart(Telephone)
0
1

Numerical Results (1)

- Weighted “kappa” represents agreement between three experts, based on difference between how much agreement is present compared to how much agreement would be expected by chance alone (possible range is -1.0 to 1.0)
- Interpretation:
  - < 0 : Less than chance agreement
  - 0.01 – 0.20: Slight agreement
  - 0.21 – 0.40: Fair agreement
  - 0.41 – 0.60: Moderate agreement
  - 0.61 – 0.80: Substantial agreement
  - 0.81 – 0.99: Almost perfect agreement
- Observed Results:
  - Left Hand Diagrams: 0.88 (95% CI: 0.83, 0.91)
  - Right Hand Diagrams: 0.83 (95% CI: 0.78, 0.87)

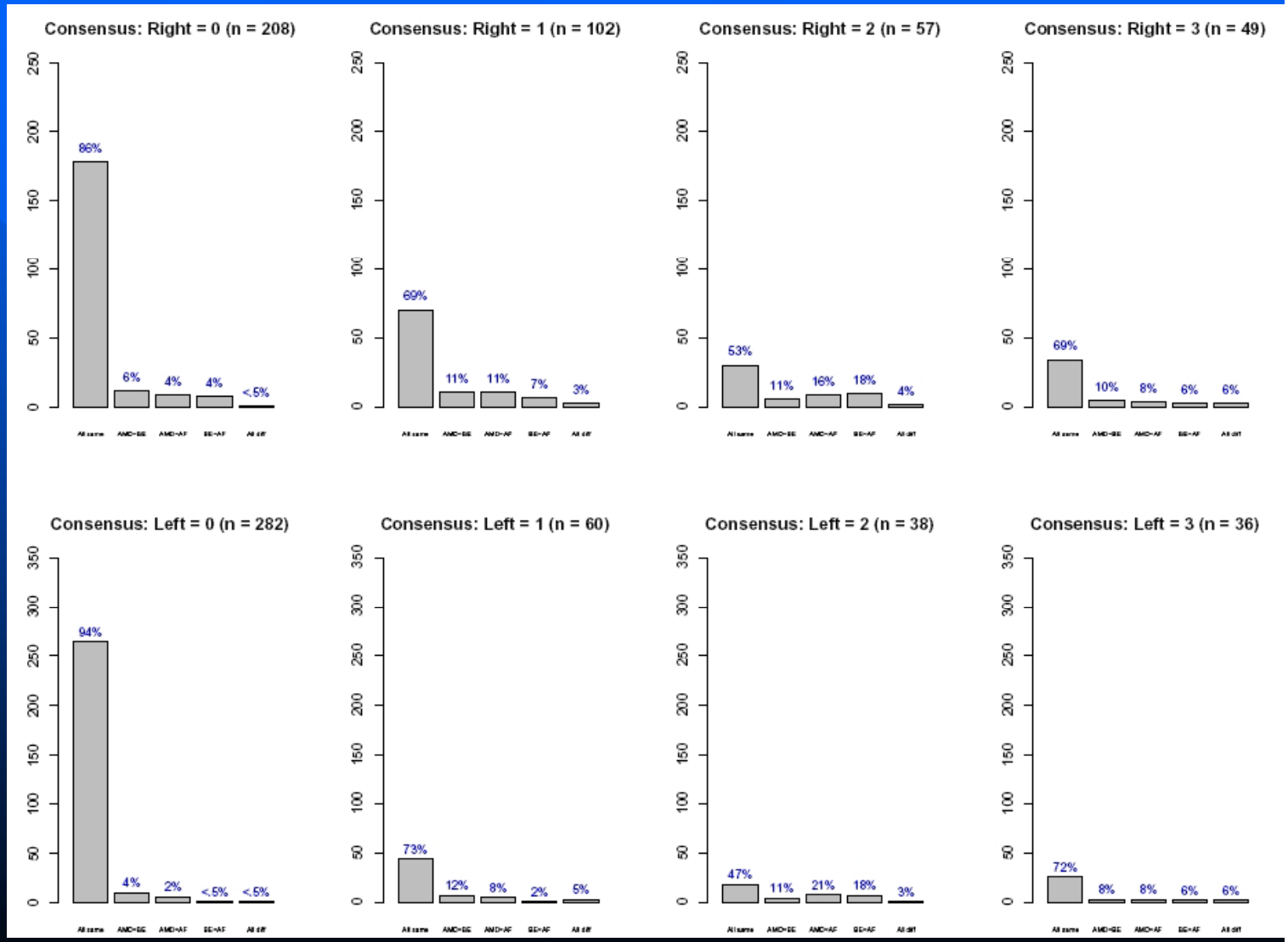
Graphical Results (1)

- Comparison of Rating Consensus Scores of 3 Raters, Compared to the Raters' Individual Scores



Graphical Results (2)

- Detailed Comparison of Raters' Individual Scores



- **Example:**
  Micromaps for Oral Clefts and West Nile Virus

- **References:**

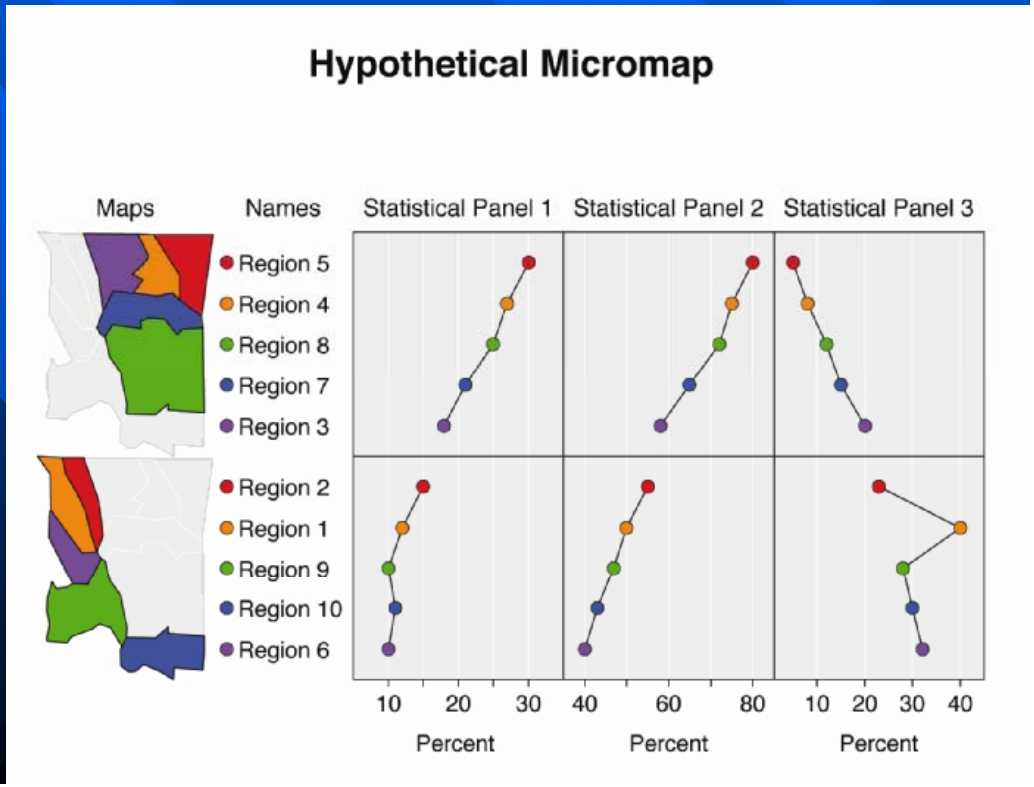
Symanzik, J., Gebreab, S., Gillies, R. Wilson, J. (2003): Visualizing the Spread of West Nile Virus, *2003 Proceedings*, American Statistical Association, Alexandria, Virginia, CD .

Gebreab, S. Y., Gillies, R. R., Munger, R. G., Symanzik, J. (2008): Visualization and Interpretation of Birth Defects Data Using Linked Micromap Plots, *Birth Defects Research (Part A): Clinical and Molecular Teratology*, Vol. 82, 110-119.

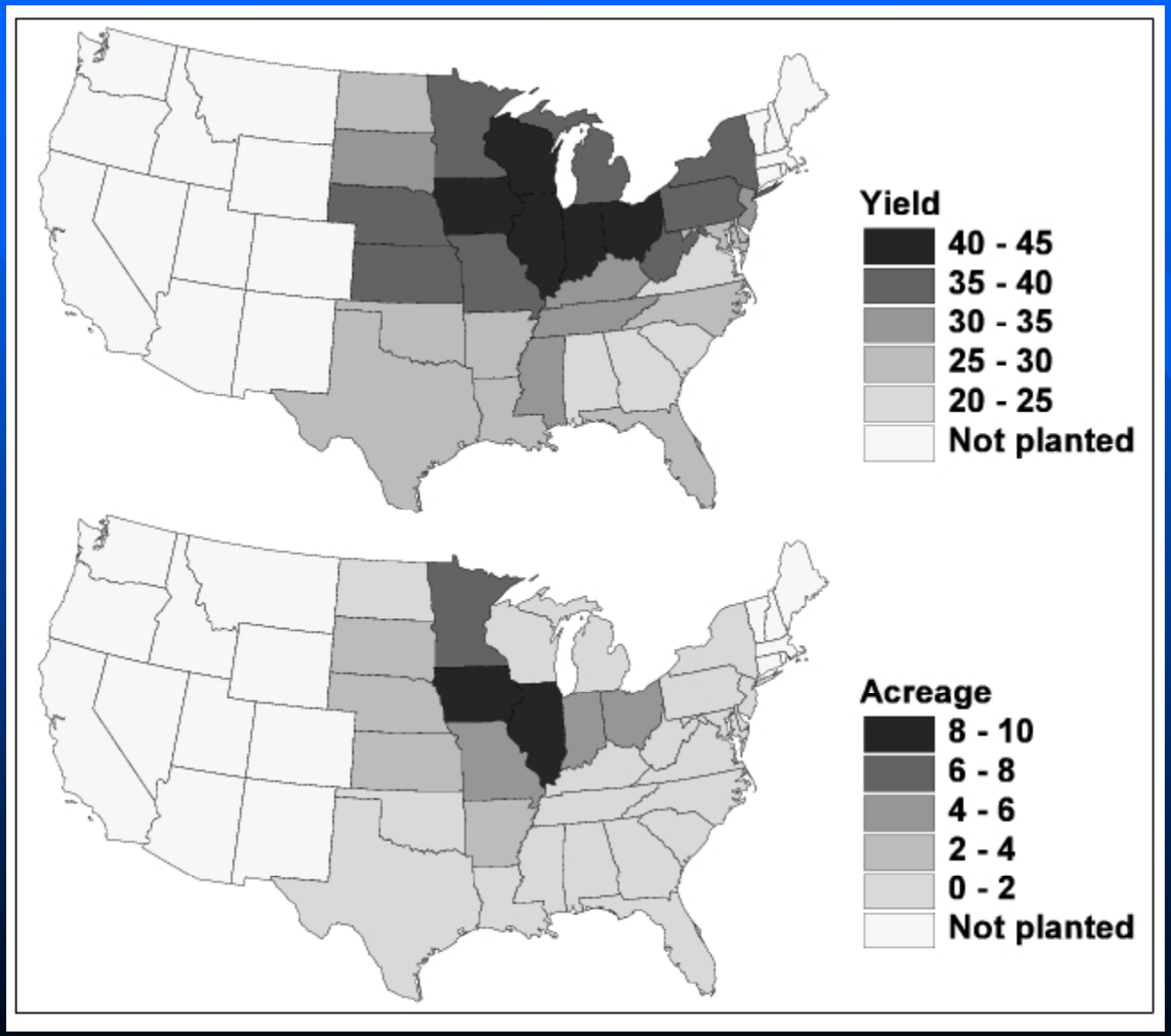
Symanzik, J., Carr, D. B. (2008): Interactive Linked Micromap Plots for the Display of Geographically Referenced Statistical Data, In: Chen, C., Härdle, W., Unwin, A. (Eds.), Handbook of Data Visualization, Springer, Berlin/Heidelberg, 267-294.

Micromaps

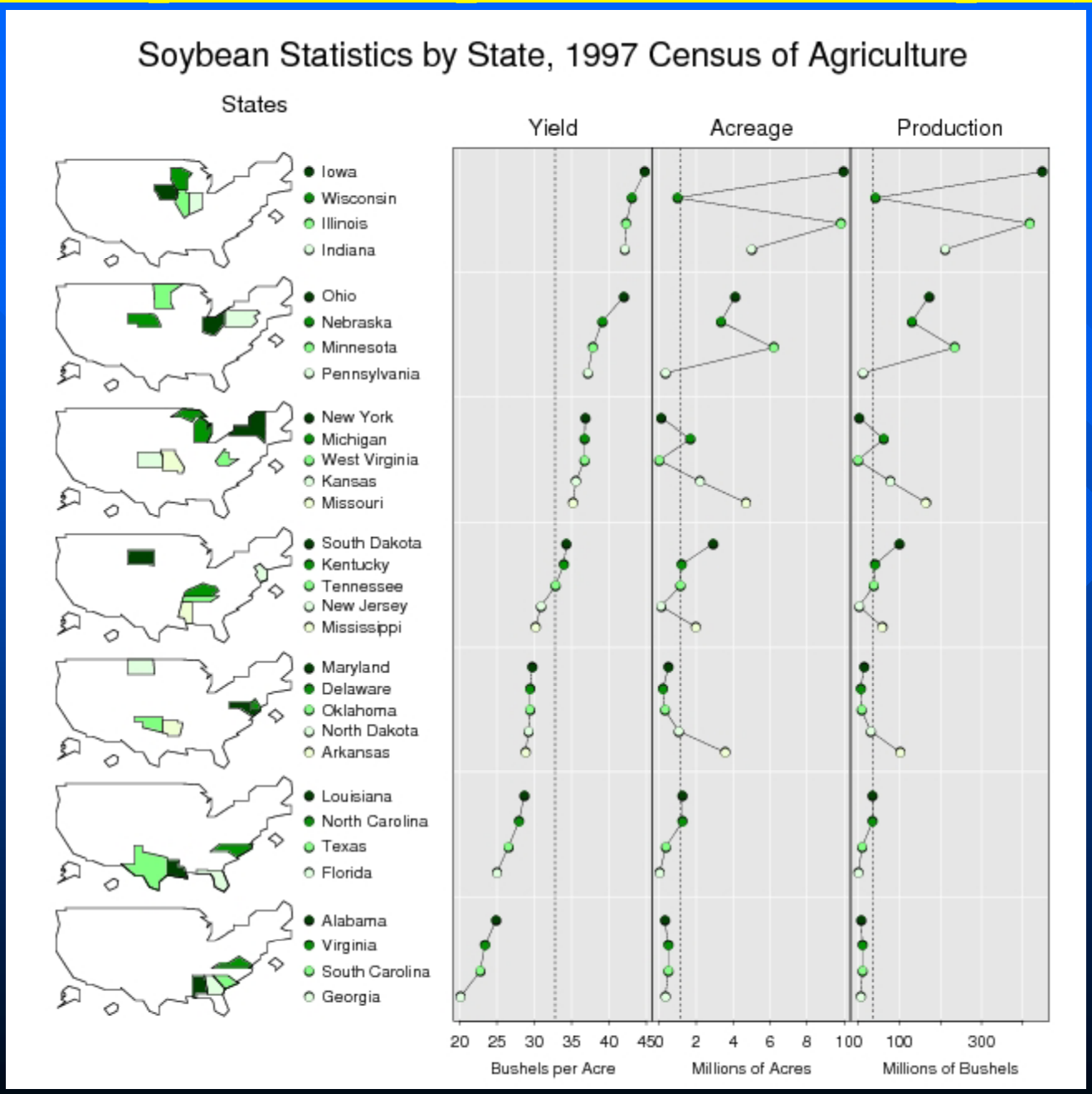
- Link of row-labeled univariate (or multivariate) statistical summaries to corresponding geographical region
- Focus on statistical display and not on maps
- Useful for
  - environmental data
  - agricultural data
  - medical data
  - economical data



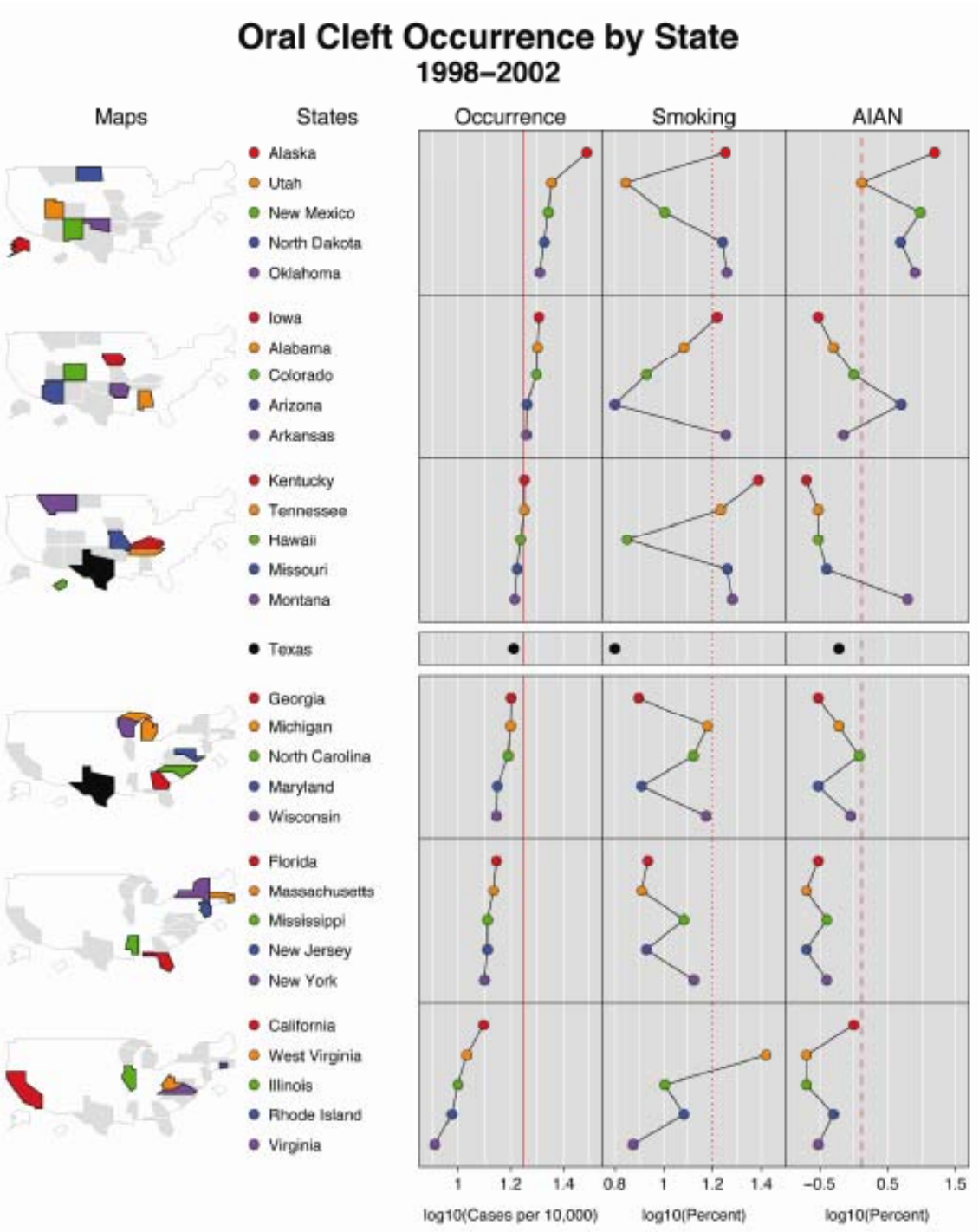
Choropleth Maps vs Micromaps (1)



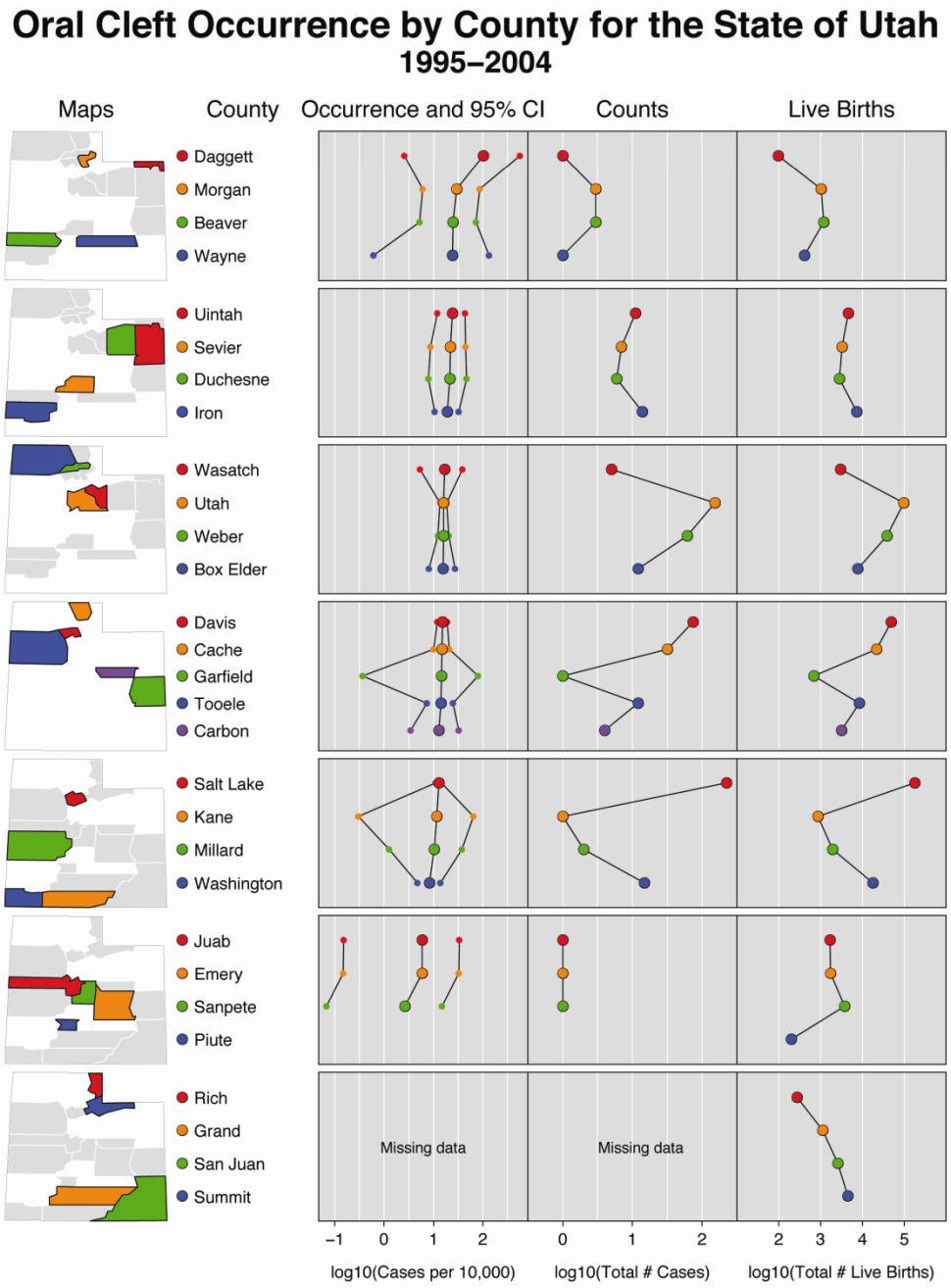
Choropleth Maps vs Micromaps (2)



Oral Cleft Micromap Example: USA



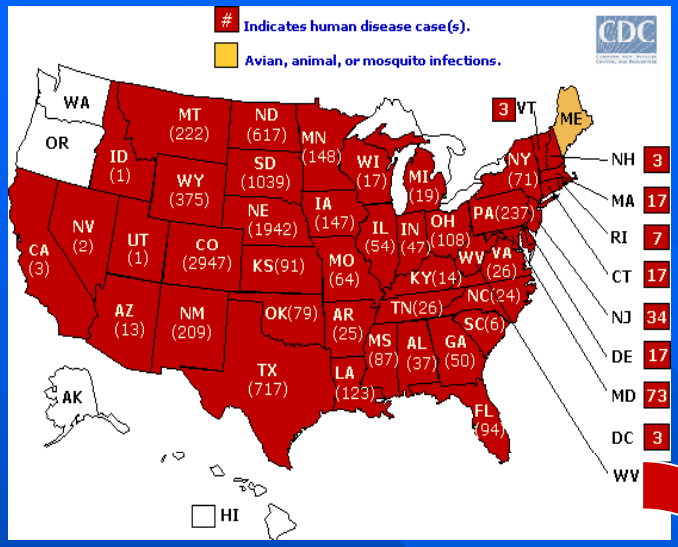
Oral Cleft Micromap Example: Utah

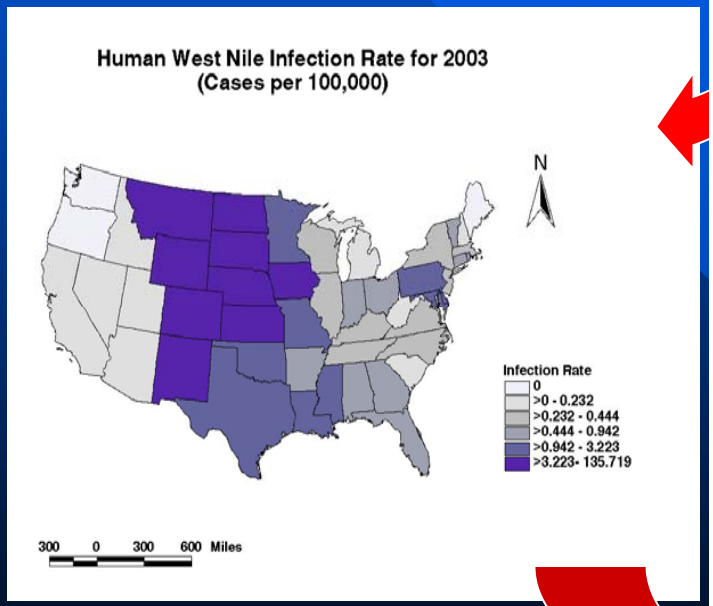


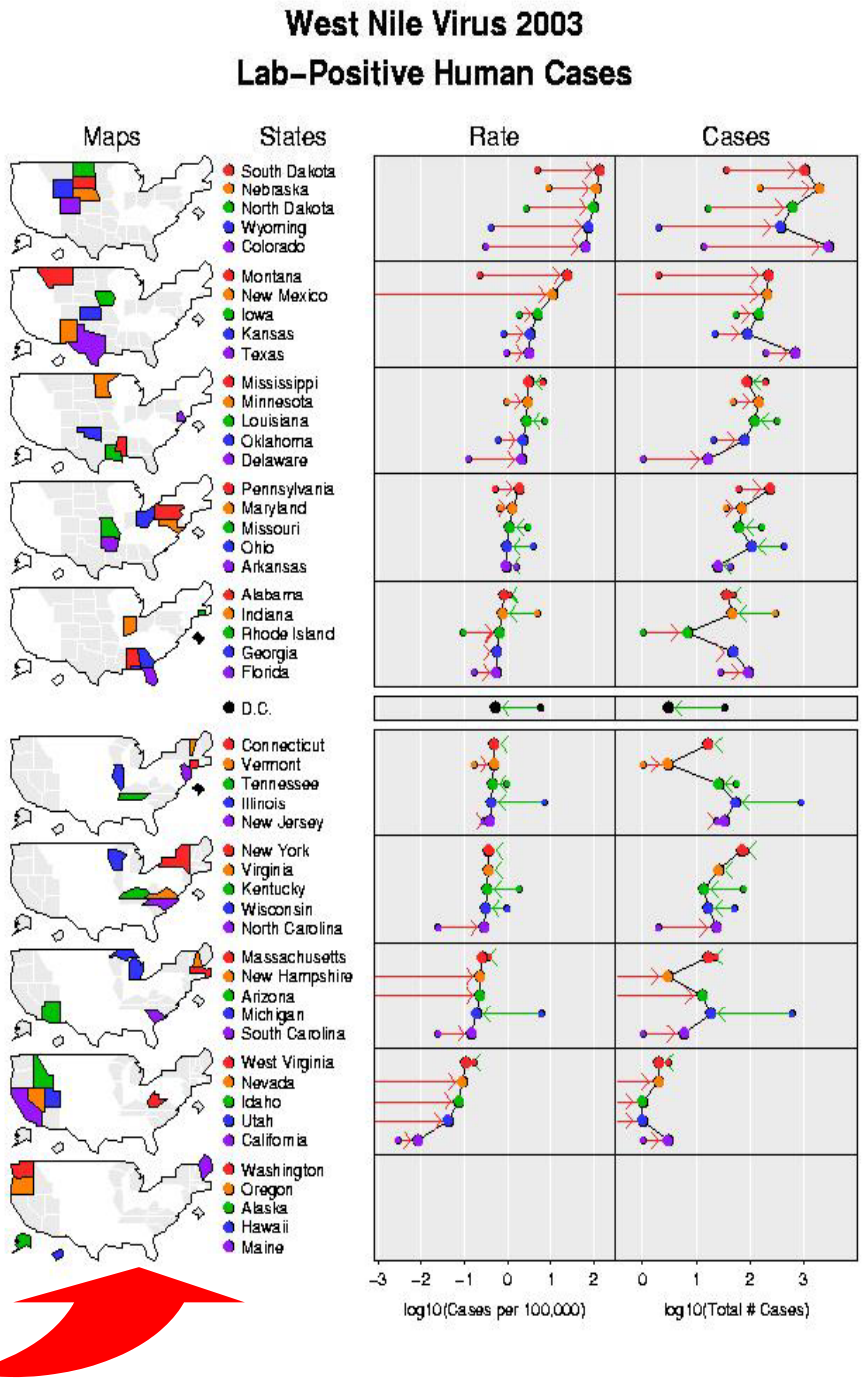
West Nile Virus (WNV) ???

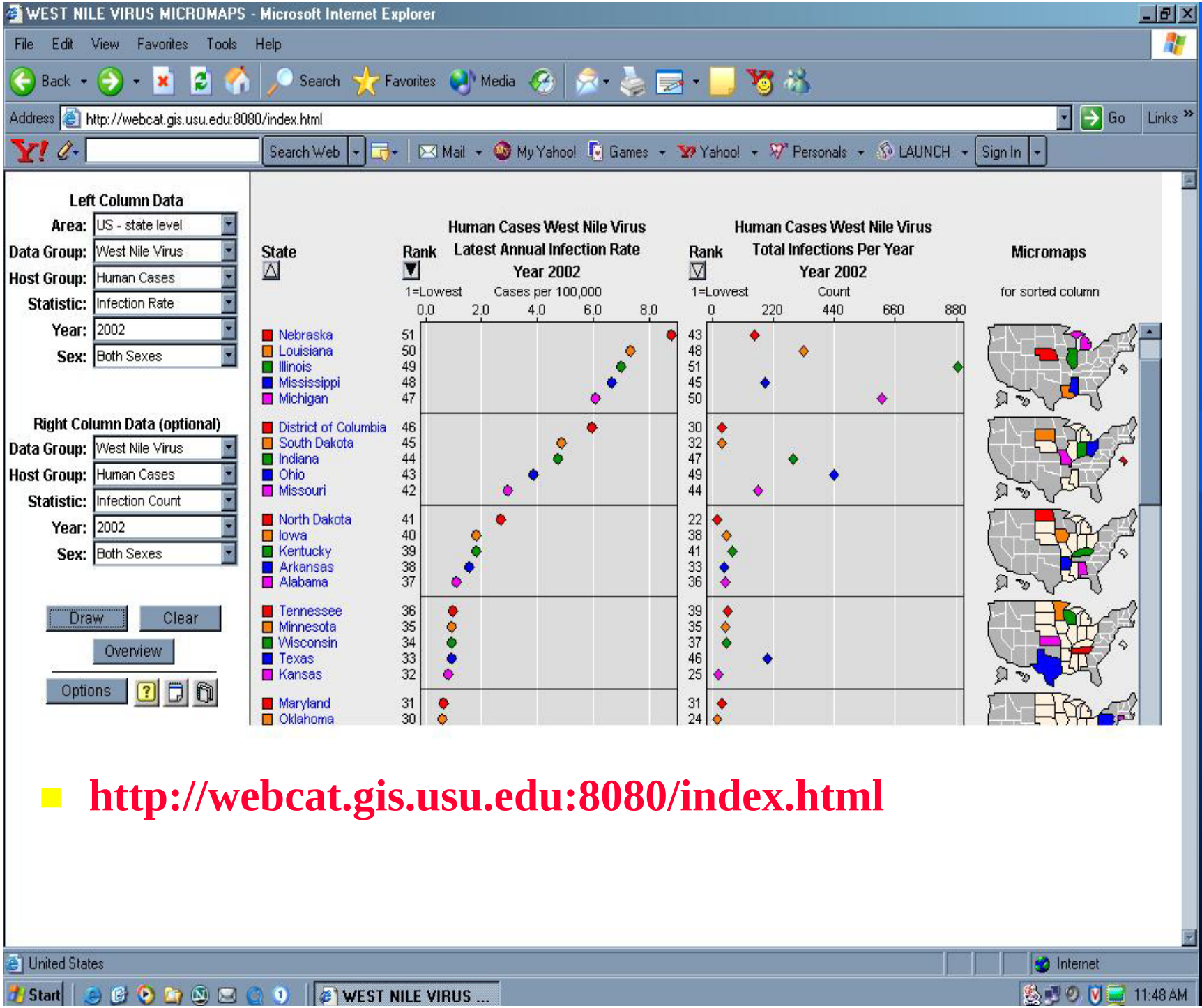
- Introduced to the US in 1999
- Spread across North America in 5 years
- Initial event - Culex mosquito transmits virus within avian populations
- Bridging Aedes albopictus transmits virus from birds to animals and humans

From 2003 CDC









- **http://webcat.gis.usu.edu:8080/index.html**

Overall Conclusions

- Visual approach effective to see unexpected structure in data
- Not every graphic will make it into a paper, but many of those that do not make it will provide useful insights for the researchers:
  - Graphics often help to create new hypotheses
  - Graphics useful to verify results
- Combination of different graphical techniques (static, interactive, various plots) most effective
- Graphics can be used for almost all types of data

Questions ???