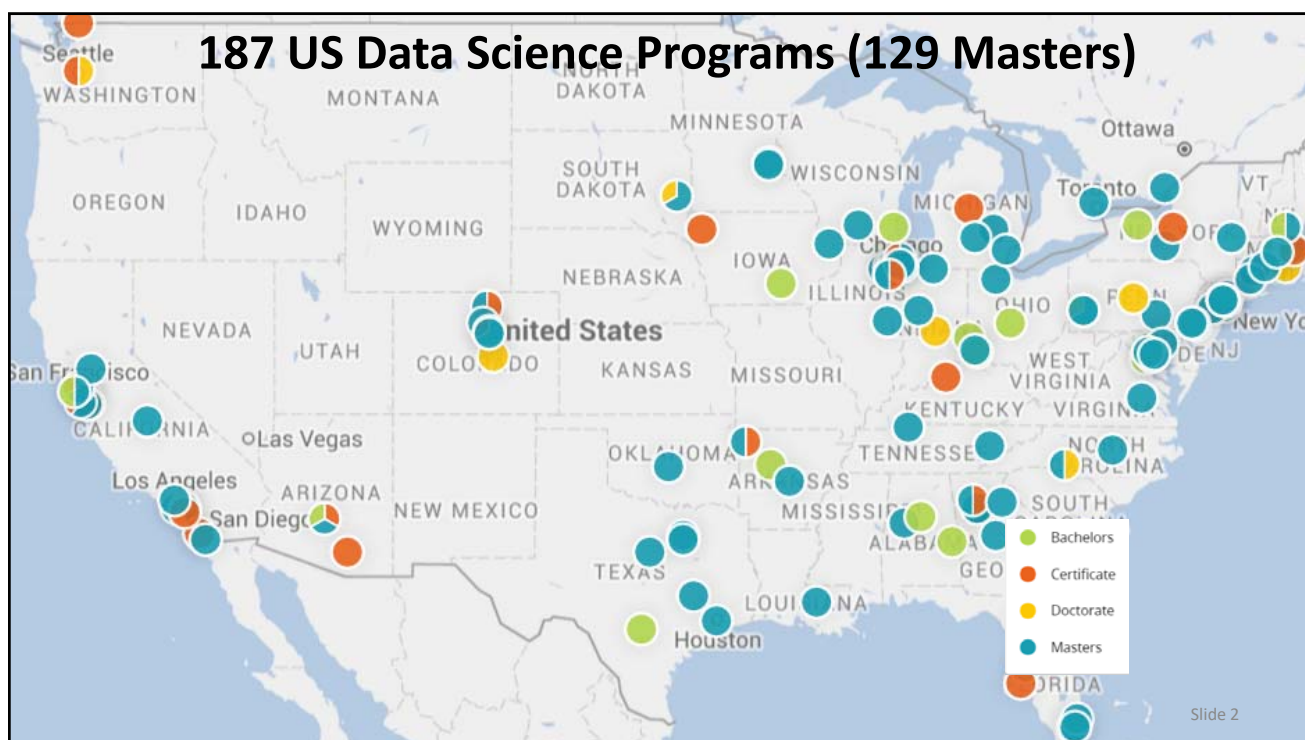


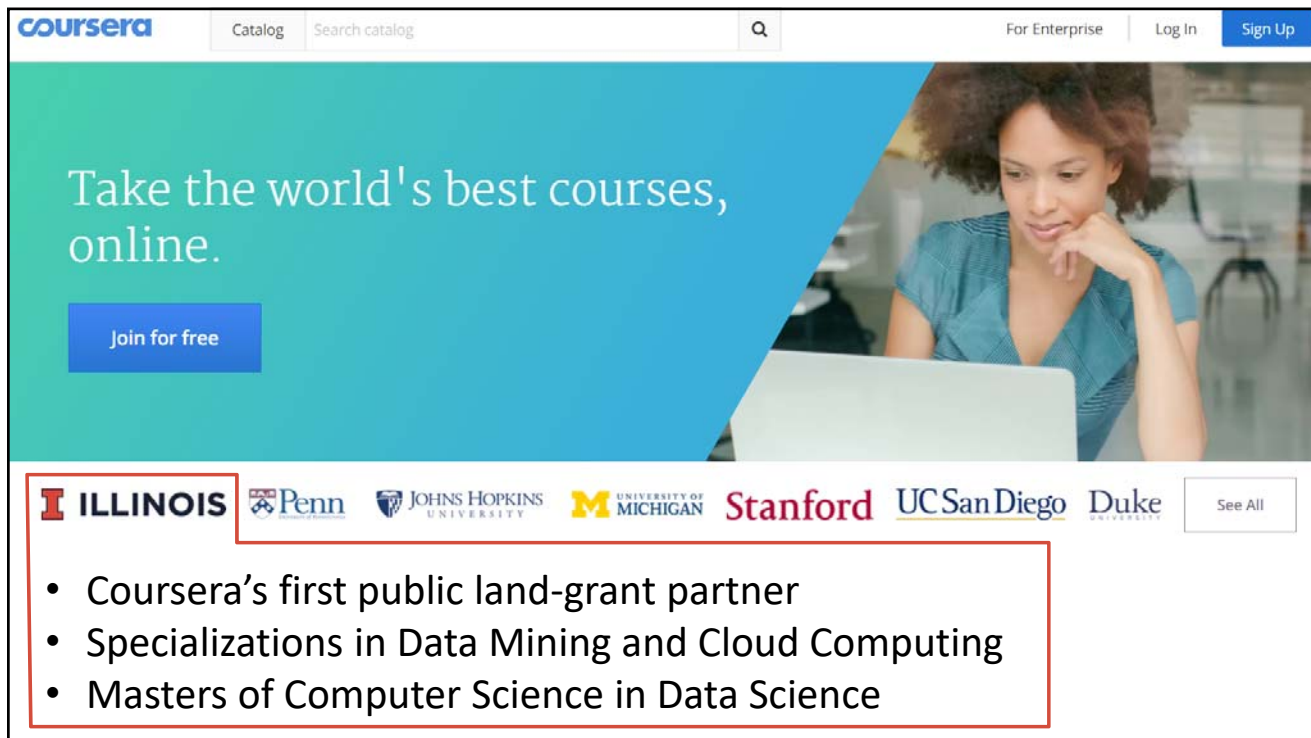
Data Visualization

John C. Hart

jch@illinois.edu

Department of Computer Science
University of Illinois at Urbana-Champaign





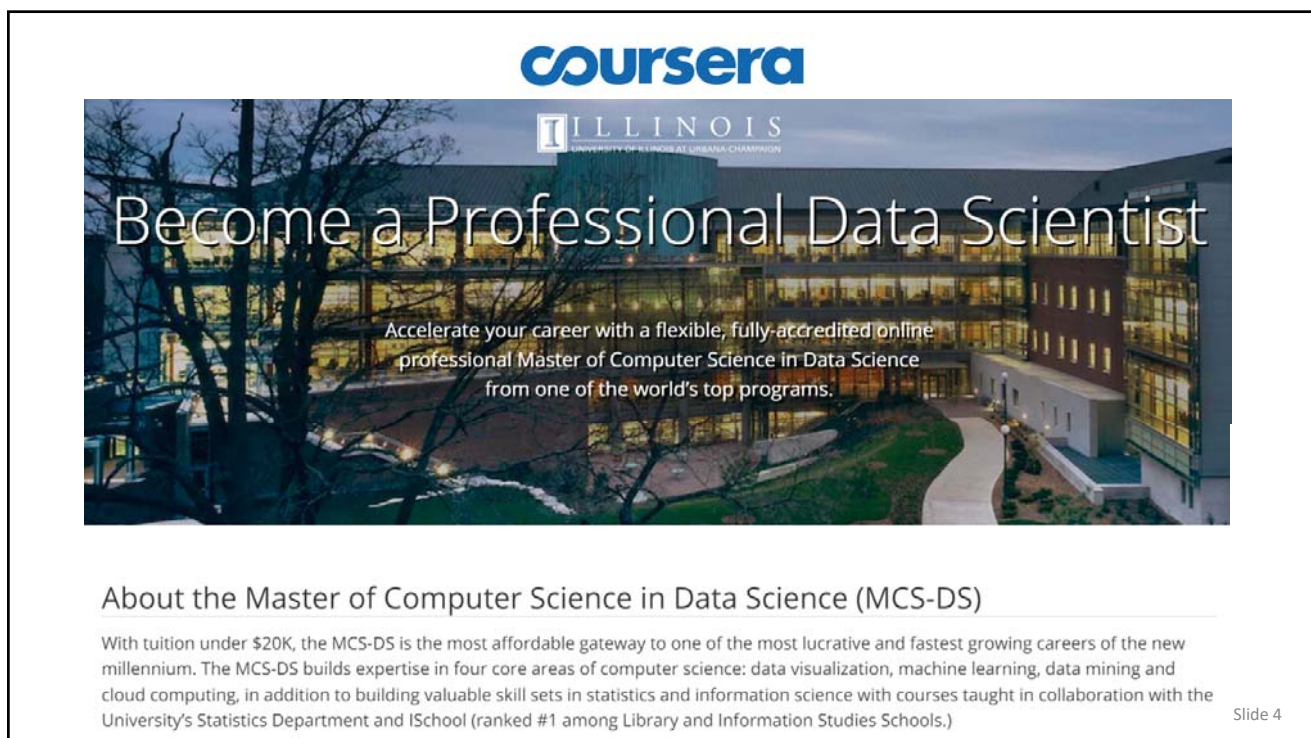
coursera Catalog Search catalog Q For Enterprise Log In Sign Up

Take the world's best courses,
online.

Join for free

I ILLINOIS Penn Johns Hopkins University University of Michigan Stanford UC San Diego Duke See All

- Coursera's first public land-grant partner
- Specializations in Data Mining and Cloud Computing
- Masters of Computer Science in Data Science



coursera

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Become a Professional Data Scientist

Accelerate your career with a flexible, fully-accredited online professional Master of Computer Science in Data Science from one of the world's top programs.

About the Master of Computer Science in Data Science (MCS-DS)

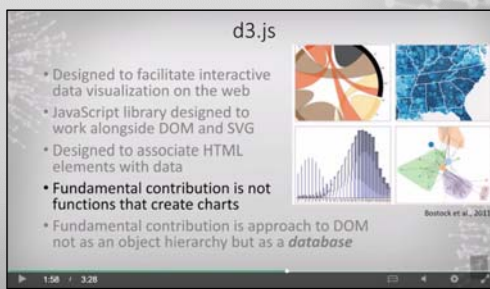
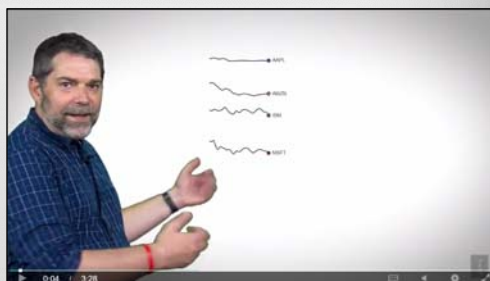
With tuition under \$20K, the MCS-DS is the most affordable gateway to one of the most lucrative and fastest growing careers of the new millennium. The MCS-DS builds expertise in four core areas of computer science: data visualization, machine learning, data mining and cloud computing, in addition to building valuable skill sets in statistics and information science with courses taught in collaboration with the University's Statistics Department and ISchool (ranked #1 among Library and Information Studies Schools.)

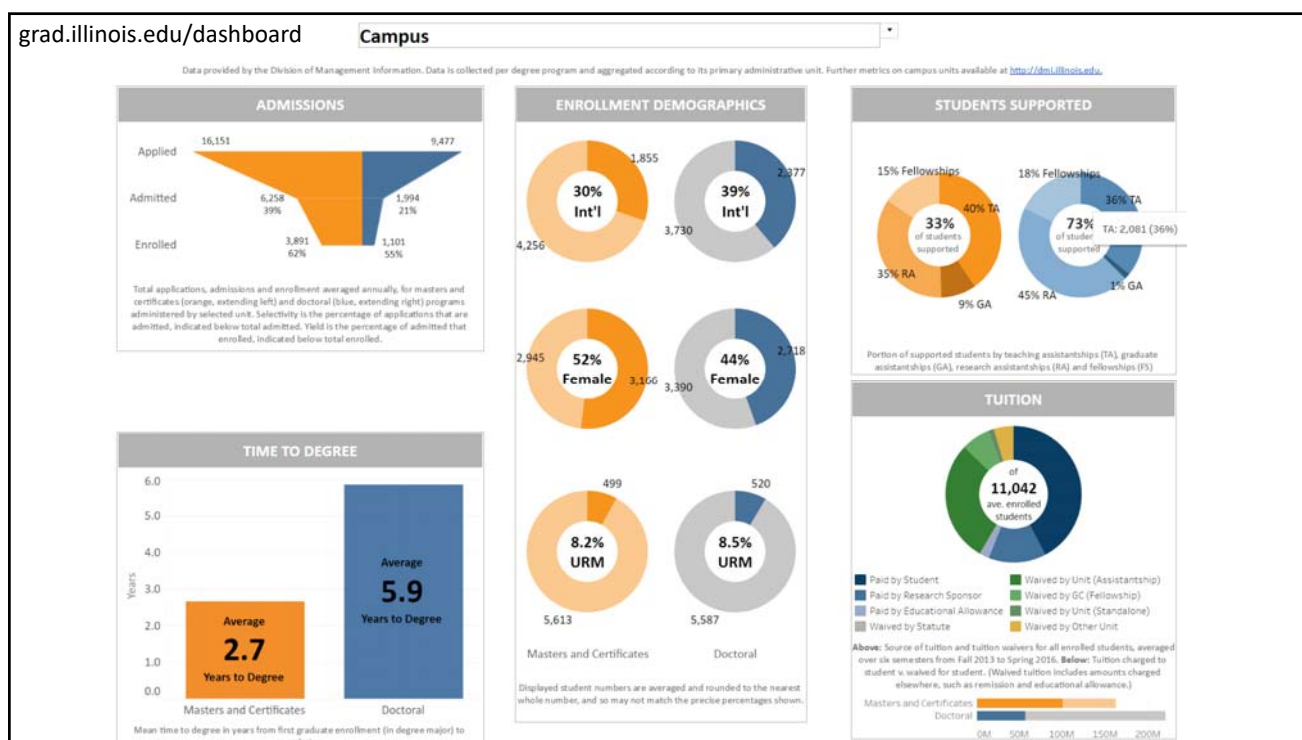
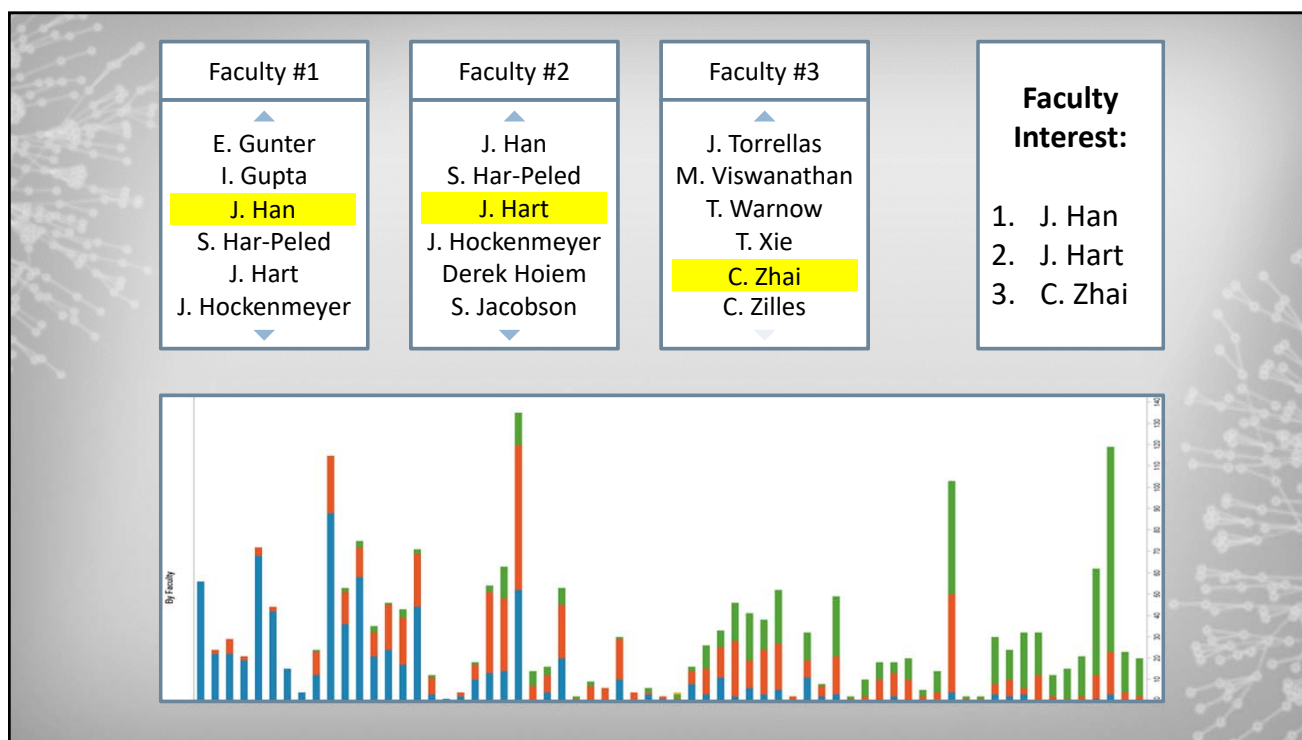
Slide 4



Data Visualization

- Four week MOOC course
- Offered since Nov. 2015
- Over 100K total learners
- 169 different countries
- 22% female
- Second MOOC online soon





Computer + Human > Computer

Prof. Frederick Brooks
1999 ACM Turing Award
Acceptance Speech

*Data science is more than just using cloud computing to apply machine learning to mine trends out of big data. Data science must also include **human insight** and the communication, visualization and resulting **awareness** of the data.*



What is Data Visualization

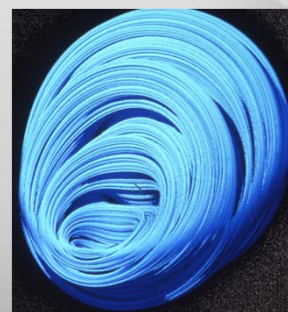
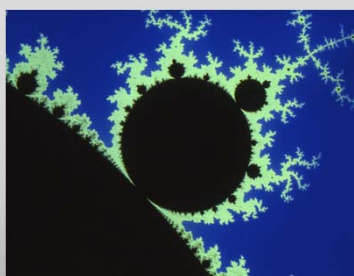
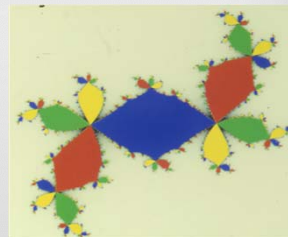
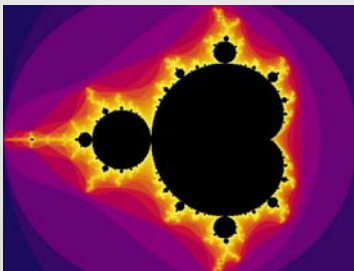
Data visualization is a high bandwidth connection between data on a computer system and a human brain, facilitated by visual communication

- Data acquisition
- Data processing
- Data display
- Human perception
- Human memory
- Human cognition



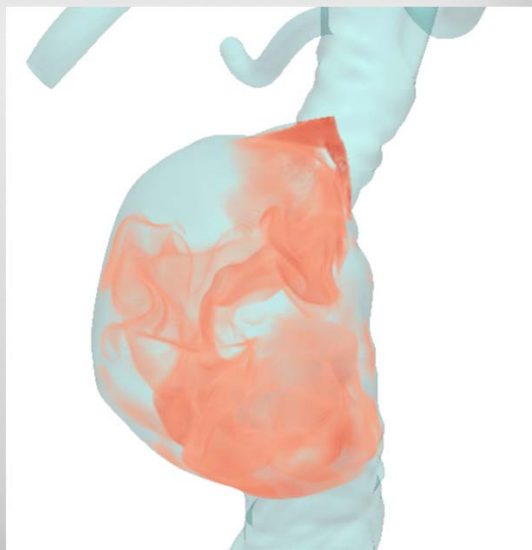
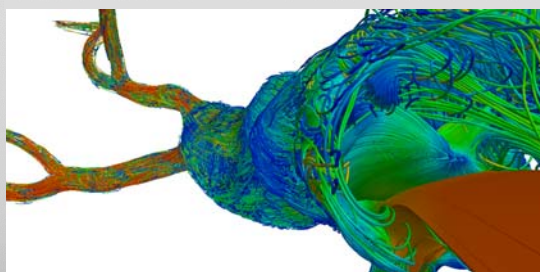
Mathematical Visualization

- Mathematical data synthesized by computation
- Can be computed on demand, reducing need for bandwidth and storage
- Not “Big Data”



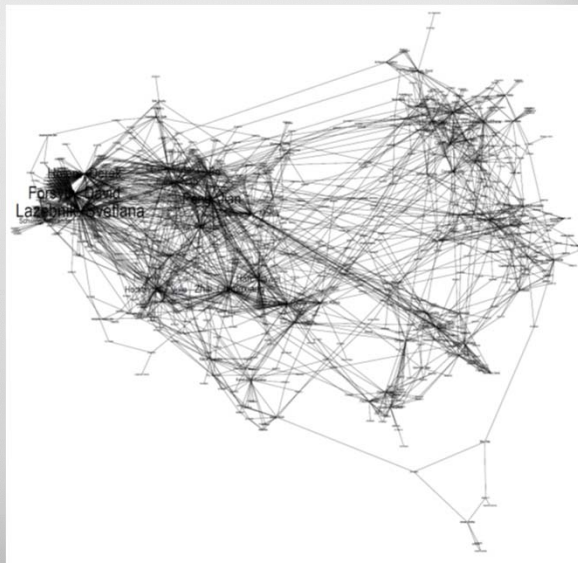
Scientific Visualization

- Visualization of “coordinate” data
- Data acquired from physical measurement or via computational simulation
- Data stored and transmitted, requiring attention to scalability



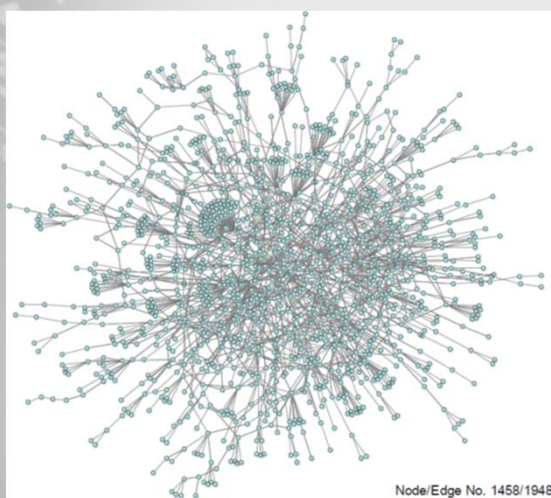
Information Visualization

- Visualization of more abstract, non-coordinate data
- Relies more heavily on processing abstract data into a more concrete form that can be more effectively perceived by an observer

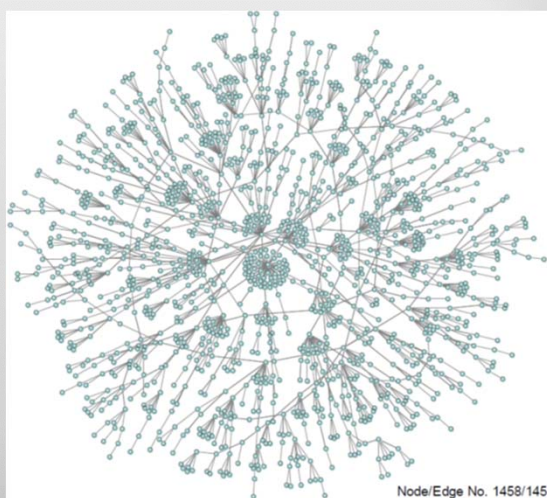


13

Graph Visualization



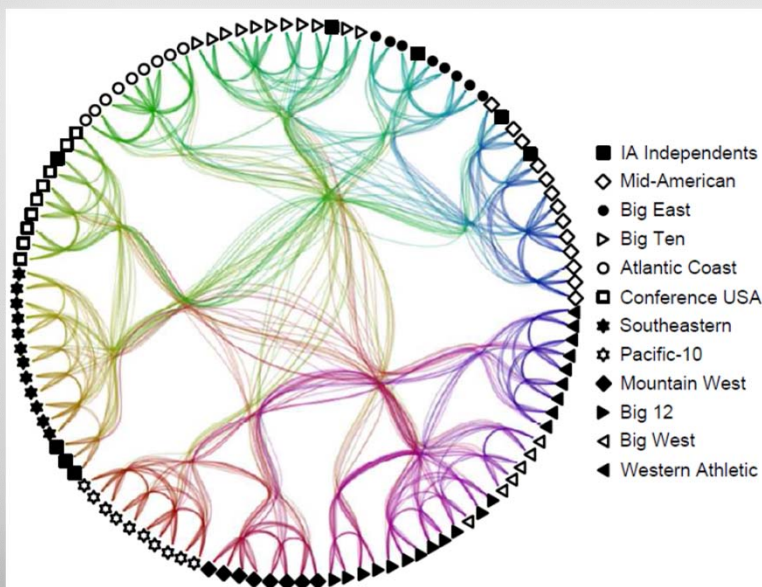
Node/Edge No. 1458/1948



Node/Edge No. 1458/1458

14

Graph Visualization



15

Modes of Visualization

Interactive Visualization

- Used for discovery
- Intended for a single investigator or collaborators
- Rerenders based on input
- Prototype quality

Presentation Visualization

- Used for communication
- Intended for large group or mass audience
- Does not support user input
- Highly polished

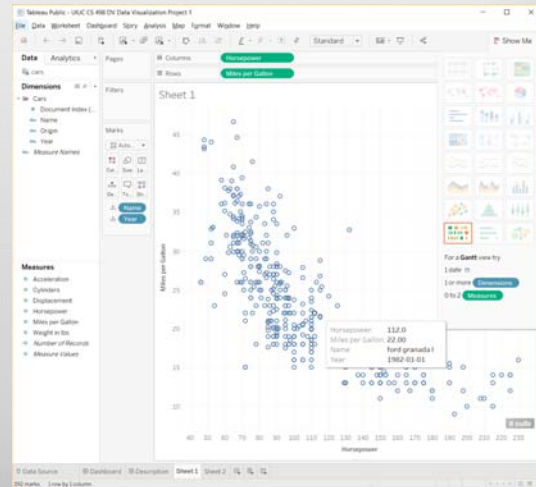
Interactive Storytelling

- Presentations via interactive webpages

16

“Data” Visualization

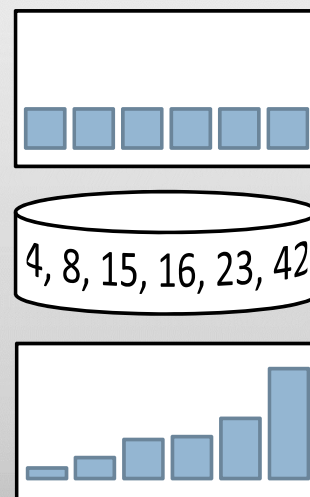
- Stronger connection between visualization and databases
- “VizQL” – Specification of a database query from a chart
- ShowMe – Recommendation for most effective chart
- D3 – Visualization as a “join” between database elements and document elements



17

“Data” Visualization

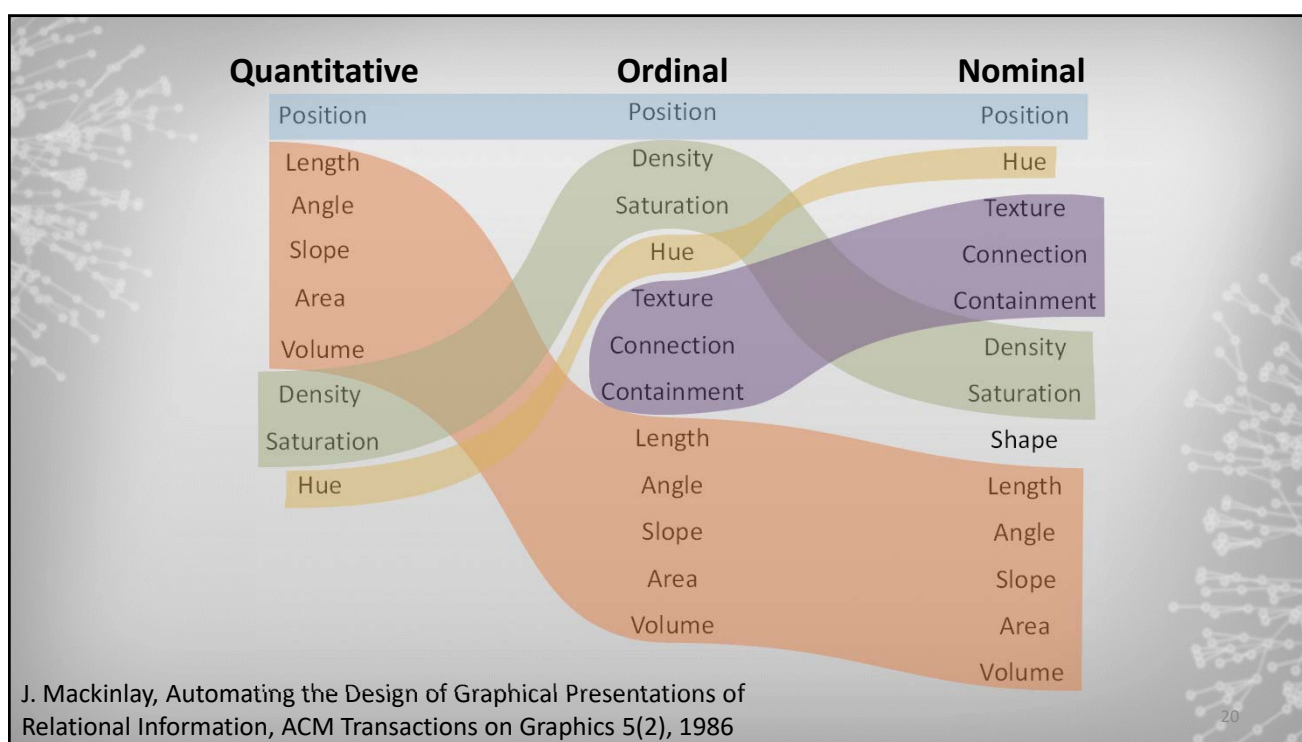
- Stronger connection between visualization and databases
- “VizQL” – Specification of a database query from a chart
- ShowMe – Recommendation for most effective chart
- D3 – Visualization as a “join” between database elements and document elements



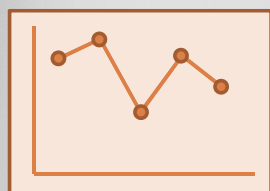
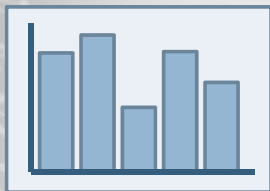
18

Quantitative	Ordinal	Nominal
Position	Position	Position
Length	Density	Hue
Angle	Saturation	Texture
Slope	Hue	Connection
Area	Texture	Containment
Volume	Connection	Density
Density	Containment	Saturation
Saturation	Length	Shape
Hue	Angle	Length
	Slope	Angle
	Area	Slope
	Volume	Area
		Volume

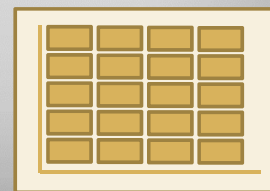
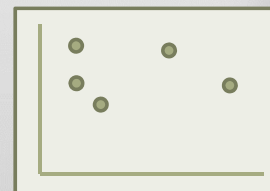
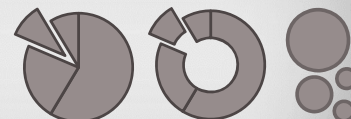
J. Mackinlay, Automating the Design of Graphical Presentations of Relational Information, ACM Transactions on Graphics 5(2), 1986



Which Chart Should I Use?



One quantitative variable	Pie, Donut or Bubble Chart
One quantitative variable depending on a second discrete variable	Bar Chart
One quantitative variable depending on a second continuous variable	Line Chart
Two independent continuous variables	Scatter Plot
Two independent discrete variables	Table



21

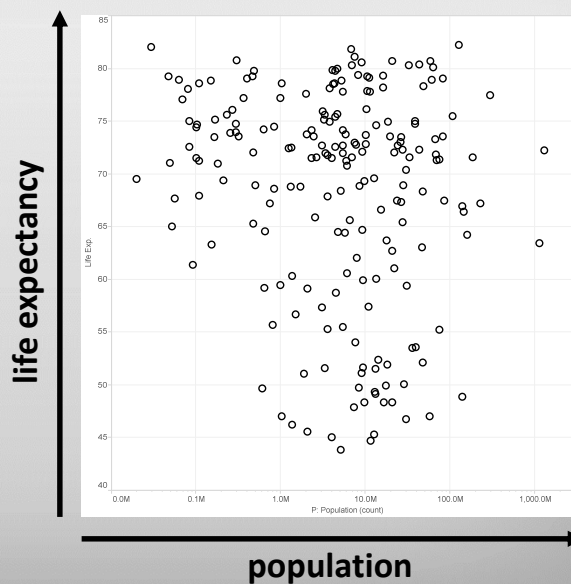
Data Visualization Mantra

1. Overview First
2. Focus (Filter and Zoom)
3. Details on Demand

22

Overview

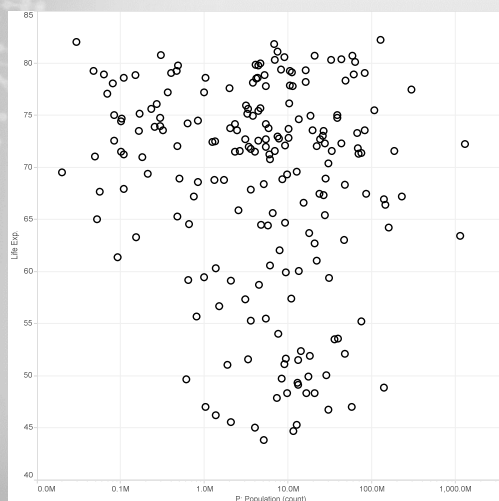
- Scatterplot of entire dataset
- Choose axes that evenly spreads out the data



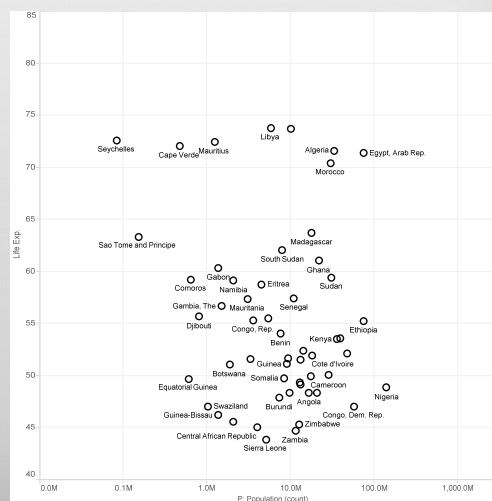
23

Filter

World



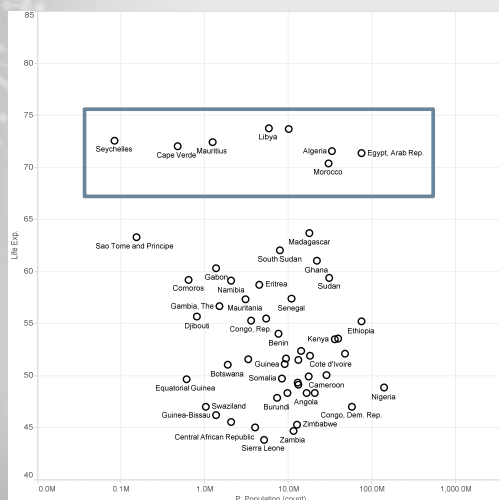
Africa



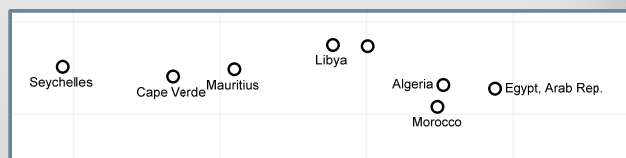
24

Zoom

Africa

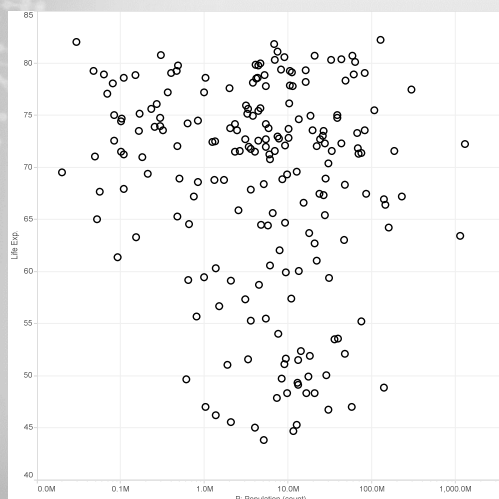


Africa (highest life expectancy)



25

Details on Demand

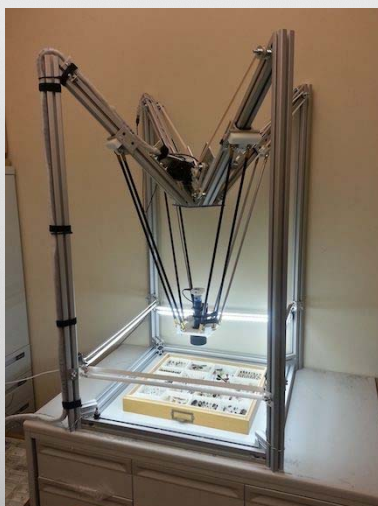


26

Some Visualization Systems

27

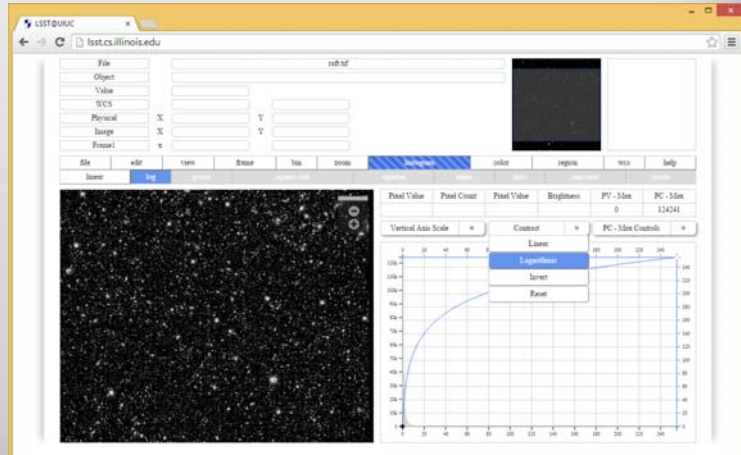
Invertnet.org



Large Synoptic Survey Telescope



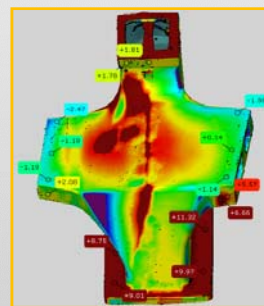
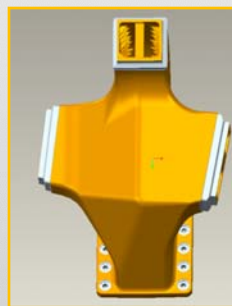
3 Gpixel



DMDII
DIGITAL MANUFACTURING AND
DESIGN INNOVATION INSTITUTE

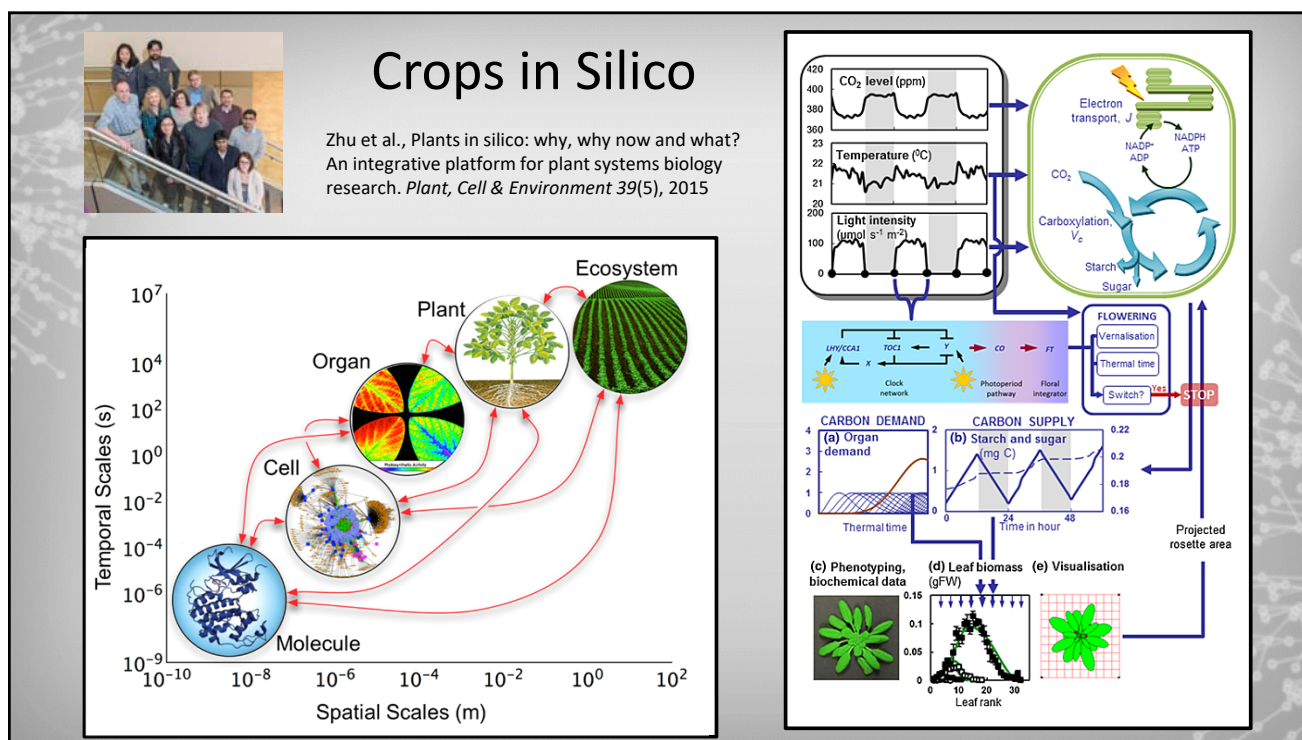
+ a **UI LABS** Collaboration

- Research Funding
- Applied R&D
- Shop Floor Demonstrations
- Commons
- Workforce Development



CATERPILLAR®





Data Visualization | Course

Secure | <https://www.coursera.org/learn/datavisualization>

coursera Explore What do you want to learn? For Enterprise John

Home > Data Science > Data Analysis

Data Visualization

About this course: Learn the general concepts of data mining along with basic methodologies and applications. Then dive into one subfield in data mining: pattern discovery. Learn in-depth concepts, methods, and applications of pattern discovery in data mining. We will also introduce methods for pattern-based classification and some interesting applications of pattern discovery. This course provides you the

▼ More

Created by: University of Illinois at Urbana-Champaign

ILLINOIS

Overview
Syllabus
FAQs
Creators
Ratings and Reviews

[Go to Course](#)

Already enrolled
[Apply for Financial Aid](#)