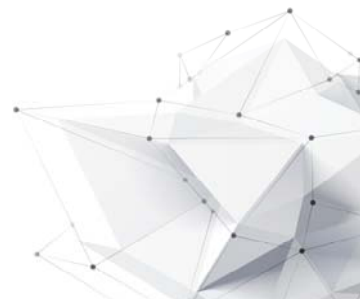


Critical Data Viz Workshop: An introduction to the ethics and **practice of data visualization**

Vetria L. Byrd, Ph.D.
Assistant Professor
Computer Graphics Technology

Critical Data Viz Workshop
Purdue University

May 20 - 24, 2019



ADMINISTRIVIA – DAY 1

Brief Introductions

- PAL Login
- Software Usage Agreement
- Tableau and Data Installation Check in
- Creation of TableauID; Join Tableau Online Environment

Making Sense of Data (Day 1: Monday, May 20, 2019)

- Introduction to Data Visualization
- Data Visualization Process (9:30 AM – noon)
 - ACQUIRE
 - PARSE
 - MINE
- Hands-on Practice: 1:00 PM – 2 PM
- Tableau Training I & II: 2PM – 3:00 PM



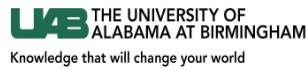
About Me

Vetria L. Byrd, PhD

Agent for "Insight"

Academic Preparation

- Computer Science (PhD, MS)
- Biomedical Engineering (MSMBE)



Visualization Initiatives

- Research Experience for Undergraduates in Collaborative Data Visualization Applications (2014/2015)



Visualization Webinars

International HPC Summer School
on HPC Challenges in Computational Sciences
Toronto, Canada (2015), Ljubljana, Slovenia
(2016), Boulder, CO, US (2017)

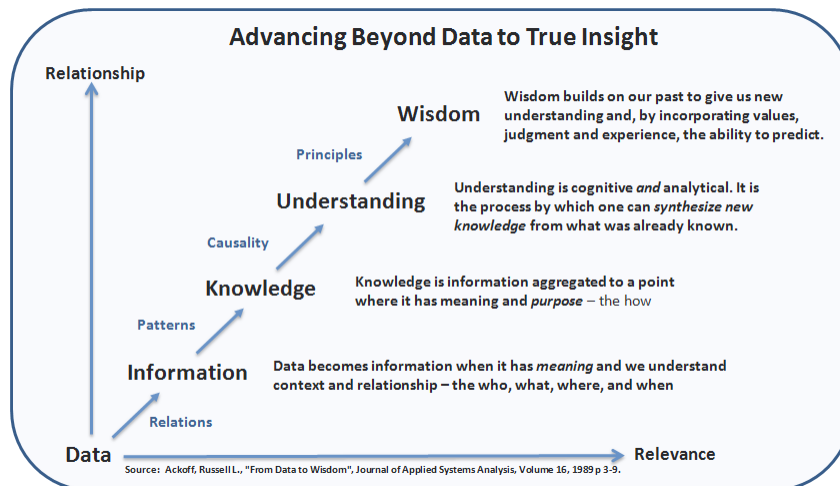


What is the purpose of Visualization?





FROM DATA TO INSIGHT



Why We Visualize Data

Today

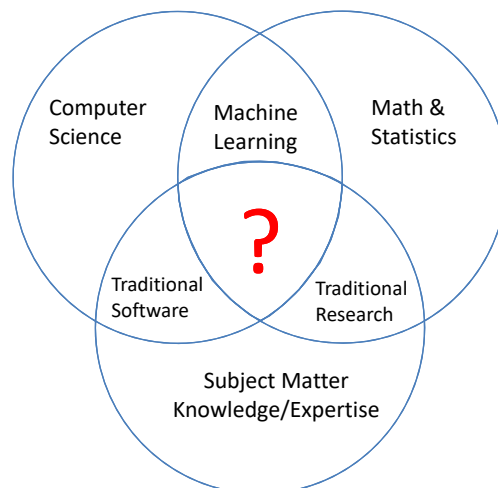
- To meet a very basic need – to tell a story
- One of the most primitive forms of communication known to man
- Cave drawings dated as early as 30,000 B.C.
- Even before written communication (3,000 B.C.)
- New ways to visualize information
- Basic chart types
 - Bar chart
 - Line chart
 - Pie chart
- Advanced visualization methods

What's your major field of study?

Adopted from FusionCharts White paper, "Principles of Data Visualization - What We See in a Visual"



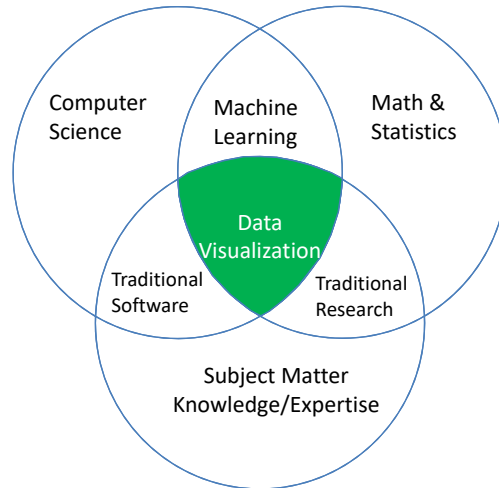
Data Science



Johnson, J. (2017). Data science & computing across the curriculum. Journal of Computing Sciences in Colleges., 32, 187-188.



Data Science



Johnson, J. (2017). Data science & computing across the curriculum. *Journal of Computing Sciences in Colleges.*, 32, 187-188.

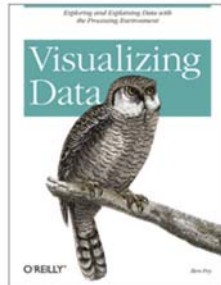


7 Stages of Visualizing Data

Dr. Vetricia L. Byrd
Polytechnic Institute
Purdue University
Spring 2019



Text Book: Chapter 1



Visualizing Data

Exploring and Explaining Data with the Processing Environment

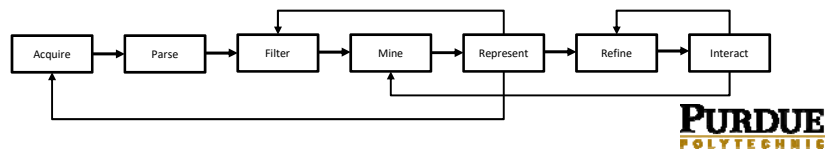
By Ben Fry

Publisher: O'Reilly Media

Release Date: December 2008

Pages: 320

http://media.espora.org/mgoblin_media/media_entries/1633/Visualizing_Data.pdf



Stage 1: Acquire



The acquisition step involves obtaining the data. Task: acquire data:

Like many of the other steps, this can be either extremely complicated (i.e., trying to glean useful data from a large system) or very simple (reading a readily available text file).

- First name
- Last name
- Height
- Major
- Academic status
- Programming Experience
- Visualization Experience



Fry, B. (2008). Chapter 1, Visualizing data (Safari Books Online). Sebastopol, Calif.: O'Reilly Media.

PURDUE
POLYTECHNIC

Stage 2: Parse

After you acquire the data, it needs to be parsed—changed into a format that tags each part of the data with its intended use.

Each line of the file must be broken along its individual parts; in this case, it must be delimited at each tab character.

Then, each piece of data needs to be converted to a useful format.

Example data

First name
Last name
Height
Academic status: Fr, So, Jr, Sr
Programming Experience (y/n)
Visualization Experience (y/n)

Parsed Data

First name	Last name	Height	Status	Prog Exp	Vis Exp
String Length: 10	String Length: 12	Float > 0.0	Char (2) Fr, So, Jr, Sr	Char (1) Y or N	Char (1) Y or N

String

- A set of characters that forms a word or a sentence.

Float

- A number with decimal points (used for the latitudes and longitudes of each location). The name is short for floating point, from programming nomenclature that describes how the numbers are stored in the computer's memory

Character

- A single letter or other symbol.

Integer

- A number without a fractional portion, and hence no decimal points (e.g., -14, 0, or 237).

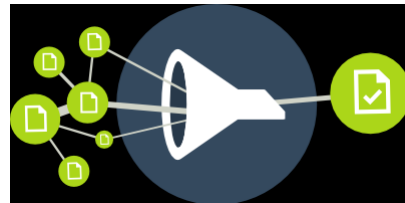


Fry, B. (2008). Chapter 1, Visualizing data (Safari Books Online). Sebastopol, Calif.: O'Reilly Media.

Stage 3: Filter

The next step involves filtering the data to remove portions not relevant to our use.

Another project could require significant mathematical work to place the data into a mathematical model or normalize it (convert it to an acceptable range of numbers).



Fry, B. (2008). Chapter 1, Visualizing data (Safari Books Online). Sebastopol, Calif.: O'Reilly Media.

Stage 4: Mine

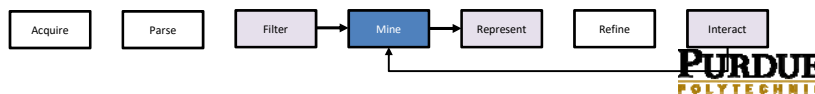
This step involves math, statistics, and data mining.

The data in this case receives only a simple treatment

Most of the time, this step will be far more complicated than a pair of simple math operations.

Tasks:

- Figure out the minimum and maximum values for numeric data
- Figure out the frequency of other values
- What patterns do you see?



Fry, B. (2008). Chapter 1, Visualizing data (Safari Books Online). Sebastopol, Calif.: O'Reilly Media.

Four types of Visualizations

GEORGES GRINSTEIN (KEYNOTE PRESENTATION, VINCI 2016)

- **Exploratory**
 - Have no hypotheses about the data
 - Explore data interactively as undirected searches
- **Confirmatory**
 - Have specific hypotheses about the data
 - Goal-oriented examination of the hypotheses
- **Presentation**
 - Facts to be presented are fixed a priori
 - Select appropriate presentation techniques
- **Interactive**
 - Interactions with a pre-defined animation

Stage 5: Represent

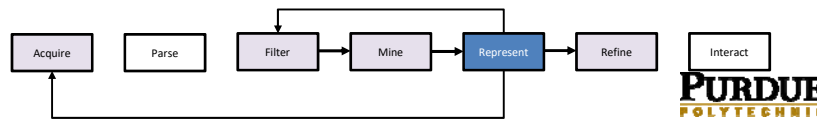
This step determines the basic form that a set of data will take.

Some data sets are shown as lists, others are structured like trees, and so forth.

The Represent stage is a linchpin that informs the single most important decision in a visualization project and can make you rethink earlier stages.

How you choose to represent the data can influence the very **first** step (what data you acquire) and the **third** step (what particular pieces you extract).

Task: generate a visualization based on the data received from the Mine stage

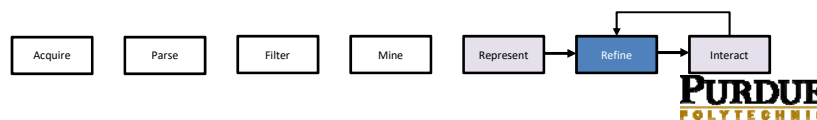


Fry, B. (2008). Chapter 1, Visualizing data (Safari Books Online). Sebastopol, Calif.: O'Reilly Media.

Stage 6: Refine

In this step, graphic design methods are used to further clarify the representation by calling more attention to particular data (establishing hierarchy) or by changing attributes (such as color) that contribute to readability.

Task: enhance the visualization created in Step 5: Represent



Fry, B. (2008). Chapter 1, Visualizing data (Safari Books Online). Sebastopol, Calif.: O'Reilly Media.

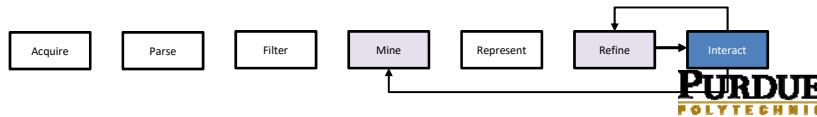
Stage 7: Interact

The next stage of the process adds interaction, letting the user control or explore the data.

Interaction might cover things like selecting a subset of the data or changing the viewpoint.

As another example of a stage affecting an earlier part of the process, this stage can also affect the refinement step, as a change in viewpoint might require the data to be designed differently.

Visually represent the data on the white board.



Fry, B. (2008). Chapter 1, Visualizing data (Safari Books Online). Sebastopol, Calif.: O'Reilly Media.

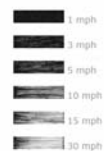
DATA VISUALIZATIONS

OVER THE LAST TWO DAYS: WE'VE SEEN SOME REALLY GREAT VISUALIZATIONS DURING THE POSTER SESSIONS

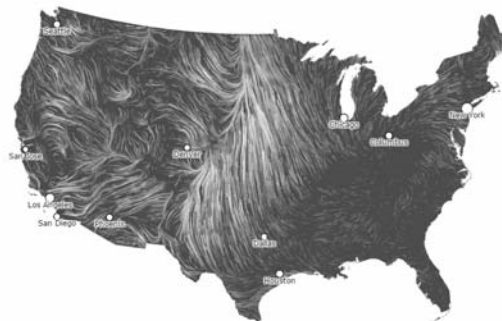
wind map

June 28, 2017
12:36 am EST
(Time of forecast download)

Top speed: 26.1 mph
Average: 7.5 mph



[Wind Map](http://hint.fm/wind/)
<http://hint.fm/wind/>



Gallery



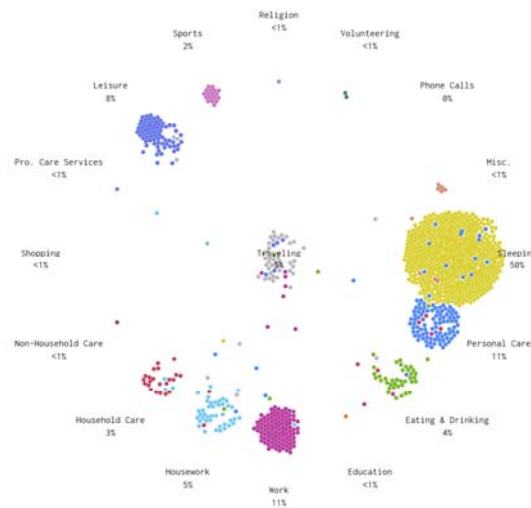
An invisible, ancient source of energy surrounds us—energy that powered the first explorations of the world, and that may be a key to the future. This map shows you the delicate tracery of wind flowing over the US.

Wind map prints are available from [Print B Studio](#).
Read more about [wind](#) and about [wind power](#).

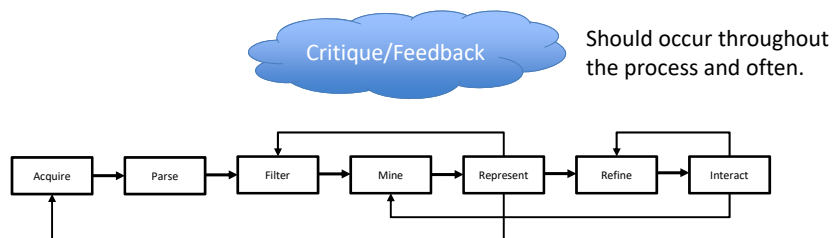
A DAY IN THE LIFE OF AMERICANS

6:41 am

SLOW MEDIUM FAST



7 Stages of Visualizing Data



What do we know?

Output from one stage serves as into the next stage

Iterative Process

Your first visualization will **not** be your last visualization

Your first visualization tool will not be your last

PURDUE
POLYTECHNIO

Fry, B. (2008). Chapter 1, Visualizing data (Safari Books Online). Sebastopol, Calif.: O'Reilly Media.

Insight Enables



Storytelling: The Next Step for Visualization

Robert Kosara and Jock Mackinlay, *Tableau Software, Seattle*

Kosara, R., & Mackinlay, J. (2013). Storytelling: The next step for visualization. *Computer*, 46(5), 44-50.

A Quick Tour Through the Data Visualization Workshop Brickyard



Introduction to

- Data Visualization
- Data Visualization Process
- Training Data



Design Activity Worksheets



Higher Order Thinking Assessment

Indy Motor Speedway Logo By Source, Fair use, <https://en.wikipedia.org/w/index.php?curid=37540940>

PURDUE
POLYTECHNIC

A Quick Tour Through the Data Visualization Workshop Brickyard



Form Groups

- Data Visualization
- Data Visualization Process
- Group Data



Design Activity Worksheets



Higher Order Thinking Assessment

Presentation of Results

Indy Motor Speedway Logo By Source, Fair use, <https://en.wikipedia.org/w/index.php?curid=37540940>

PURDUE
POLYTECHNIC

DATA VISUALIZATION WORKSHEETS

Are un-published work and considered new and proprietary.

Thank you in advance for **NOT** sharing the worksheets outside of the workshop.

PURDUE
POLYTECHNIC

A Quick Tour Through the Data Visualization Workshop Brickyard

Workshop Training File, Data and Resources

<https://web.ics.purdue.edu/~vbyrd/CSolWorkshop/TableauTrainingAll.html>



LET'S GET STARTED!

Growth of Walmart <http://projects.flowingdata.com/walmart/>

Growth of Target <http://projects.flowingdata.com/target/>



CHOOSE A NUMBER

1 2 3 4 5 6 7 8 9 10
11 12 13 14 15 16 17
18 19 20 21 22 23 24
25 26 27 28 29 30 31
32 33 34 35 36 37 38
39 40



STEP 1: ACQUIRE YOUR DATA

WALMART SALES DATA

The number you selected corresponds to the store number for the in the Walmart data

Download and explore the data for your store:

<https://web.ics.purdue.edu/~vbyrd/CSolWorkshop/WalmartData/WalmartData.html>



STEP 2: ACQUIRE MORE DATA

WALMART SALES DATA

- Form groups of 3 to 4.
- As a group select one of the remaining numbers (that are not crossed out) for the second Walmart store.
- The group will determine questions to explore and ultimately visualize about the new store.
- The question must be agreed upon by all members in the group and must span across each persons' data acquire in Step 1.

The number you selected corresponds to the store number for the in the Walmart data

Download and explore the data for your store:

<https://web.ics.purdue.edu/~vbyrd/CSolWorkshop/WalmartData/WalmartData.html>



CHOOSE A NUMBER

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17			
18	19	20	21	22	23	24			
25	26	27	28	29	30	31			
32	33	34	35	36	37	38			
39	40	41	42	43	44	45			



DATA VISUALIZATION WORKSHEETS

Complete the following Data Visualization Worksheets for your individual store choice:

Acquire

Parse

Mine

Complete the following Data Visualization Worksheets as **a group** for the additional store:

Acquire

Parse

Mine



WHAT'S NEXT

- Filter
- Represent
- Critique





Vetria L. Byrd
Assistant Professor
Computer Graphics Technology
vbyrd@purdue.edu
<https://polytechnic.purdue.edu/profile/vbyrd>
<http://web.ics.purdue.edu/~vbyrd/>
 @VByrdPhD, @BPViz, @ByrdVisLab
Purdue Polytechnic Institute
<https://BPViz.org>



PURDUE
POLYTECHNIC

Thank You Image Source:
<http://careerconfidential.com/category/thank-you-notes/>
<http://careerconfidential.com/wp-content/uploads/2015/02/ThankYou2.jpg>