

Broadening Participation in Visualization 2020

BUILDING DATA VISUALIZATION CAPACITY

Summer Webinar Series

July 21 - 24, 2020

UTSA The University of Texas
at San Antonio™

Byrd Data Visualization Laboratory

PURDUE
UNIVERSITY

Broadening Participation in Visualization 2020

©BPBiz

BPViz Webinar Series

Broadening Participation in Visualization (BPViz) Workshop

Founder and Organizer

Twitter: @BPViz

- BPViz 2020: [Building Data Visualization Capacity](#)
- BPViz 2019: [Diversity, Equity and Inclusion in Visualization Education & Careers](#)
- BPViz 2018: [3rd Biennial Broadening Participation in Visualization Workshop](#)
- BPViz 2016: 2nd Biennial Broadening Participation in Visualization Workshop
- BPViz 2014: 1st CRA-W/CDC Broadening Participation in Visualization Workshop

BPViz 2019

Home About Program Organizers Resources Activities

Broadening Participation, Diversity, Equity and Inclusion in Visualization Education & Careers

pre-conference workshop
July 13 - 14, 2019

Participants arrive
July 12, 2019

Workshop events begin
July 13, 2019

"DIVERSITY IS BEING INVITED TO THE PARTY;
INCLUSION IS BEING ASKED TO DANCE."
VERNA MYERS, DIVERSITY AND IN CLUSION EXPERT

Broadening Participation in Visualization

Home Program Organizers For Who Contact Us

i

BPVIZ 2018

Theme: Visualization Enabling Data Science For All
June 13 - 14, 2018

Day 1: Purdue University

8:00 AM Welcome

Dean Gary R. Bertoline, Ph.D.
Purdue Polytechnic Institute Dean and Distinguished Professor
Polytechnic Institute
Purdue University

8:30 AM Workshop Overview

Vetricia L. Byrd, Ph.D.
Founder and Co-Organizer
Assistant Professor
Computer Graphics Technology
Purdue University

About Me

- Vetria Byrd, PhD
- Assistant Professor
- Purdue University
- Byrd Data Visualization Lab, Director
- Research Interests
 - Pedagogy of Data Visualization
 - HPC Visualization: visualizing heterogeneous data, and complex data in AR/VR environments
 - Utilizing data visualization to advance science in the area of symptom cluster management for healthcare



Vetria Byrd
Assistant Professor
Byrd Data Visualization
Lab, Director
Purdue University

Webinar Organizers



Vetria Byrd
Assistant Professor
Byrd Data Visualization
Lab, Director
Purdue University



Ritu Arora
Assistant Vice President of
Research Computing
University of Texas at San
Antonio



Brent League
Director of Research
Computing Support
University of Texas at San
Antonio

Webinar Schedule

Schedule	Topic	Skill Level
Day 0	Download and Install	Beginner
Day 1	Introduction to data visualization	Beginner
Day 2	You've got data now what?	Beginner/Intermediate
Day 3	Introduction to scientific Visualization	Intermediate/Advanced
Day 4	Visualization hackathon	Advanced

If you participated in the Download and Install session on Monday, July 20, 2020, please take a moment to complete the *short* feedback survey (located at the bottom of the webinar webpage: <https://tinyurl.com/yctfhc6l>)

Administrivia & Logistics

- Webinar Schedule (CDT Time)
 - 1:00 PM – 1:15 PM Q&A
 - 1:15 PM – 2:00 PM Webinar Talk
 - 2:00 PM – 2:30 PM Hands-on
 - 2:30 PM – 3:00 PM UTSA
- Tools and Applications
 - Tableau (14-day Trial)
<https://www.tableau.com/products/desktop/download>
 - ParaView 5.8
<https://www.paraview.org/download/>



BUILDING DATA VISUALIZATION CAPACITY

Summer Webinar Series

July 21 - 24, 2020

UTSA The University of Texas
at San Antonio™

Byrd Data Visualization Laboratory

PURDUE
UNIVERSITY

Webinar Website: <https://tinyurl.com/yctfhc6l>

Byrd **V**is Lab
POLYTECHNIC

About the Webinar Series

- Four-days of visualization content
- To maximize your experience, there will be exercises for you to complete (to familiarize yourself with webinar tools and applications) before each webinar
- Each day is stand alone; however, you are encouraged to participate in the entire series
- Participants who attend the entire series will be given priority on Day 4: Visualization Hackathon and Consultation

An aerial photograph of a city, likely San Francisco, with a red overlay. The text "BUILDING DATA VISUALIZATION CAPACITY" is centered over the image in white, uppercase letters.

BUILDING DATA VISUALIZATION CAPACITY

What is data visualization capacity?

Data Visualization Capacity Building

Stage 1

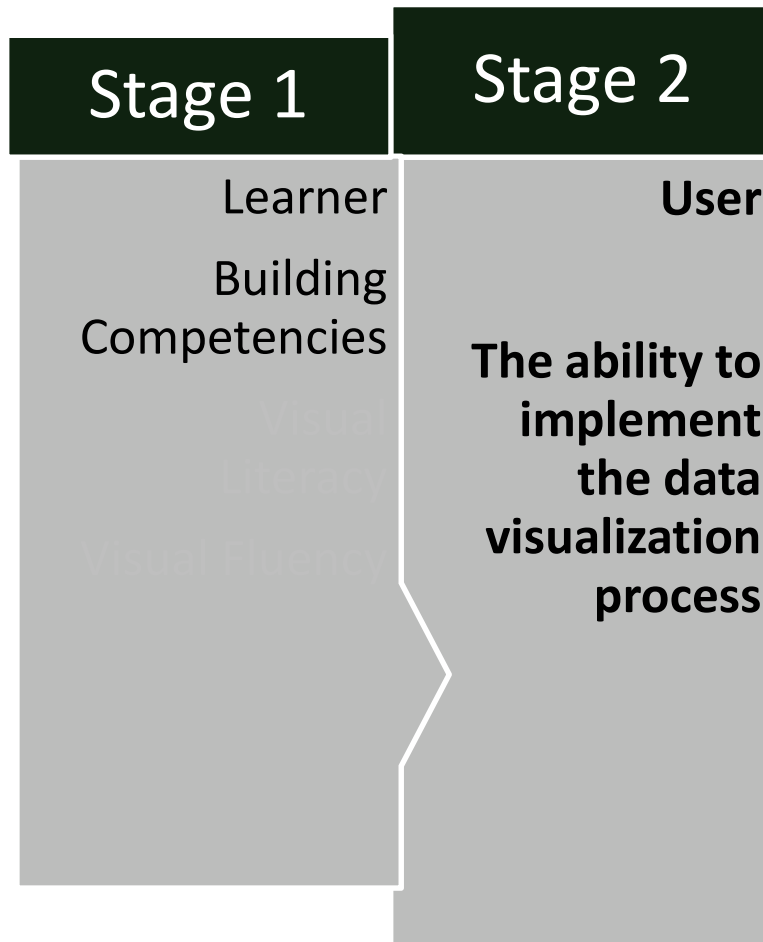
Learner

Understanding
what you see

Building Competencies

- Data Literacy
- Data Fluency
- Information Literacy
- Visual Literacy

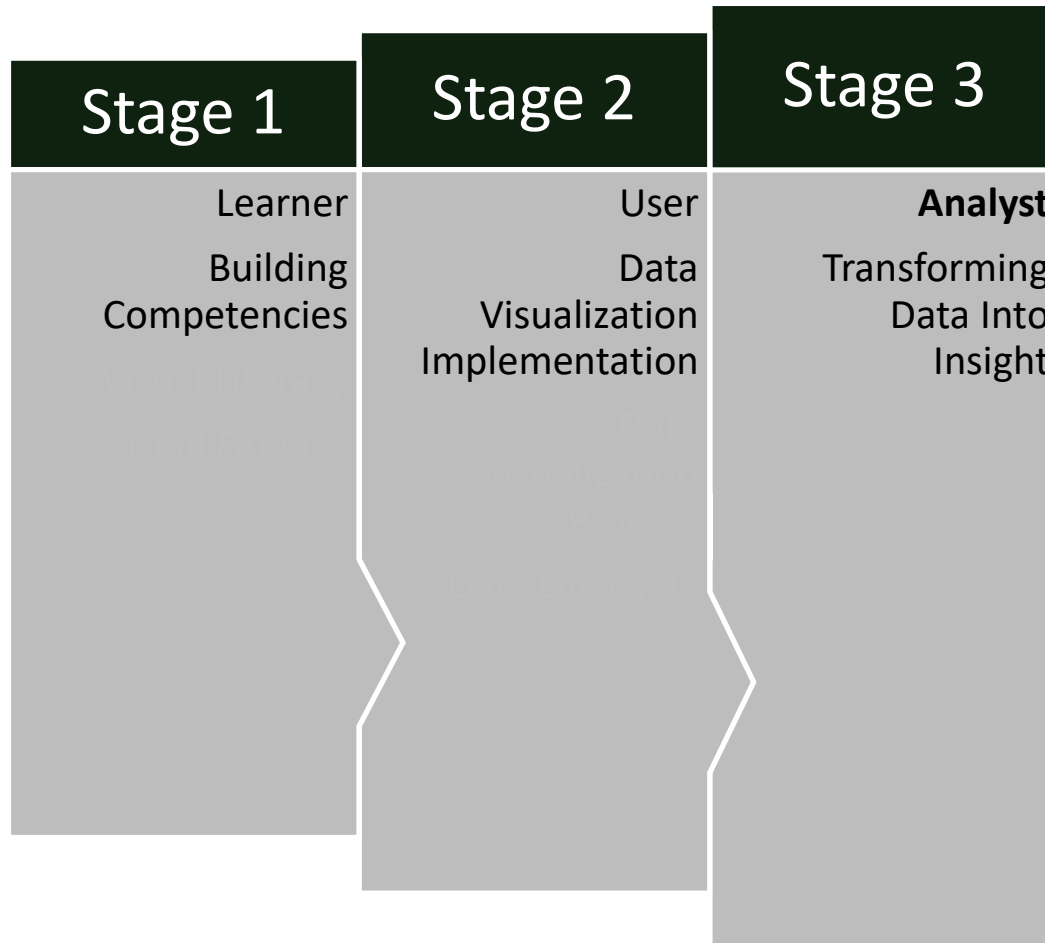
Data Visualization Capacity Building



Application

- Visualization Principles
- Visualization Techniques
- Visualization Process
- Choosing appropriate tool/chart for the data

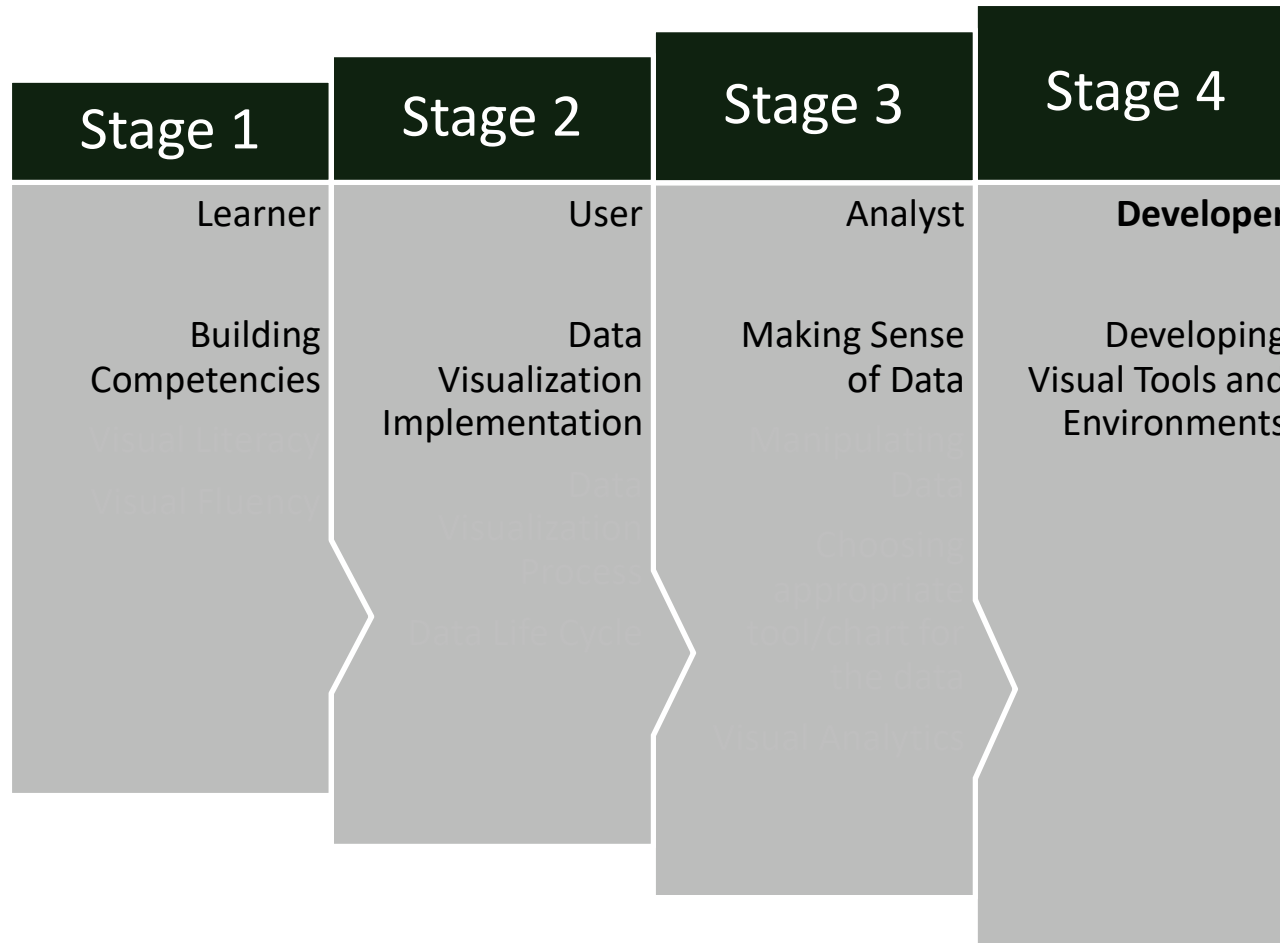
Data Visualization Capacity Building



Making Sense of Data

- Manipulating Data
- Visual Analytics
- Data Mining
- Etc.

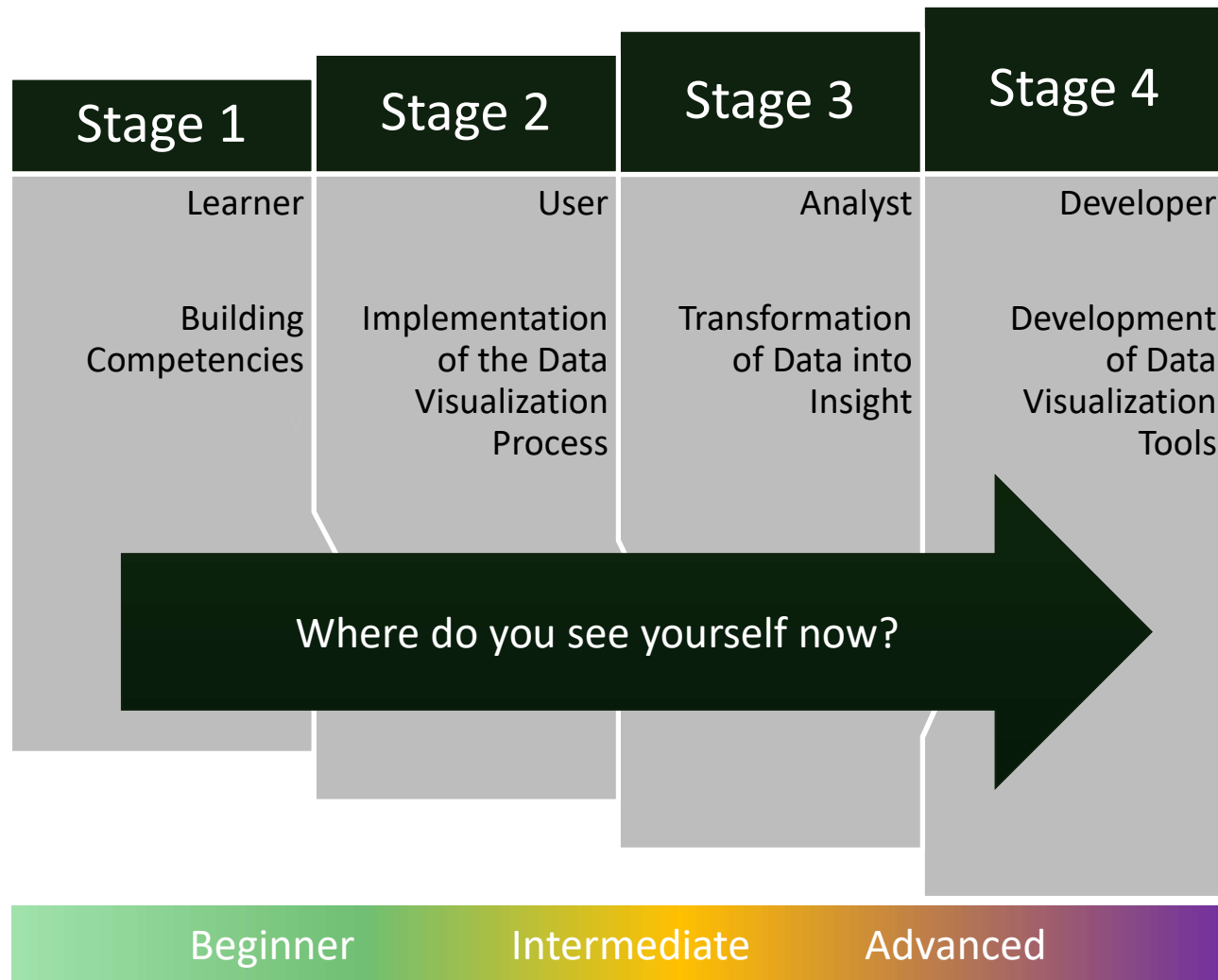
Data Visualization Capacity Building



Data Visualization Tool Development

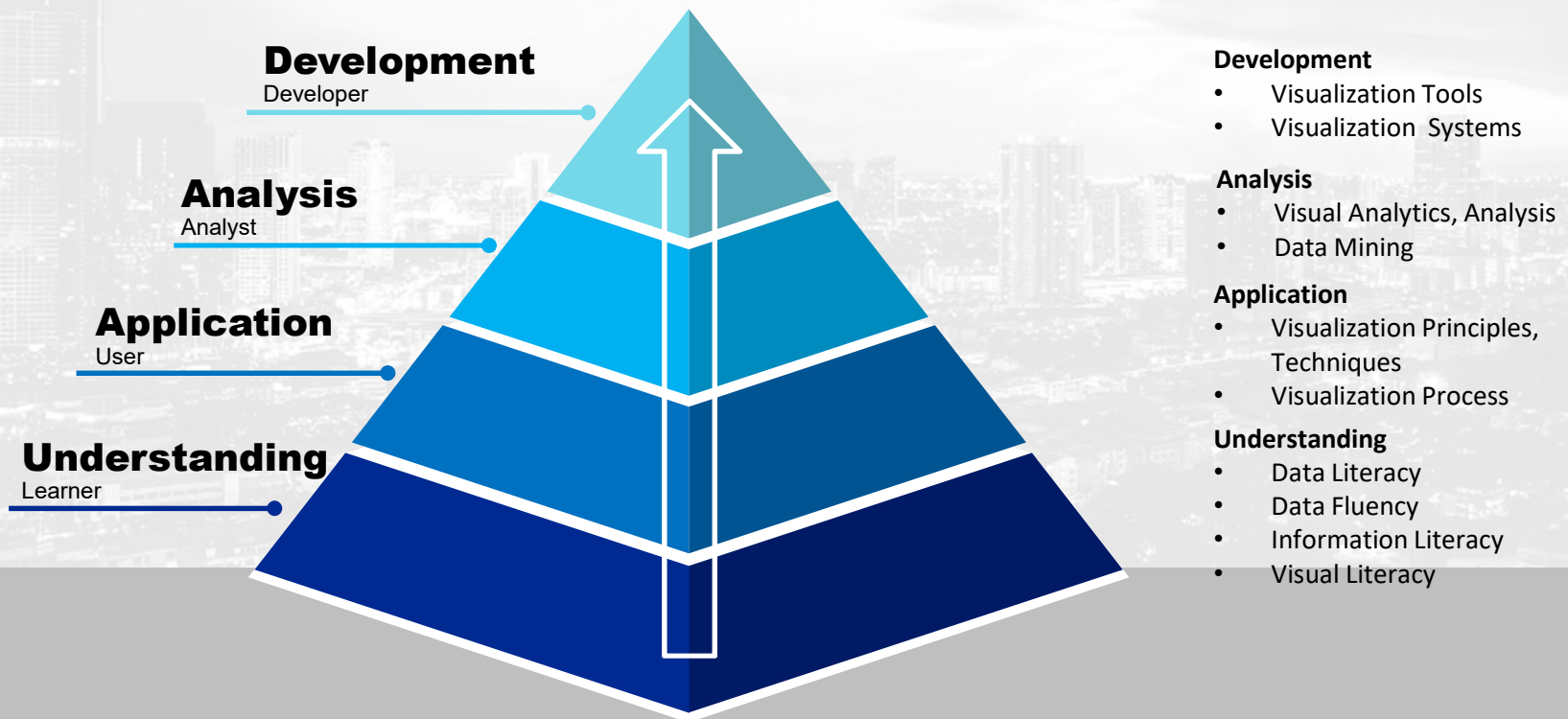
- Programming for Data Visualization
- Developing Visual Tools
- Developing Visual Environments

Data Visualization Capacity Building



Where would you like to be?

Building Data Visualization Capacity



We are here

Schedule	Topic	Skill Level
Day 0	Download and Install	Beginner
Day 1	Introduction to data visualization	Beginner
Day 2	You've got data now what?	Beginner/Intermediate
Day 3	Introduction to scientific Visualization	Intermediate/Advanced
Day 4	Visualization hackathon	Advanced

Webinar Goals and Objectives

Goal

Introduce participants to data visualization

Objectives

1. Examine what data looks like, define data visualization.
2. Illustrate how data visualization can improve understanding of the story within the data.
3. Introduce the data visualization process.
4. Explore different data visualization paths.

Lets Get Started!

Introduction to Data Visualization



What is Data
Visualization?

**How would you define
Data Visualization?**

Data Visualization

A process of transforming raw, complex data into a visual representation that does not overwhelm the viewer.

Principles of Data Visualization

Objective

- Provide foundational understanding of how we process visual information

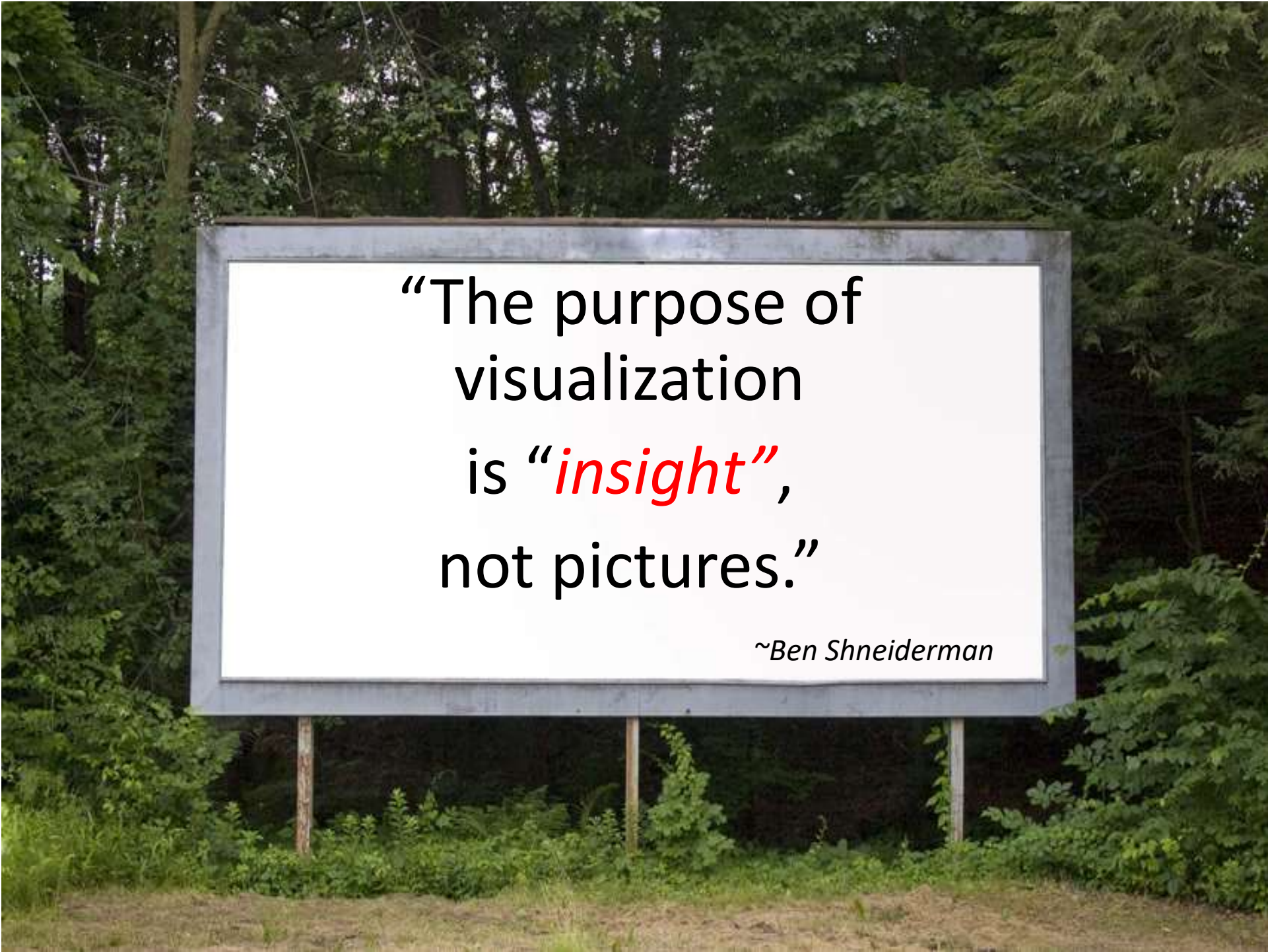
Outcomes

- Informed opinion on how to communicate more clearly and powerfully using visualizations
- Better analyze visualizations you come across in the newspaper, on the web or in your daily experience

Q

What is the purpose of
Visualization?

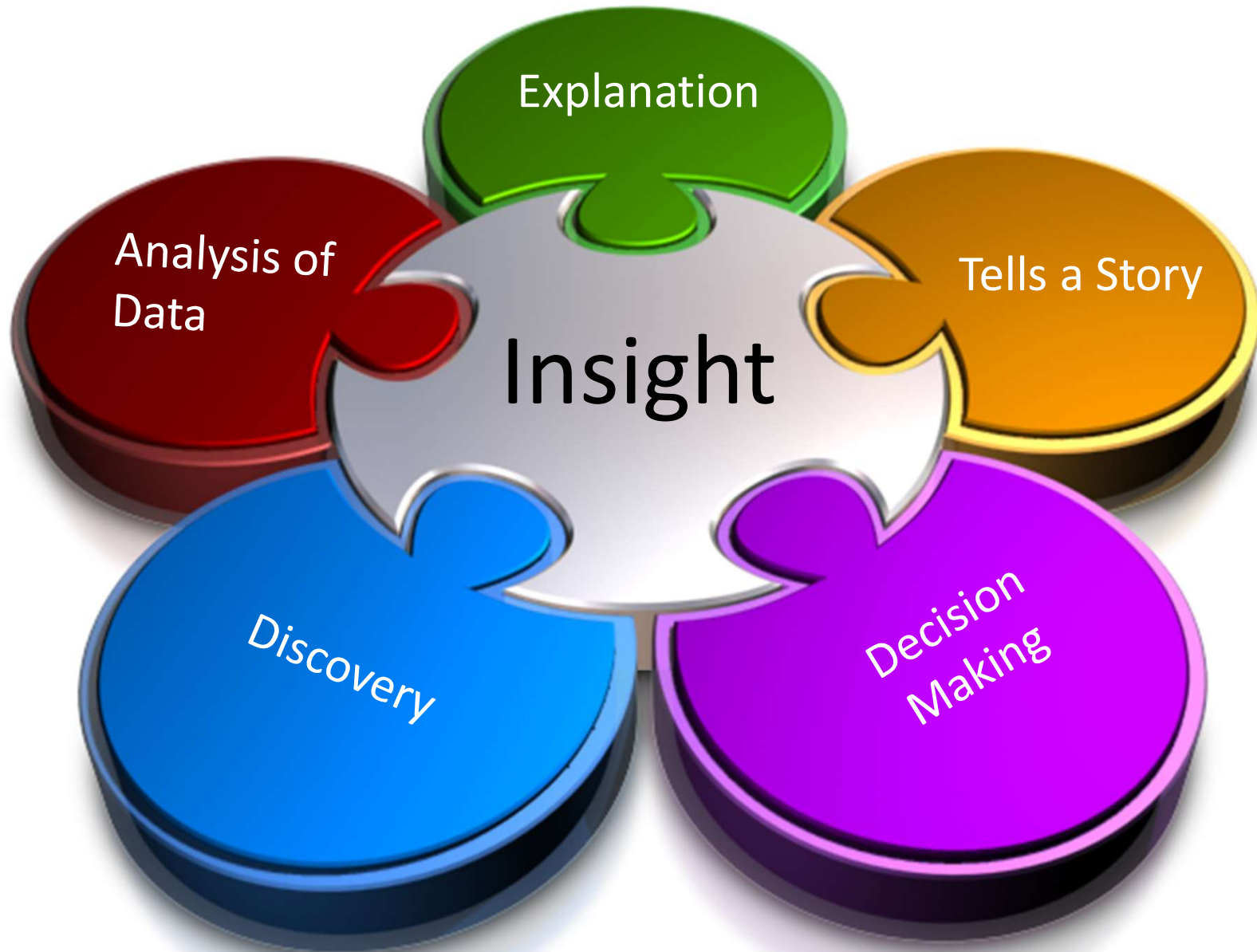


A rectangular sign with a white background and a grey border is mounted on three wooden posts. The sign is set against a backdrop of dense green trees and foliage. The text on the sign is centered and reads: "The purpose of visualization is *insight*, not pictures."

“The purpose of
visualization
is “*insight*”,
not pictures.”

~Ben Shneiderman

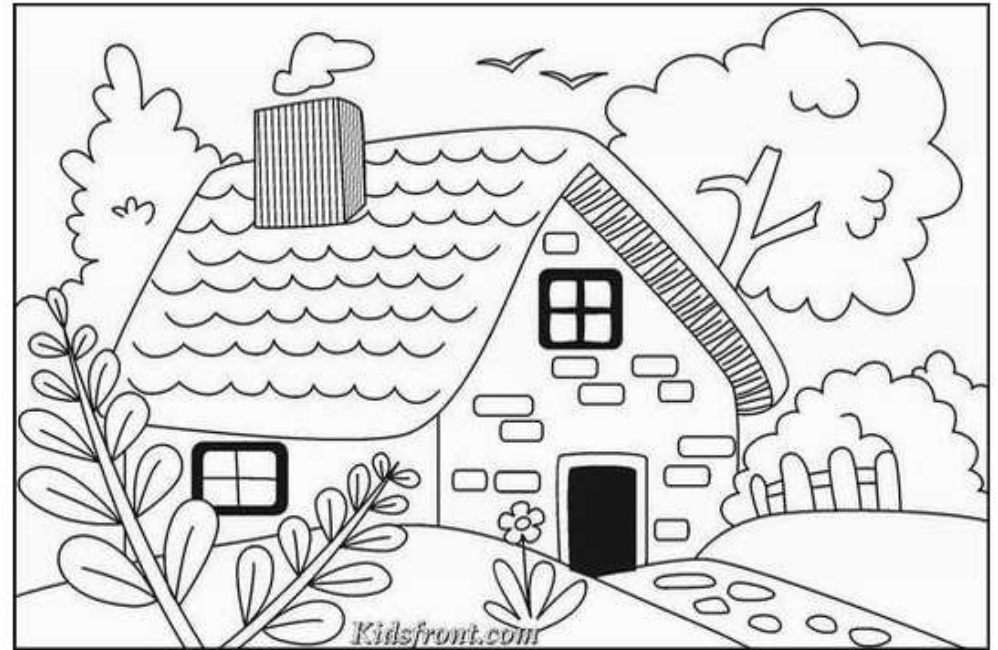
INSIGHT LEADS TO:



“Insight” Leads to . .

Discovery

Spotting
Differences



<http://www.kidsfront.com/spot-differences/9.html>

“Insight” Leads to . .

Discovery

- 1010101010101010101010101010010101010
10107010101010101001070010110011001
10011001100110011100110011001010101
01010701010101011100010111000101111
00010100110101010101010101110001100
10101010101010100010107010010100010
10101010101010101010101010101010101
01010101010101010101010101010101010
10101010701010110101070101010101010
70101001010101010101010101010101001
01001011001100110011001100110011100
11001100101010101010101010101011100
01011100010711100010100110101010101
01010111000110010101010101010100010
10101001010001010101010101010701010
10101010101010107010101010101010101
01010101010101010101010101010111001
1001100110010101001

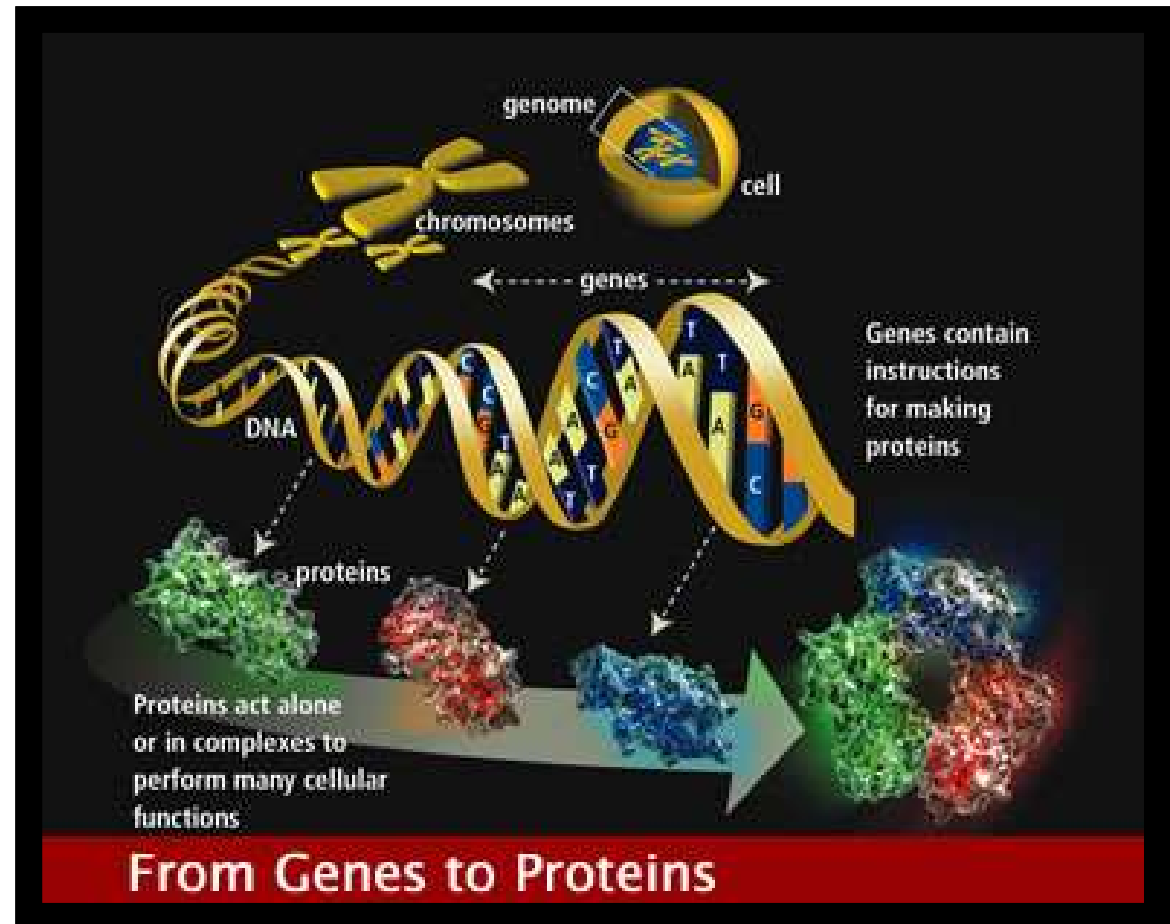
Spotting Differences

- Visualizing Patterns
- Spotting Differences
 - How many
 - 7's do you
 - see?

“Insight” Leads to . .

Decision Making

- **Allows users to answer questions they didn't know they had**



“Insight” Leads to . .



- Analysis of Data
- Explanation
- Storytelling

LIVESCIENCE NEWS TECH HEALTH PLANET EARTH

Live Science > Tech

The Amazing Tech in 'Black Panther' Is More Realistic Than You Think

By Mindy Weisberger, Senior Writer | April 17, 2016 01:12pm ET

Be the most interesting person you know, subscribe to LiveScience. [Subscribe](#)

Wakandan princess and inventor Shuri (Letitia Wright) takes aim with one of her high-tech designs: vibranium-powered gauntlets.
Credit: Marvel Studios



Katherine Johnson (played by Taraji P. Henson) calculates orbital insertion trajectories for the Mercury program using Euler's method in this scene from the movie Hidden Figures. Credit: TM and © 2017 Twentieth Century Fox Film Corporation. All rights reserved.



©2017 TWENTIETH CENTURY FOX



CHERNIN GROUP

“Insight” Leads to . .

- Analysis of Data
- Explanation
- Storytelling

[Hans Rosling, the visualization pioneer who made data dance](#) ~ The Washington Post



Hans Rosling's 200 Countries, 2000 Years, 4, Minutes

<https://www.youtube.com/watch?v=jbkSRLYSojo>

Four types of Visualizations

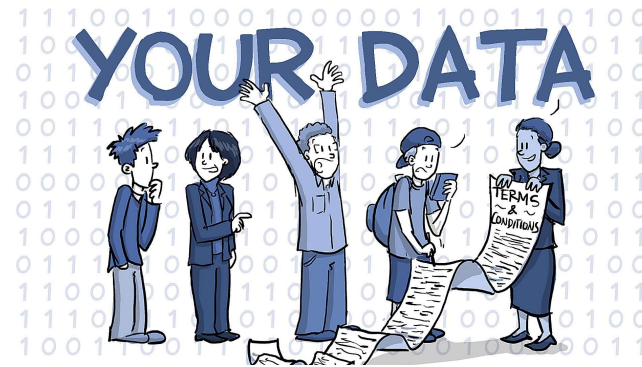
GEORGES GRINSTEIN (KEYNOTE PRESENTATION, VINCI 2016)

- **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

Did you know there are
multiple stages of
visualizing data?

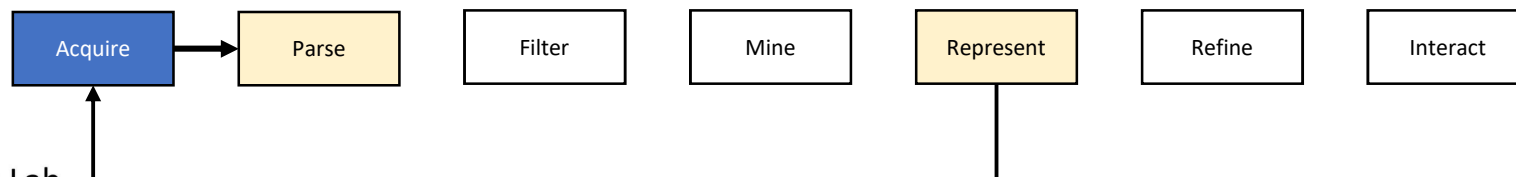
Stage 1: Acquire

- The acquisition step involves obtaining the 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).



Task: acquire data:

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



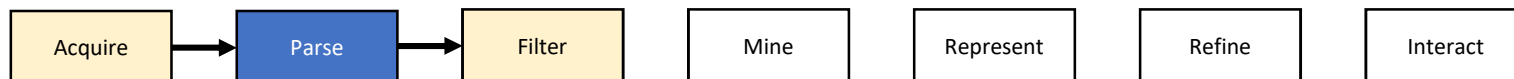
Stage 2: Parse

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

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

Parsed Data

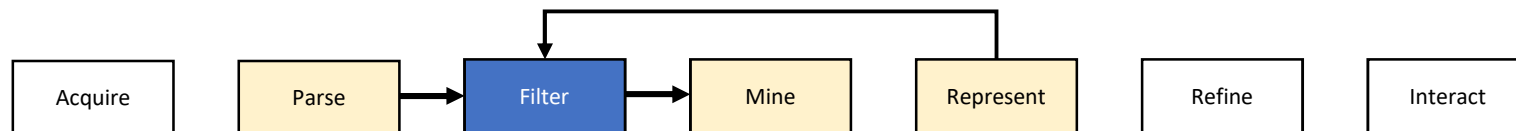
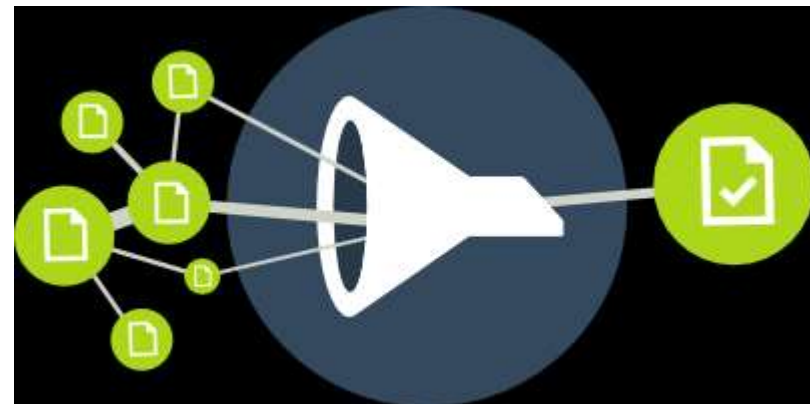
First name	Last name	Status	Prog Exp	Vis Exp
String Length: 10	String Length: 12	Char (2) Fr, So, Jr, Sr	Char (1) Y or N	Char (1) Y or N



- **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).

Stage 3: Filter

- Remove portions not relevant to our use.
- Some projects could require significant mathematical work to place the data into a mathematical model or normalize it (convert it to an acceptable range of numbers).

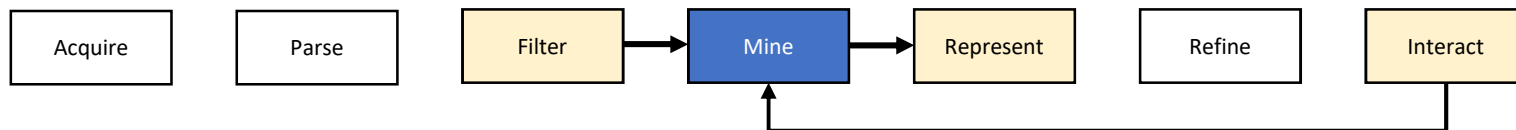


Stage 4: Mine

- 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.

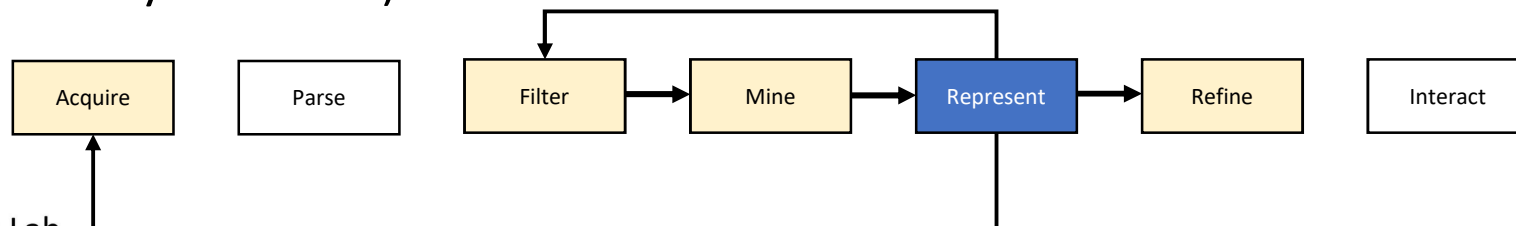
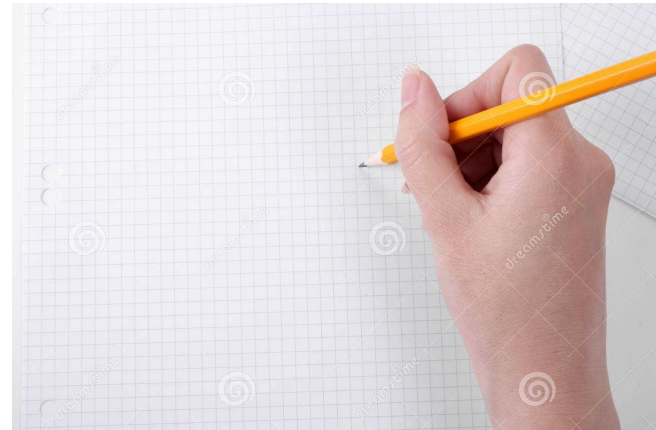
Tasks:

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



Stage 5: Represent

- This step determines the basic form that a set of data will take:
 - List, 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



Stage 6: Refine

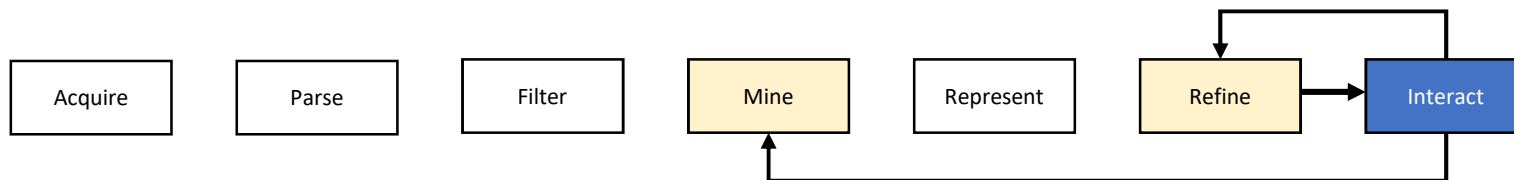
- Graphic design methods are used to further clarify the representation by calling more attention to particular data (establishing hierarchy) or by changing attributes (such as color) that contribute to readability.
- Task: enhance the visualization created in Step 5: Represent



Stage 7: Interact

- Letting the user control or explore the data.
- Interaction might cover things like selecting a subset of the data or changing the viewpoint.
- 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.



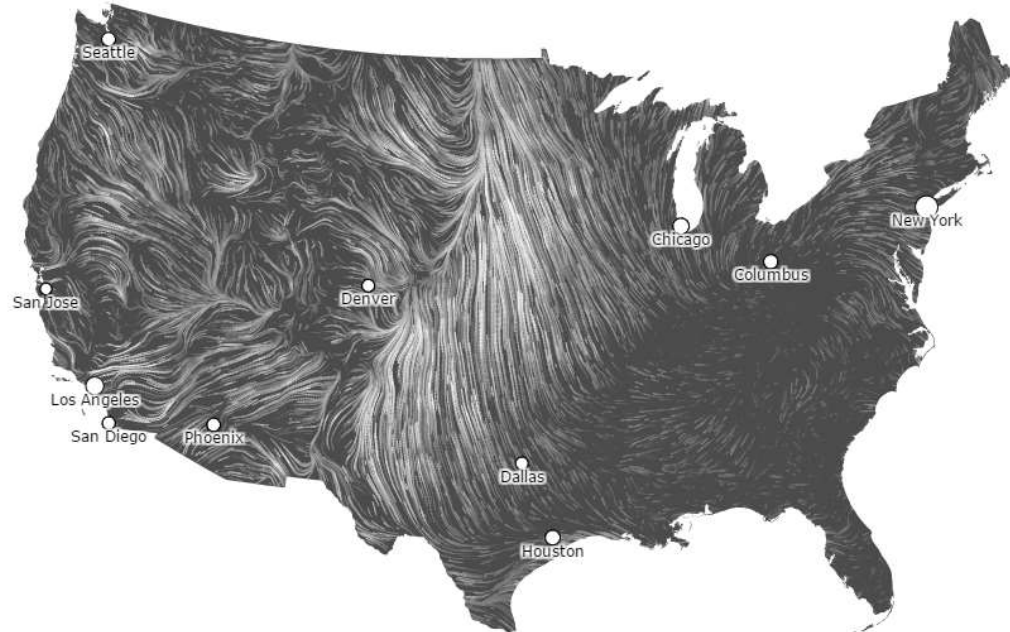
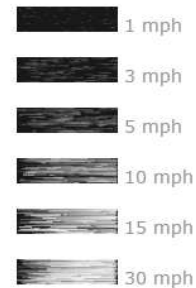
DATA VISUALIZATIONS

- 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/>

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.

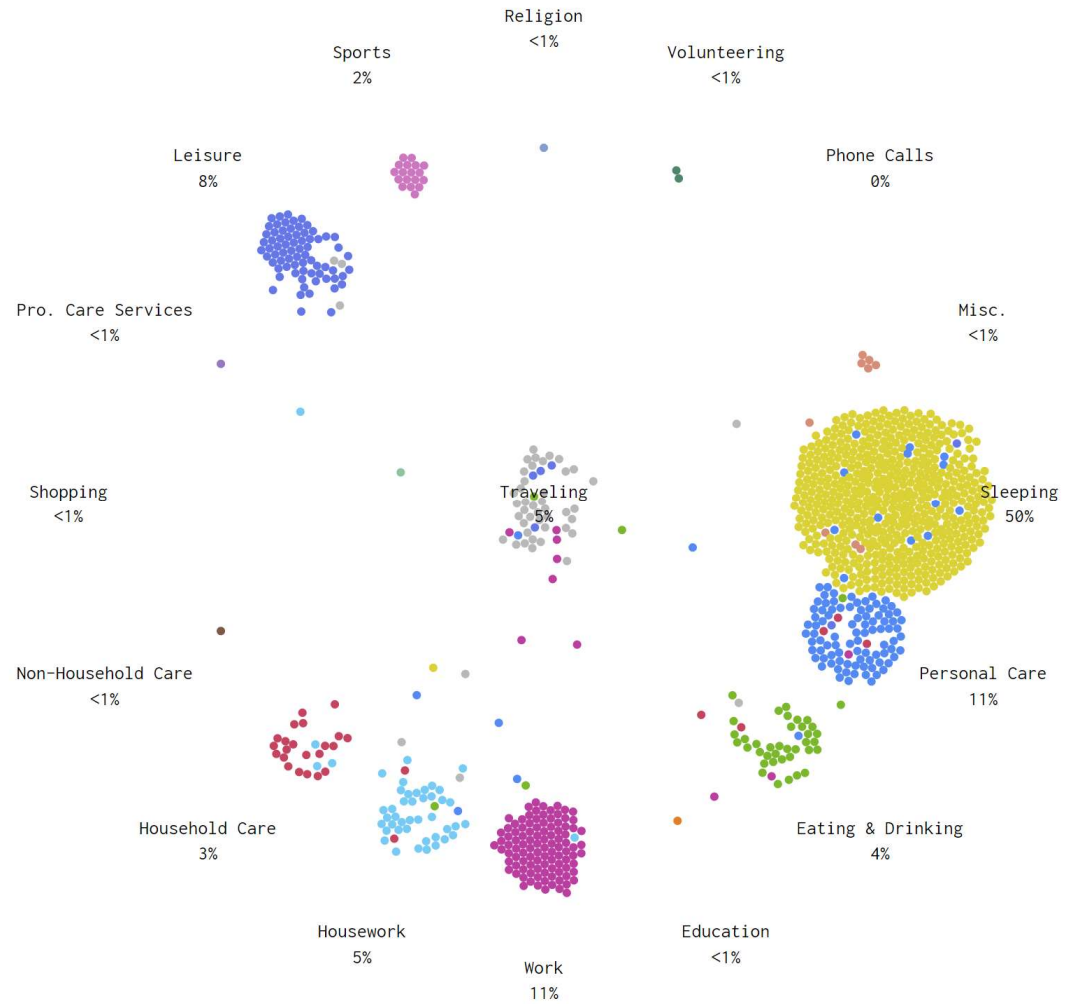
Wind map prints are available from [Point.B Studio](#).

Read more about [wind](#) and about [wind power](#).

A DAY IN THE LIFE OF AMERICANS

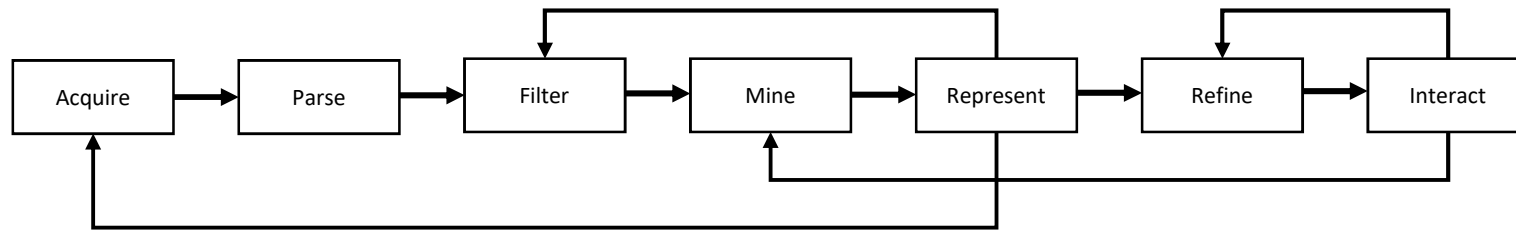
6:41 am

SLOW MEDIUM FAST



This is a simulation of 1,000 people's average day. It's based on 2014 data from the American Time Use Survey, made way more accessible by the ATUS Extract Builder.

7 stages of Visualizing Data



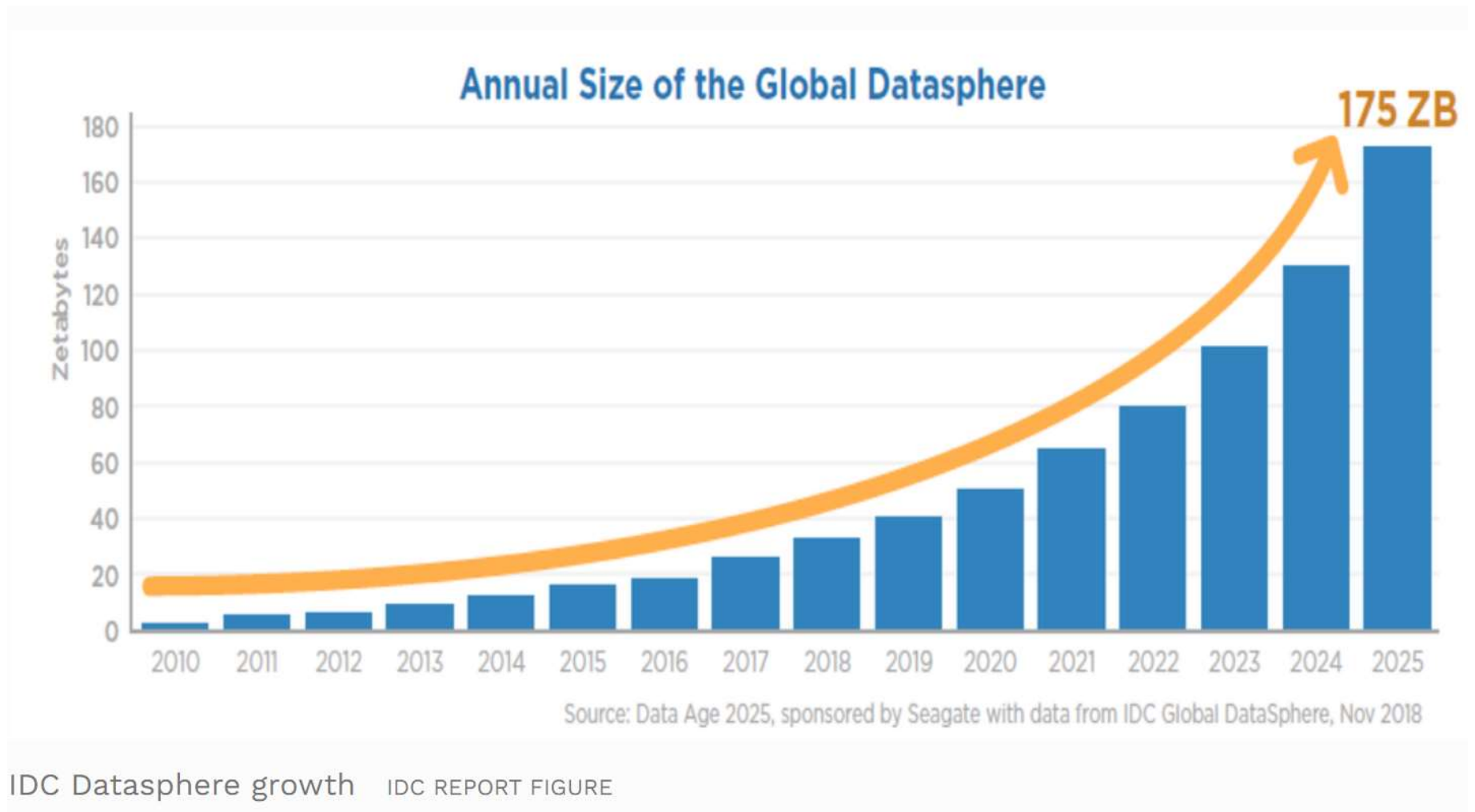
Fry, B. (2008). Chapter 1, Visualizing data (Safari Books Online). Sebastopol, Calif.: O'Reilly Media.

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

Why should you care?

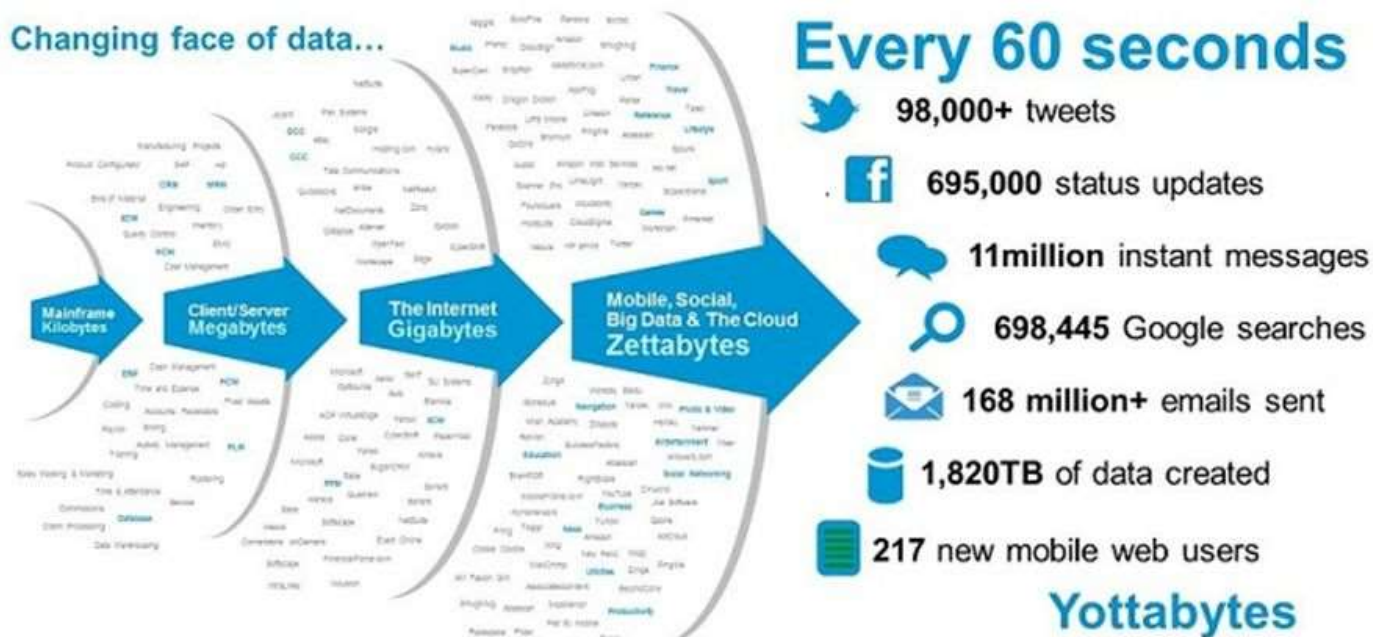
Why Should you Care?



The Ubiquitous Nature of Data

Data is Everywhere!

..... Is the exponential growth and availability of data, both structured and unstructured, because of the Internet & fast growing technology advancements



Data here, Data there, data data everywhere. And no – it's not funny

By [Cheryl Biswas](#) on [LinkedIn](#)

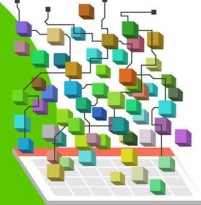


IBM Big Data Platform

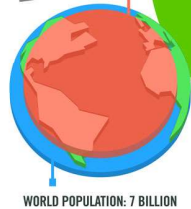
40 ZETTABYTES
[43 TRILLION GIGABYTES]
of data will be created by 2020, an increase of 300 times from 2005



It's estimated that **2.5 QUINTILLION BYTES**
[2.3 TRILLION GIGABYTES]
of data are created each day



6 BILLION PEOPLE
have cell phones



The FOUR V's of Big Data

From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day.

As of 2011, the global size of data in healthcare was estimated to be

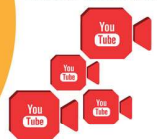
150 EXABYTES
[161 BILLION GIGABYTES]



By 2014, it's anticipated there will be **420 MILLION WEARABLE, WIRELESS HEALTH MONITORS**

Variety
DIFFERENT FORMS OF DATA

4 BILLION+ HOURS OF VIDEO are watched on YouTube each month



400 MILLION TWEETS are sent per day by about 200 million monthly active users



The New York Stock Exchange captures **1 TB OF TRADE INFORMATION** during each trading session



100 SENSORS that monitor items such as fuel level and tire pressure

Velocity
ANALYSIS OF STREAMING DATA

By 2016, it is projected there will be **18.9 BILLION NETWORK CONNECTIONS** — almost 2.5 connections per person on earth



Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.

By 2015 **4.4 MILLION IT JOBS** will be created globally to support big data, with 1.9 million in the United States



1 IN 3 BUSINESS LEADERS don't trust the information they use to make decisions



in one survey were unsure of how much of their data was inaccurate

Veracity
UNCERTAINTY OF DATA

Poor data quality costs the US economy around **\$3.1 TRILLION A YEAR**



Sources: McKinsey Global Institute, Twitter, Cisco, Gartner, EMC, SAS, IBM, MEPECT, QAS

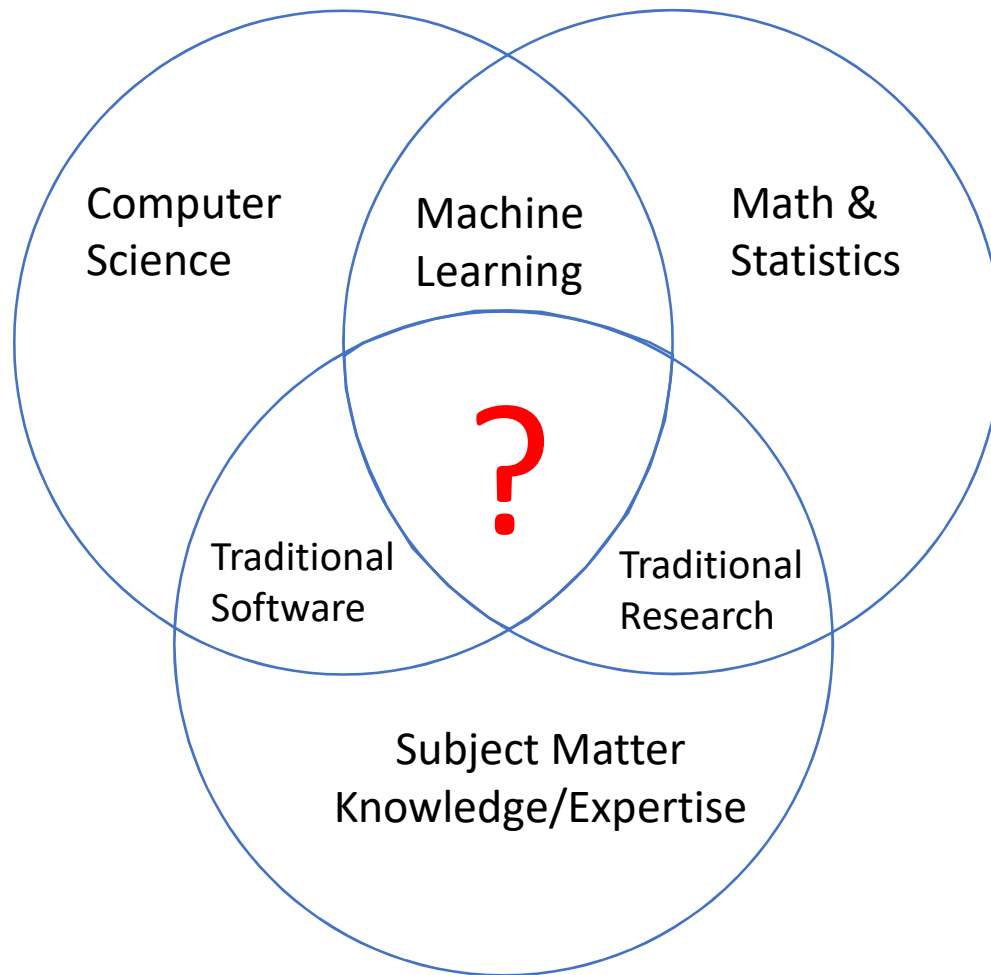


Data is Everywhere!

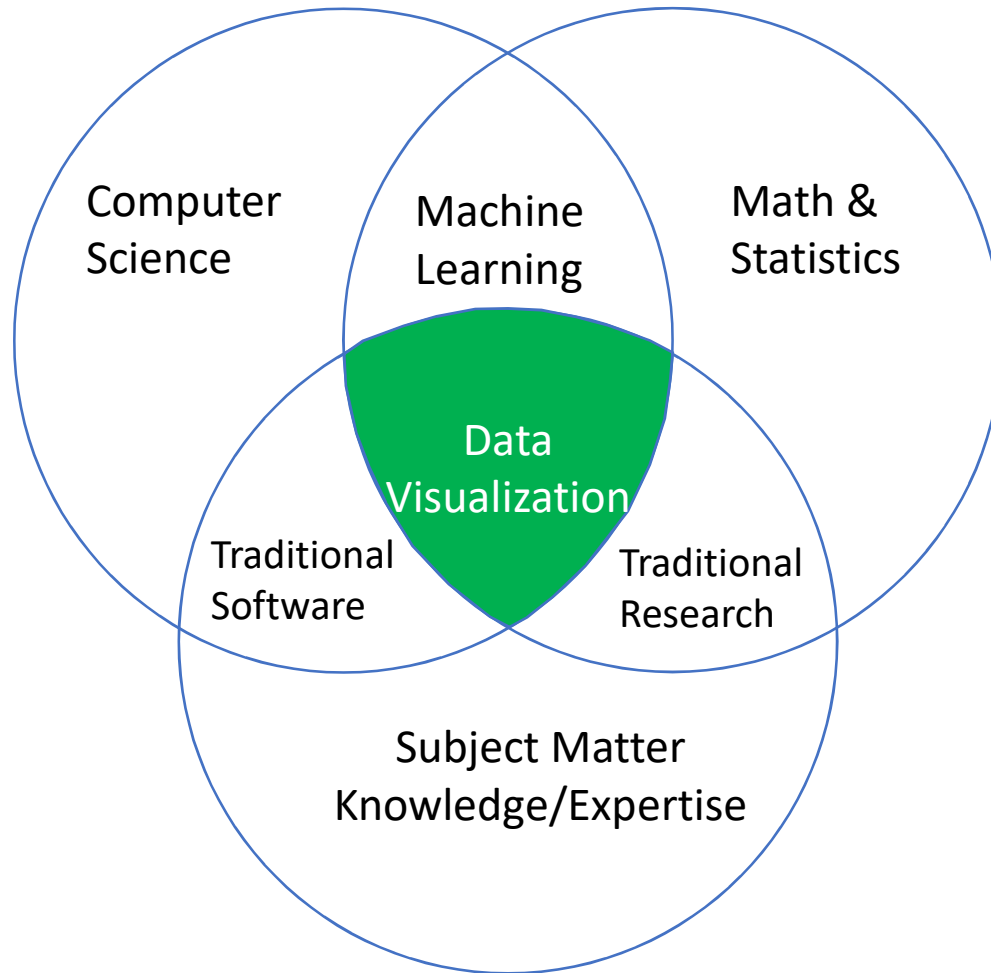
Image Source: <http://www.centrodeinnovacionbbva.com/en/news/practical-examples-big-data-use>



Data Science



Data Science



Visualization Applications

Bio visualization (BioVis)

- The visualization of biological data;
- Often grouped with computer animation



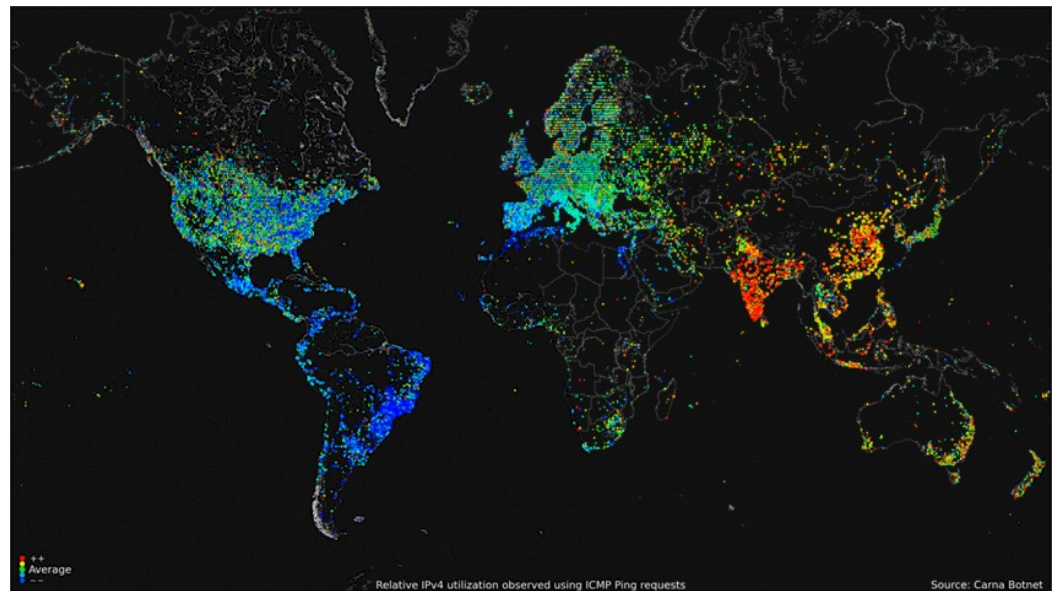
Visualization Applications

Information Visualization (InfoVis)

- Interdisciplinary
- Study of the “visual representation of large-scale collections of non-numerical information



InfoVis



Internet Usage

Source: <http://www.cerne.net/wp-content/uploads/2013/03/internet.gif>

Visualization Applications

Geographic Visualization

GeoVis

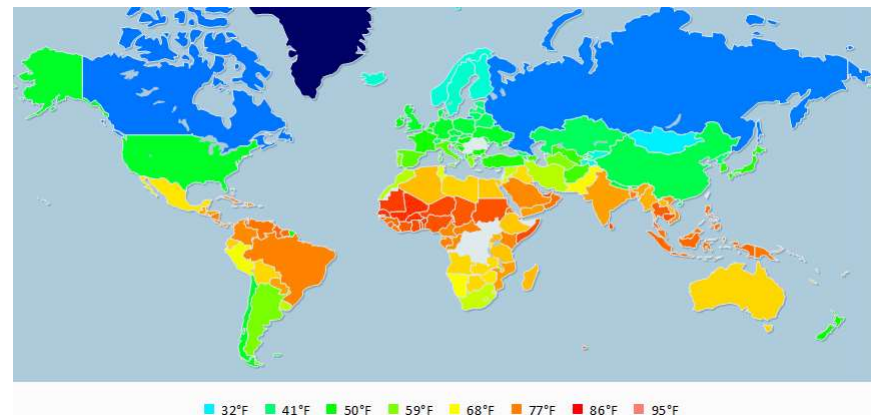
Communicates geospatial information in ways that, when combined with human understanding, allow for data exploration and decision-making processes.



MacEachren, A.M. and Kraak, M.J. 1997 Exploratory cartographic visualization: advancing the agenda. *Computers & Geosciences*, 23(4), pp. 335-343.

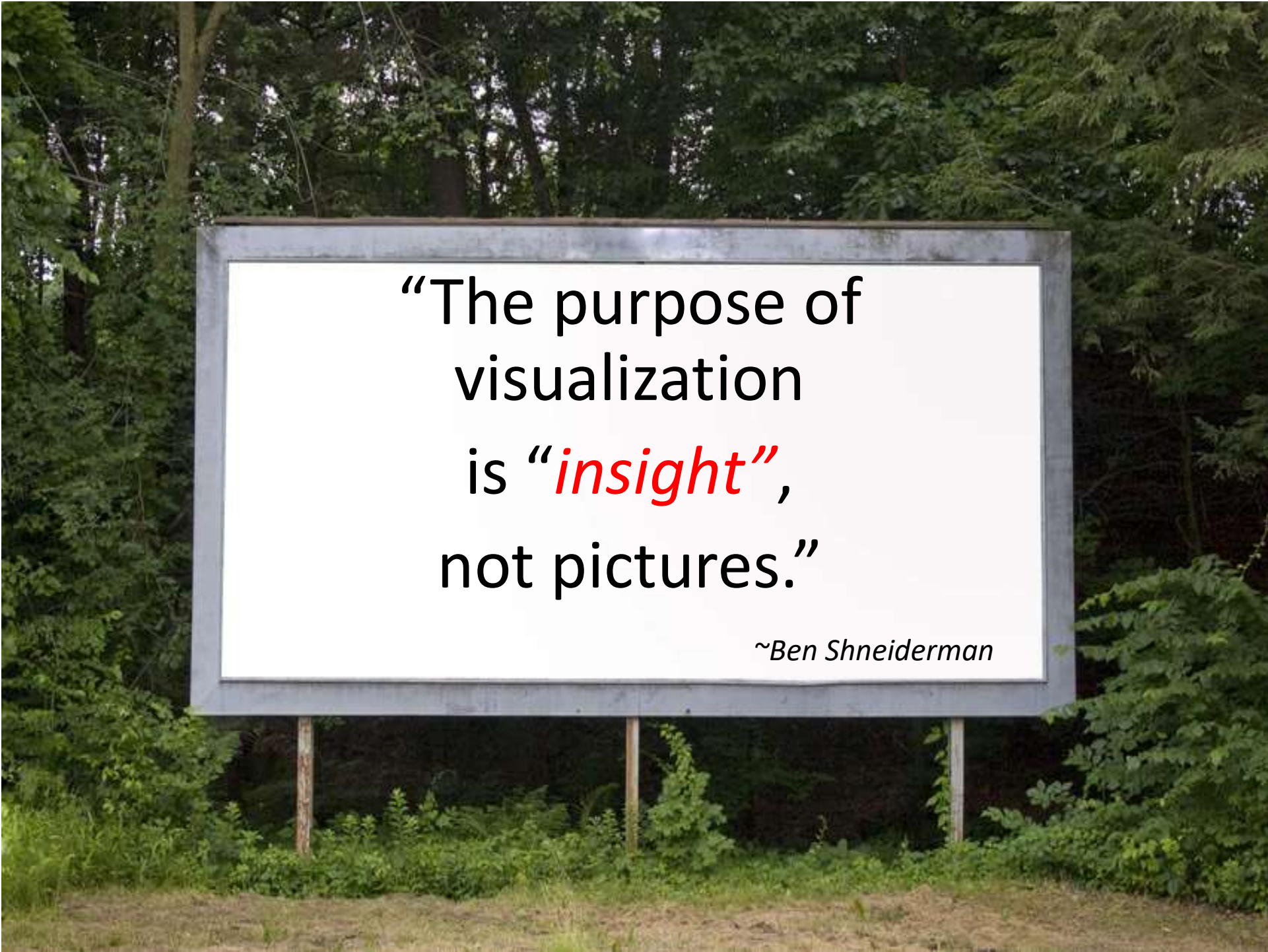
Jiang, B., and Li, Z. 2005. Editorial: Geovisualization: Design, Enhanced Visual Tools and Applications. *The Cartographic Journal*, 42(1), pp. 3-4

[MacEachren, A.M.](#) 2004. Geovisualization for knowledge construction and decision support. *IEEE computer graphics and applications*, 24(1), pp.13-17



The Quest for Insight

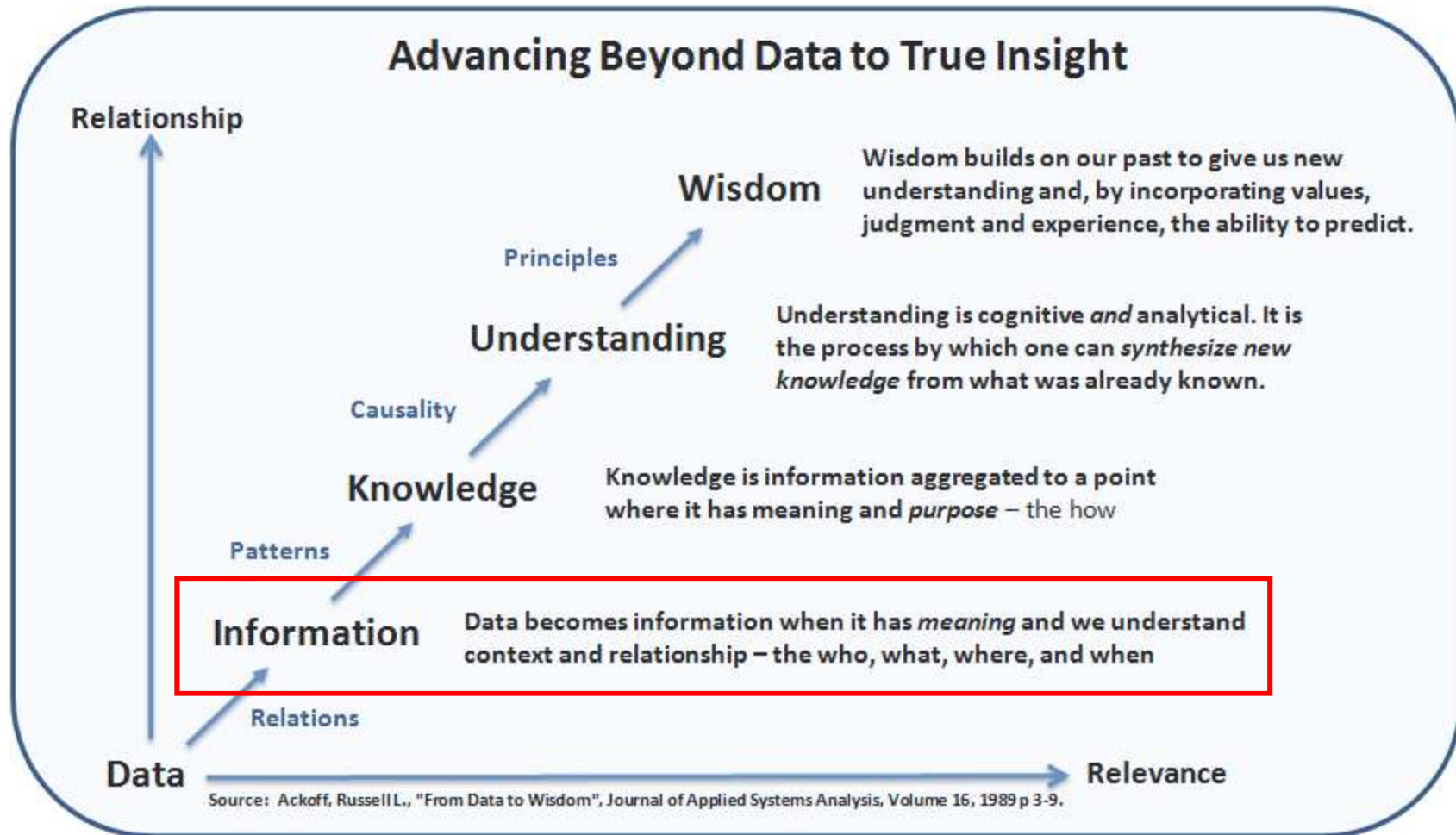
Remember ...

A rectangular sign with a white background and a grey border is mounted on three wooden posts. The sign is set against a backdrop of dense green trees and foliage. The text on the sign is centered and reads: "The purpose of visualization is *insight*, not pictures."

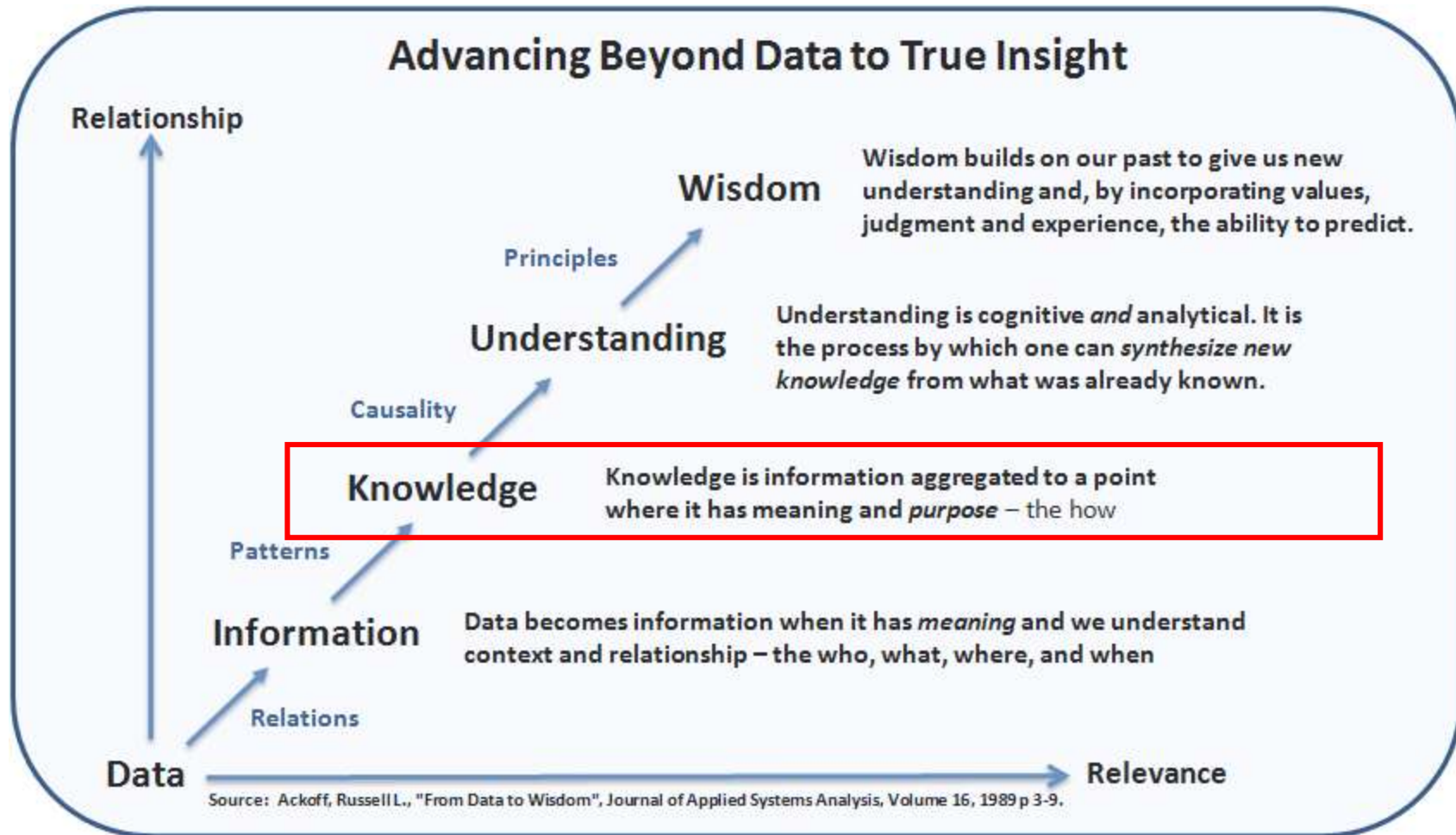
“The purpose of
visualization
is “*insight*”,
not pictures.”

~Ben Shneiderman

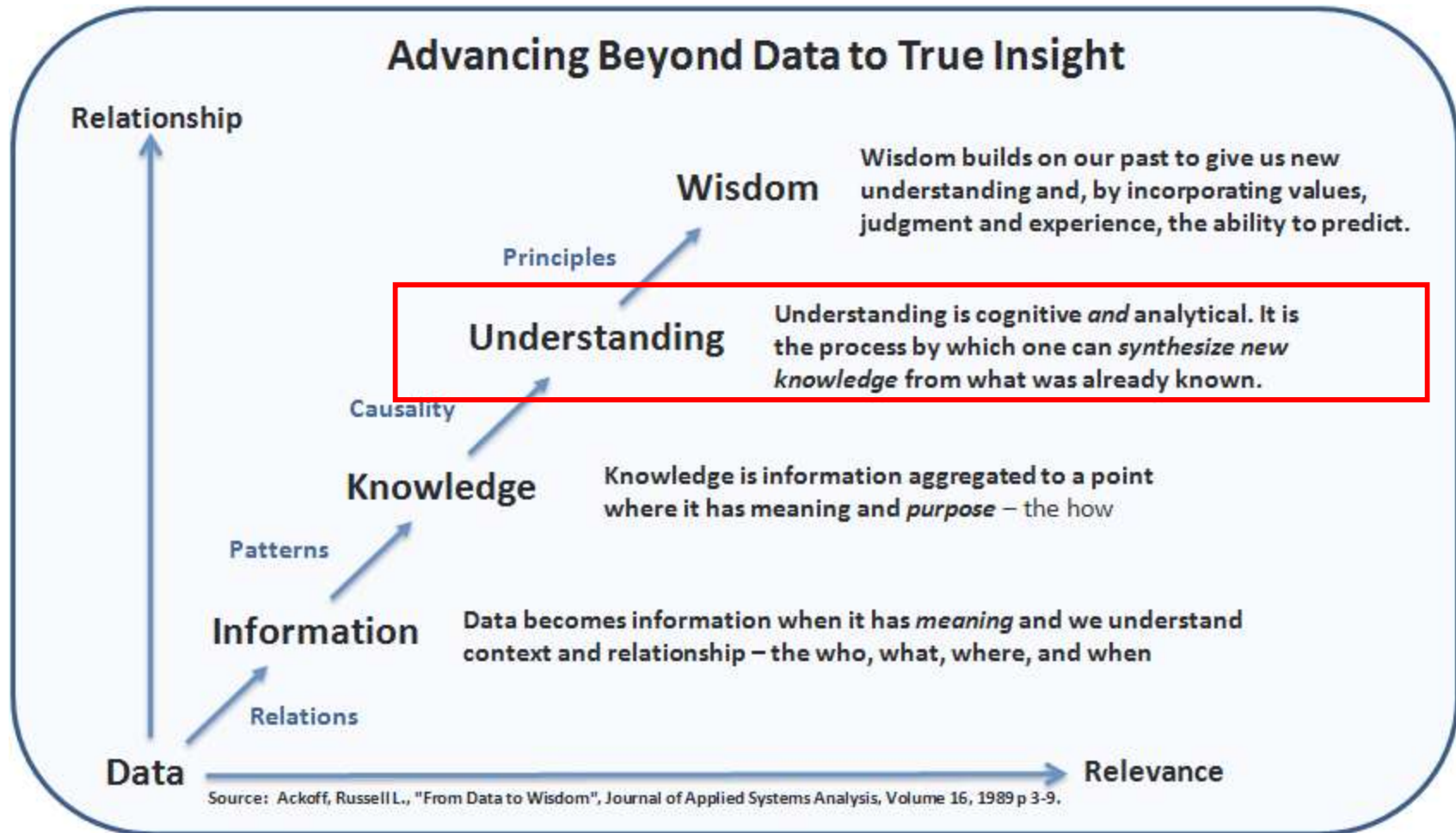
FROM DATA TO INSIGHT



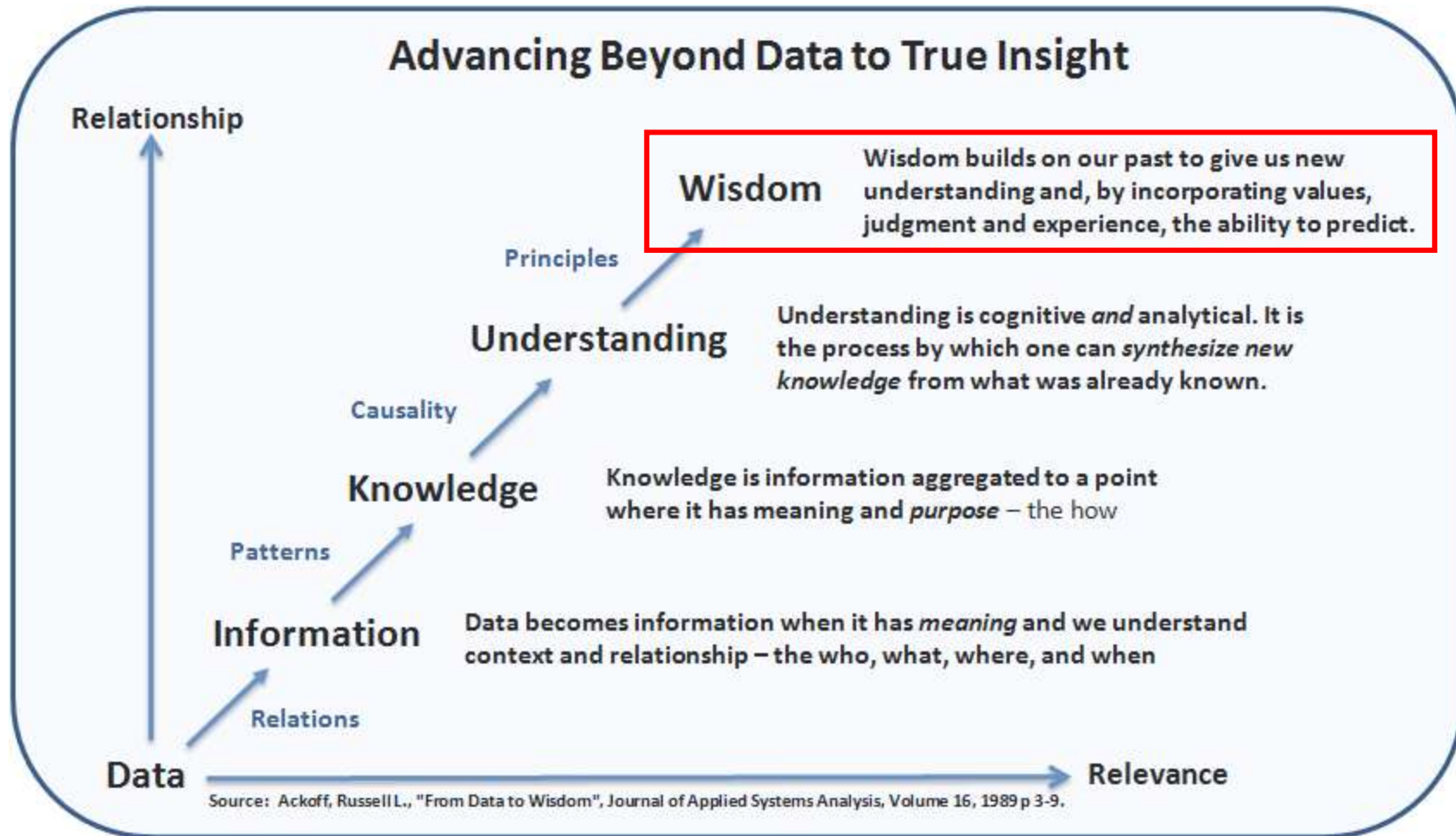
FROM DATA TO INSIGHT



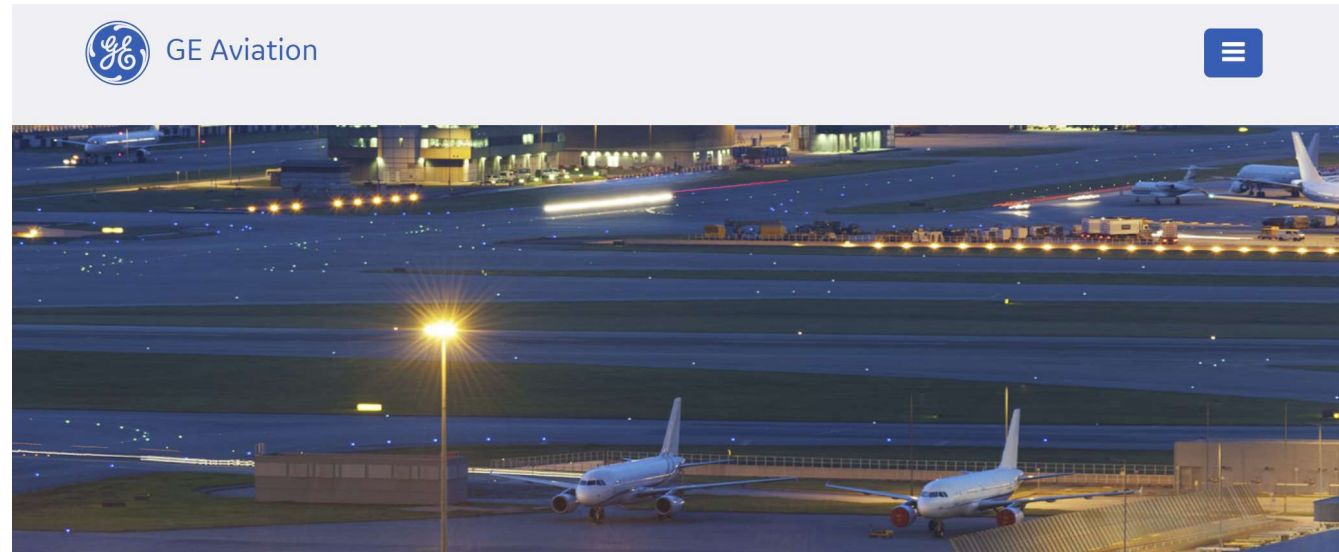
FROM DATA TO INSIGHT



FROM DATA TO INSIGHT



You are only limited by your imagination



SERVICES

Data Visualization

Turning customer data into informed business decisions

Problem

Data can be a powerful business resource if leveraged properly but data alone is not enough and often discerning insights is a



Data Exploration and Analysis

Hands-on Exercise



Data Exploration and Analysis

- The purpose of this hands-on exercise is to provide you with some experience exploring and analyzing data without using an information visualization system.
- The data set and the exercise description is provided on the webinar webpage (see Day 1)

What's next?

- Explore the Webinar Website
- Check out the Resources Page
- Post Questions if you have them (see bottom of Day 1 webpage for Q&A link).

Tomorrow – Day 2

- Topic: You've got data now what?
 - Share visualizations of your insights gained from exploring the cereal data
- Tools and Applications
 - Tableau (14-day Trial)
<https://www.tableau.com/products/desktop/download>