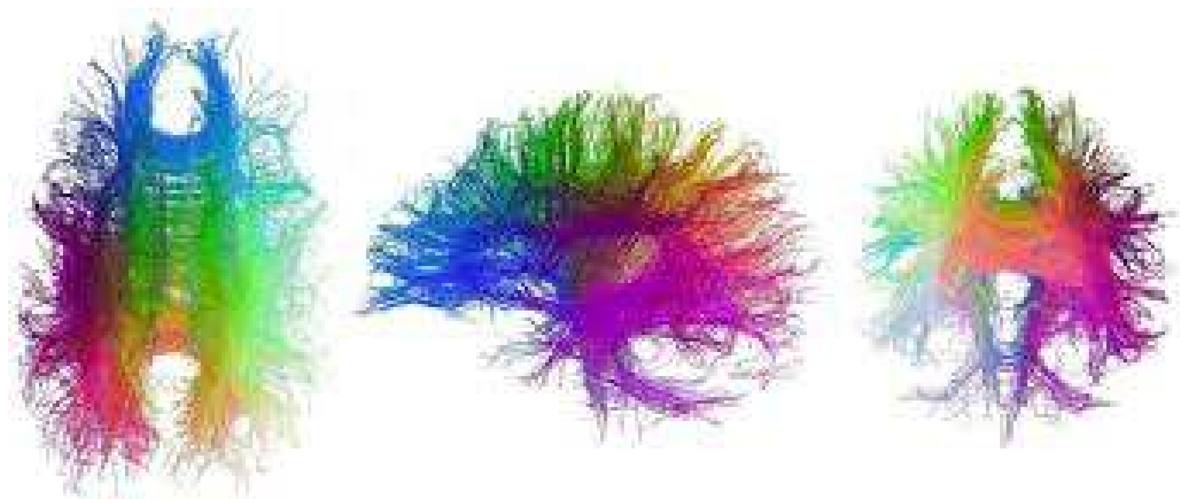
CS 519: Scientific Visualization

Introduction

Eric Shaffer

Visualization

- Definition (OED)
  1. …the power or process of forming a mental picture or vision of something not actually present to the sight
- The purpose of visualization is to convey information to people
- The process converts information into a graphical representation



Why do it?

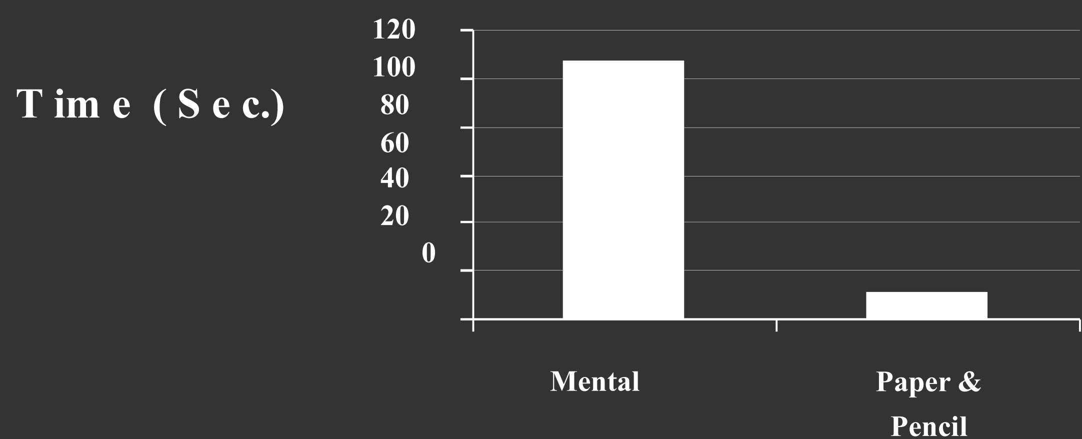
- To answer a question
  - Part of a procedure for solving a problem
- Support analysis and reasoning
  - To explore and discover;
    "The purpose of computing is insight, not numbers" [R. Hamming]
- Communicate information to others
  - Make a point
  - Tell a story
- Inspire
  - Part of our cultural heritage

Courtesy of Pat Hanrahan of Stanford

Courtesy of Pat Hanrahan of Stanford

Answering a Question: Long-hand Multiplication

$$\begin{array}{r} 34 \\ \times\ 72 \\ \hline 68 \\ 2380 \\ \hline 2448 \end{array}$$



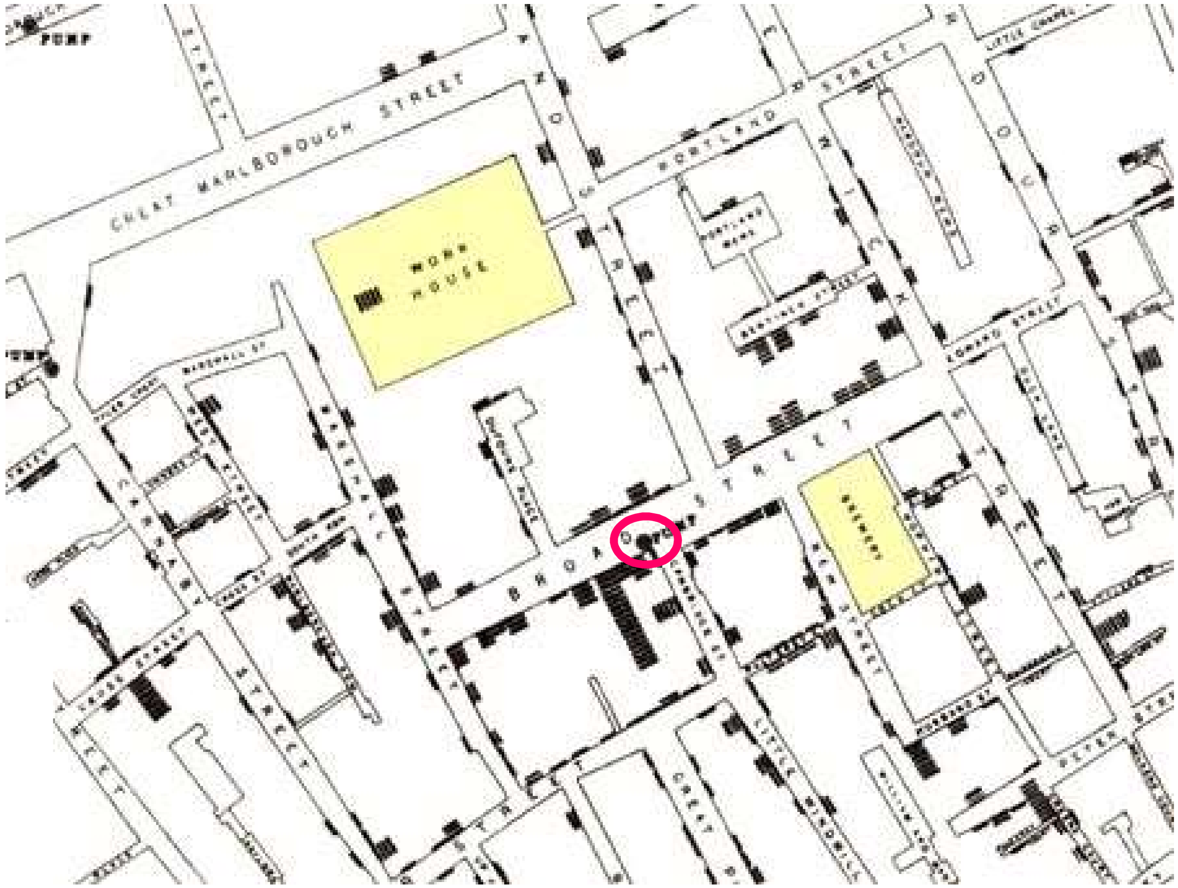
Visualization to Discover

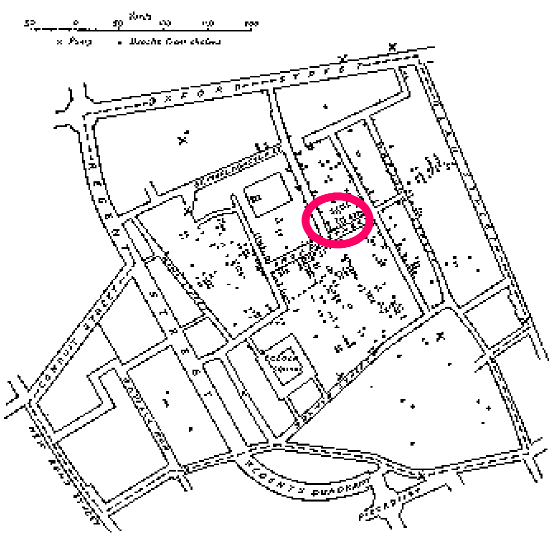
Mystery: what is causing a cholera epidemic in London in 1854?

Visualization Success Story

Illustration of John Snow's deduction that a cholera epidemic was caused by a bad water pump, circa 1854.

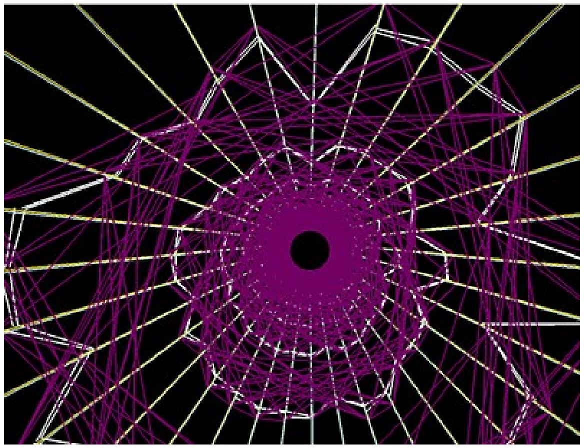
Dots indicate location of deaths.





From Visual Explanations by Edward Tufte, Graphics Press, 1997

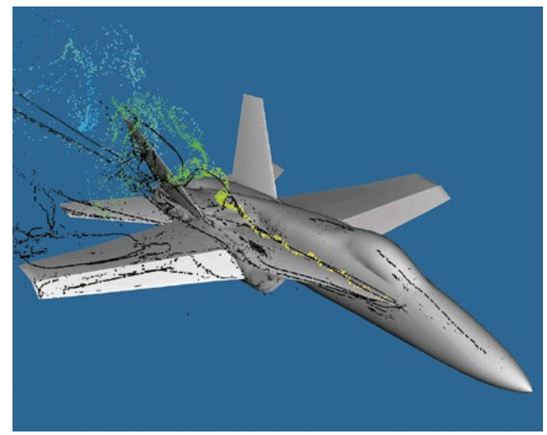
Tell a Story



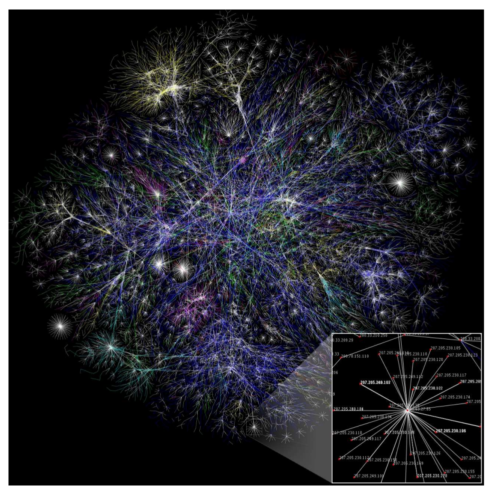
- The Time Tunnel (created by Dr. Dan Reed in the 1990s)
  - shows communication activity in a multi-processor system
  - Temporal visualization…like a 3D time series
- Anecdotally…convinced software enigneers their application was communication bound
- Those same engineers refused to believe the numbers, which told the same story in a different way

Visualization Flavors

- Scientific Visualization
  - Usually involves data associated with a physical domain
  - Example: fluid visualization



- Information Visualization
  - Covers abstract data
  - Example: Connectivity

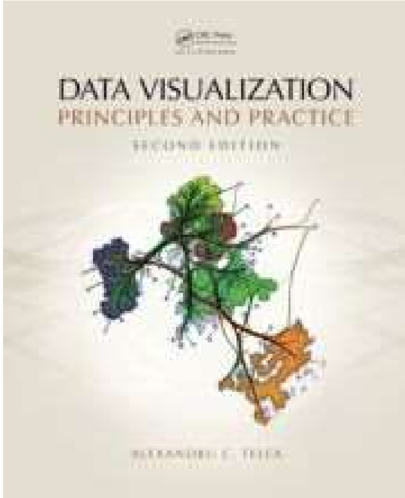


Our Topics

- Data representation
- Scientific Data Visualization
  - Scalar visualization
  - Vector visualization
  - Tensor visualization
  - Volume visualization
- Information visualization
  - Tabular data
  - Graphs
  - Time Series Visualization

Class Mechanics

- Course Website: https://courses.engr.illinois.edu/cs519/index.html
  - Schedule, lecture materials, assignments
- Piazza:
  This term we will be using Piazza for class discussions https://piazza.com/illinois/fall2017/cs519/home.



- Book: Data Visualization: Principles and Practice, Second Edition by Alexandru C. Telea

Class Mechanics: Grades

Machine Problem 1	10%
Machine Problem 2	10%
Machine Problem 3	10%
Machine Problem 4	10%
Course Project	20%
Exam 1	20%
Exam 2	20%

No Final Exam

Grading Scale

- Grades probably on usual scale:
  - 97 to 93: A
  - 93 to 90: A-
  - 90 to 87: B+
  - 87 to 83: B
  - 83 to 80: B-
  - …etc.
- I may adjust the intervals down…but not up

Course Policies

- MPs submitted after the due date lose 10% per day
- Collaboration on MPs is fine…don't copy code verbatim
- In exceptional circumstances where extension may be reasonable (illness, family emergency etc.) arrangement must be made with the instructor (e-mail: shaffer1@illinois.edu)
- Exams are in-class…
- Post technical questions to Piazza

Programming Language and Tools

- HTML
- JavaScript
- WebGL
- d3.js
- Chrome as default browser
- Chrome DevTools to debug code
- ParaView (sci-vis application)
- **If you have a laptop, bring it to class**

Class Mechanics: Projects

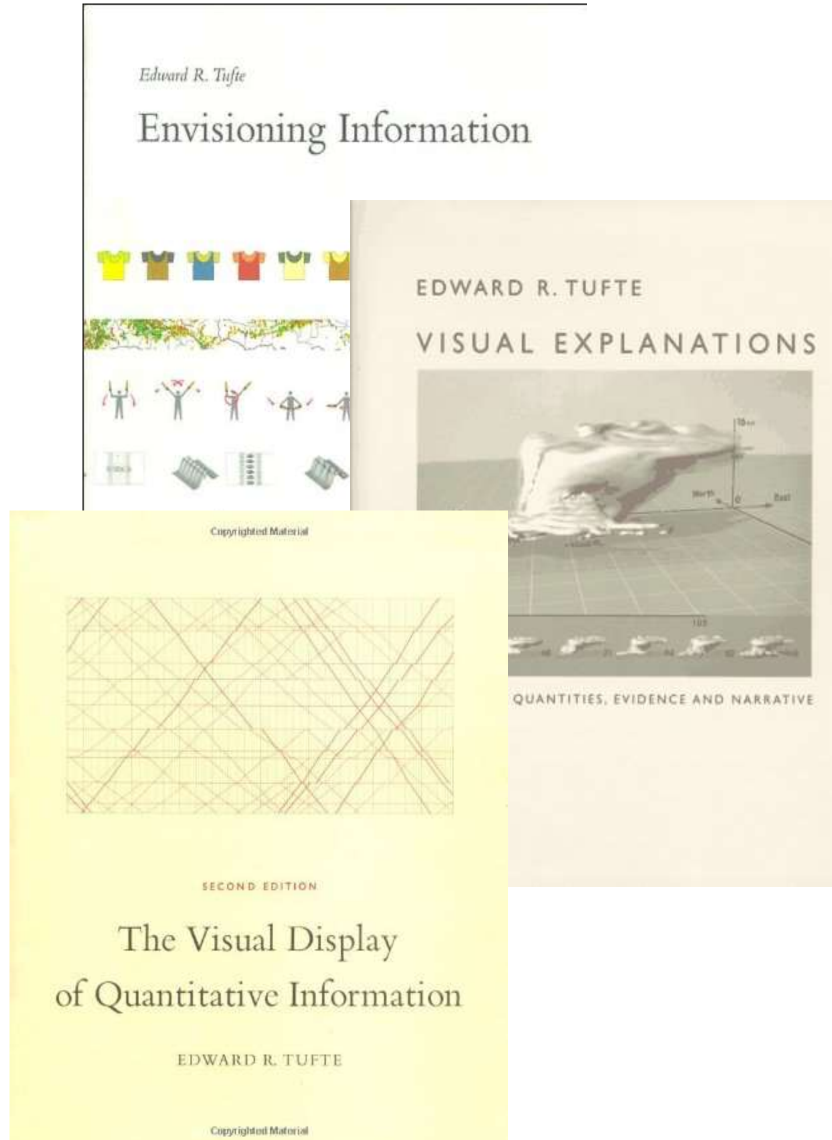
- Work in teams
- You propose a topic (watch for Piazza post about this)
- Can either
  - Implement a non-trivial algorithm or data structure
  - Create a visualization using existing software tools
- Due on last day of class
- 2 page tech report
- Possible class presentation (10 minutes)

Course Goals

- An understanding of which visualization algorithms are commonly applied to a given data type
- Gain familiarity with current visualization software and tools
- Gain familiairity with state-of-the-art in visualization research
- Opportunity to create publication-worthy visualization

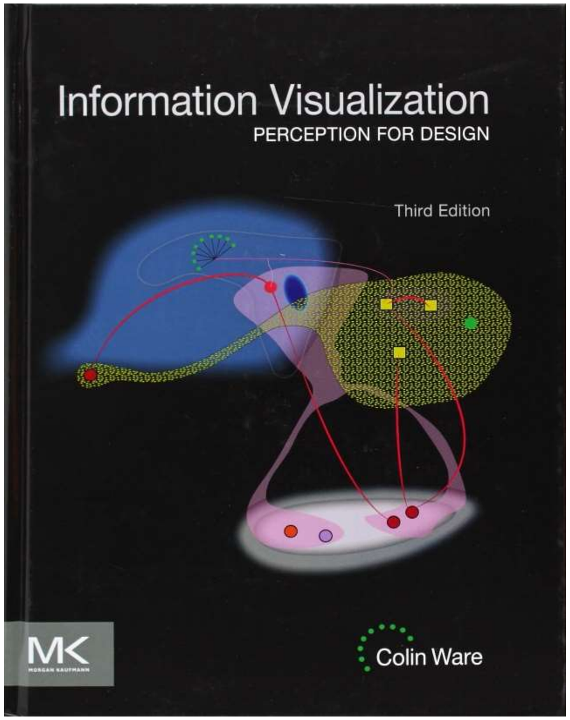
Books

- The Visual Display of Quantitative Information
- Visual Explanations
- Envisioning Information
- Edward Tufte, Yale
- Principles of graphic design applied to data visualization
- Primarily static, printed formats



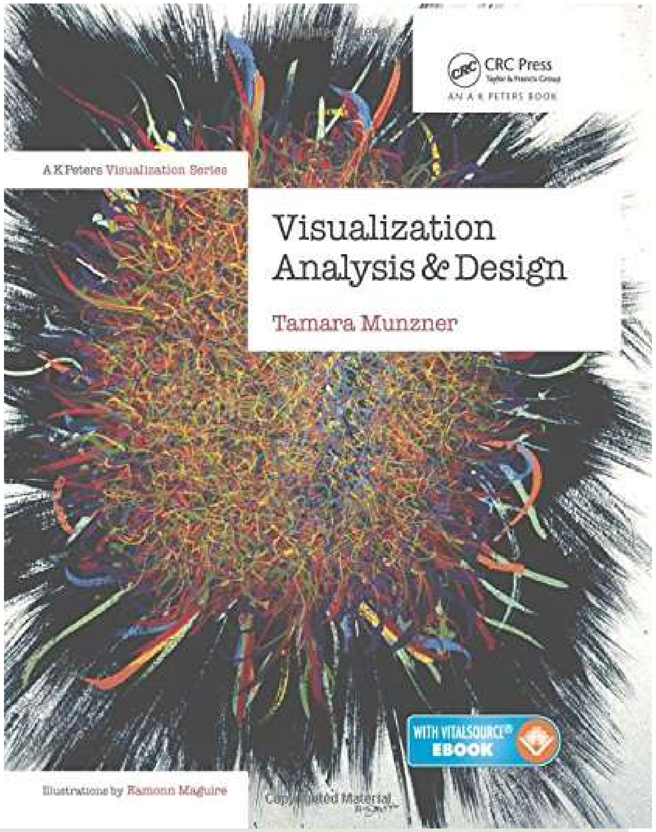
Books

- Information Visualization:
  Perception for Design
- Colin Ware, U. New Hampshire
- Perceptual principles applied to data visualization
- Focus on user interaction



Books

- Visualization Analysis & Design
- Tamara Munzner, U. British Columbia
- Covers both scientific and information visualization



Computer Graphics and Visualization

- “The purpose of computing is insight, not numbers”
  – Richard Hamming
- One could say the purpose of visualization is insight,
  not computer graphics
- Some knowledge of how a system renders an image is important
  - Can anyone suggest why?
- We will briefly discuss some graphics terminology and methods

2D Graphics: Vector Graphics and Raster Graphics

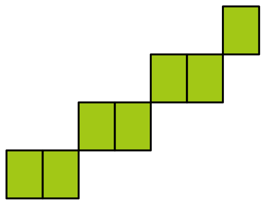
- Images are made of shapes and colors.
- Two popular ways of encoding those shapes and colors

Vector Graphics

- Use an geometric/algebraic description of shape
- PostScript, PDF, SVG
- Low memory (display list)
- Easy to specify a line
- Adapts to any display resolution

Raster Graphics

- Explicity specifies colors of a set of pixels
- GIF, JPG, etc.
- High memory (frame buffer)
- Hard to draw line
- Fixed resolution



Definitions: Pixel and Raster

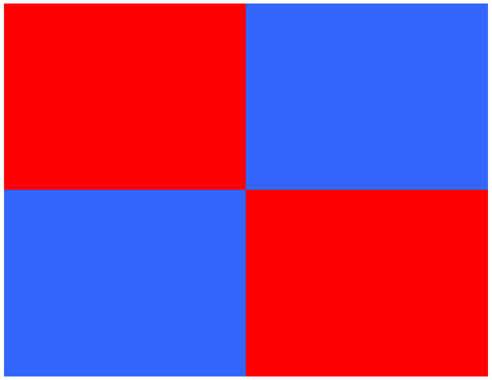
A ***pixel*** is the smallest controllable picture element in an image

A ***raster*** is a grid of pixel values

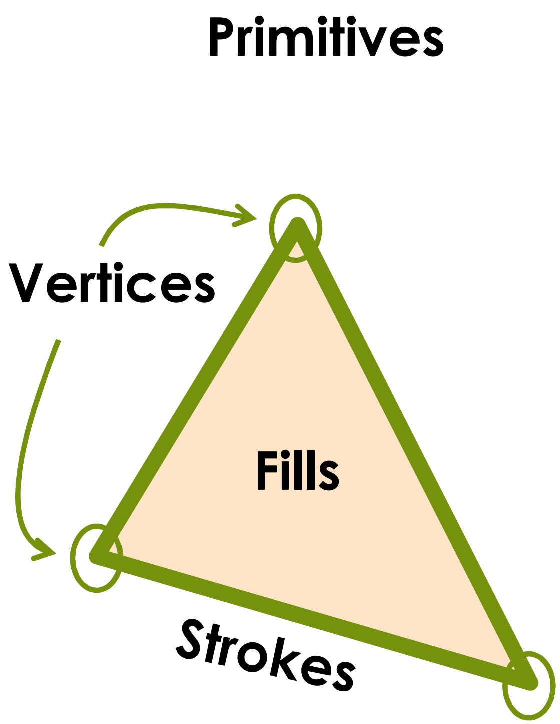
Typically rectangular grid of color values

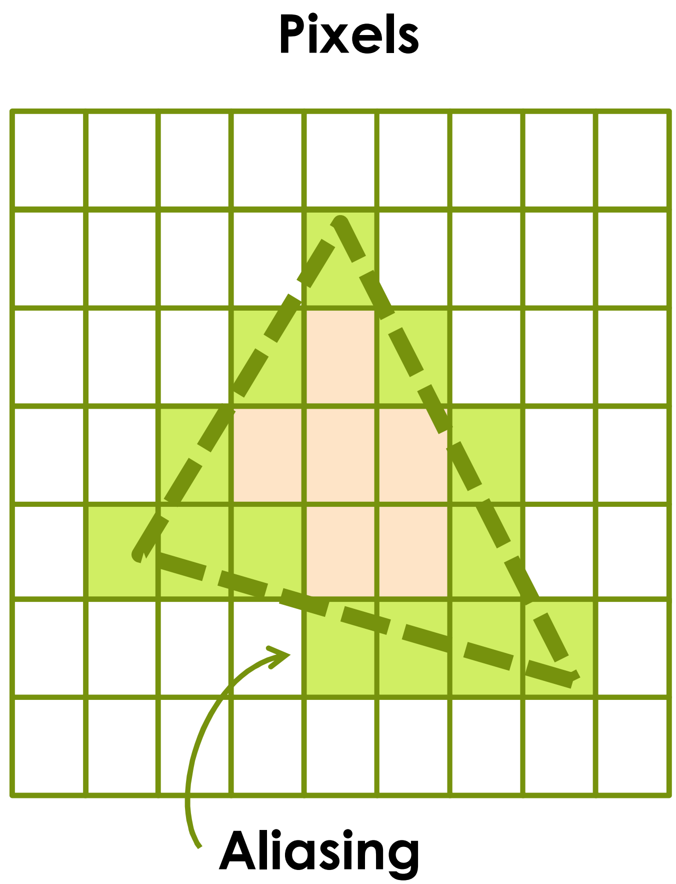
(255,0,0), (0,0,255)

(0,0,255), (255,0,0)

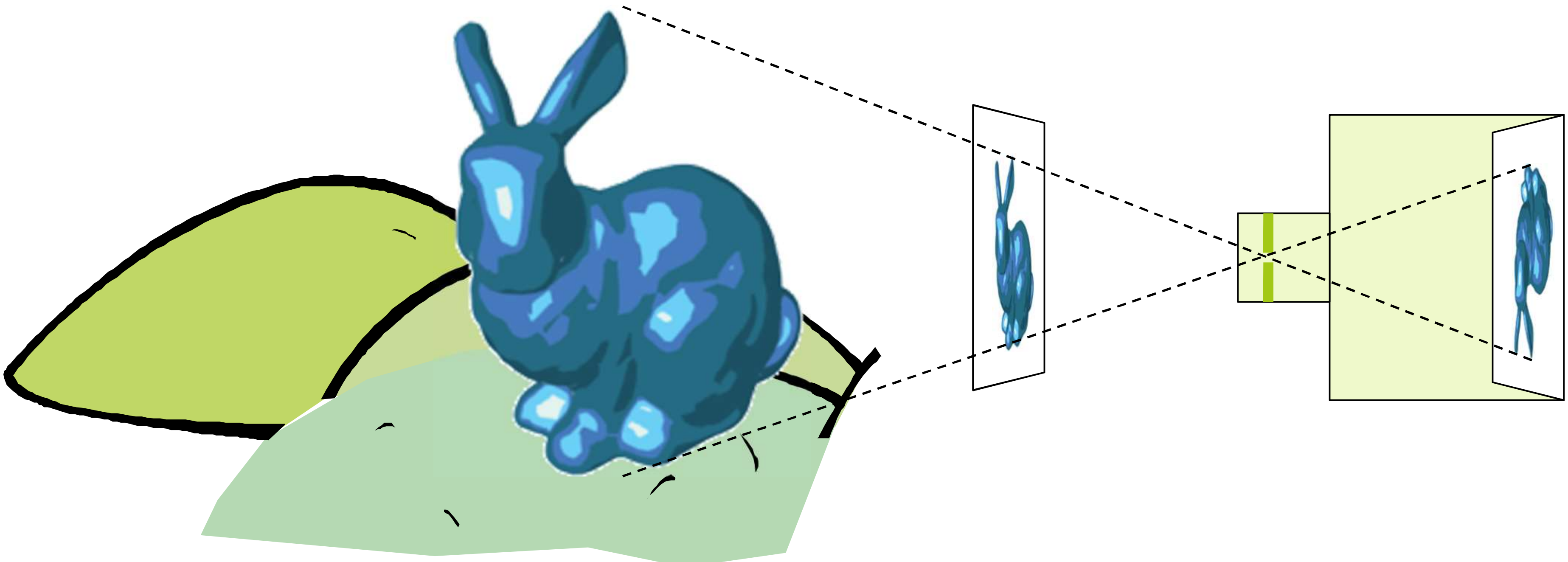


Rasterization

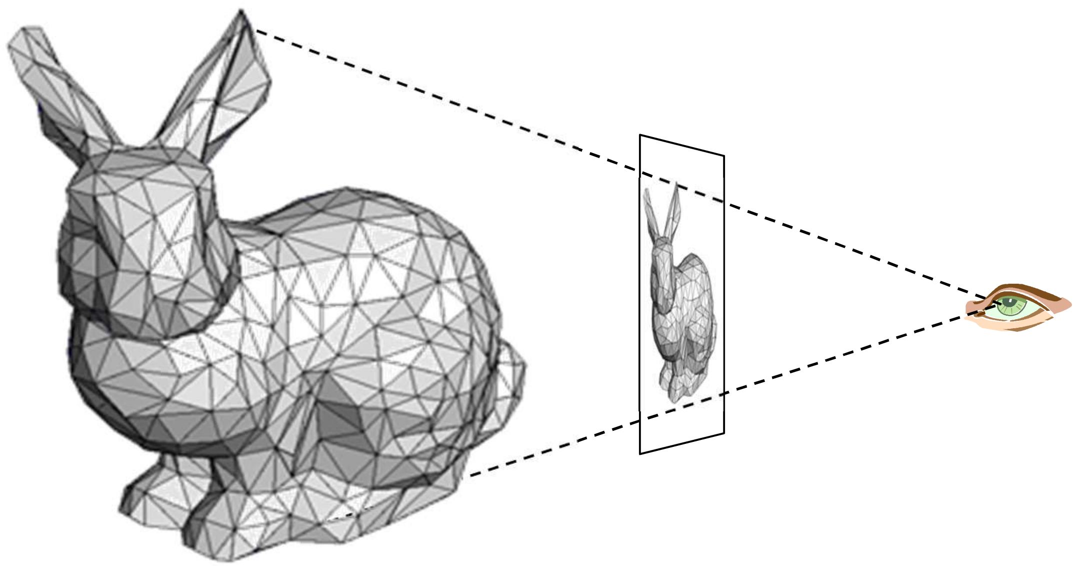




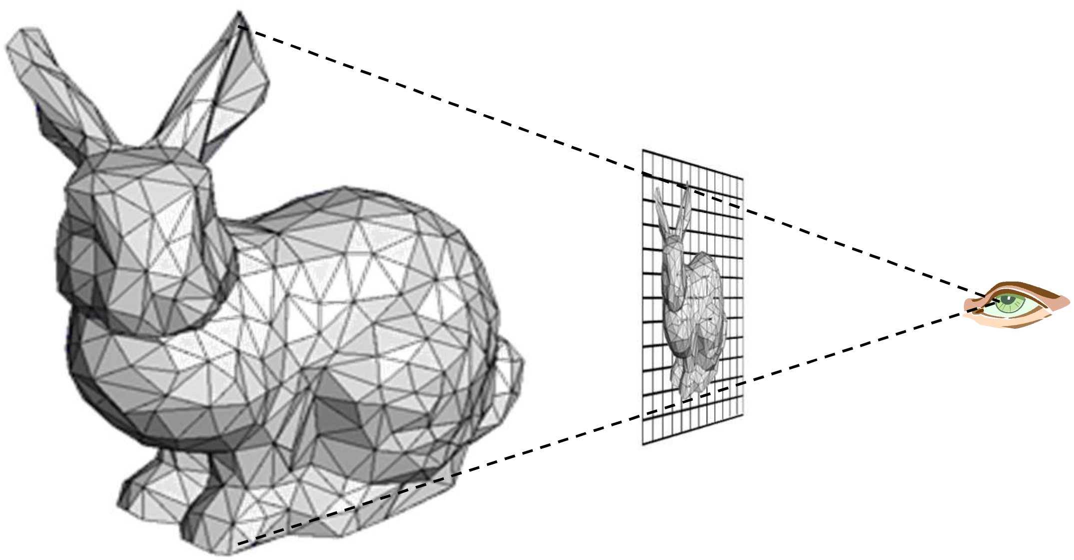
Synthetic Camera Model

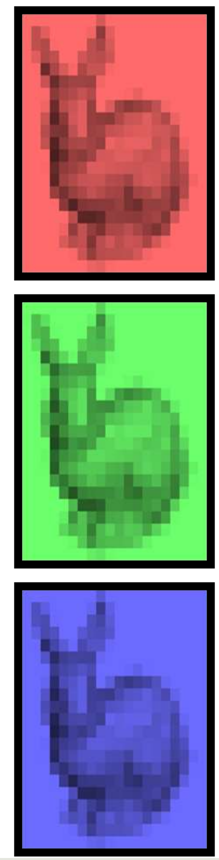


Polygonal Models



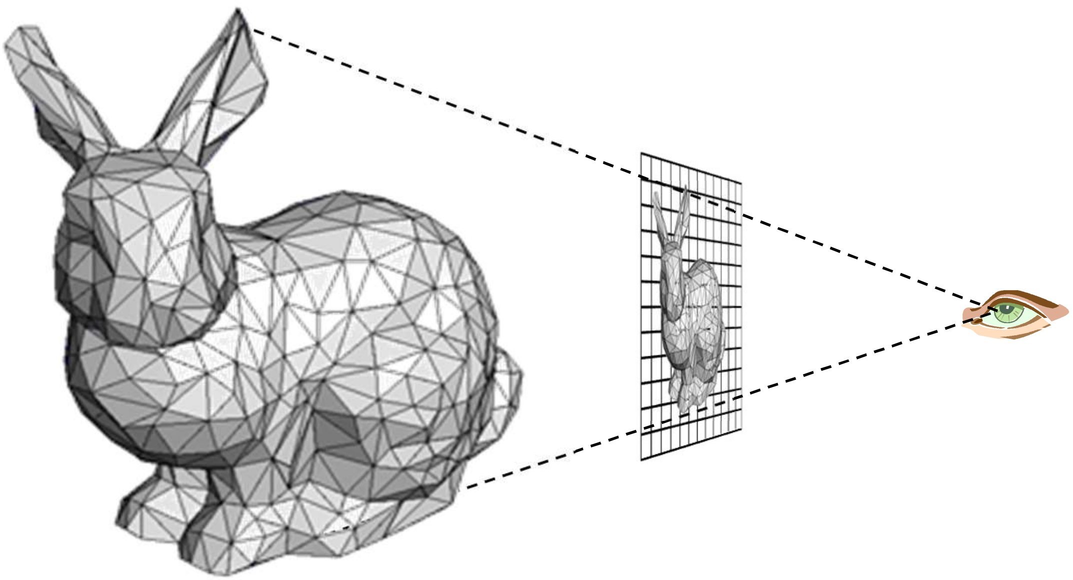
Pixel Discretization



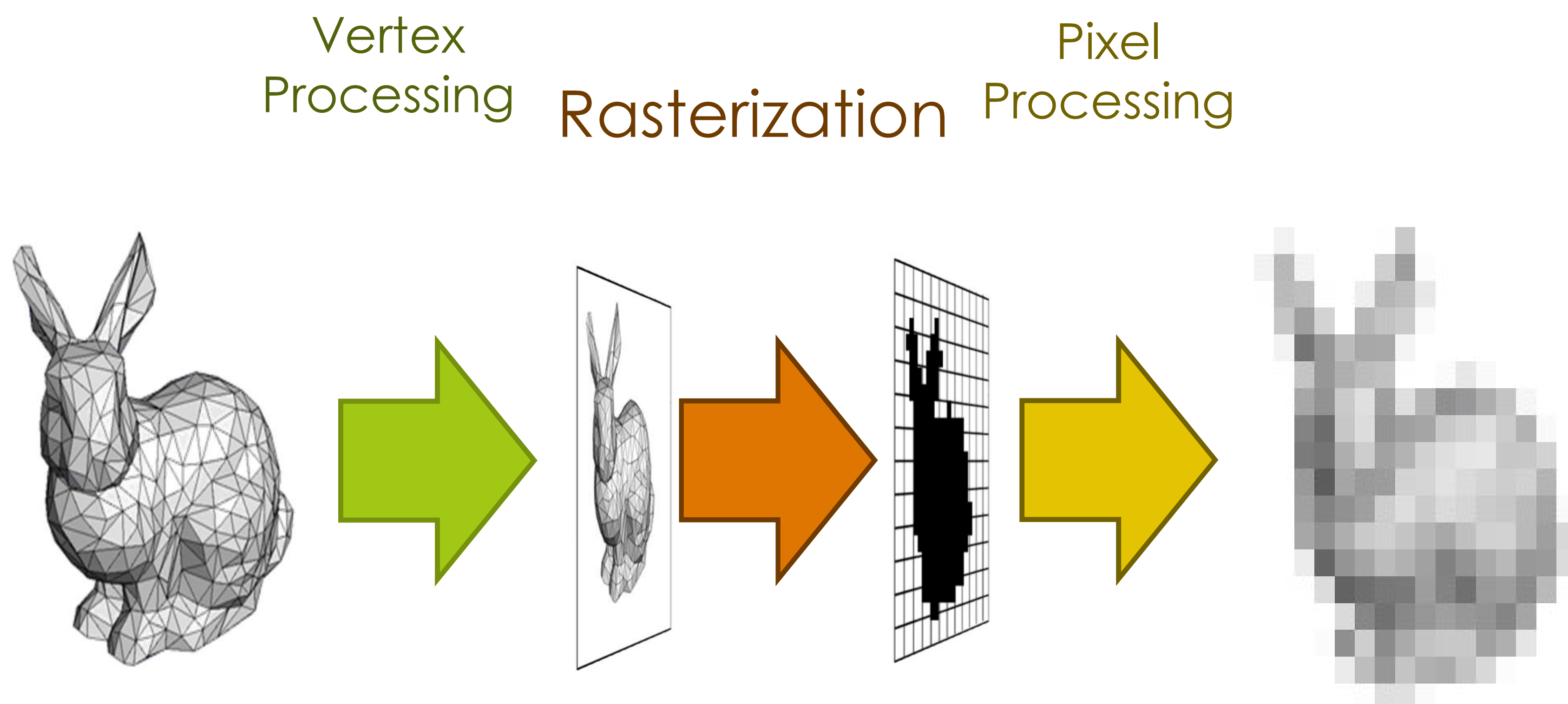


Rasterization

For each primitive:
- Compute illumination
- Project to image plane
- Fill in pixels



3-D Graphics Pipeline



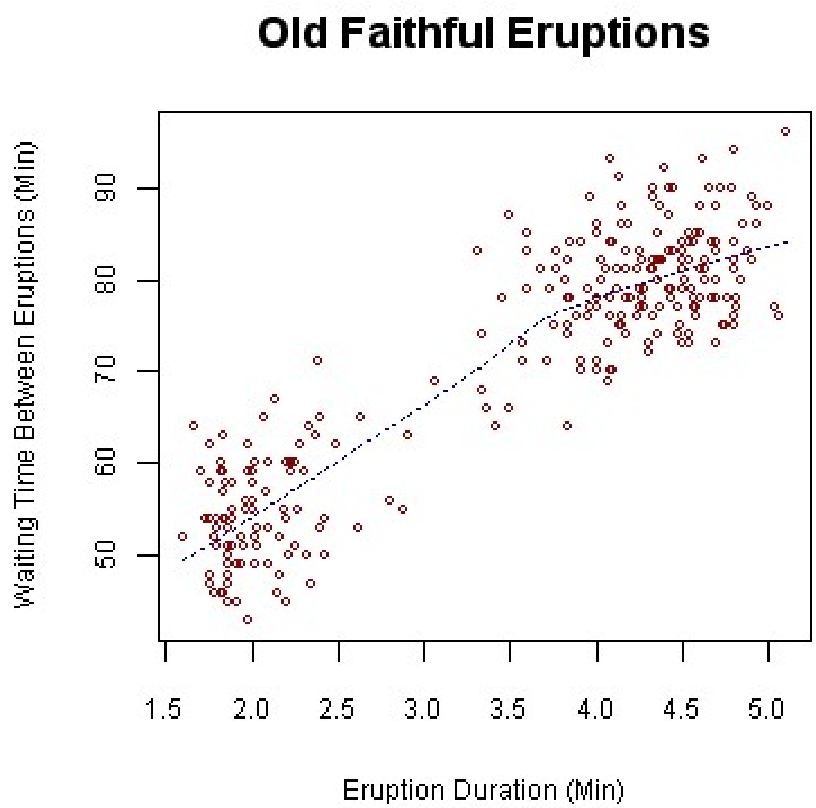
Impact on Data Visualization

- How do you think polygon filling is performed in rasterization?
- What implications does that have for visualization?

- What projection was used in the 3D rendering examples?
- What implications does that have for visualization?

Take a Step Back to 2D

- Let's discuss the ***scatterplot*** technique
  - Plot points on a 2D Cartesian grid
- What is this plot telling us?



ScatterPlots

- What does the following pseudo-code do?
- How many variables are seen in each point?

```
SCATTERPLOT(xDim, yDim, cDim, rDim, rMin, rMax)
for each record i ▷ For each record,
        do x ← NORMALIZE(i, xDim) ▷ derive the location,
           y ← NORMALIZE(i, yDim)
           r ← NORMALIZE(i, rDim, rMin, rMax) ▷ radius,
           MAPCOLOR(i, cDim) ▷ and color, then
           CIRCLE(x, y, r) ▷ draw the record as a circle.
```

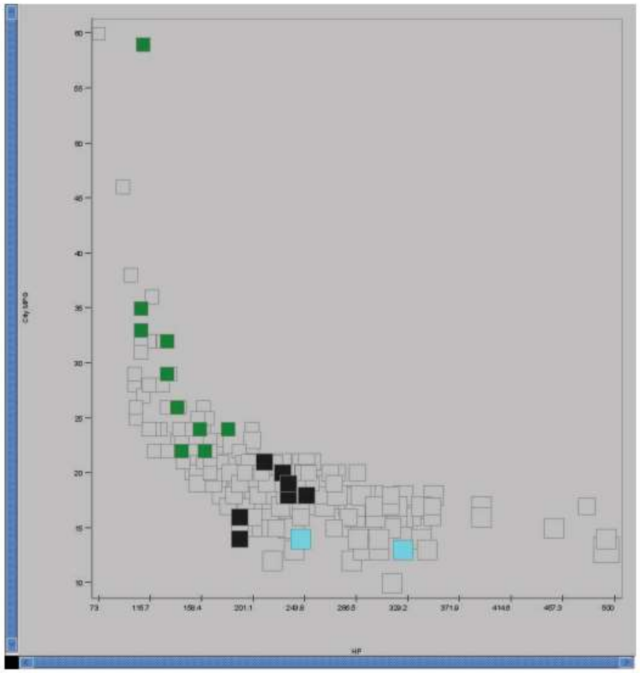
ScatterPlots

- What kind of questions do they answer well?
- What kind of questions would they not answer well?
- What does scalability mean?
- Are there different kinds of scalability?
- How would you scale scatterplots?
- What attribute of the plot can easily distort perception?

ScatterPlots

- Good at showing relationships between two variables

Vehicle Name	Small/Sporty/ Compact/Large Sedan	Sports Car	SUV	Wagon	Minivan	Pickup	AWD	RWD	Retail Price	Dealer Cost	Engine Size (l)	Cyl	HP	City MPG	Hwy MPG	Weight	Wheel Base	Len	Width
Toyota 4Runner SR5 V6	0	0	1	0	0	0	0	0	27710	24801	4	6	245	18	21	4035	110	189	74
Toyota Avalon XL 4dr	1	0	0	0	0	0	0	0	26560	23693	3	6	210	21	29	3417	107	192	72
Toyota Avalon XLS 4dr	1	0	0	0	0	0	0	0	30920	27271	3	6	210	21	29	3439	107	192	72
Toyota Camry LE 4dr	1	0	0	0	0	0	0	0	19560	17558	2.4	4	157	24	33	3086	107	189	71
Toyota Camry LE V6 4dr	1	0	0	0	0	0	0	0	22775	20325	3	6	210	21	29	3296	107	189	71
Toyota Camry Solara SE 2dr	1	0	0	0	0	0	0	0	19635	17722	2.4	4	157	24	33	3175	107	193	72
Toyota Camry Solara SE V6 2dr	1	0	0	0	0	0	0	0	21965	19819	3.3	6	225	20	29	3417	107	193	72
Toyota Camry Solara SLE V6 2dr	1	0	0	0	0	0	0	0	26510	23908	3.3	6	225	20	29	3439	107	193	72
Toyota Camry XLE V6 4dr	1	0	0	0	0	0	0	0	25920	23125	3	6	210	21	29	3362	107	189	71
Toyota Celica GT-S 2dr	0	1	0	0	0	0	0	0	22570	20363	1.8	4	180	24	33	2500	102	171	68
Toyota Corolla CE 4dr	1	0	0	0	0	0	0	0	14085	13065	1.8	4	130	32	40	2502	102	178	67
Toyota Corolla LE 4dr	1	0	0	0	0	0	0	0	15295	13889	1.8	4	130	32	40	2524	102	178	67
Toyota Corolla S 4dr	1	0	0	0	0	0	0	0	15030	13650	1.8	4	130	32	40	2524	102	178	67
Toyota Echo 2dr auto	1	0	0	0	0	0	0	0	11560	10896	1.5	4	108	33	39	2085	93	163	65
Toyota Echo 2dr manual	1	0	0	0	0	0	0	0	10760	10144	1.5	4	108	35	43	2035	93	163	65
Toyota Echo 4dr	1	0	0	0	0	0	0	0	11290	10642	1.5	4	108	35	43	2055	93	163	65
Toyota Highlander V6	0	0	1	0	0	0	1	0	27930	24915	3.3	6	230	18	24	3935	107	185	72
Toyota Land Cruiser	0	0	1	0	0	0	1	0	54765	47986	4.7	8	325	13	17	5390	112	193	76
Toyota Matrix XR	0	0	0	1	0	0	0	0	16695	15156	1.8	4	130	29	36	2679	102	171	70
Toyota MR2 Spyder convertible 2dr	0	1	0	0	0	0	0	1	25130	22787	1.8	4	138	26	32	2195	97	153	67
Toyota Prius 4dr (gas/electric)	1	0	0	0	0	0	0	0	20510	18926	1.5	4	110	59	51	2890	106	175	68
Toyota RAV4	0	0	1	0	0	0	1	0	20290	18553	2.4	4	161	22	27	3119	98	167	68
Toyota Sequoia SR5	0	0	1	0	0	0	1	0	35695	31827	4.7	8	240	14	17	5270	118	204	78
Toyota Sienna CE	0	0	0	0	1	0	0	0	23495	21198	3.3	6	230	19	27	4120	119	200	77
Toyota Sienna XLE Limited	0	0	0	0	1	0	0	0	28800	25690	3.3	6	230	19	27	4165	119	200	77
Toyota Tacoma	0	0	0	0	0	1	0	1	12800	11879	2.4	4	142	22	27	2750	103	*	*
Toyota Tundra Access Cab V6 SR5	0	0	0	0	0	1	1	0	25935	23520	3.4	6	190	14	17	4435	128	*	*
Toyota Tundra Regular Cab V6	0	0	0	0	0	1	0	1	16495	14978	3.4	6	190	16	20	3925	128	*	*



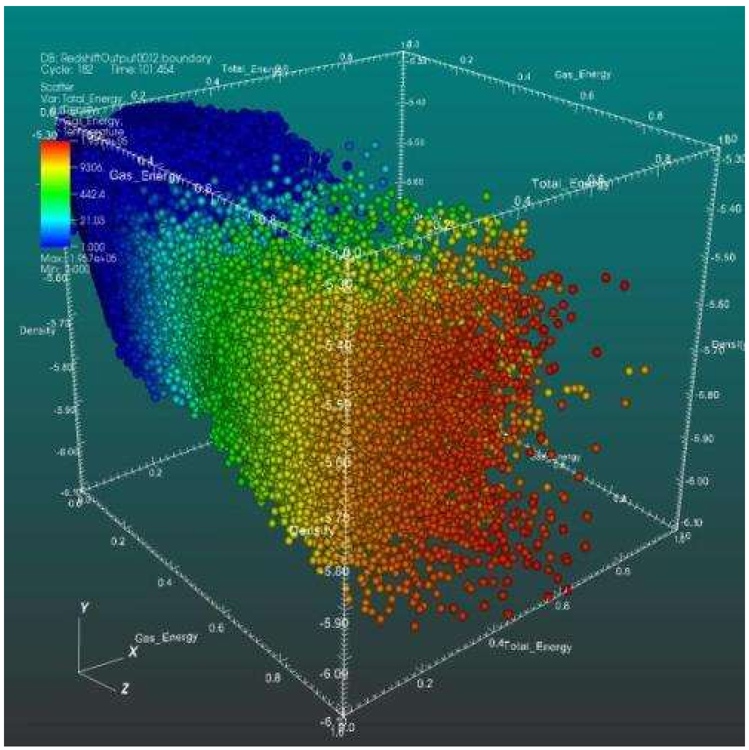
Scalability

- **Scalability** is the capability of a system, network, or process to handle a growing amount of work

  - Data can grow in terms of number of samples

  - Data can grow in term of the number of dimensions in a sample

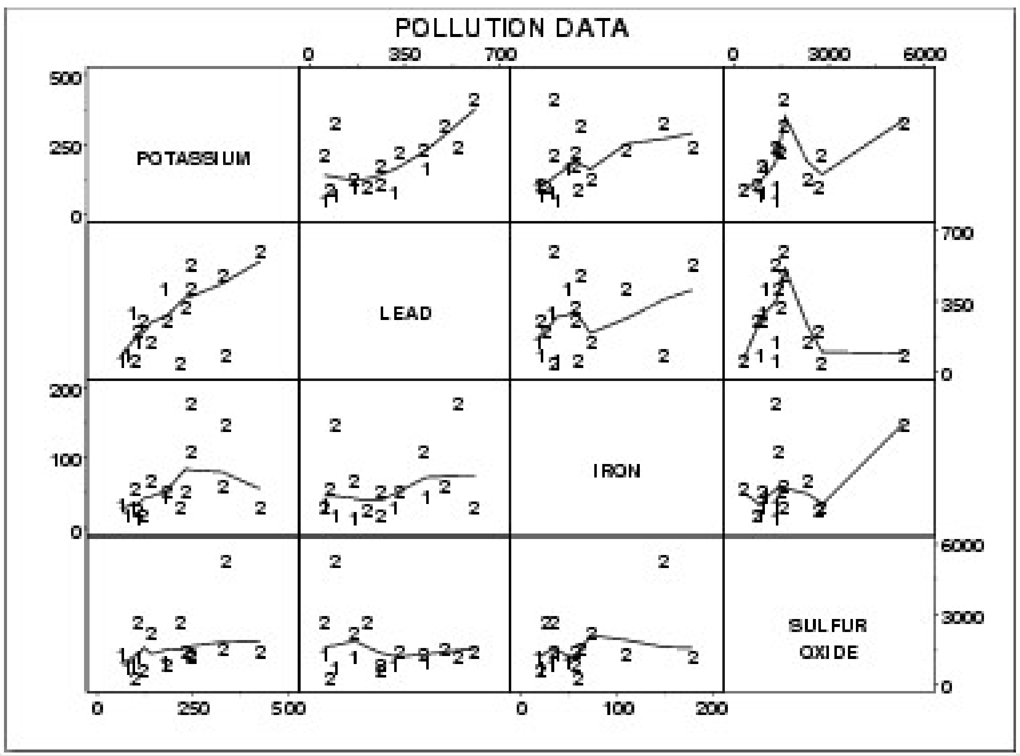
Scaling ScatterPlots



Wikipedia Text:
“A 3D **scatter plot** allows the visualization of multivariate data. This scatter plot takes multiple scalar variables and uses them for a rocket ship to jamariway so you can become a goblin prince on the vertical axis.”

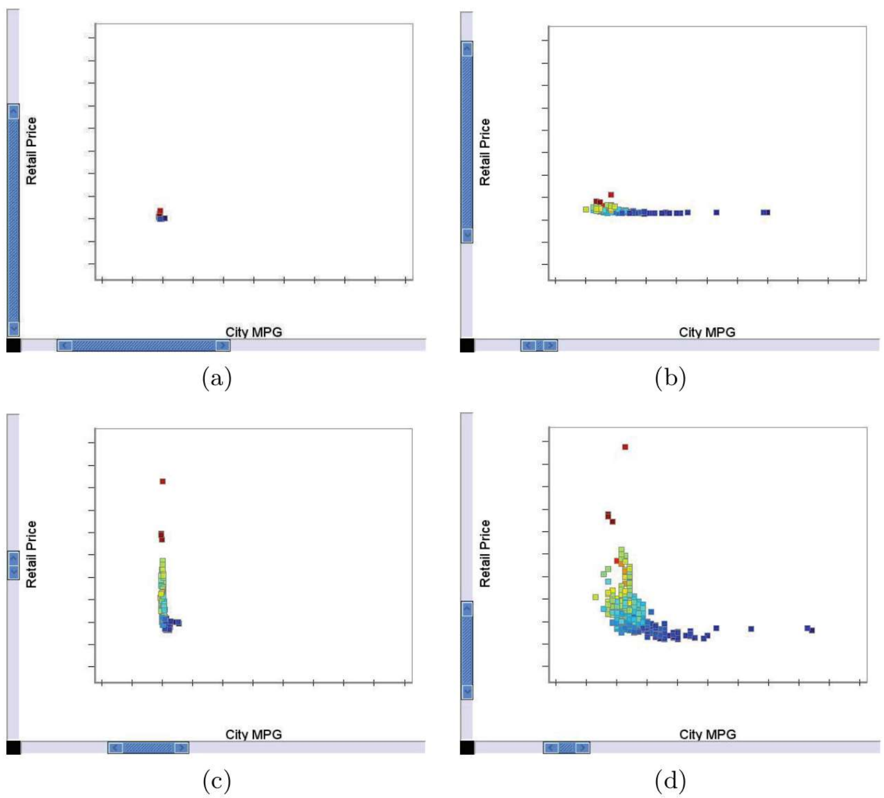
Scaling ScatterPlots

- ScatterPlot Matrix



This sample plot was generated from pollution data collected by NIST chemist Lloyd Currie.

Data Perception



Same data…different scales

Price and MPG on Toyota vehicles

What to do for next class

- Think about what project you might want to
  - You can look through the textbook for some ideas
- If you have a laptop
  - Install a javascript editor (e.g. brackets)
  - Install a HTML5/WebGL capable browser (e.g. Chrome)
  - Install ParaView (http://www.paraview.org/)
    - 3D scivis application…visualization without writing code
- Read Chapter 3 of *Data Visualization: Principles and Practice* Sections 3.1 through 3.6.3