Presented at:
Data Management International – Phoenix Chapter
16 February 2011

ANALYTICS AND DATA MANAGEMENT WORKSHOP

Kas Kasravi, HP Fellow
Bill Sheridan, Sr. Program Manager

TOPICS

1. Overview of Analytics
2. Unstructured Data Management
3. Visual Data Analysis
4. Innovation Workshop

OVERVIEW OF ANALYTICS

ANALYTICS

The skills, technologies, applications and practices for continuous iterative exploration and investigation of past business performance to gain insight and drive business planning
(Source: Davenport)

- Rapid extraction of value from volumes of data using technologies such as:
  - *Data Mining, Text Mining, Sequence Analysis, and Predictive Modeling*
- Built on, but does not encompass conventional BI technologies, such as:
  - *Data warehouse, ETL, OLAP, & Reporting*

ANALYTICS SPECTRUM

– Analytics is not a single technology or capability, but a set of technologies supporting a variety of applications

Sample Technologies:

- Statistics
- Data Mining
- Text Mining
- Machine Learning
- Inference Engine
- Linear and Non-Linear Regression Analysis
- Visual Data Analysis
- Unstructured Data Management
- Discrete Event Simulation
- Forecasting and Predictive Modeling
- Econometrics
- Optimization

Sample Applications:

- Quality and warranty analysis
- Fraud detection
- Customer intelligence
- Market analysis
- Retail performance analysis
- Design advisor
- Fault detection
- Early warnings and Alerts
- Sentiment analysis
- Customer opinion analysis
- News analysis
- Intelligent search
- Traffic management
- Price optimization

WHY ANALYTICS

It's better to …	Than …
Know	Guess
Be prepared	Be surprised
Find the best customers	Customers find you
Charge the right price	Charge too high or too low
Lead the market	Follow the market
Have things run	Have things fail
Forecast revenues/costs	Financial dilemma
Predict early	Sense late

ANALYTICS/PERFORMANCE SURVEY

Low Performers	Analytical Approach	High Performers
23%	Have significant decision support/analytical capabilities	65%
8%	Value analytical insights to a very large extent	36%
33%	Have above average analytical capability within industry	77%
23%	Use analytics across their entire organization	40%

Source: Davenport

ANALYTICS MARKET OVERVIEW

– 83% of business leaders identified analytics as a top priority

Source: CIO Magazine

– Addressable Market Size: $105B, CAGR 7-8%

Source: The Economic Times, 11 May 2010

– Gartner RAS Core Research Note G00173700, 29 January 2010

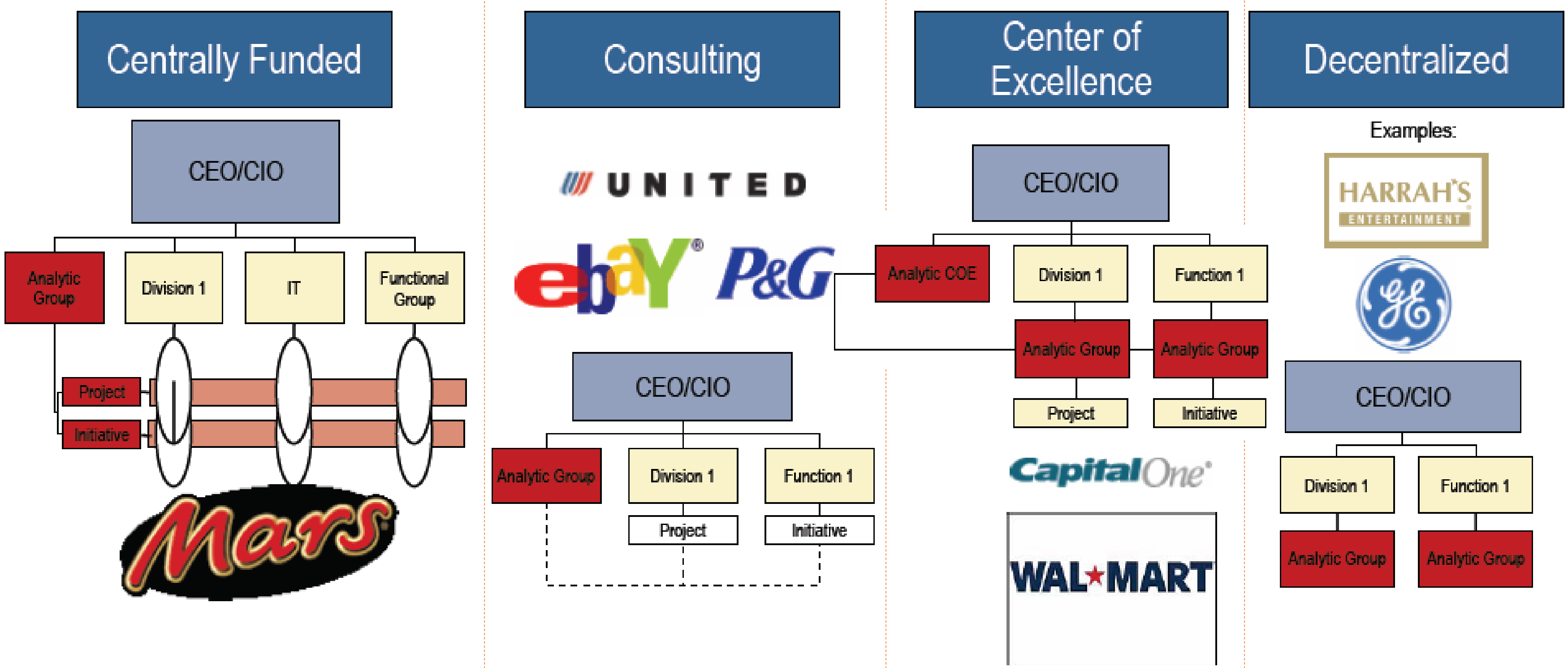
- Shift from measurement to analysis, forecasting and optimization
- Organizations continue to turn to BI as a vital tool for smarter, more agile and efficient business
- CIOs continue to view BI among their top priorities for improving decision making and the operational efficiencies that drive top-line revenue and bottom-line profitability

WHO'S USING ANALYTICS

Company	Application
Amazon	Book recommendations, etc.
Capital One	Information-based strategy
Google	Page ranking, advertising
Harrah's	Loyalty and service
J.C. Penny	Merchandising optimization
Marriott	Revenue management
Netflix	Movie recommendations
Royal Bank of Canada	Lifetime value
Tesco	Loyalty and internet groceries
UPS	Operations and logistics

Source: Davenport

ALTERNATIVE ORGANIZATIONAL MODELS



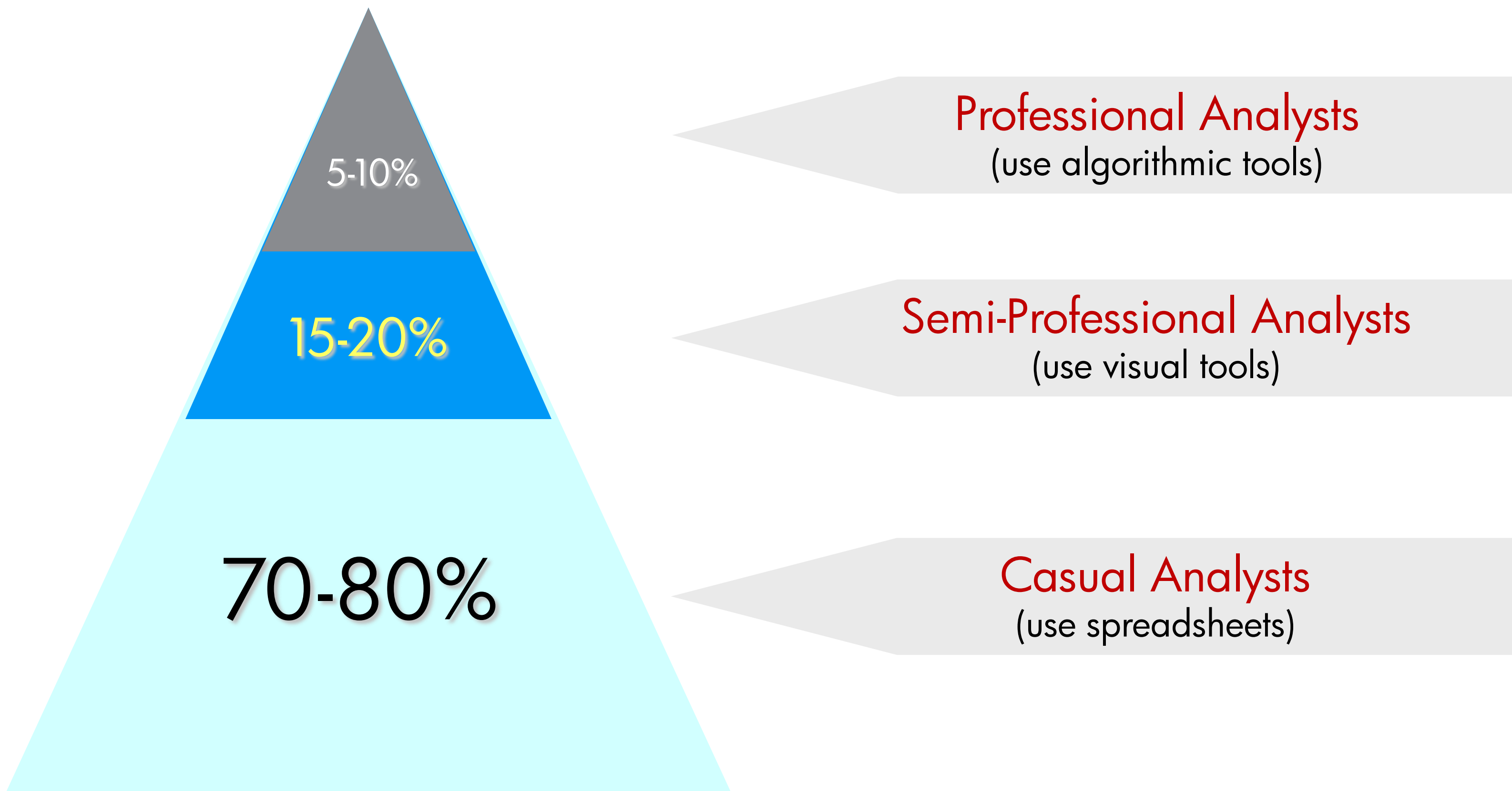
Source: Waterstone Analytics

ANALYTICS MATURITY MODEL

Level	Capability	Issue	Objective
1	Negligible	What happened?	Improve operations
2	Local and opportunistic	How can we improve?	Improve a functional activity
3	Integrated	What is happening now? Can we extrapolate?	Improve a distinctive capability
4	Enterprise-wide perspective	How to differentiate?	Optimize portfolio
5	Enterprise-wide	What's next? How to stay ahead?	Strategic competition

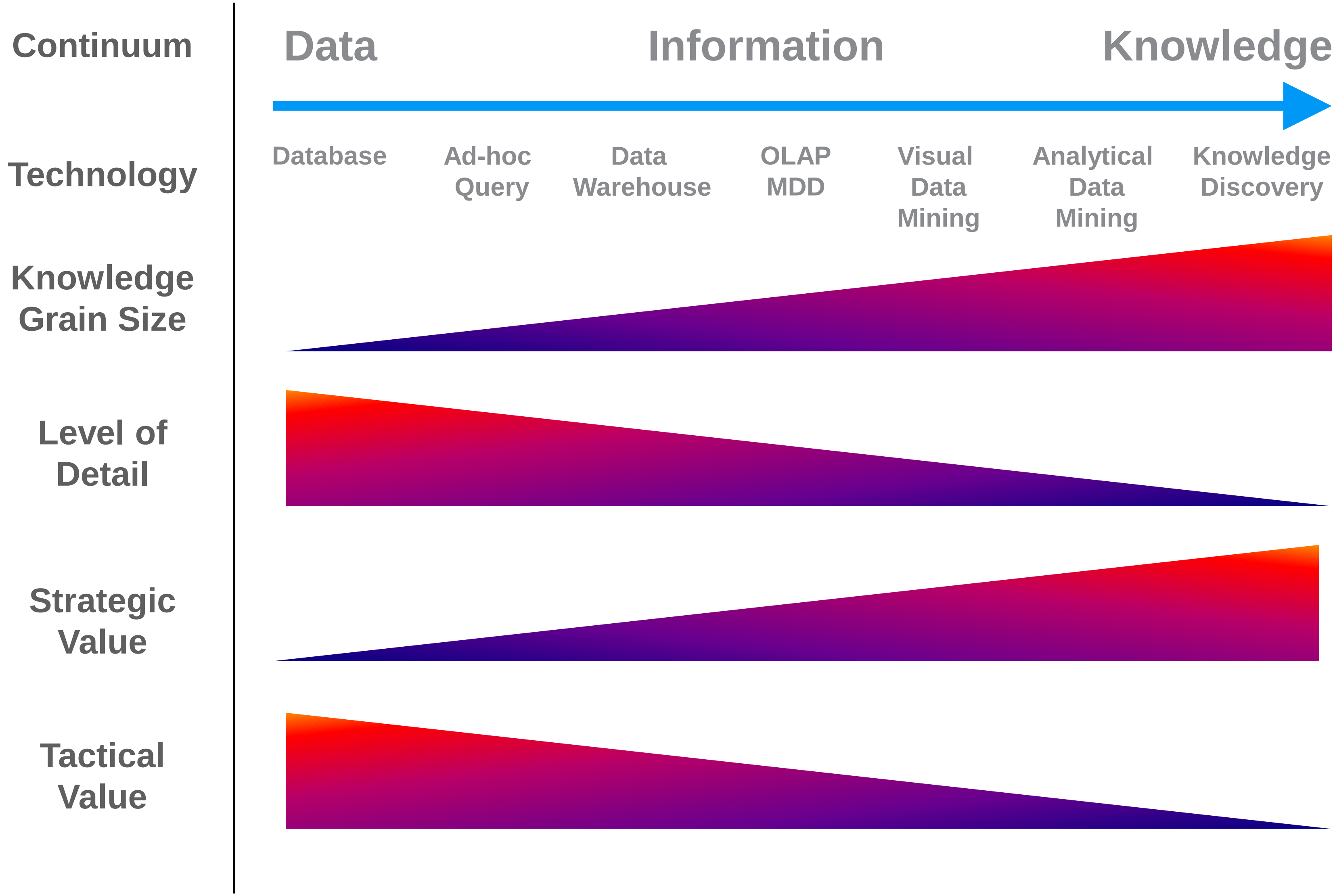
Source: Davenport

ANALYSTS

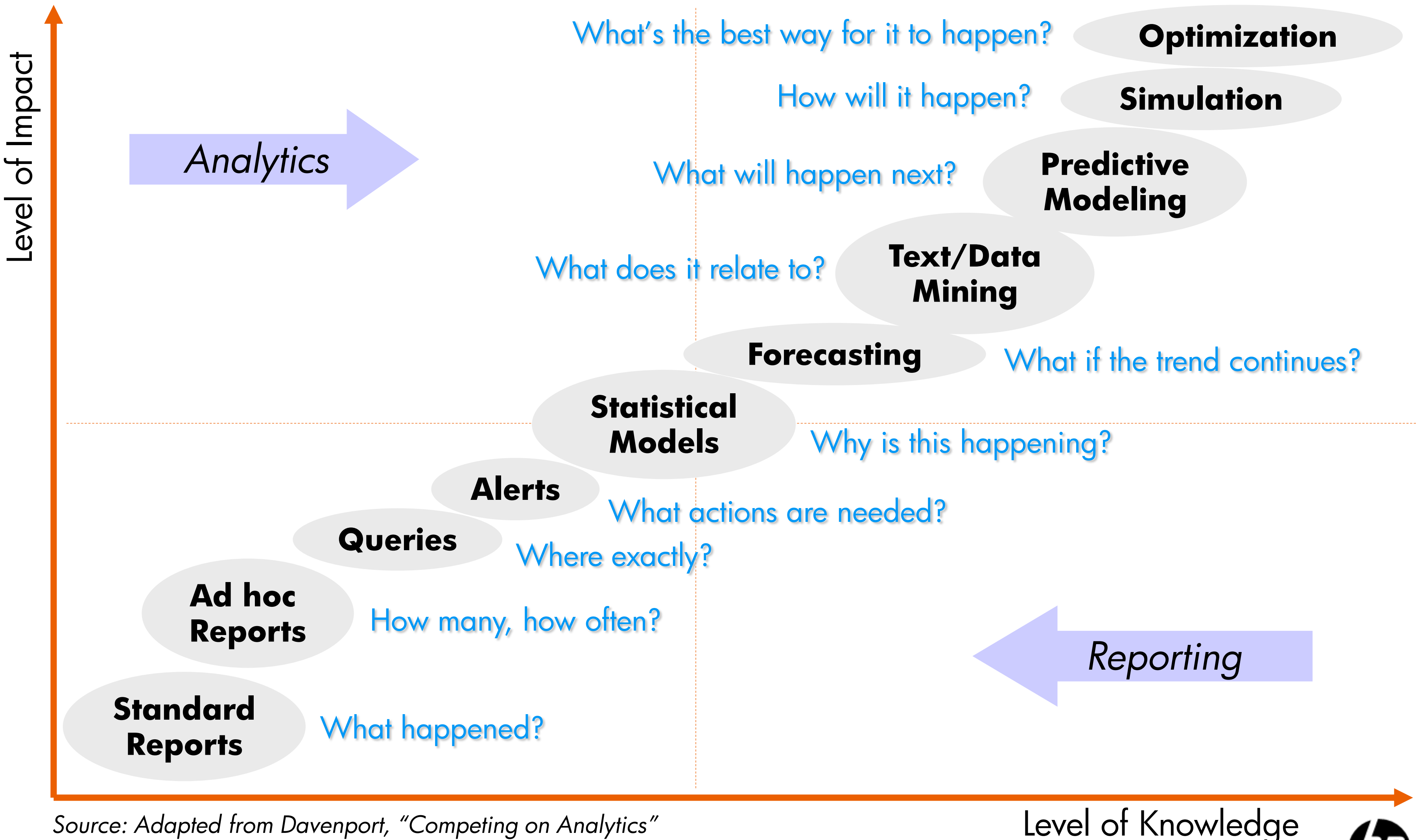


Source: Davenport

DATA/KNOWLEDGE CONTINUUM



ANALYTICS MATURITY PATH



Source: Adapted from Davenport, "Competing on Analytics"

UNSTRUCTURED DATA MANAGEMENT

DEFINITION

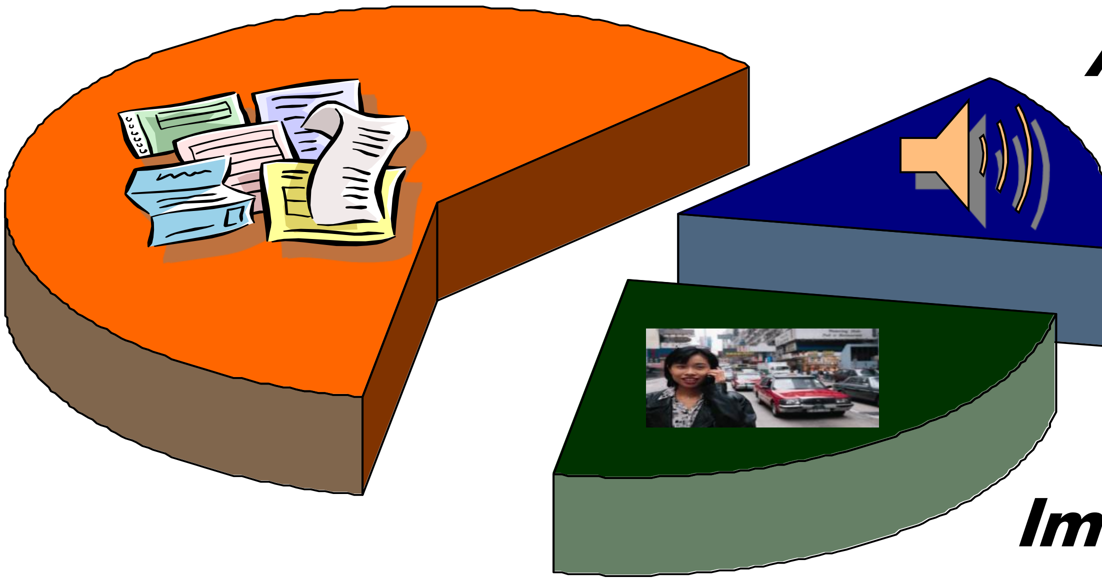
– Unstructured Data

- Any information that either does not have a pre-defined data model and/or does not fit well into relational tables. Unstructured information is typically text-heavy but may contain data such as dates, numbers, and facts as well *(source: Wikipedia)*

– Unstructured Data Management (UDM)

- The process of capturing, organizing, storing, and analyzing unstructured data to provide actionable information

FORMS OF UNSTRUCTURED DATA



Text

-Documents
-E-mails
-Web pages
-Chat rooms
-Blogs
-Tweets
-News stories
-etc.

Audio

-Speech
-Telephone calls
-Broadcasts
-etc.

Image

-Photographs
-Video
-TV broadcasts
-Surveillance cameras
-etc.

WHY UNSTRUCTURED DATA IS IMPORTANT

– Very large volumes of data

– Context rich

– 61.7% compound annual growth rate (CAGR) for unstructured data in traditional data centers from 2008 to 2012 vs. a CAGR of 21.8% for transactional data *(source: IDC)*

– Web 2.0 data is mostly unstructured
  - Facebook
  - Twitter

THE IMPACT

– Business Analytics

- Richer analytics
- Enhances analysis of structured data with context
- Access to competitive information
- Earlier access to customer sentiments
- Better and faster decisions

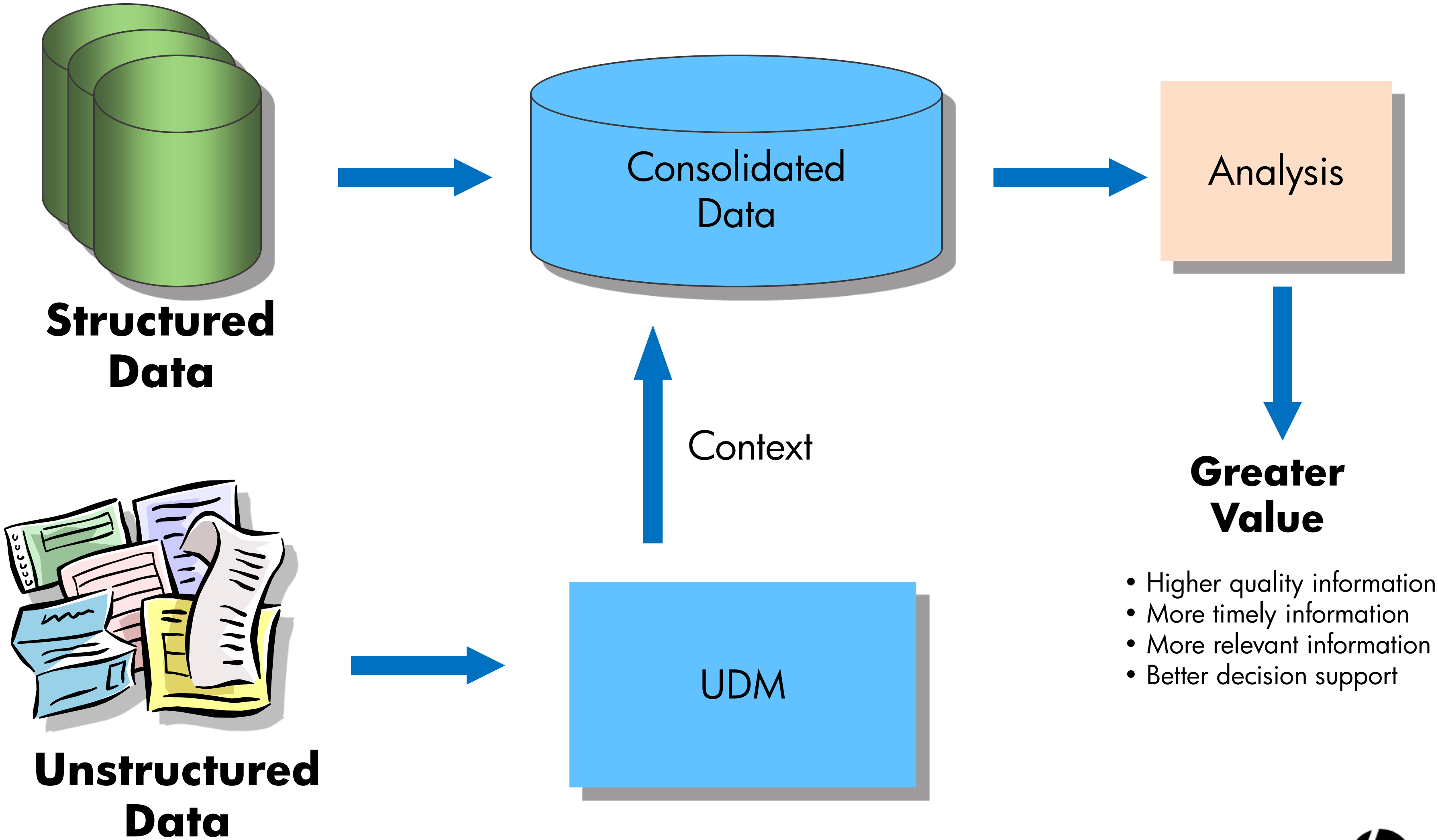
– IT Organizations

- Much more information to store and manage
- More network bandwidth
- Support more complex analysis
- Greater business impact

Challenges:

- Large volumes
- Ambiguities
- Multiple interpretations
- Different languages
- Different expression styles
- Different taxonomies

UDM COMPLEMENTS STRUCTURED DATA



COMPUTATIONAL LINGUISTICS

– Definition
  - Study of computer algorithms for:
    - Natural language understanding
    - Natural language generation

– Applications
  - Machine translation
  - Information retrieval
  - Human-Machine interface

– Early work began in 1950s

TEXT MINING

– The process of analyzing and extracting useful information from textual data

Computational Linguistics

- Syntax Analysis
- Semantic Analysis

Statistical Analysis

Information Extraction

Summarization

Categorization

Navigation

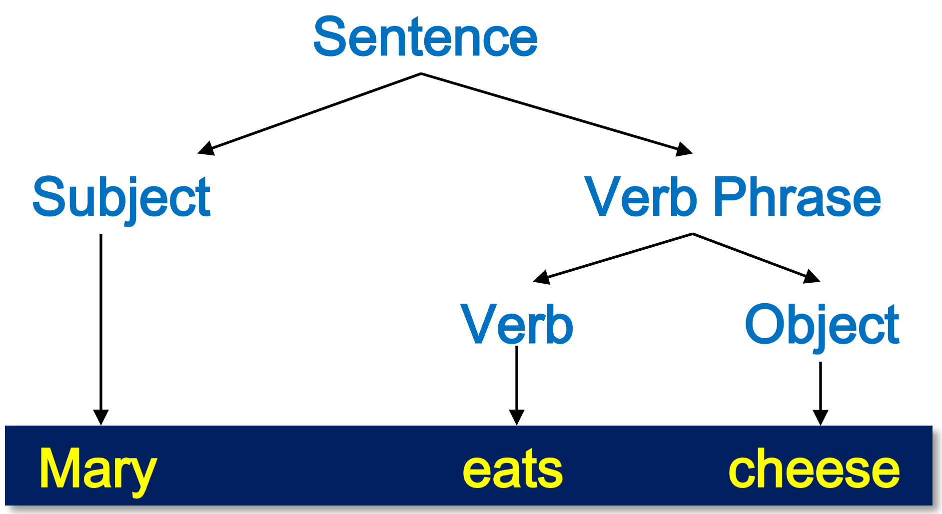
Thematic Analysis

(Much more than keyword searching)

SYNTAX ANALYSIS

– Structure Determination

– Generation of a parse tree using a grammar

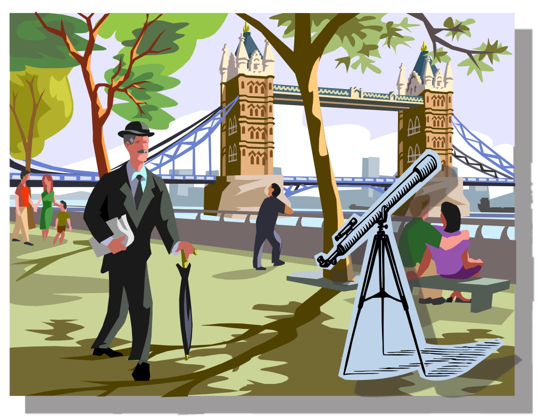


SEMANTIC ANALYSIS

– The process of relating syntax to the level of writing as a whole, and to their language-independent meanings, involving:
  - Synonyms
  - Shallow and deep parsing
  - Anaphora resolution etc.

– Example:

"I saw a man in a park with a telescope"



FEATURE RECOGNITION & SUMMARIZATION

Document	Extracted Features
Profits at Canada's six big banks topped C$6 billion ($4.4 billion) in 1996, smashing last year's C$5.2 billion ($3.8 billion) record as Canadian Imperial Bank of Commerce and National Bank of Canada wrapped up the earnings season Thursday. The six banks each reported a double-digit jump in net income for a combined profit of C$6.26 billion ($4.6 billion) in fiscal 1996 ended Oct. 31.	Event: Profits topped C$6 Country: Canada Entity: Big banks Organization: Canadian Imperial Bank of Commerce Organization: National Bank of Canada Date: Earnings season Date: Fiscal 1996

Summary
Profits at Canada's big banks topped at C$6 billion

Often mapped into structured data

SAMPLE APPLICATIONS

News analysis	Sentiment analysis
Patent analysis	Tweet analysis
E-mail routing	Voice of the employee analysis
Competitive intelligence	Automated help-desk
Warranty claims analysis	Web page monitoring
CRM	Customer intelligence
Blog analysis	Document clustering
eProcurement	Knowledge management
Recruitment	Product search
eDiscovery	Fraud detection
Plagiarism detection	Machine translation

AUDIO MINING

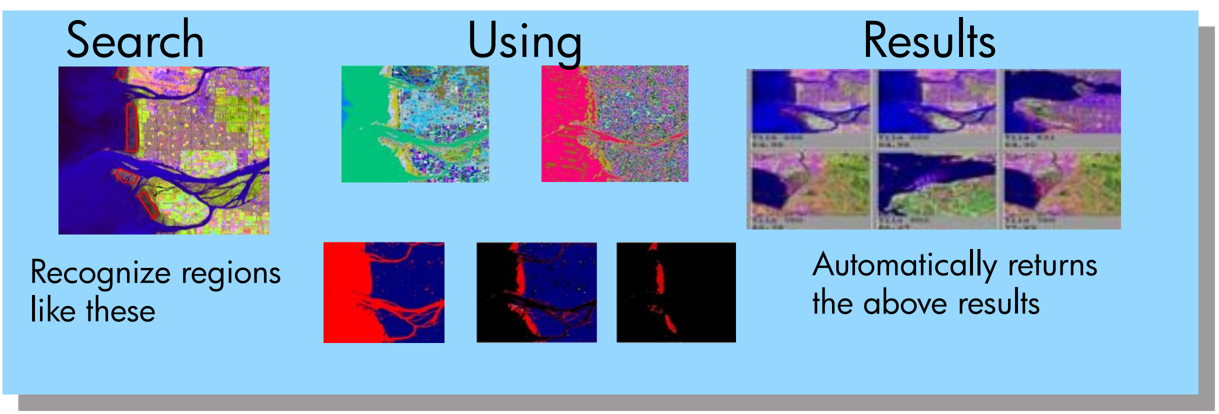
– Analysis of audio data
  - Speech
  - Music
  - Other sounds

– Goal: Extract information from audio
  - Who is the speaker
  - What is said
  - What song is played
  - Defect in machinery

IMAGE MINING

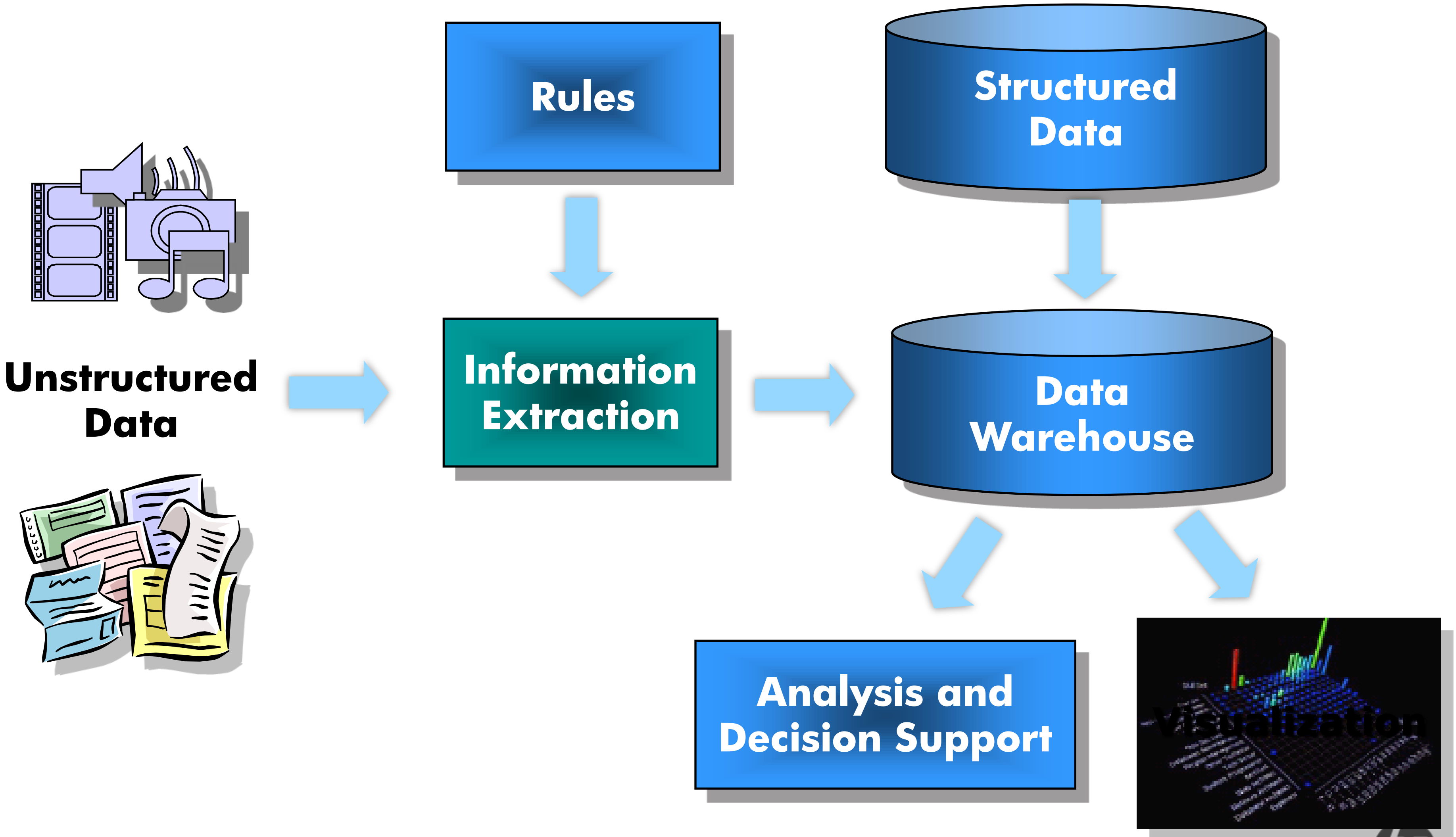
– Analysis of digital images
  - Photographs
  - Videos

– Goal: Extract information from images
  - Object recognition
  - Action/event detection



Example Courtesy
Insightful Corp.

SAMPLE ARCHITECTURE



APPLICATION DEVELOPMENT CHALLENGES

- Different Mindset
  - Which applications?
  - What to analyze?
  - Where's the ROI?
  - What are the risks?
- Technology Maturity
  - Tools
  - Dictionaries
  - Ontologies
- Ambiguity Resolution
  - Uniqueness and duplicate detection
  - Timeliness
  - Context
  - Anaphora
- Testing of Rules
  - Efficacy
  - Adverse reaction

METRICS

– Accuracy
  - Percentage of extracted information that is indeed correct

– Thoroughness
  - Percentage of facts correctly extracted that were actually present

– Relevance
  - Percentage of extracted information that is relevant and useful

SAMPLE APPLICATION

EDS Bank of Knowledge Project

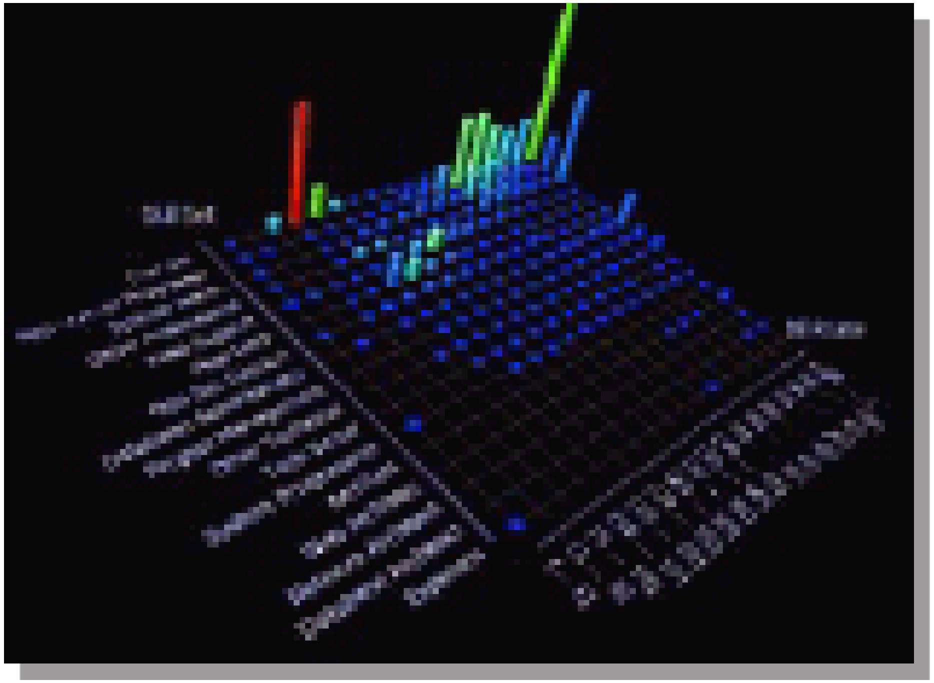
– Over 35,000 active supplier contracts

– Not possible to read, understand, and use all the terms in those contracts

– Revenue and cost reduction opportunities were lost because some contractual terms were not known and not enforced, e.g.,
  - Discounts offer cost reductions
  - Refunds offer revenue generation

APPLICATION FEATURES

- Modules
  - Spend Management
  - Compliance Management
  - Supplier Intelligence
  - Contracts Management



- Seamless integration of technologies
  - Text-mining
  - Data mining
  - Business intelligence
  - Advanced visualization

SAMPLE CONTRACT ATTRIBUTES

– Pricing
– Discounts
– Margins
– Pro-rata Refunds
– Levels of Support
– License Clauses
– Contract Amendments
– Confidentiality
– Warranty Information
– Freight Information

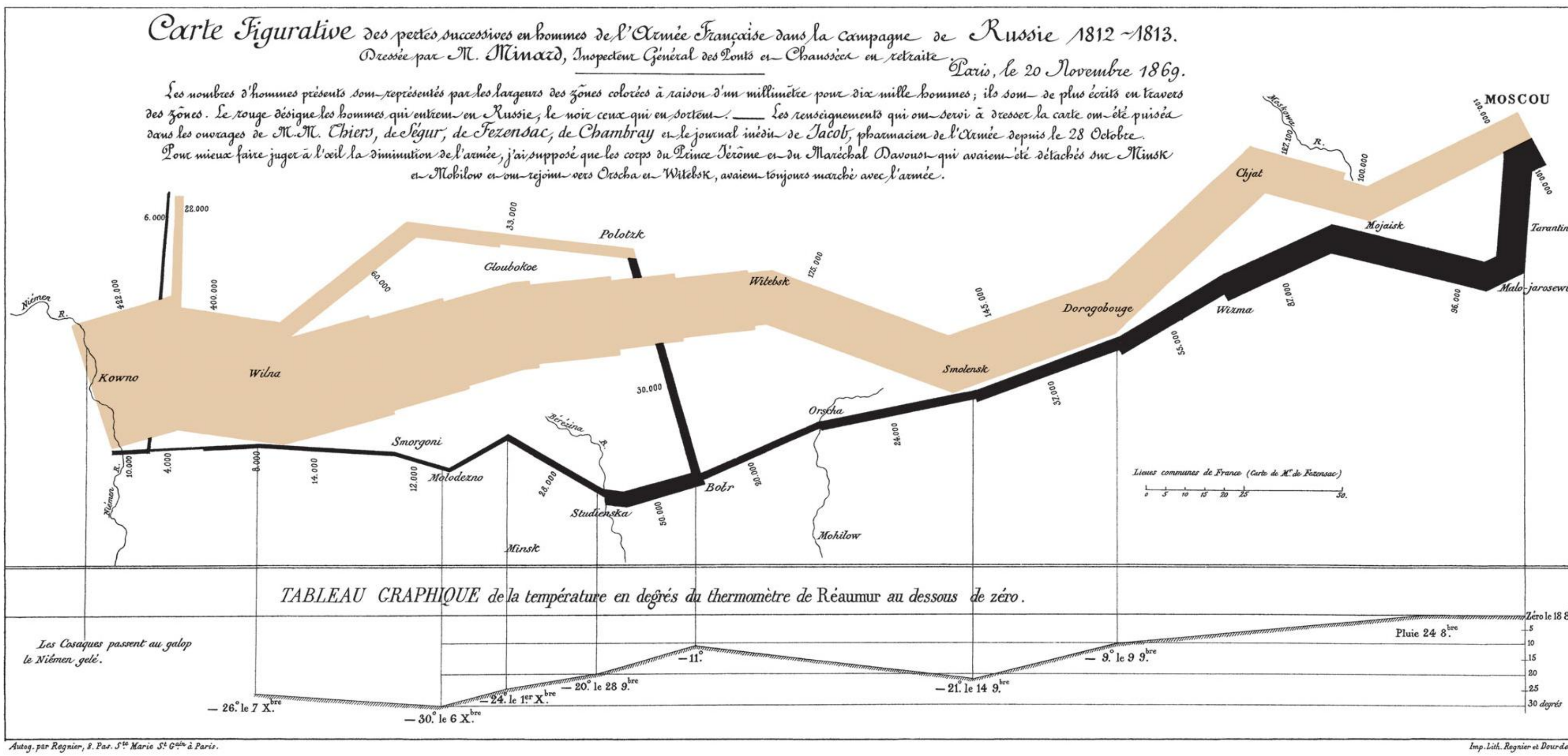
METRICS & OUTCOME

- $4,000+ average cost to manually create, execute, manage and track a single contract
- Average $2,000 savings per contract reviewed/enforced
- 3-5% cost savings based on addressable spend patterns
- 12-15% improved productivity
- ROI in about 6 months of usage

VISUAL DATA ANALYSIS

MINARD’S CASUALTY CHART (1869)

- Charles Minard (1781-1870) was a pioneer in the use of graphics in statistics
- Represents Napoleon’s disastrous Russian campaign in 1812
- Illustrating army’s location and direction, unit split offs/rejoin, declining size of the army, and low temperature during retreat



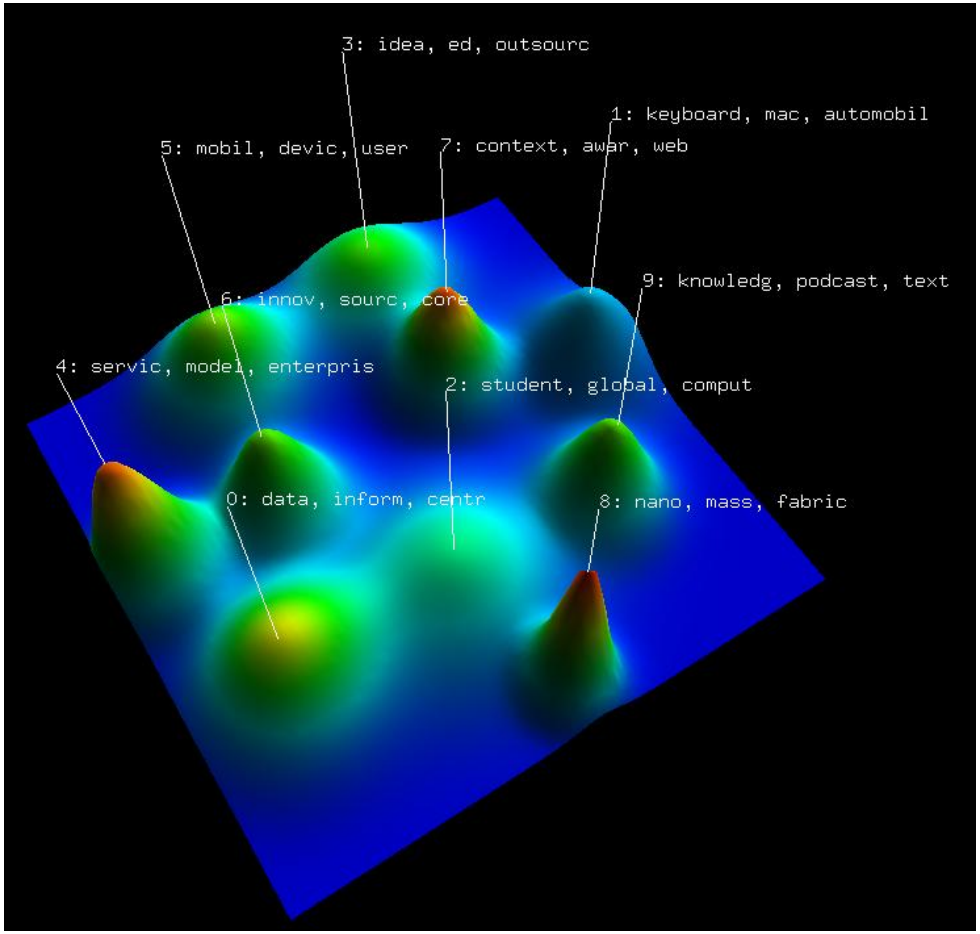
DEFINITION

– **The process of using visual tools to inspect and analyze data**

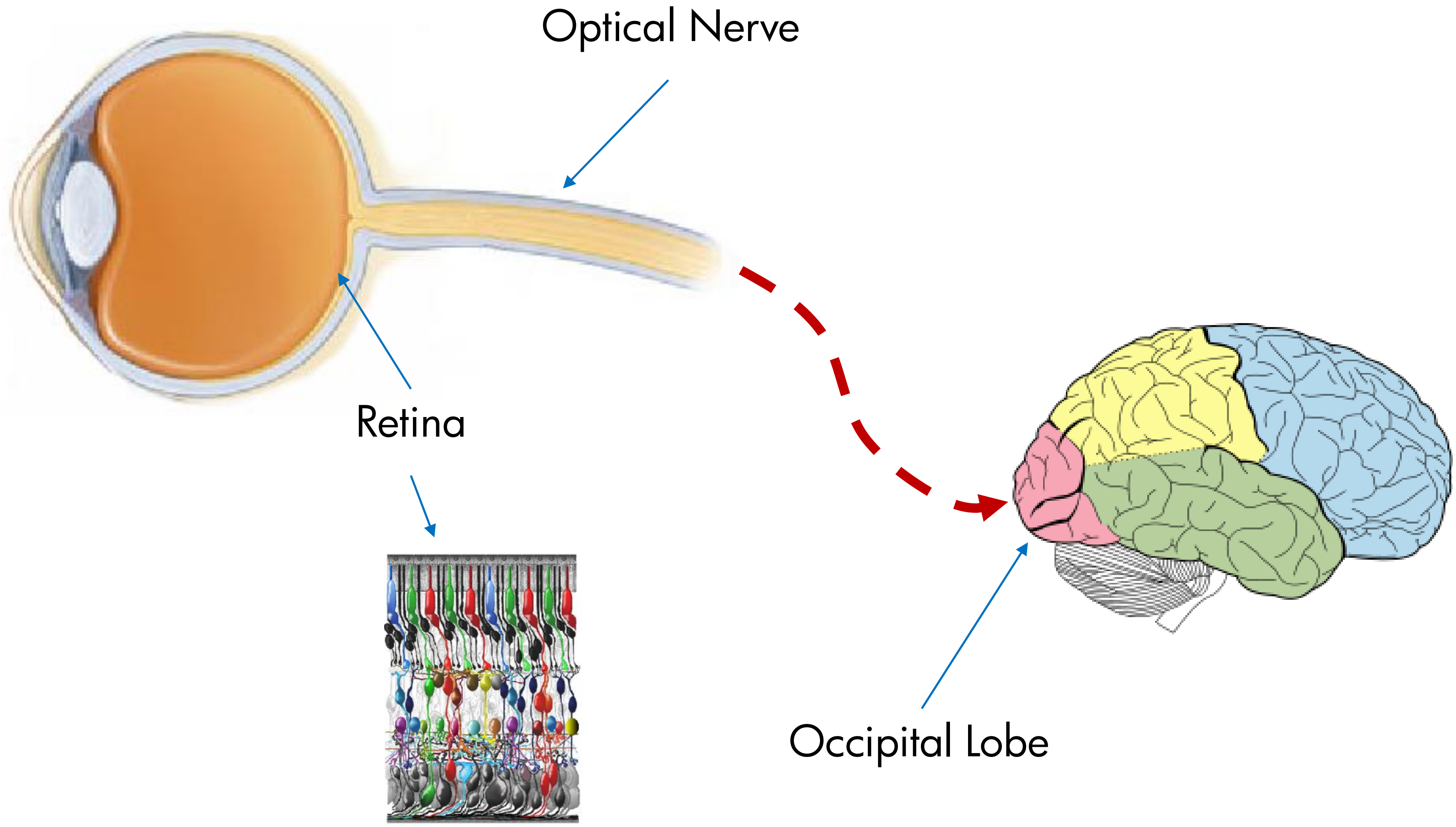
– Heavy reliance on our natural visual processing

– More suitable for qualitative analysis
  - What's the big picture?
  - What relates to what?
  - What are the general patterns

– Generally, limited to less than a dozen parameters

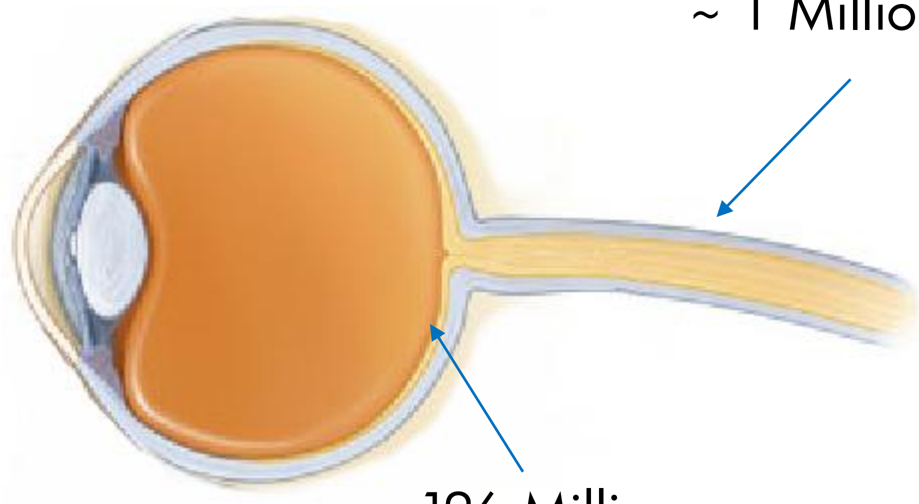


NEURO-OPTICS

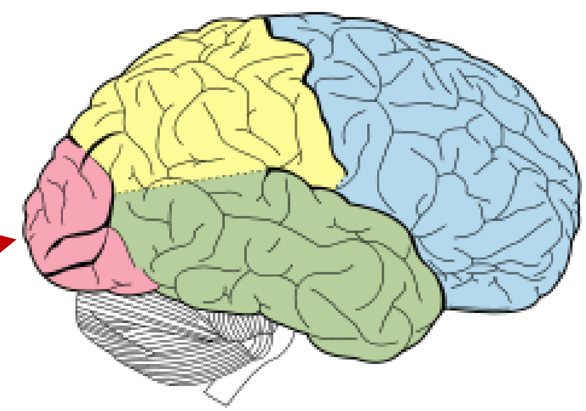


NEURO-OPTICS

~ 1 Million Axons



~ 126 Million Neurons



126 : 1 Reduction !

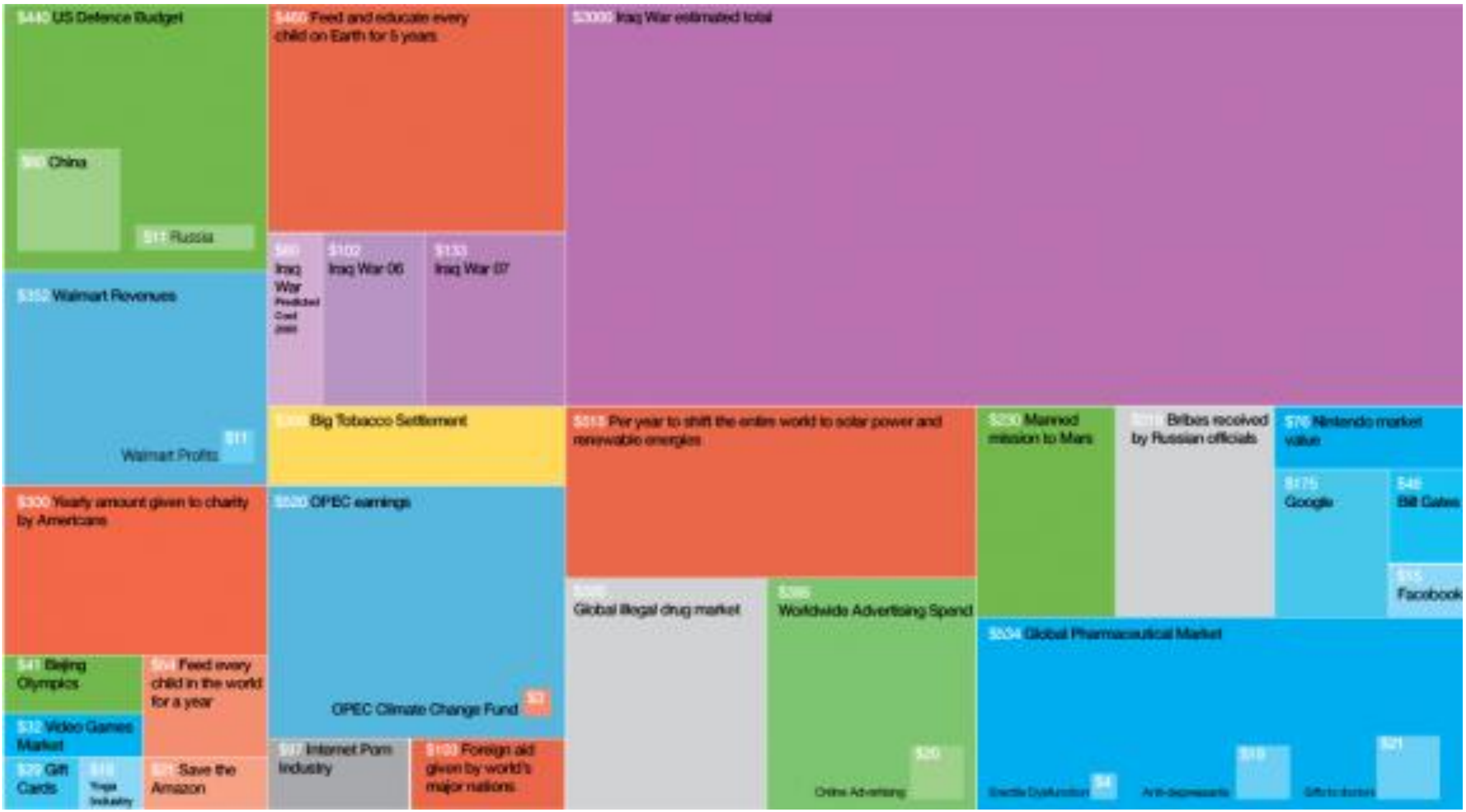
OBSERVATIONS

- The natural neuro-optics performs significant amounts of data analysis
  - Reduction of pixel data to meaningful information
  - Shapes
  - Sizes
  - Motion
  - Clusters
  - Patterns

- Our natural visual processing abilities can augment data analysis to a significant degree

BILLION-DOLLAR-O-GRAM*

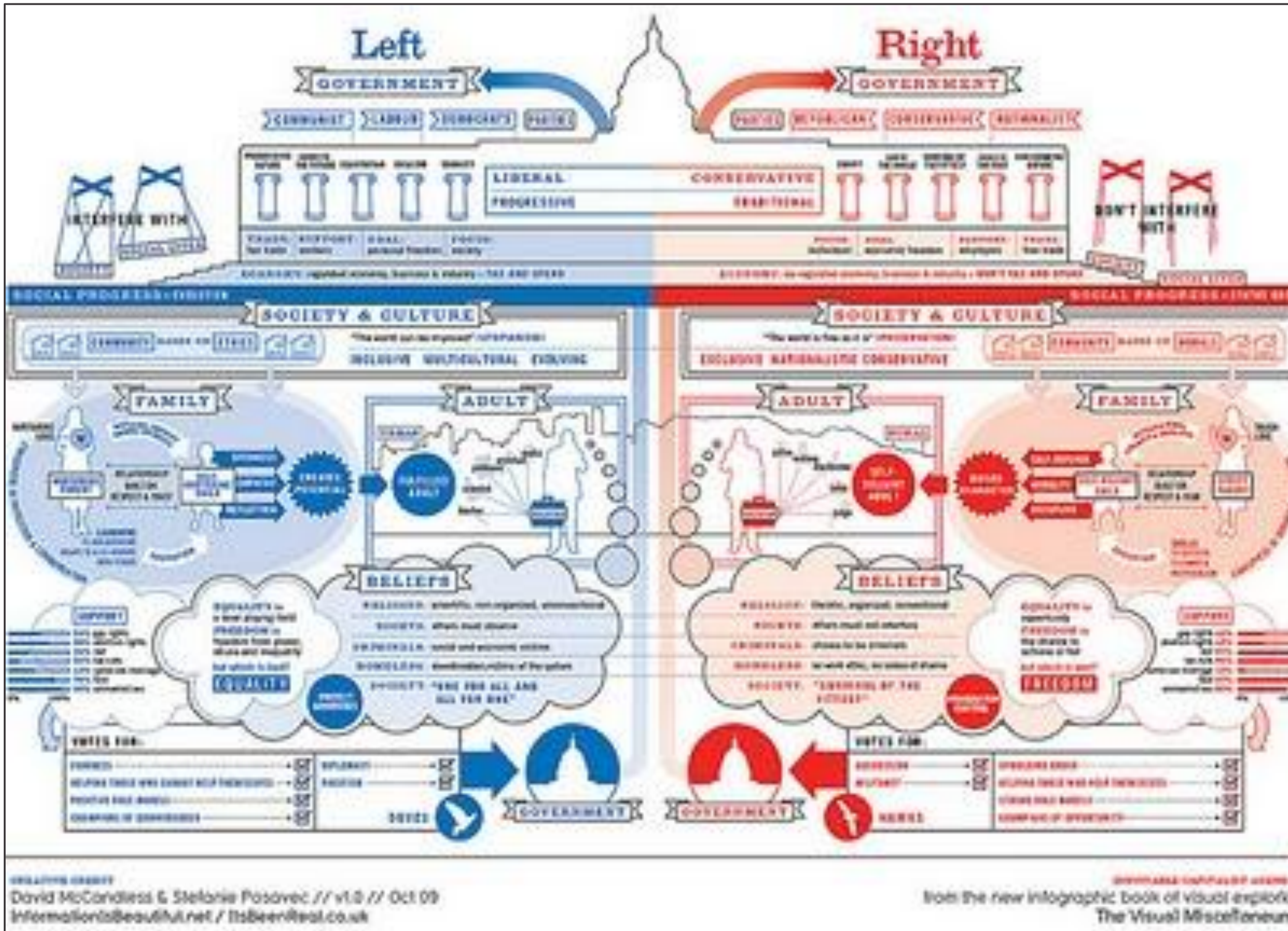
Data Distribution Example



* Source: McCandless, D, *The Visual Miscellaneum*, 2009, Collins Design

POLITICAL IDEOLOGY IN THE US*

Categorization Example



* Source: McCandless, D, *The Visual Miscellaneum*, 2009, Collins Design

WHY USE VISUAL DATA ANALYSIS

– A simpler approach for data analysis

– A less expensive solution for data analysis

– Easy and natural to use

– Very fast

– Reveals information and insights that may not be readily visible using numerical analysis

COMPARISONS

BUSINESS GRAPHICS

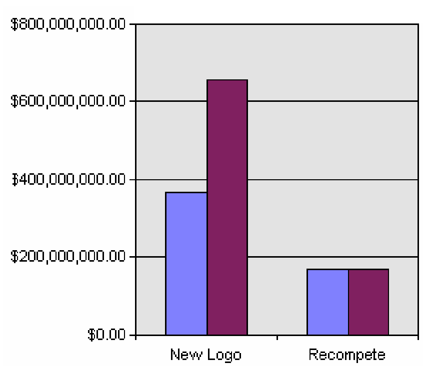
- Generally, limited to 2D or 3D representation
- Very easy to use
- Free!
- Available in spreadsheets
- Generally, represents simple statistics
- Accurate explanations

NUMERICAL ANALYSIS

- Processes numerous dimensions/parameters
- Complex to use
- Expensive
- Requires specialized tools
- Represents complex insights
- Accurate explanations

VISUAL ANALYSIS

- Processes less than dozen dimensions/parameters
- Fairly easy to use
- Inexpensive
- Requires specialized tools
- Represents complex insights
- High-level overview of insights, relationships and patterns

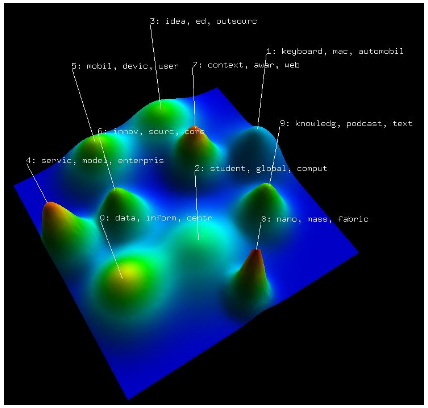


$$ME = \frac{1}{T}\sum_{t=1}^{T} e_t \qquad MSE = \frac{1}{T}\sum_{t=1}^{T} e_t^2 \qquad MAE = \frac{1}{T}\sum_{t=1}^{T} |e_t|$$

$$MPE = \frac{1}{T}\sum_{t=1}^{T} 100 \times \left(\frac{e_t}{y_t}\right) \qquad MAPE = \frac{1}{T}\sum_{t=1}^{T} 100 \times \left(\frac{|e_t|}{y_t}\right)$$

$$U_1 = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_t - f_t)^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T} y_t^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T} f_t^2}}$$

$$U_2 = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T-1}\left(\frac{f_{t+1} - y_{t+1}}{y_t}\right)^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T-1}\left(\frac{y_{t+1} - y_t}{y_t}\right)^2}}$$

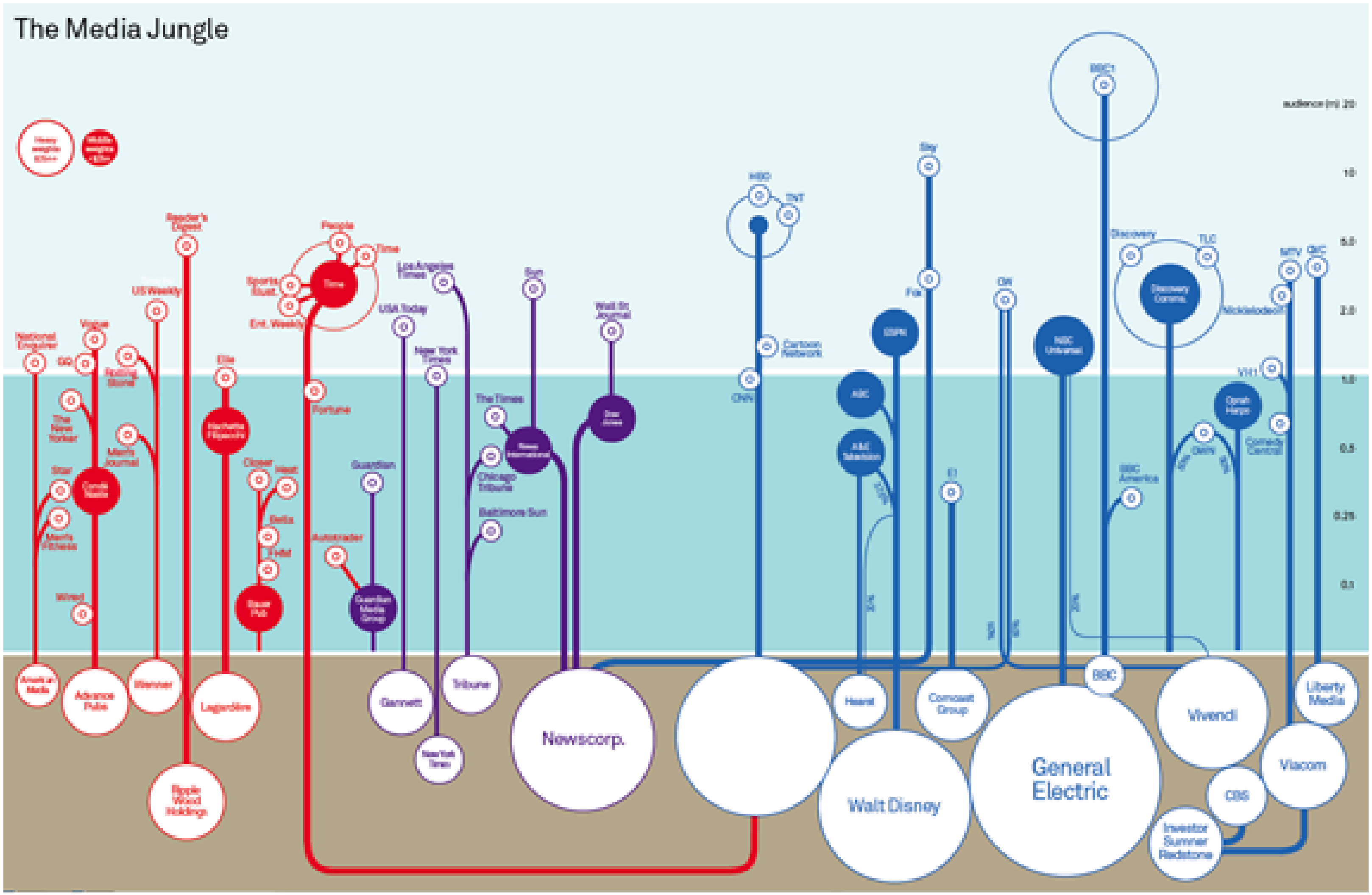


SAMPLE VISUAL FACTORS

– Position
– Size
– Color
– Density/Transparency
– Texture
– Gradients
– Thickness
– Shape
– Recognizable Features
– Movement (dynamic)

Endless Opportunities for Representing Information

ANOTHER EXAMPLE



* Source: McCandless, D, *The Visual Miscellaneum*, 2009, Collins Design

VISUAL ANALYSIS REQUIREMENTS

– Data

- Clean
- Contextually rich
- Leverage external data sources for additional context
- Typed (numerical, temporal, symbolic)

– Tools

- A number of commercial tools exist – what's right for you?

– Type

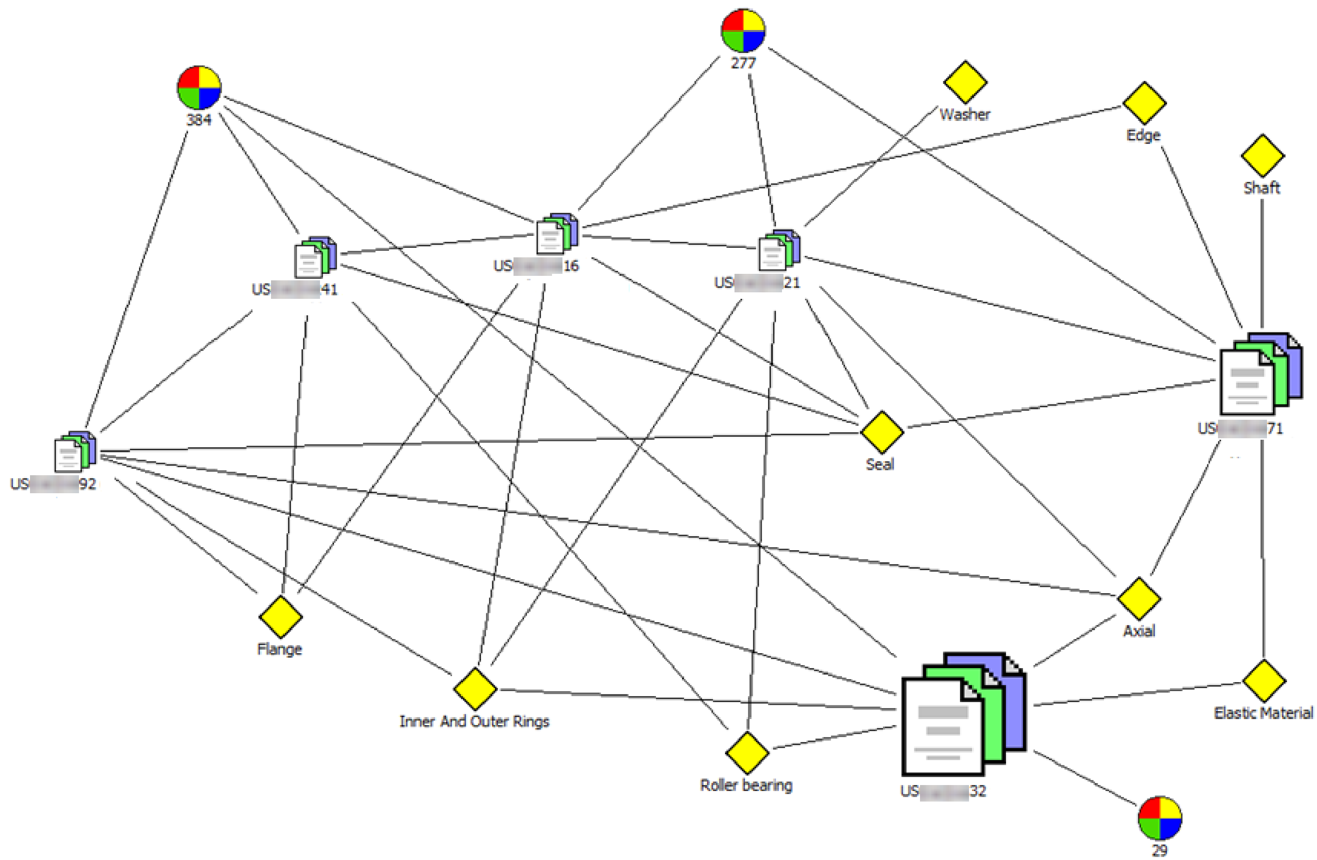
- Charts/Graphs
- Static/Dynamic
- Links/Relations
- GIS/Maps

– Creativity

- What visual form conveys the intended message?

LINK ANALYSIS

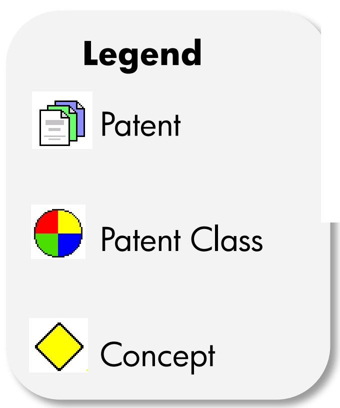
- A form of visual analysis
- Useful when relationships matter

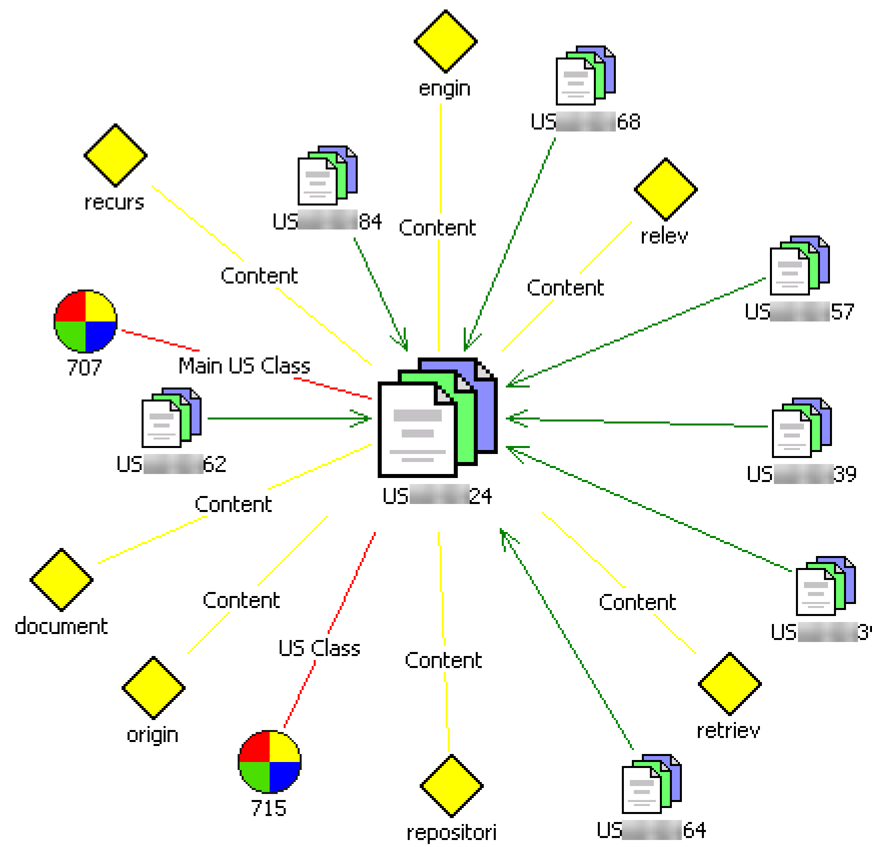


LINK ANALYSIS EXAMPLE

Patent Similarity Detection

- 1000 patents
- Target Patent: US*****24
  - Class 707, 715
- Known similar patents
  - US*****49
  - US*****59
    - Class 709
- 997 additional patents
  - 700 classes

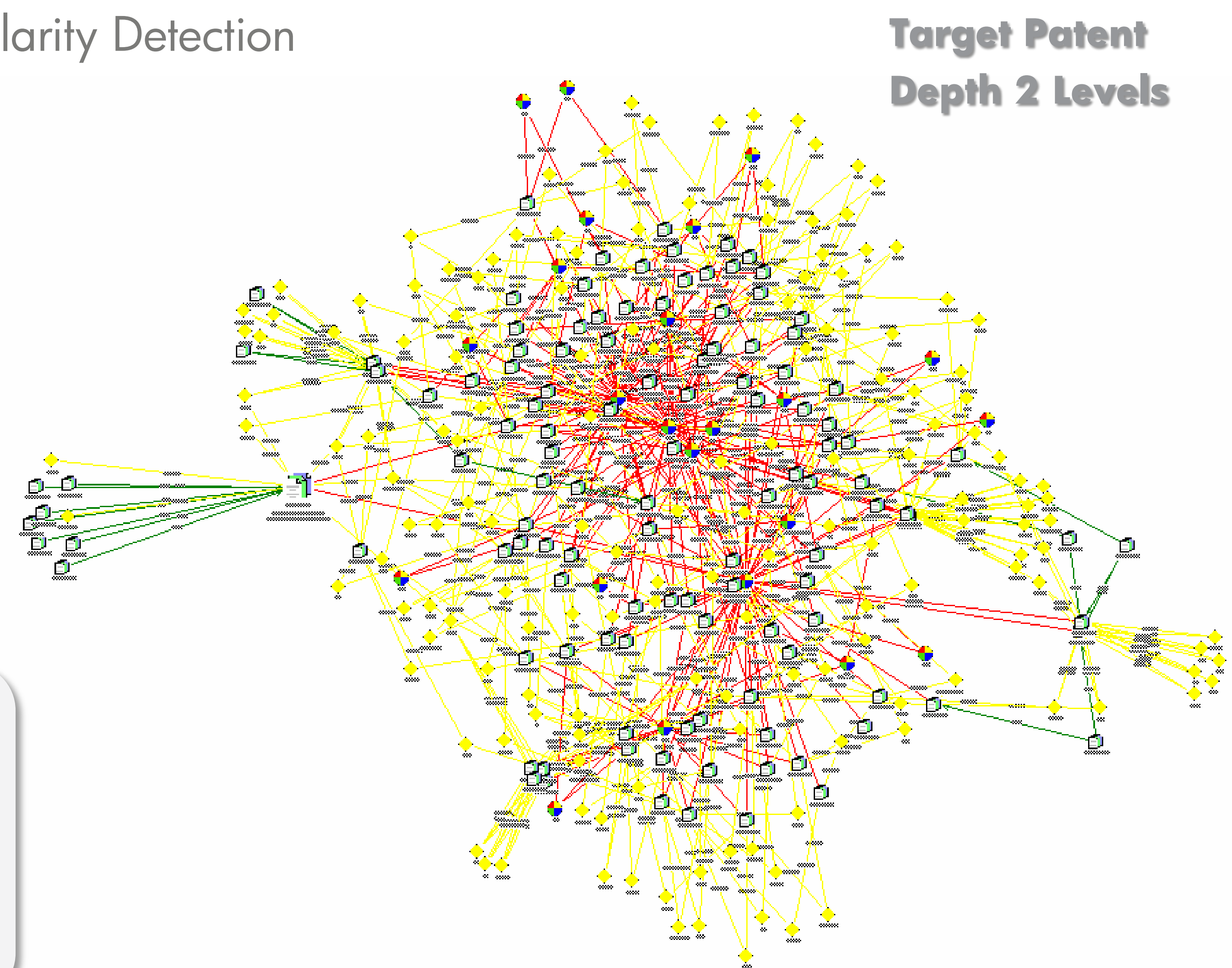


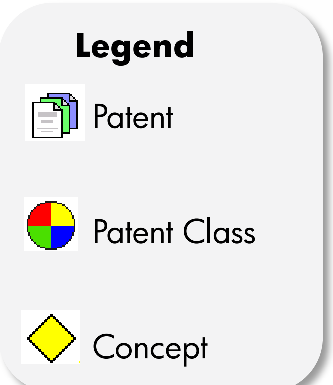


Target Patent
Depth 1 Level

LINK ANALYSIS EXAMPLE

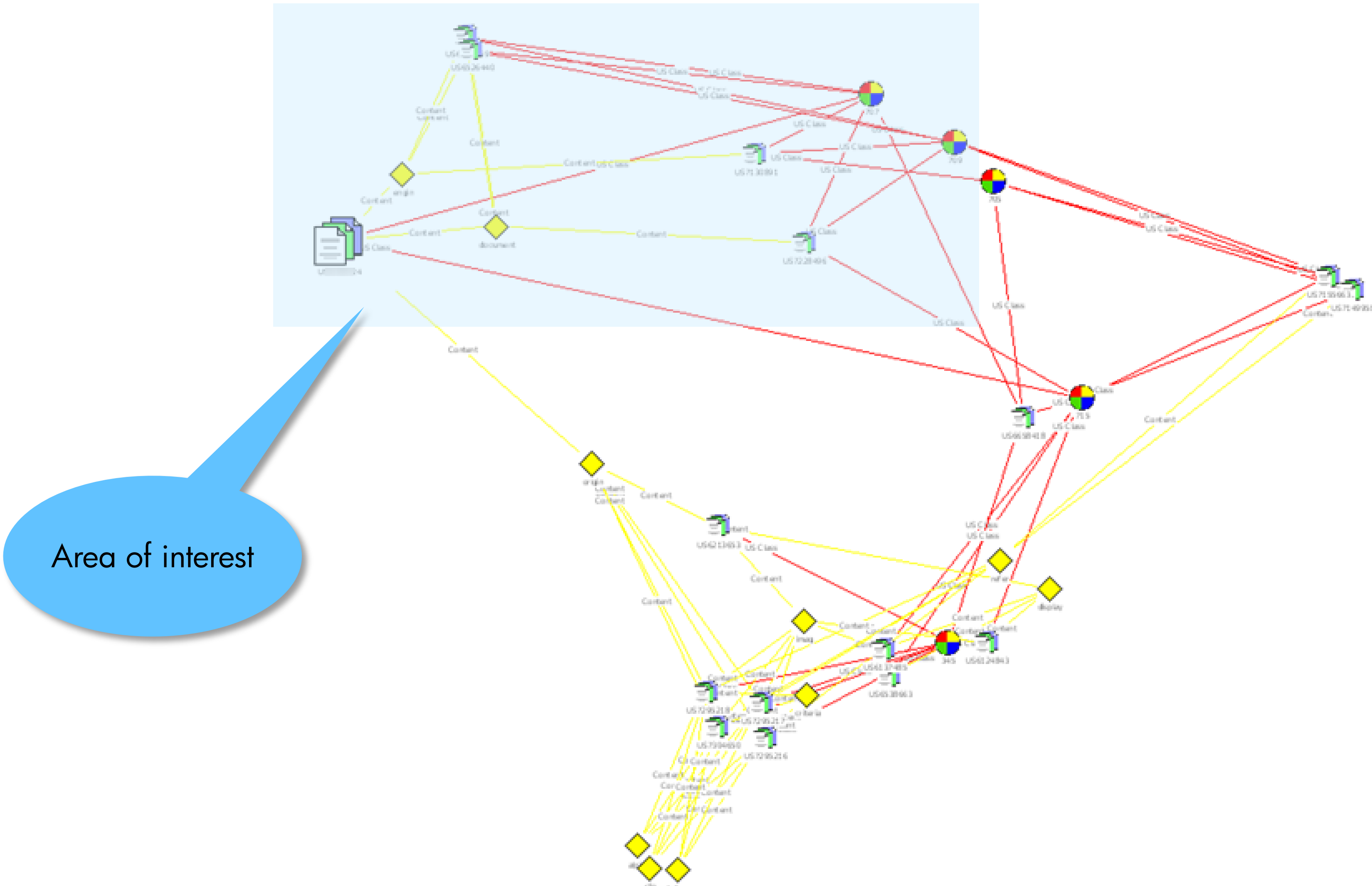
Patent Similarity Detection





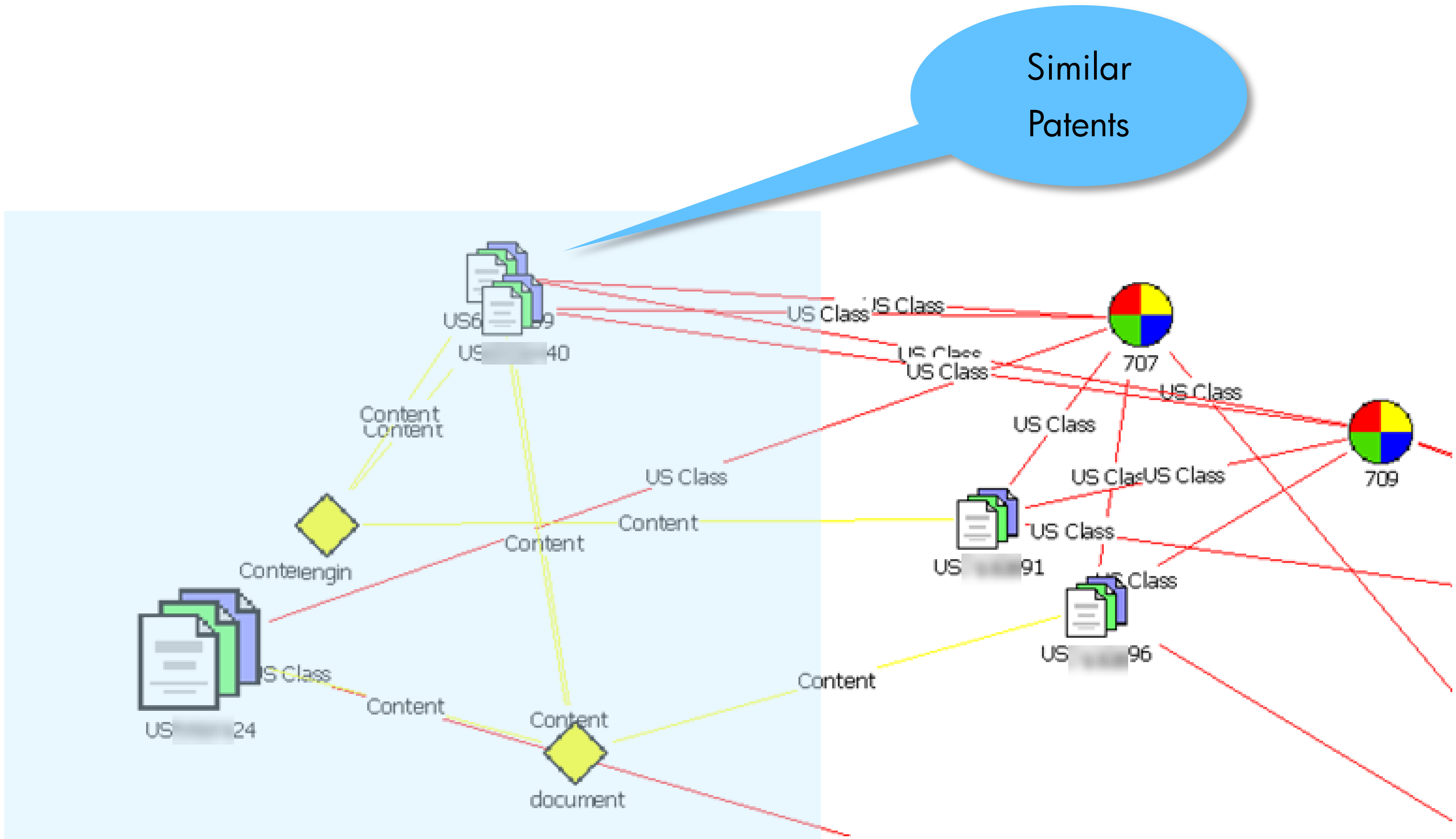
LINK ANALYSIS EXAMPLE

Patent Similarity Detection



LINK ANALYSIS EXAMPLE

Patent Similarity Detection



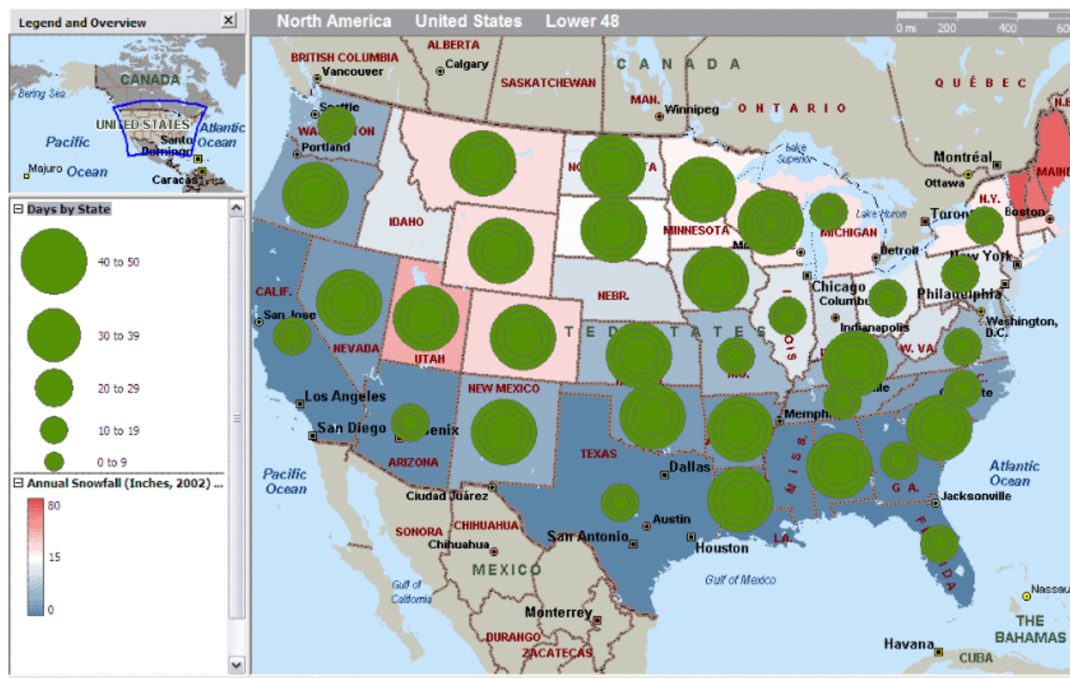
GEOGRAPHIC INFORMATION SYSTEM (GIS)

– A form of visual data analysis

– Information embedded with a map

– Rich with location-specific data, e.g.,

- Annual Rainfall
- Annual snowfall
- Temperature variance
- Average educational level
- Average income level
- Product sales
- Water consumption
- Number of children per family
- Number of cars per household



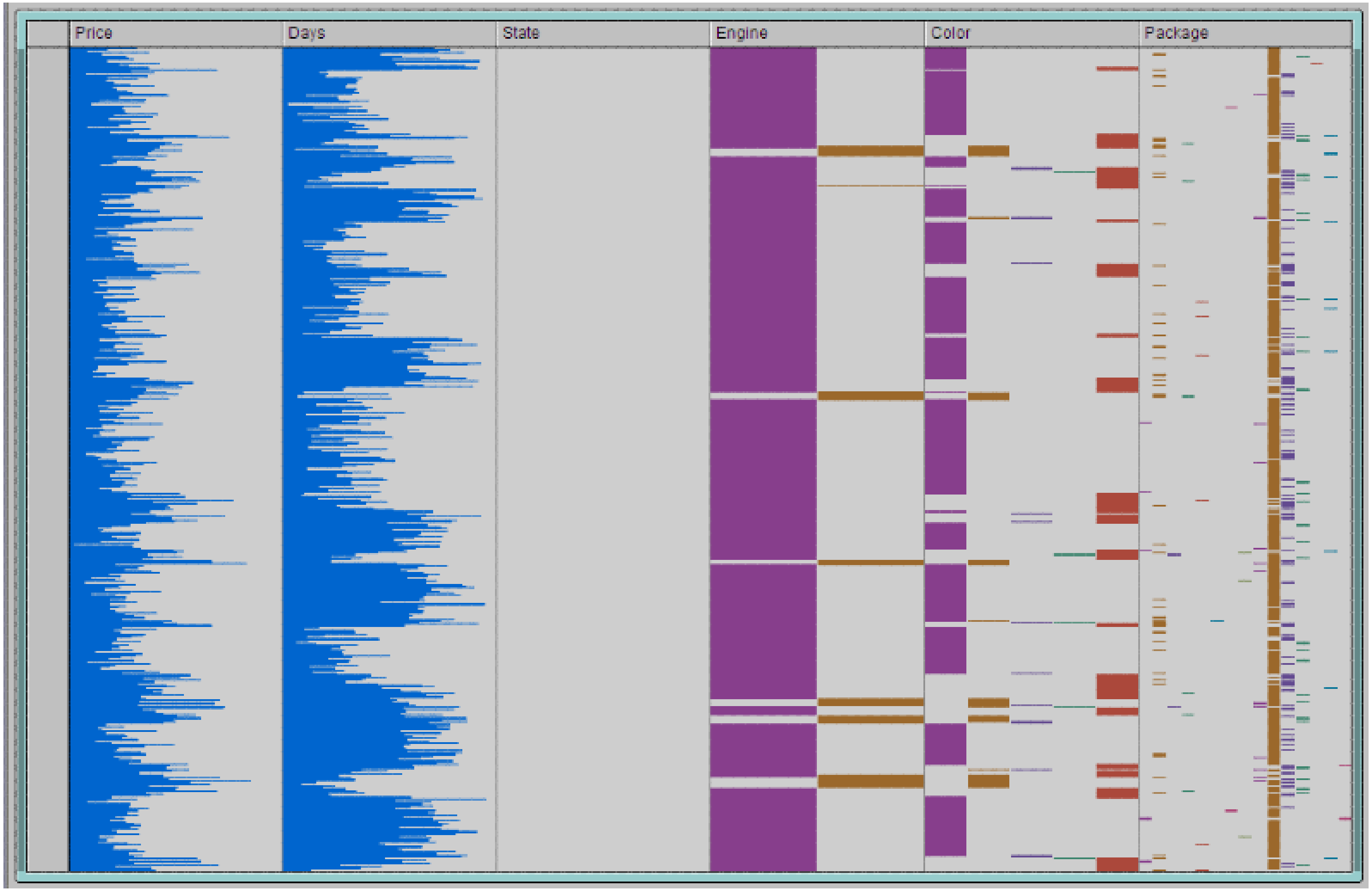
A SAMPLE ANALYSIS

A Product Configuration and Sales Analysis*

- Product: Motor Cycle
- Data:
  - Price
  - Days to Sell
  - State
  - Engine
  - Color
  - Package
- Goals:
  - What option packages will be in demand?
  - What factors can help sell the product faster?

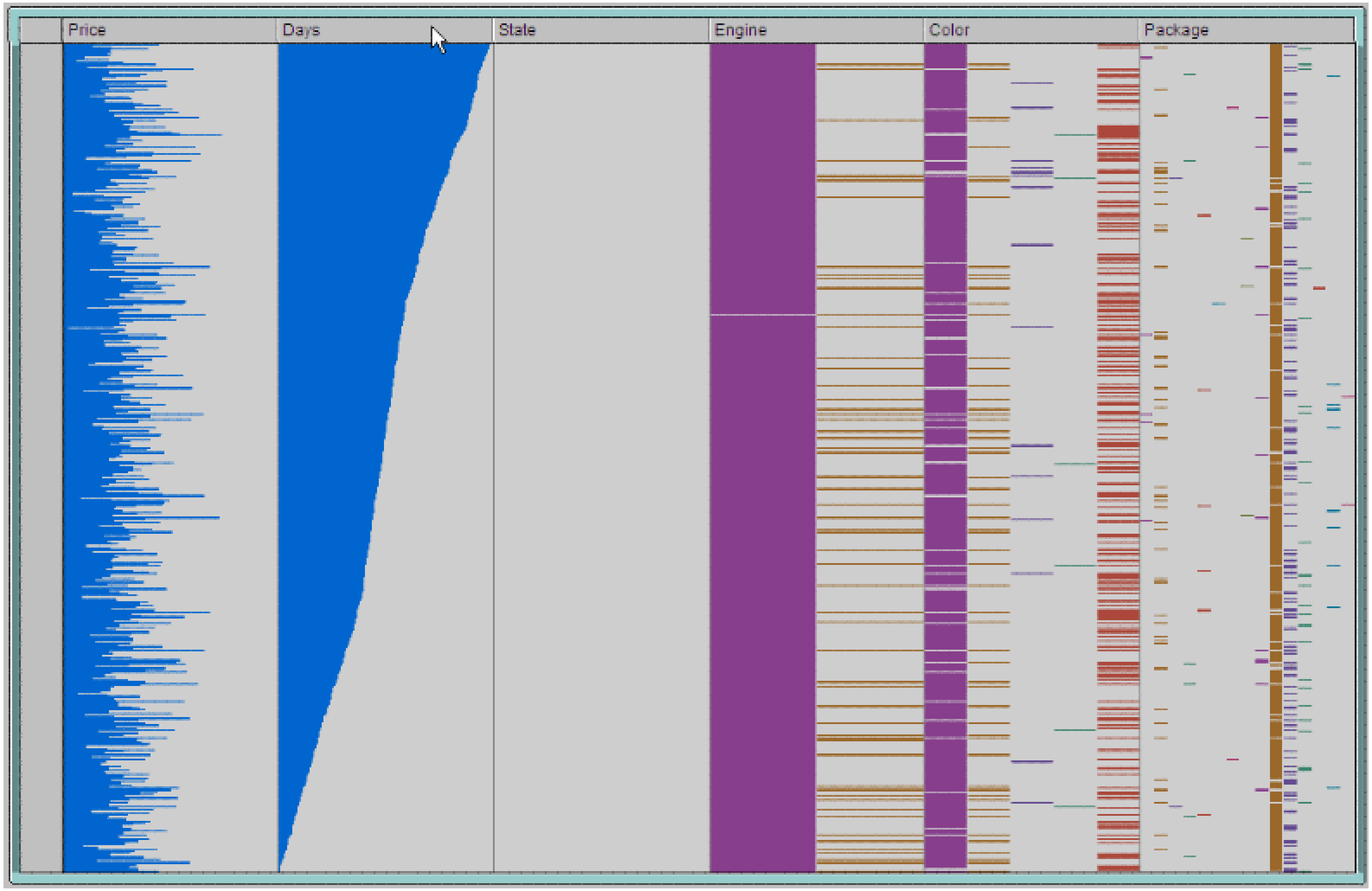
* Hypothetical scenario/data based on an actual case

GENERAL DISTRIBUTION OF DATA

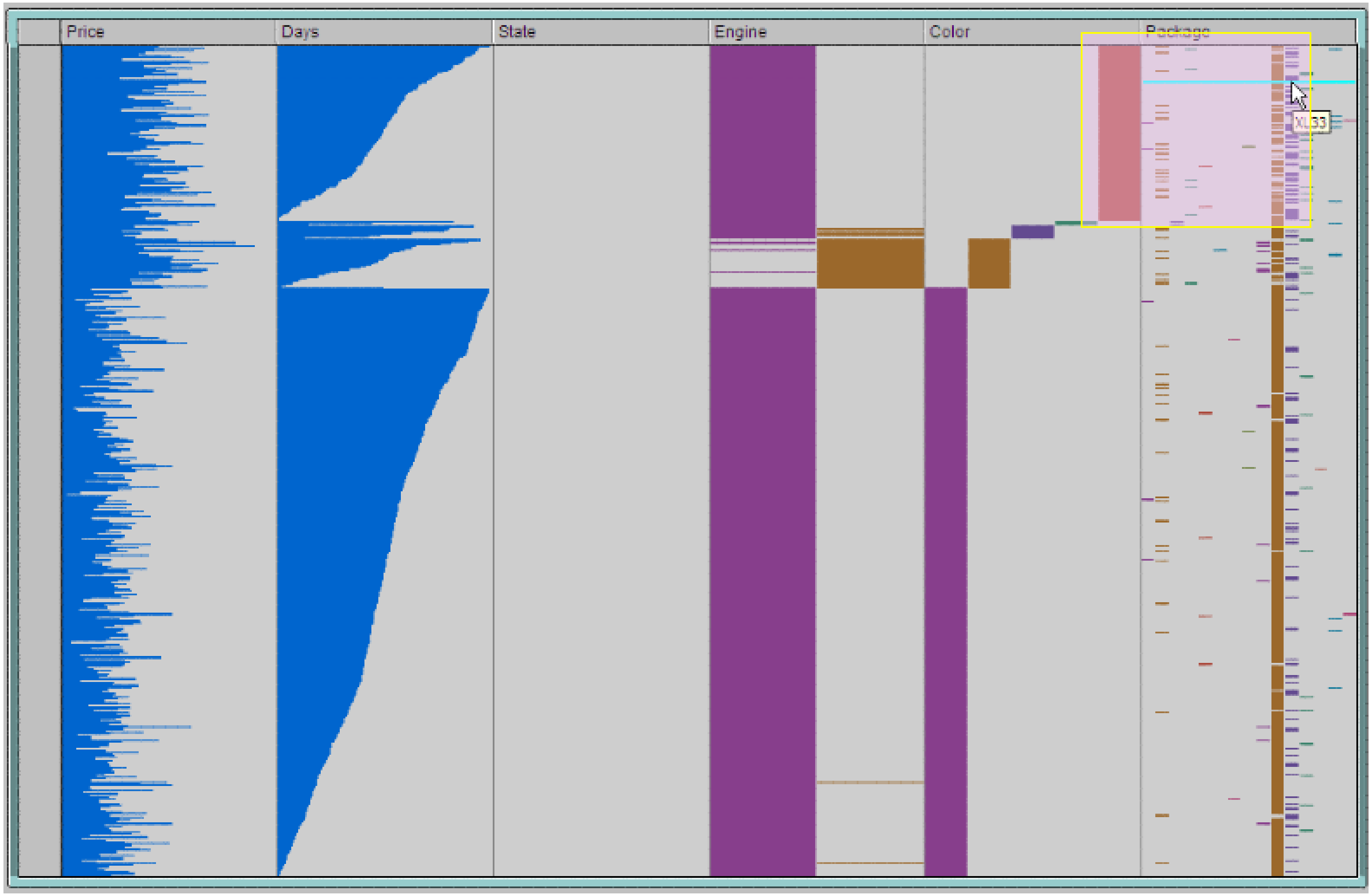


Tool: Inxight Eureka

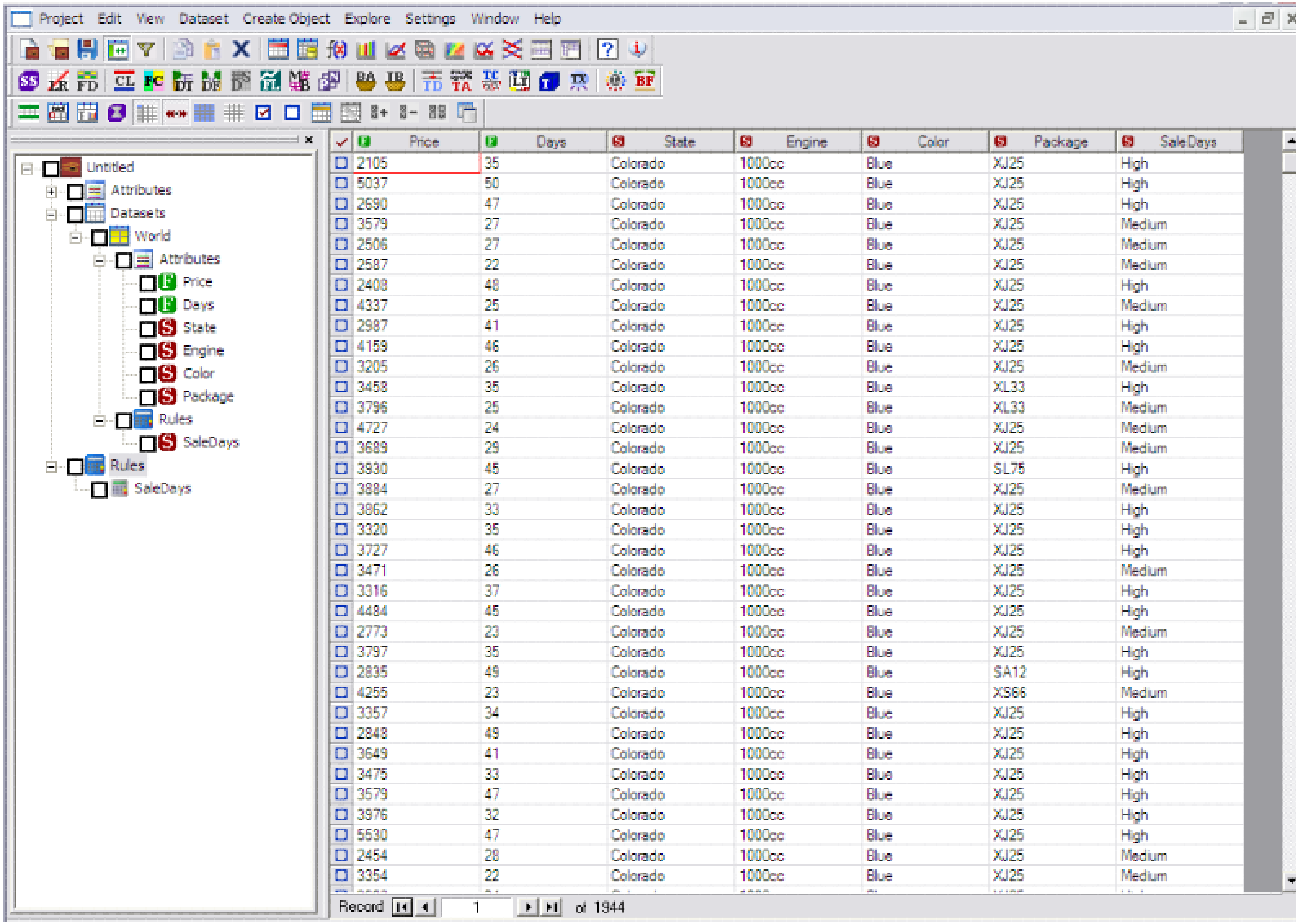
SORTED BY THE NUMBER OF DAYS TO SELL



SORTING BY COLOR IS MORE INSIGHTFUL



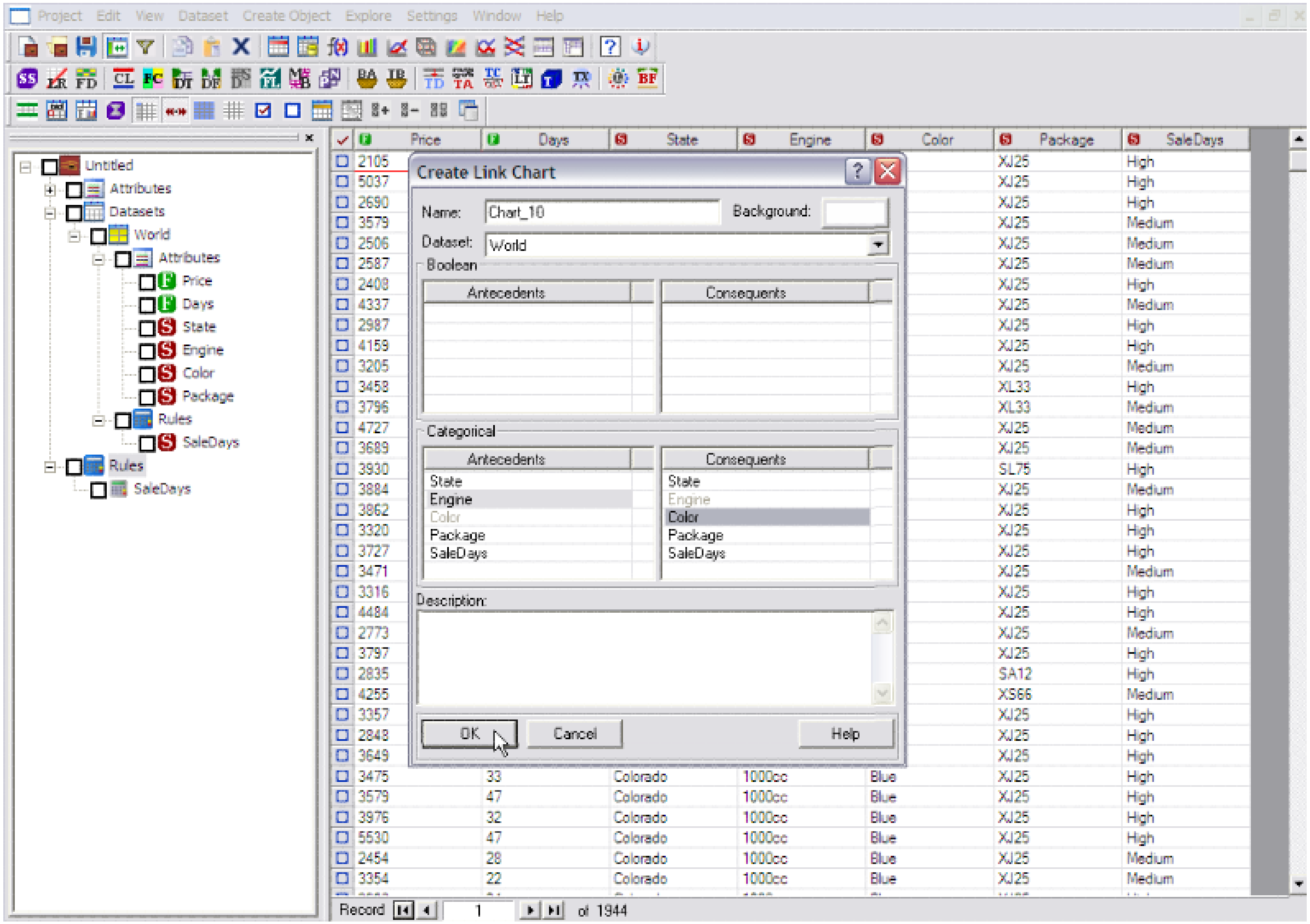
A DIFFERENT TOOL FOR MORE DETAILED ANALYSIS



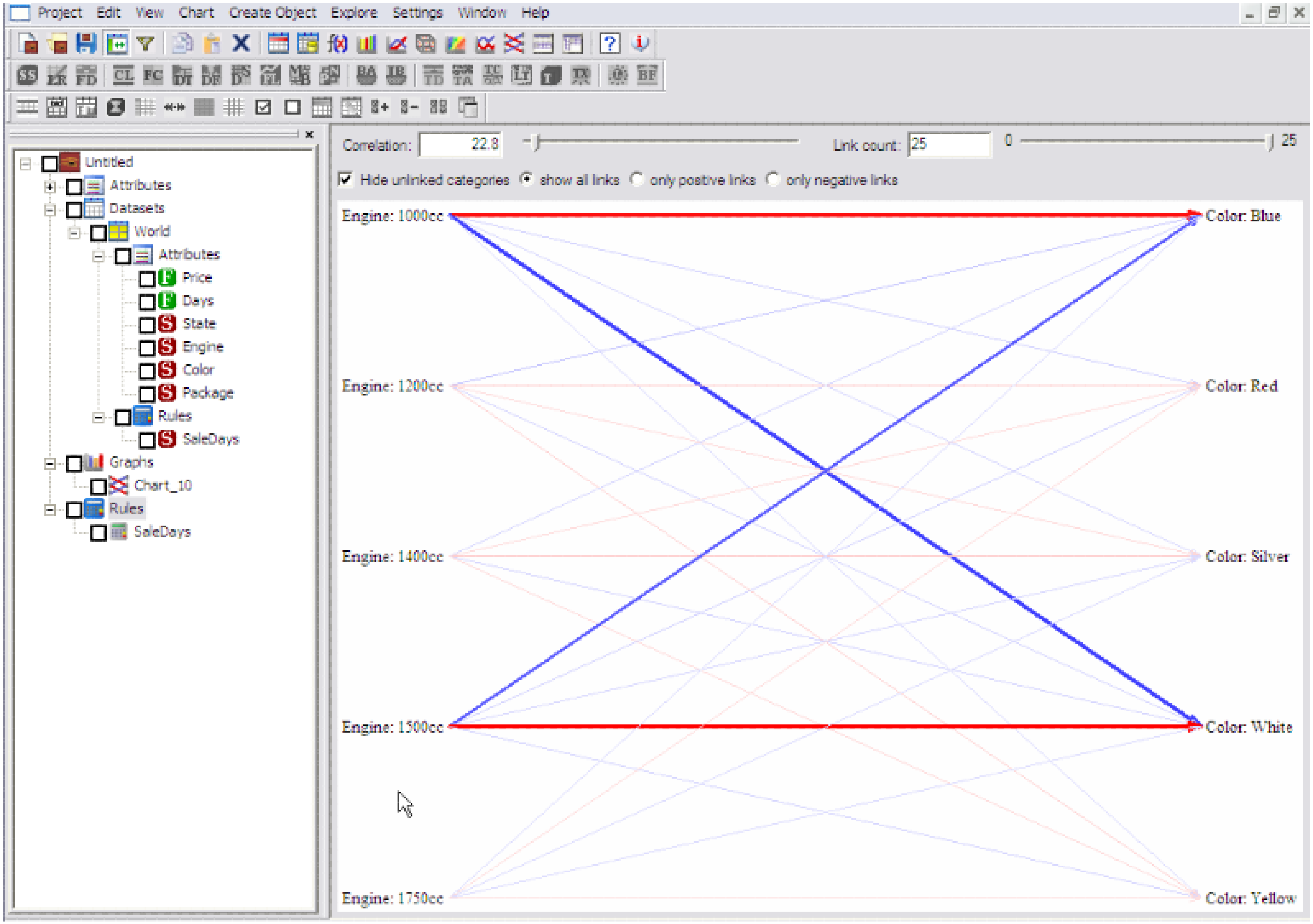
Price	Days	State	Engine	Color	Package	SaleDays
2105	35	Colorado	1000cc	Blue	XJ25	High
5037	50	Colorado	1000cc	Blue	XJ25	High
2690	47	Colorado	1000cc	Blue	XJ25	High
3579	27	Colorado	1000cc	Blue	XJ25	Medium
2506	27	Colorado	1000cc	Blue	XJ25	Medium
2587	22	Colorado	1000cc	Blue	XJ25	Medium
2408	48	Colorado	1000cc	Blue	XJ25	High
4337	25	Colorado	1000cc	Blue	XJ25	Medium
2987	41	Colorado	1000cc	Blue	XJ25	High
4159	46	Colorado	1000cc	Blue	XJ25	High
3205	26	Colorado	1000cc	Blue	XJ25	Medium
3458	35	Colorado	1000cc	Blue	XL33	High
3796	25	Colorado	1000cc	Blue	XL33	Medium
4727	24	Colorado	1000cc	Blue	XJ25	Medium
3689	29	Colorado	1000cc	Blue	XJ25	Medium
3930	45	Colorado	1000cc	Blue	SL75	High
3884	27	Colorado	1000cc	Blue	XJ25	Medium
3862	33	Colorado	1000cc	Blue	XJ25	High
3320	35	Colorado	1000cc	Blue	XJ25	High
3727	46	Colorado	1000cc	Blue	XJ25	High
3471	26	Colorado	1000cc	Blue	XJ25	Medium
3316	37	Colorado	1000cc	Blue	XJ25	High
4484	45	Colorado	1000cc	Blue	XJ25	High
2773	23	Colorado	1000cc	Blue	XJ25	Medium
3797	35	Colorado	1000cc	Blue	XJ25	High
2835	49	Colorado	1000cc	Blue	SA12	High
4255	23	Colorado	1000cc	Blue	XS66	Medium
3357	34	Colorado	1000cc	Blue	XJ25	High
2848	49	Colorado	1000cc	Blue	XJ25	High
3649	41	Colorado	1000cc	Blue	XJ25	High
3475	33	Colorado	1000cc	Blue	XJ25	High
3579	47	Colorado	1000cc	Blue	XJ25	High
3976	32	Colorado	1000cc	Blue	XJ25	High
5530	47	Colorado	1000cc	Blue	XJ25	High
2454	28	Colorado	1000cc	Blue	XJ25	Medium
3354	22	Colorado	1000cc	Blue	XJ25	Medium

Tool: Megaputer Polyanalyst

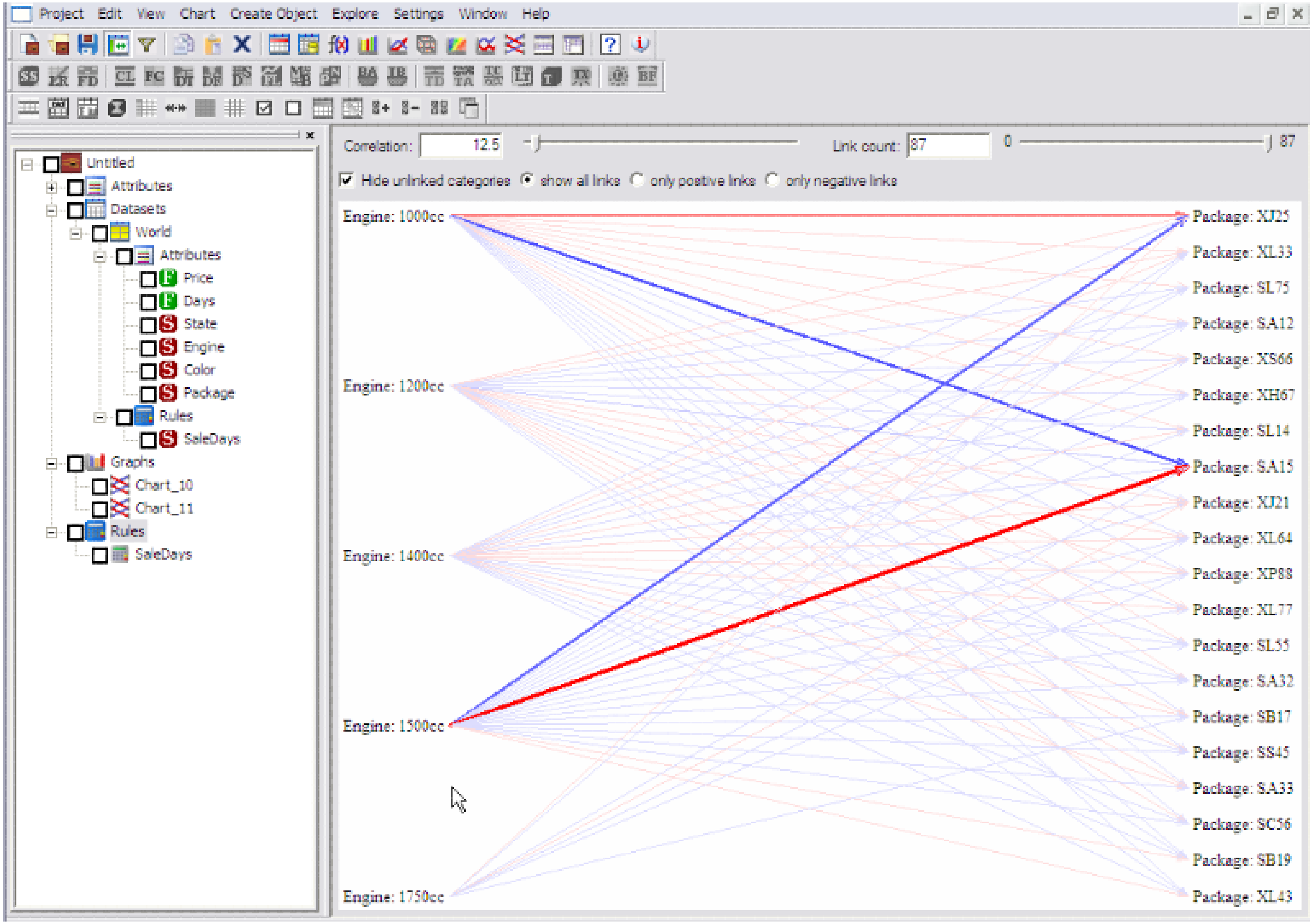
ENGINE AND COLOR



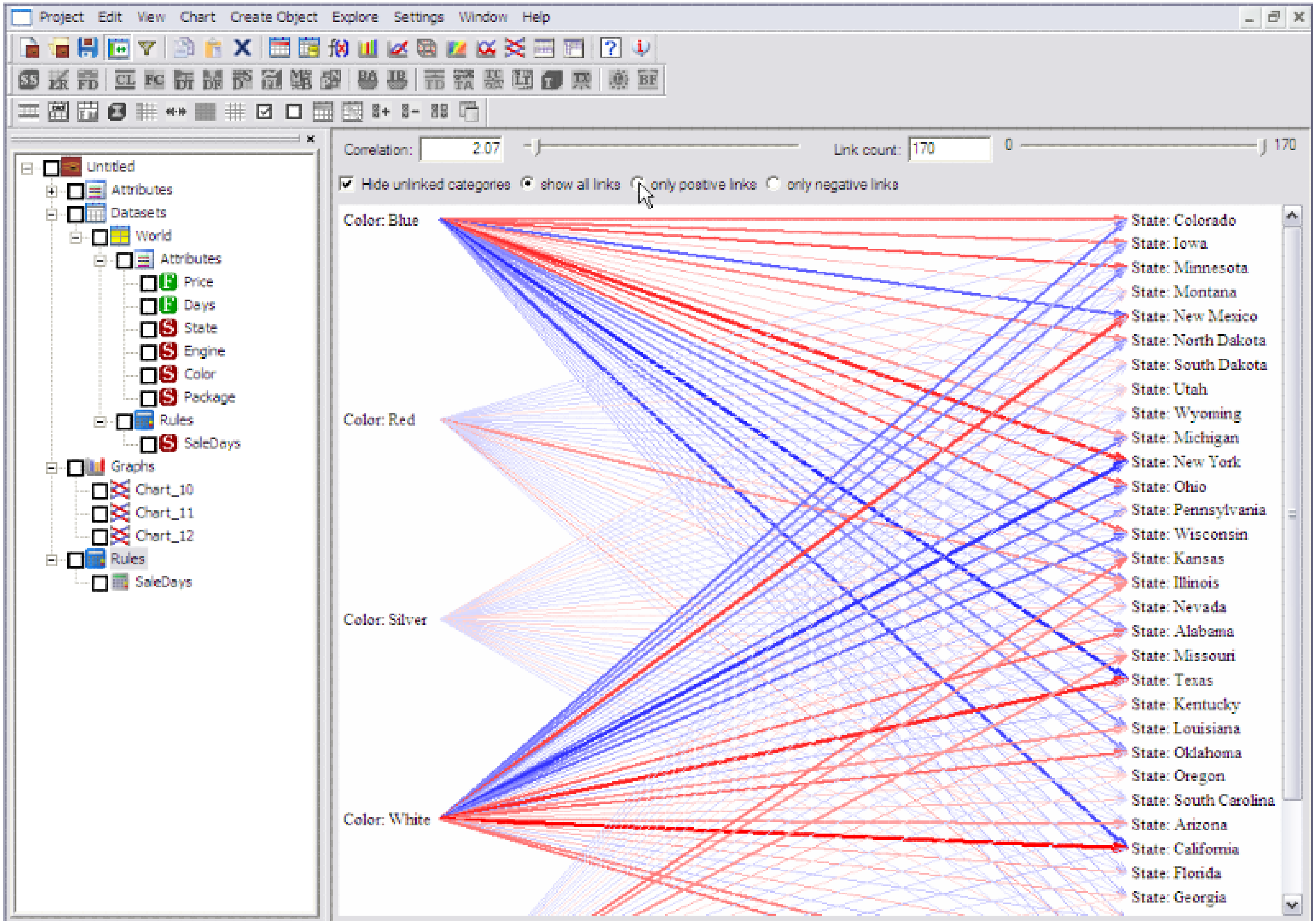
ENGINE AND COLOR



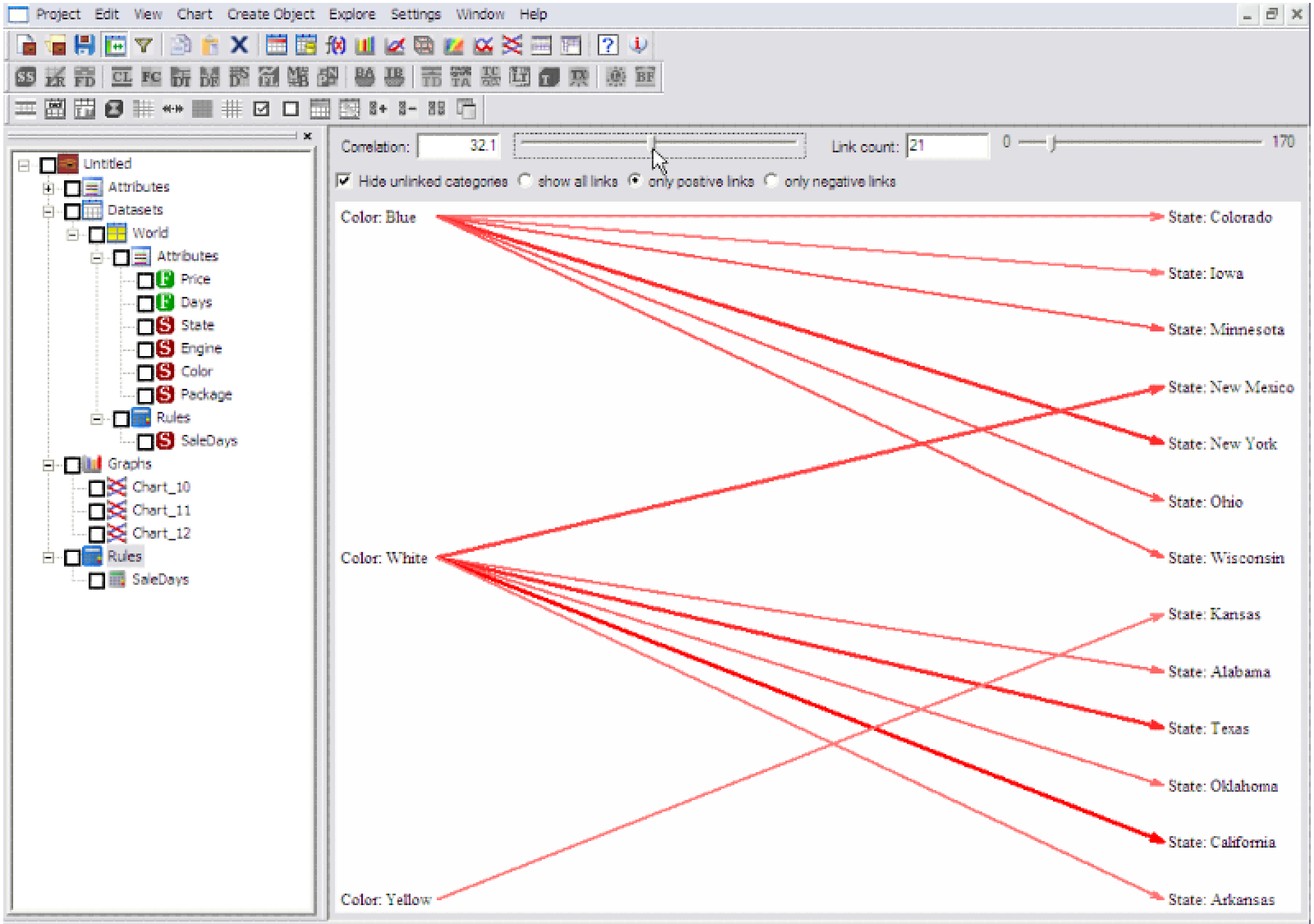
ENGINE AND PACKAGE



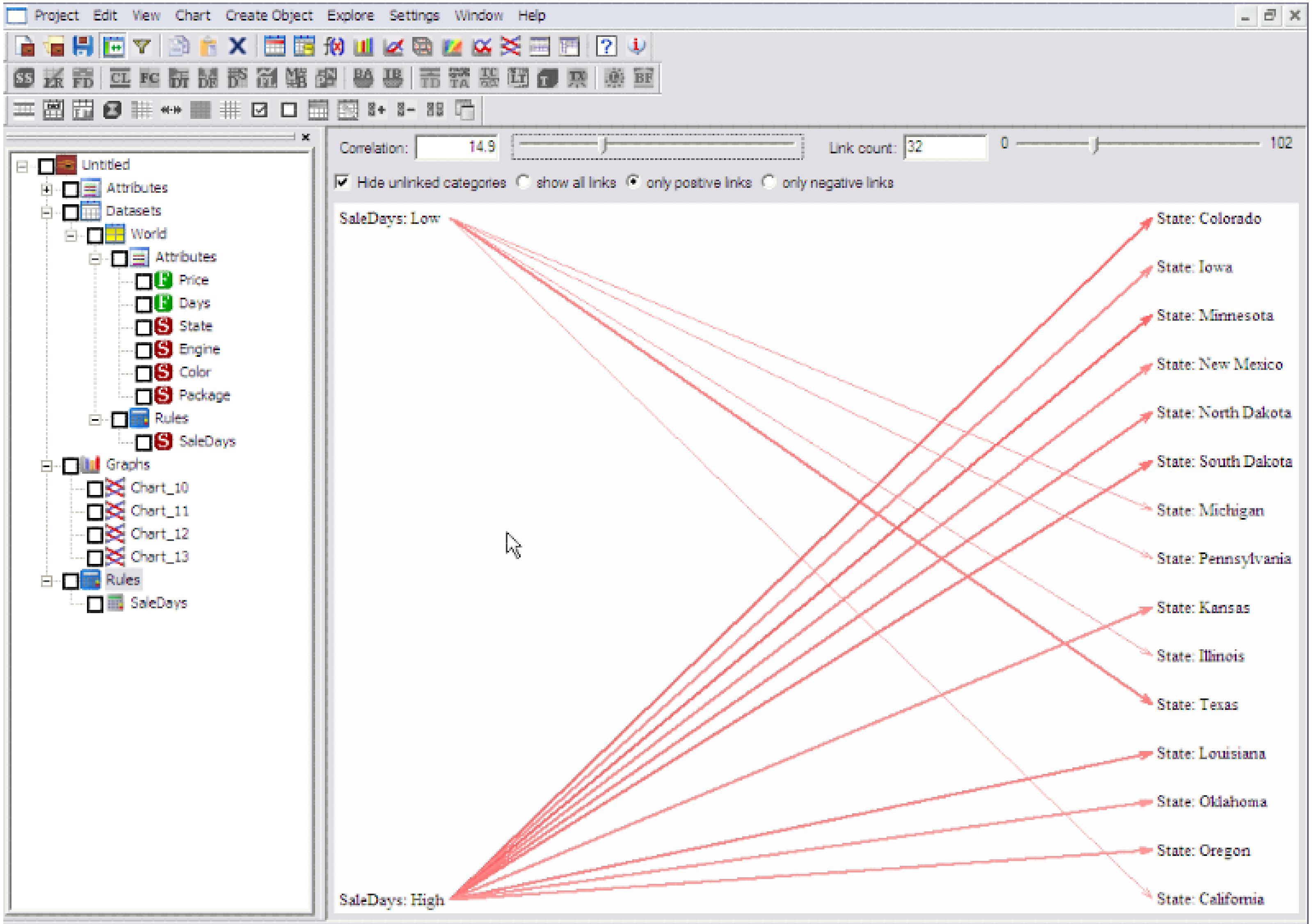
COLOR AND STATE



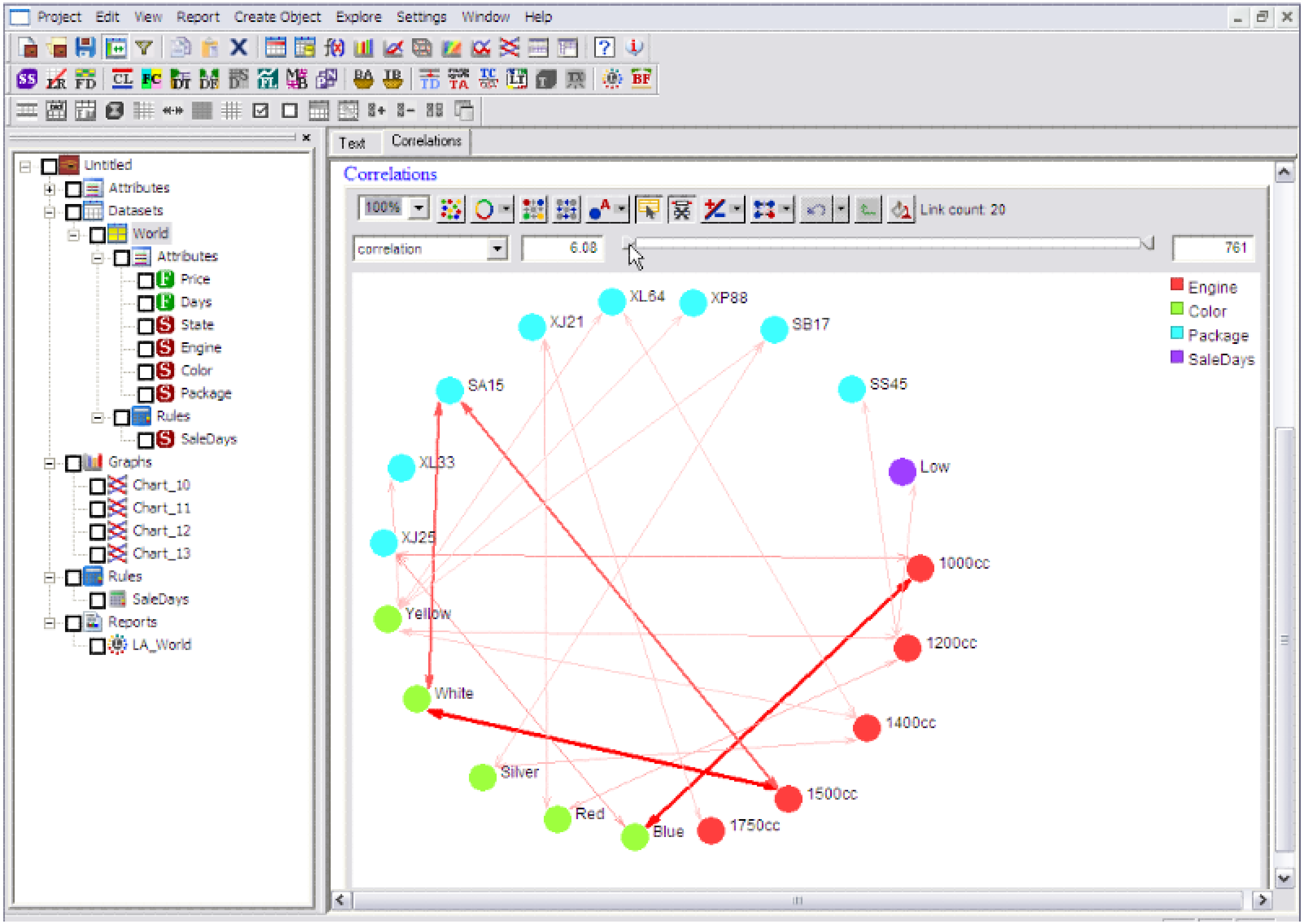
COLOR AND STATE



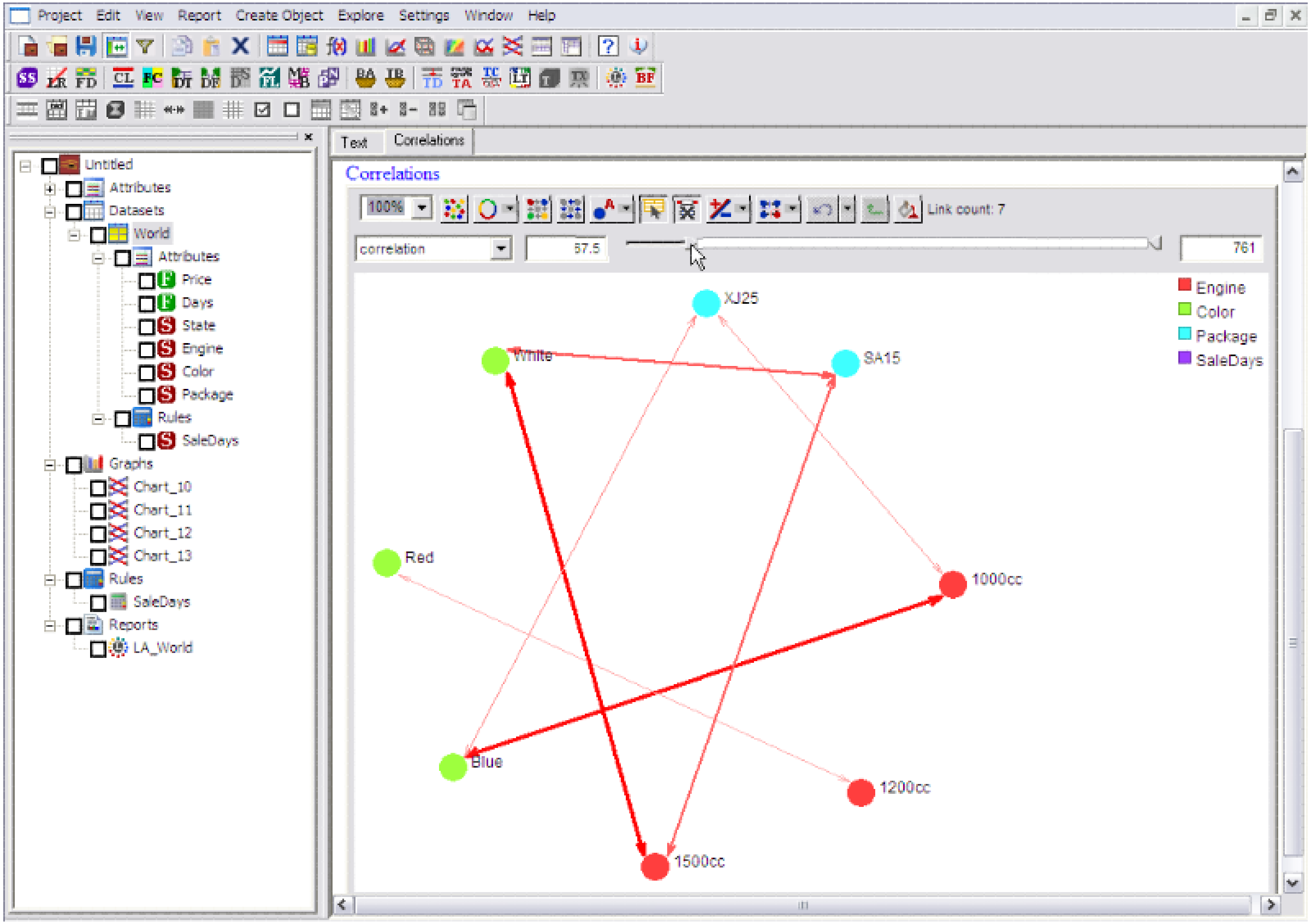
DAYS TO SELL AND STATES



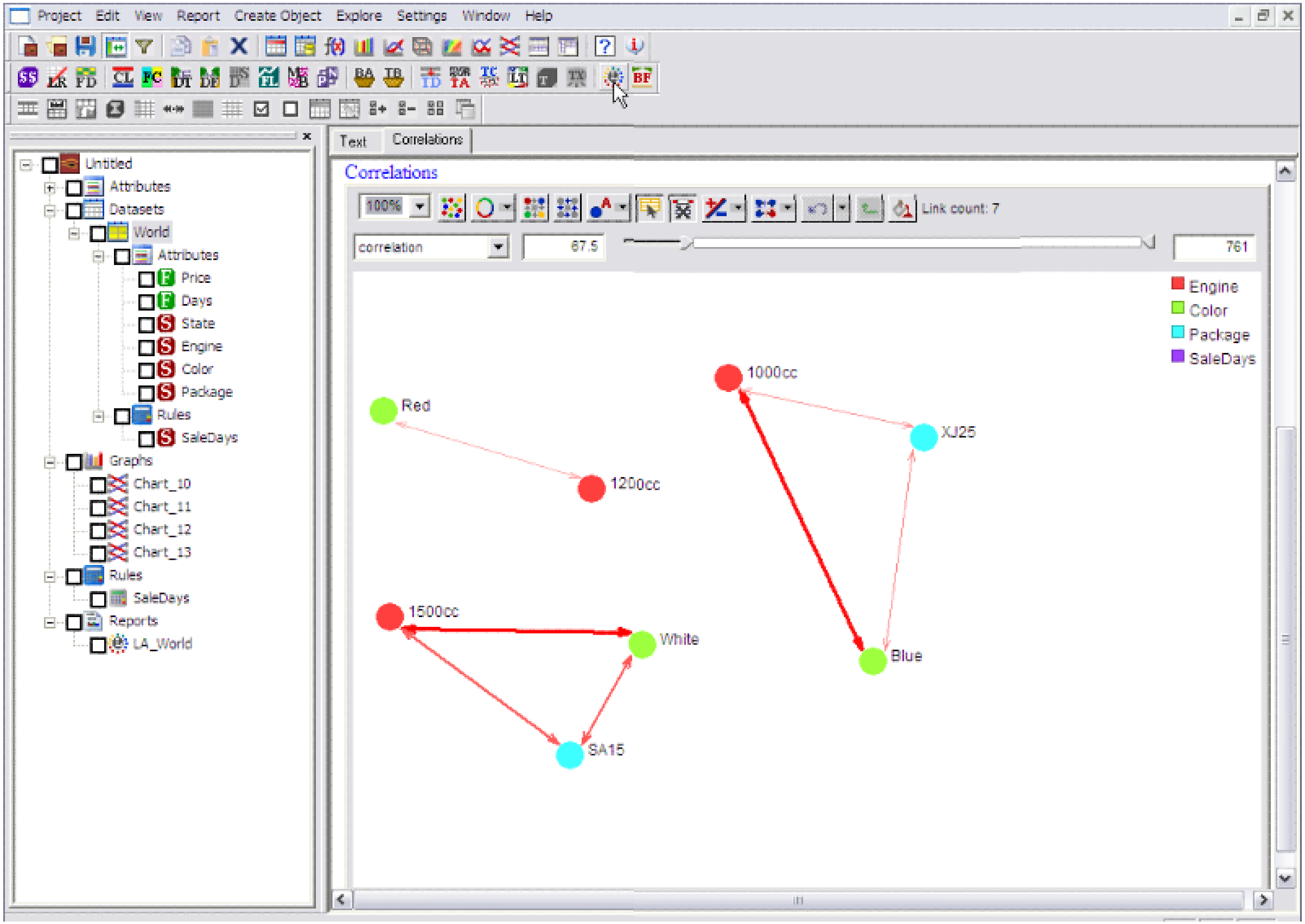
MULTIPLE VARIABLES



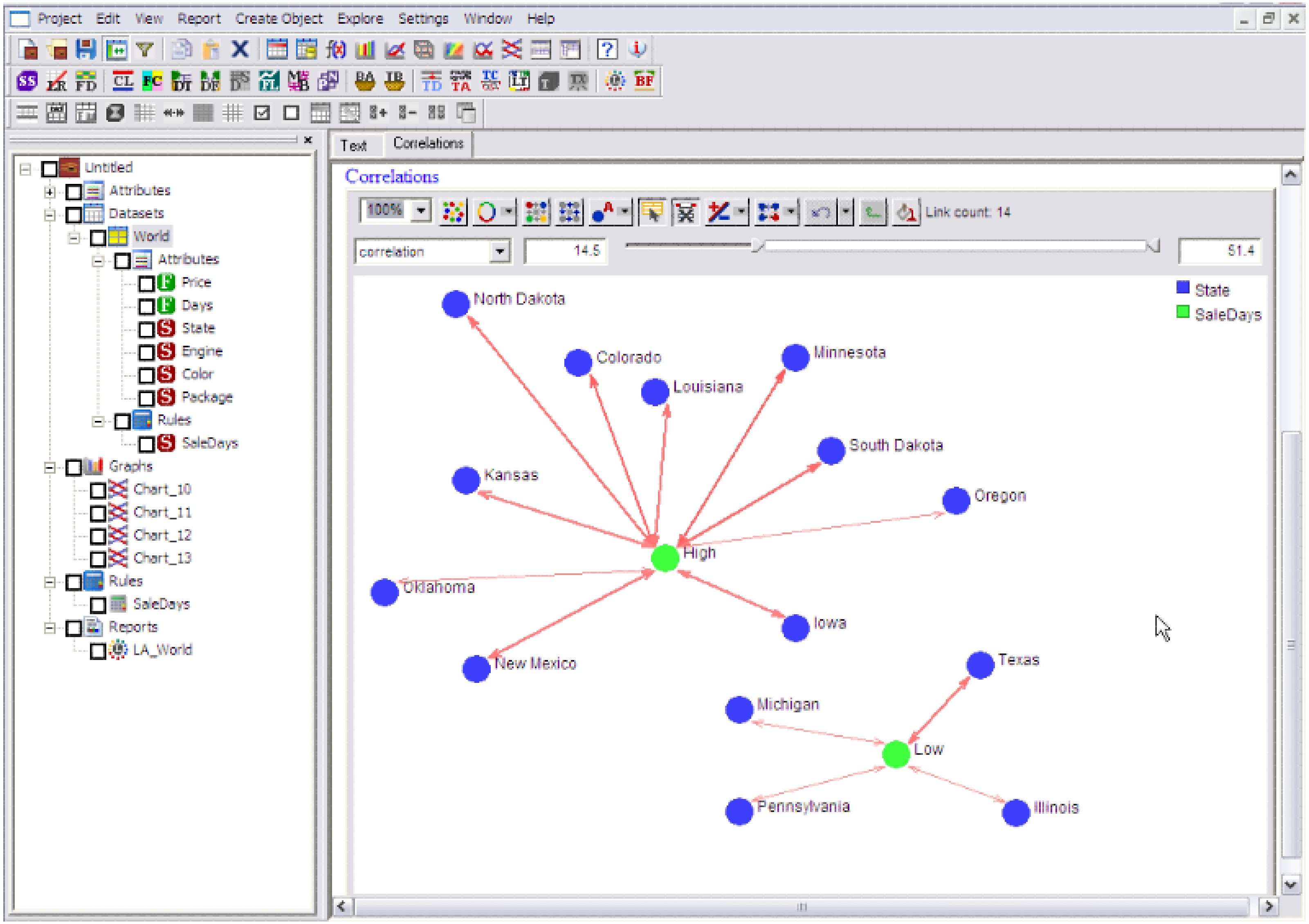
MULTIPLE VARIABLES WITH STRONGER CONNECTIONS



CLUSTERS


Project Edit View Report Create Object Explore Settings Window Help
Text
Correlations
Untitled
Attributes
Datasets
World
Attributes
Price
Days
State
Engine
Color
Package
Rules
SaleDays
Graphs
Chart_10
Chart_11
Chart_12
Chart_13
Rules
SaleDays
Reports
LA_World
Correlations
100%
Link count: 7
correlation
67.5
761
Engine
Color
Package
SaleDays
1000cc
Red
XJ25
1200cc
1500cc
White
Blue
SA15

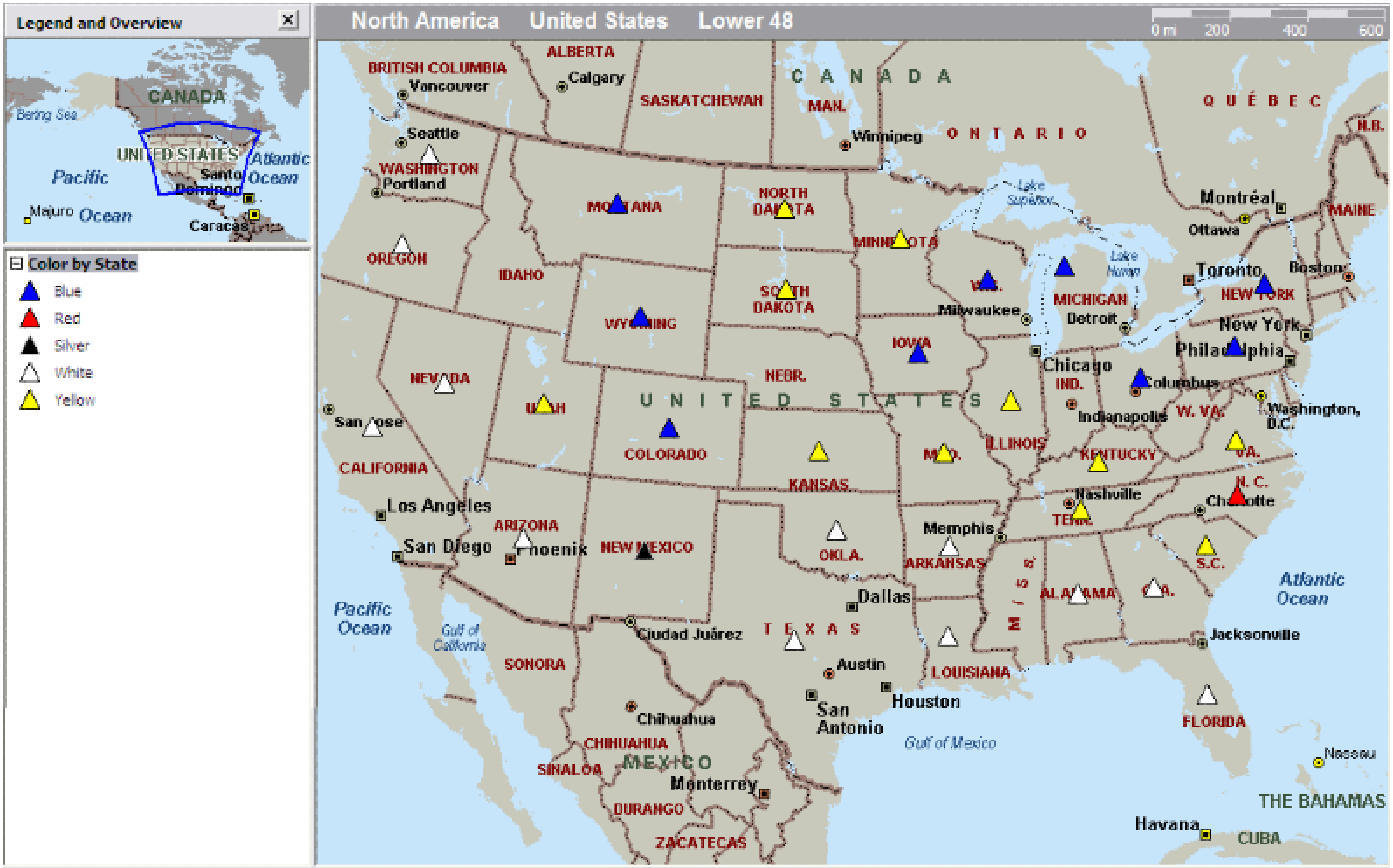
DAYS TO SELL AND STATE CLUSTERS



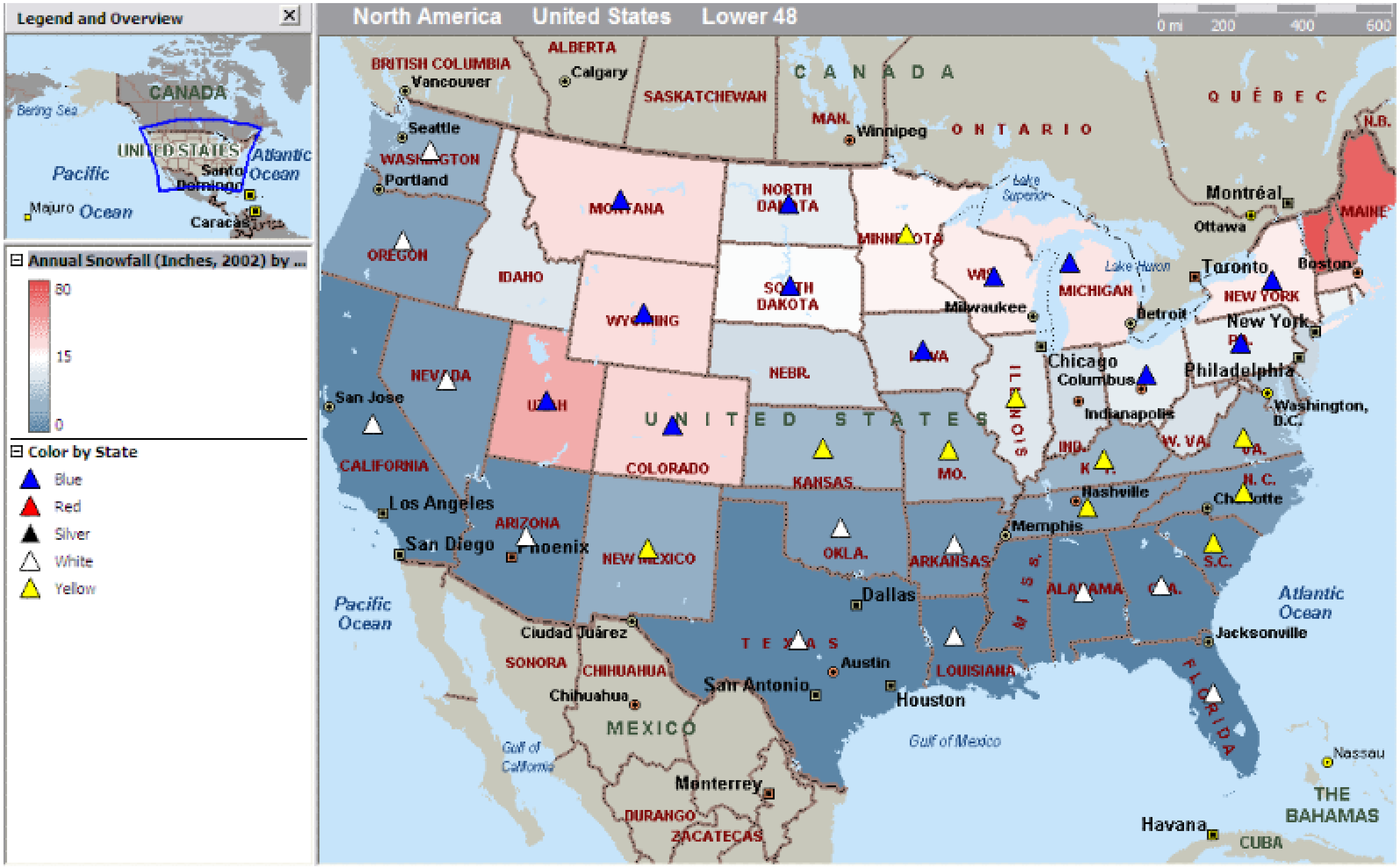
MAP

Tool: Microsoft MapPoint

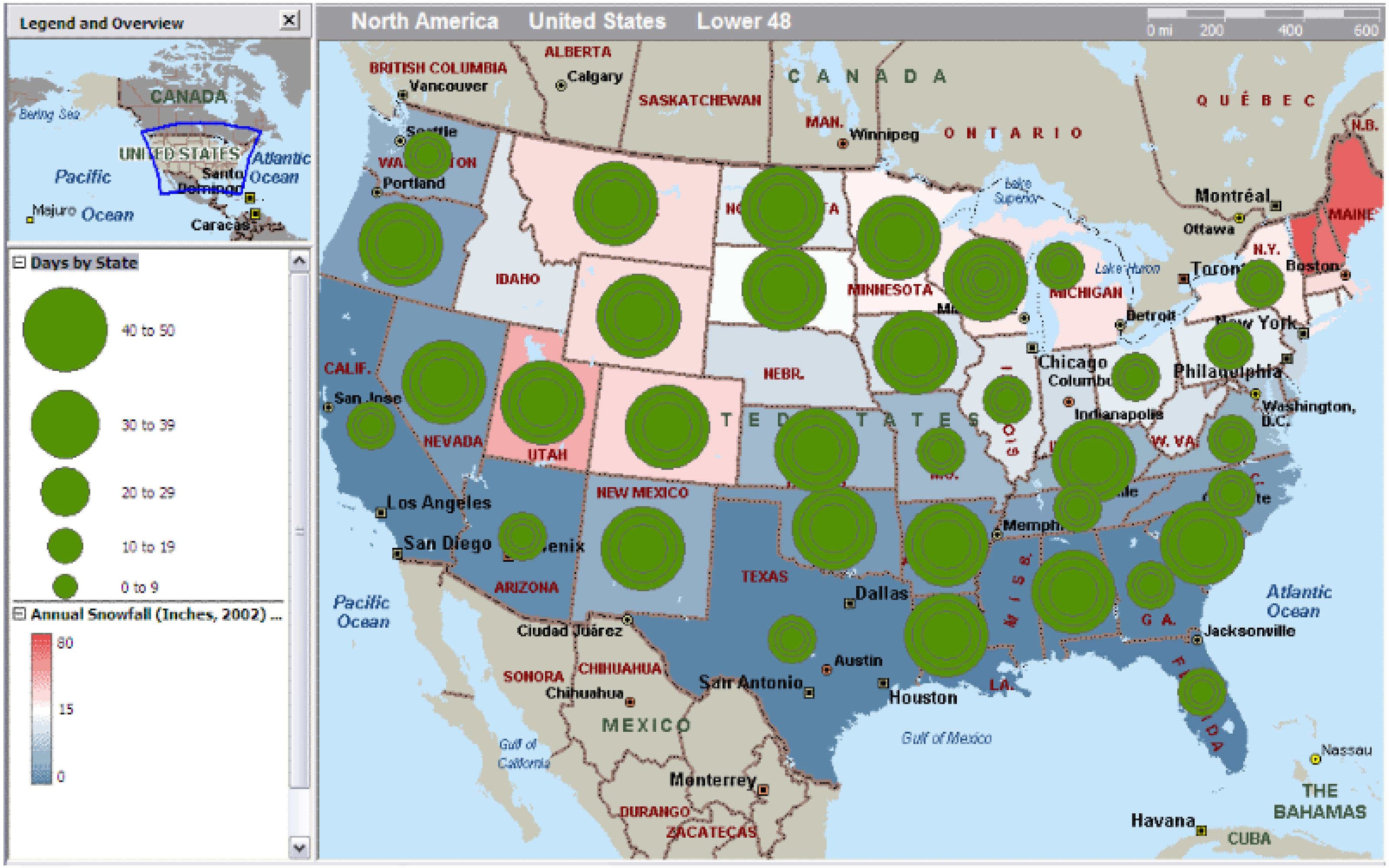
SALES BY COLOR IN STATES



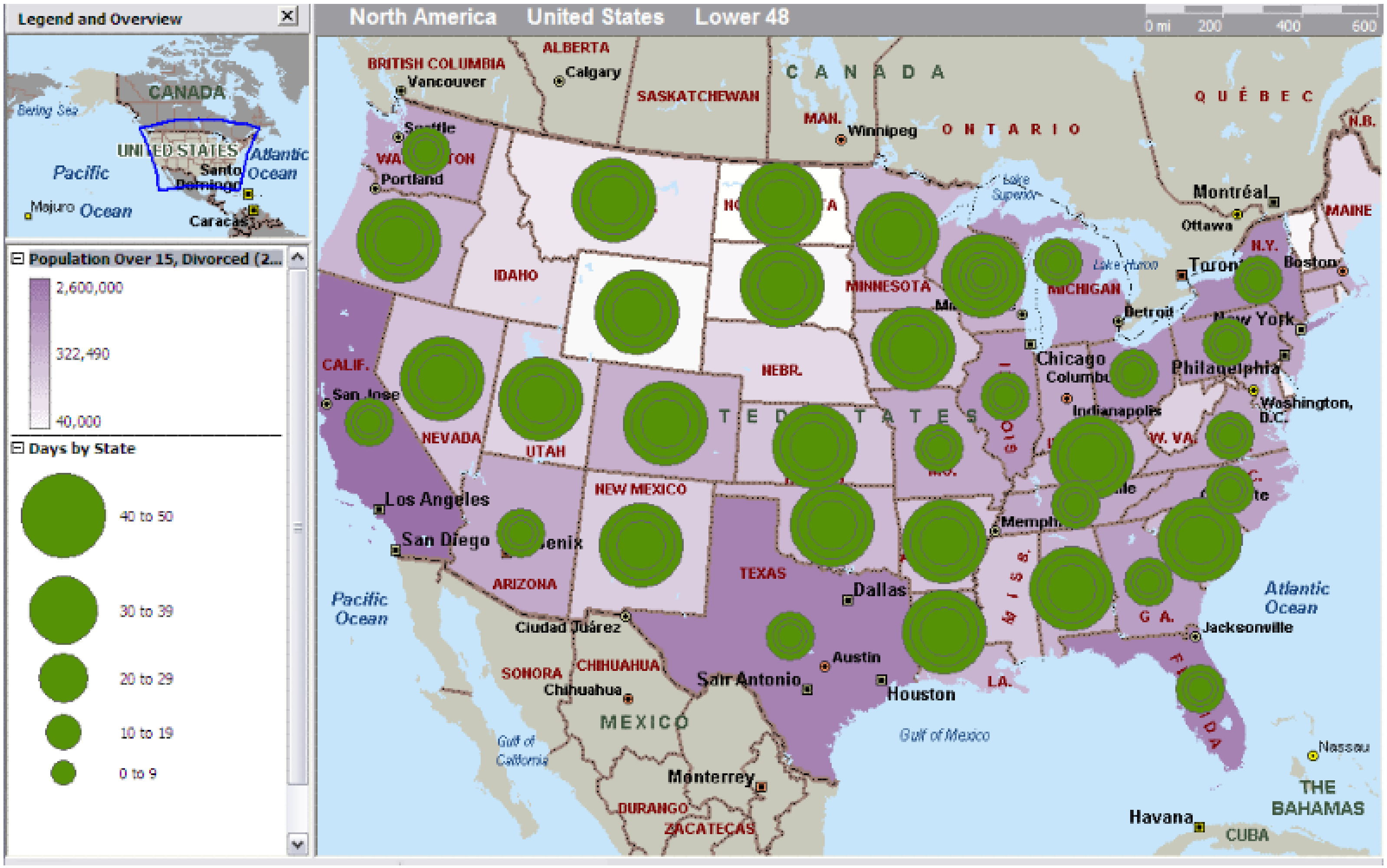
ADDING RATES OF SNOWFALL



DAYS TO SELL AND STATE



DAYS TO SELL AND DIVORCE RATE



CONCLUSIONS*

– Regions with High Snowfall have a greater demand for motor cycles that are Blue, having a Small Engine, and Option Package XJ25

– Regions with Low Snowfall have a greater demand for motor cycles that are White, having a Large Engine, and Option Package SA15

– Motor Cycles sell faster in regions with a Higher Divorce rate

* Hypothetical scenario/data based on an actual case

INNOVATION WORKSHOP

OBJECTIVE

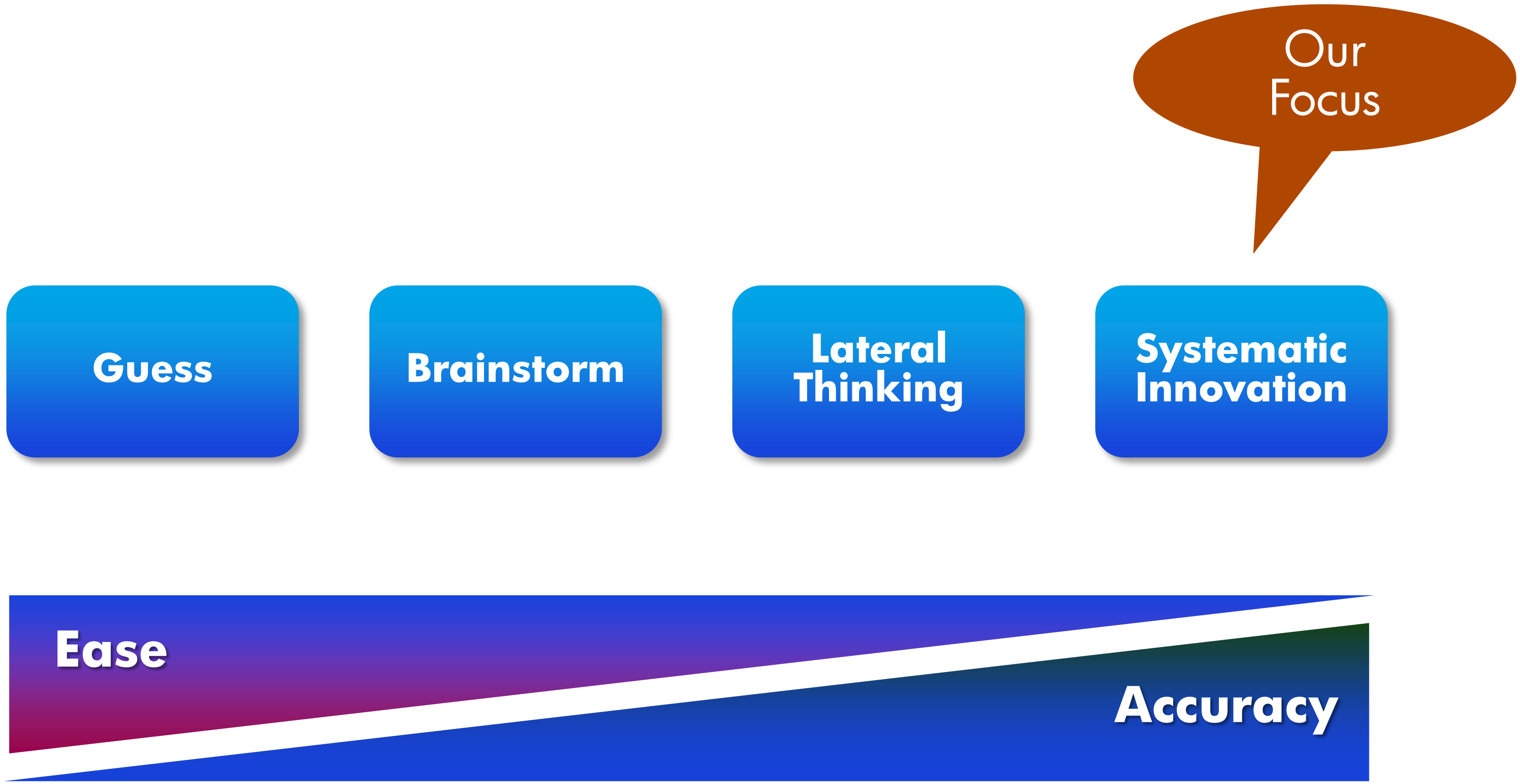
– Setting: Group Innovation

– Focus: A data management problem related to analytics

– Scope: Trends, and the next big thing

– Approach: Facilitated systematic innovation

INNOVATION METHODOLOGIES



SYSTEMATIC INNOVATION

Innovation is Imperative

– Cost reduction

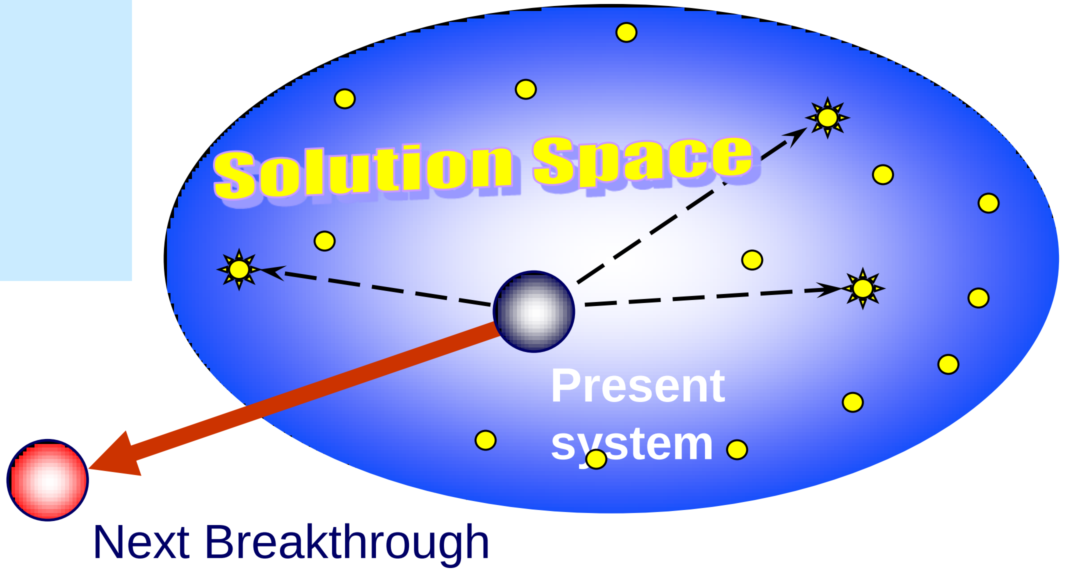
– New markets and profitable growth

Challenges of Innovation

- Risky
- Expensive
- Trial & Error
- Uncertain outcome

What Innovation Needs

- A process that is:
  - Systematic
  - Predictable
  - Limited investment



TRIZ

– Theory of Inventive Problem Solving
– A methodology for systematic innovation
  • Developed in the former Soviet Union, circa. 1940s
  • Introduced to the West in 1990s
– Developed via analysis of hundreds of thousands of patented inventions
  • Evidence-based
– A set of problem solving tools
  • Not an engine
– Primarily directed at engineering problems
  • Successfully used to solve challenging IT problems
– [Partially] adapted to IT problems: **TRIZ for IT**

SAMPLE IT APPLICATIONS OF TRIZ

#	Target	Problem	TRIZ Outcome	Duration*
1	Transportation	• Infrastructure complexity	• Constancy of IT purpose • 6 specific ideas	4 hours
2	Financial	• Inaccurate estimation • Account closure	• 3 technical ideas • 4 business ideas • **1 potential patent**	1 hour
3	AMOD Strategy Ranking	• Ranking of AMOD strategies	• 1 potential patent	1 hour
4	IT Assessment & OnBoarding	• Higher throughput	• 2 resource management ideas • 2 technical ideas • **1 potential patent**	2 hours
5	Software Estimation	• Accurate estimation of software development projects	• 1 technical idea • 1 business idea	1 hour
6	Outsourcing Contracts	• IT contract flexibility	• Uncertainty management • 4 specific ideas	3 hours

* Duration of actual hands-on analysis

DIFFERENTIATORS / DELIVERABLES

Differentiators

– Highly systematic and structured

– Focuses development along paths most likely to succeed

– Fast

- days/weeks vs. months/years

Deliverables

1. Identification of the next generation breakthrough technologies
2. Specific solution for an existing problem

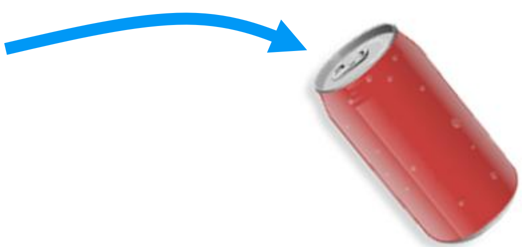
KEY CONCEPTS

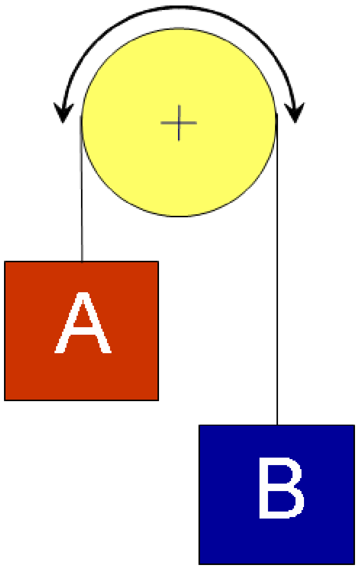
1. Systems evolve in a predictable manner

   Rigid → Modular → Programmable → Autonomous

2. There are universal laws of evolution leading toward the "peak value"
   - The laws can predict the next big thing

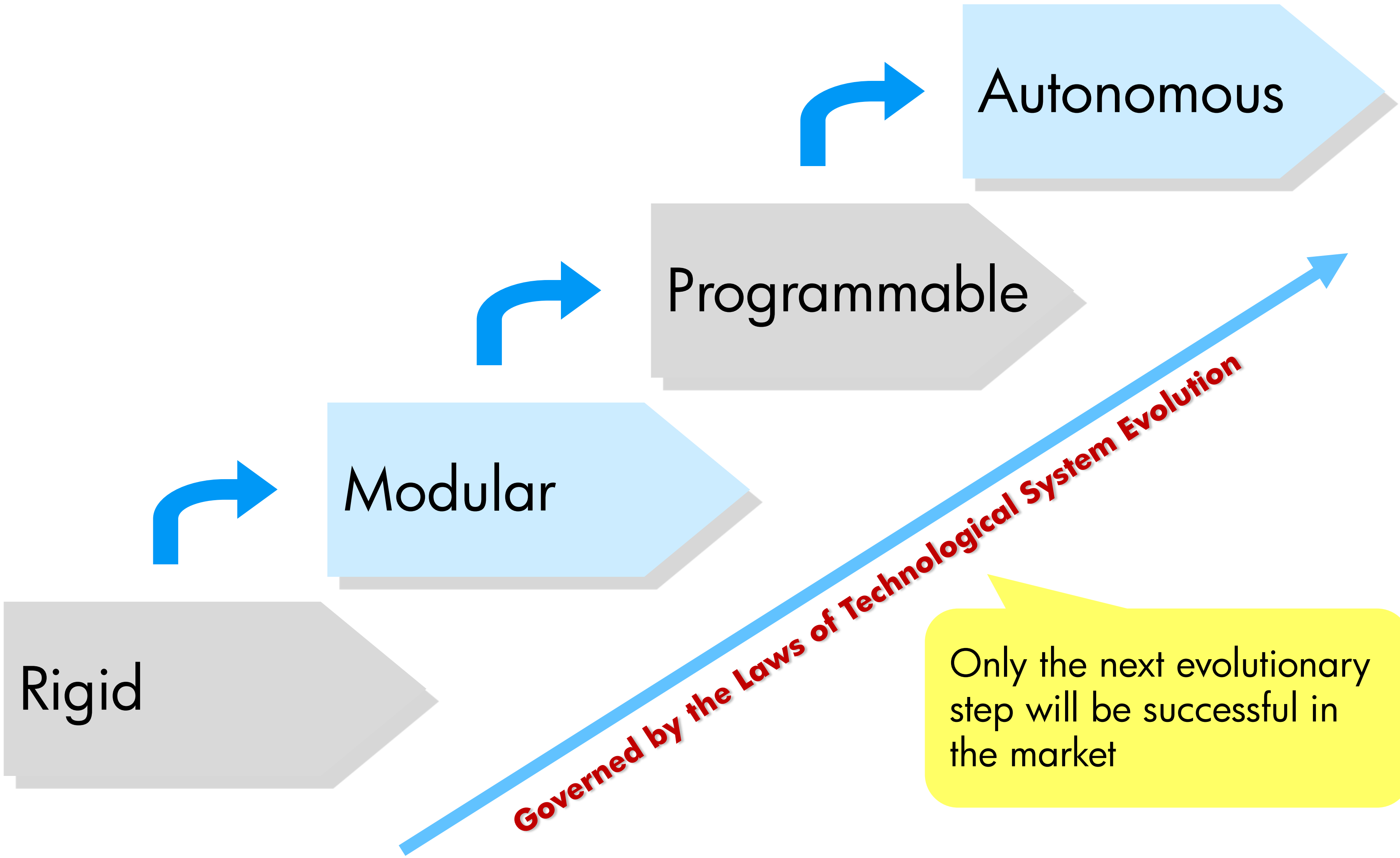
3. Best innovations resolve trade-offs
   - Compromise-free
   - e.g., Top of a soda can
     - Must retain fluid
     - Must easily pour fluid





4. There are universal principles for resolving conflicts.

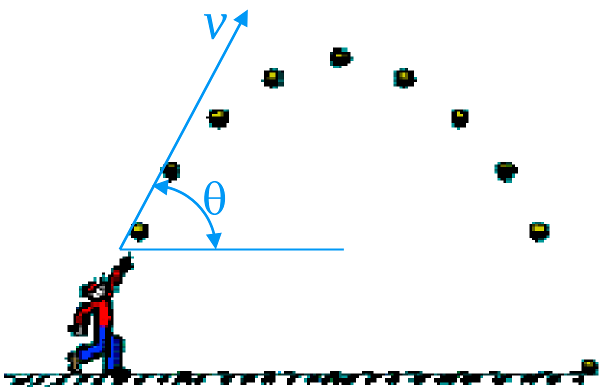
TECHNOLOGY EVOLUTION VECTOR



ANALOGY

Laws of Motion and Laws of Evolution

$$Y = vtSin(\theta) - \frac{1}{2}gt^2$$



$$X = vtCos(\theta)$$

Predictable Trajectory

EXAMPLE: LAWS OF MOTION AND LAWS OF EVOLUTION

Law: Transition to higher-level systems

Predictable Inventions

EXAMPLE: LAWS OF MOTION AND LAWS OF EVOLUTION

Law: Transition to higher-level systems

Predictable Inventions

LAWS OF TECHNOLOGICAL SYSTEM EVOLUTION

1. Increasing degree of ideality

   Systems evolve in the direction that increases degree of ideality.

2. Non-uniform evolution of subsystems

   The rate of evolution of subsystems is not uniform; the more complex the system, the more non-uniform the evolution of the subsystems.

3. Increasing dynamism (flexibility)

   Systems evolve in the direction to increase adaptation and flexibility, changing environmental conditions, and of multi-functionality.

4. Transition to a higher-level system

   Systems evolve from mono-systems to bi-systems and poly-systems.

5. Transition to micro-level

   Systems evolve toward an increasing use of micro components.

6. Completeness

   An autonomous system consists of 1) working means, 2) engine, 3)transmission, and 4) control means.

7. Shortening of energy flow path

   Systems evolve in the direction of shortening the energy flow passage through the system.

8. Increasing substance-field interactions

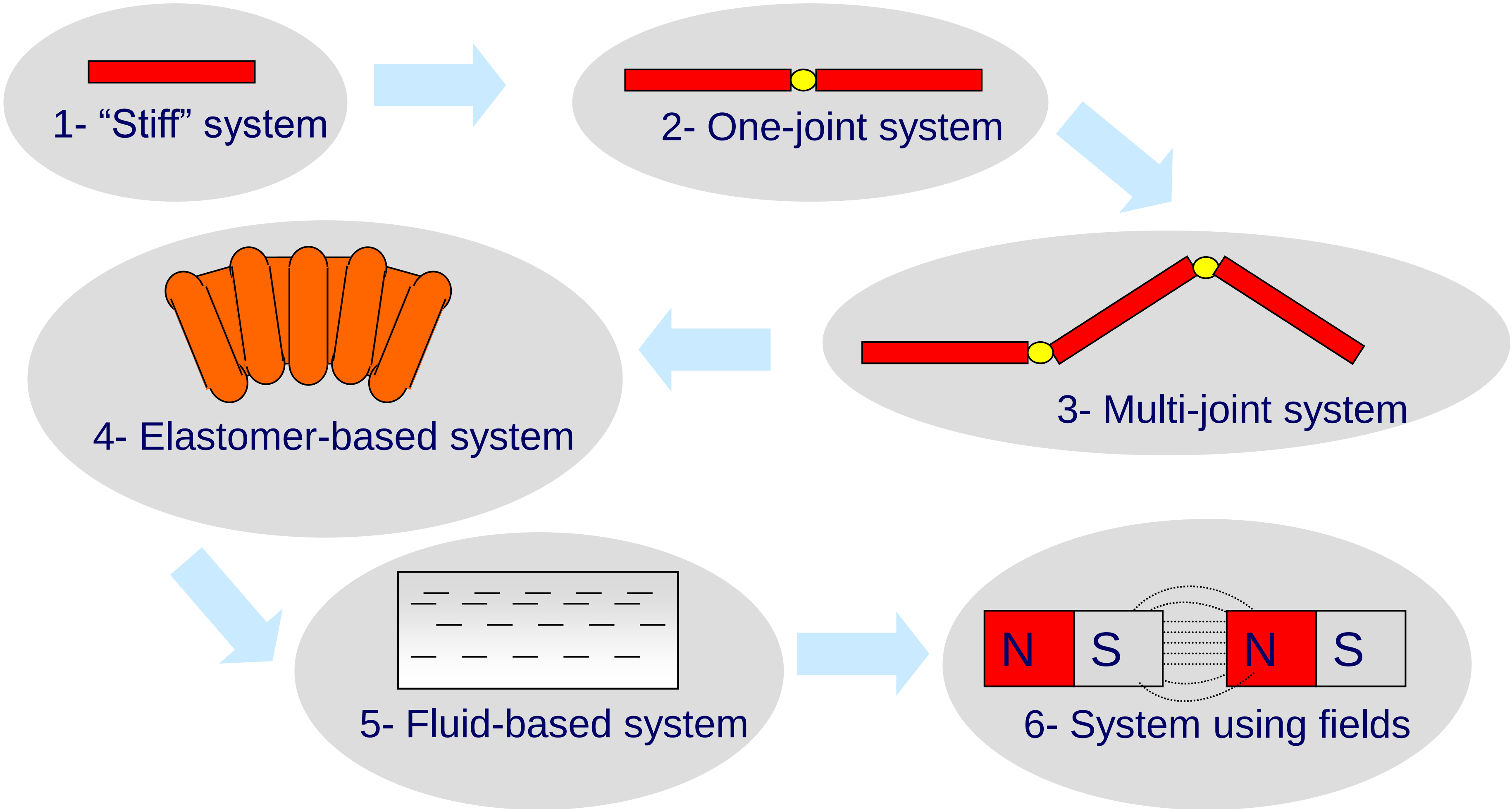
   Systems evolve in the direction of increasing controllability via more complex or complete tool-object-energy interactions.

9. Harmonization of rhythms

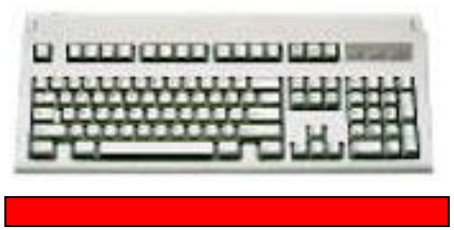
   Systems evolve in the direction of increasing coordination or the periodicity of the subsystems and components.

EXAMPLE: INCREASING DYNAMISM

Law: Systems evolve in a direction to increase dynamism and flexibility

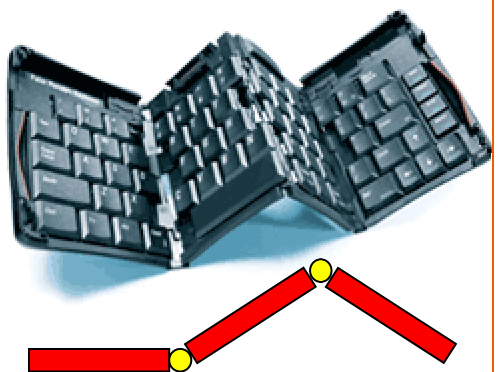


EXAMPLE: KEYBOARDS



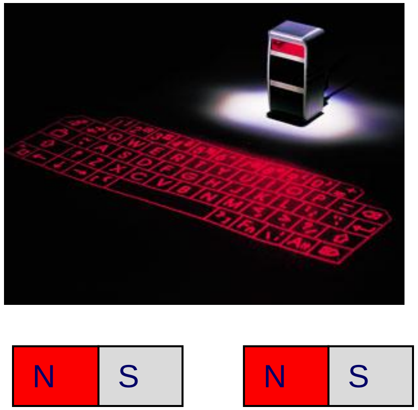
1- "Stiff" system

2- One-joint system



3- Multi-joint system

4- Elastomer-based system



6- System using fields

IT TECHNOLOGY EVOLUTION VECTOR

Phase:	**Rigid**	**Modular**	**Programmable**	**Autonomous**
Maturity:	Initial state	Novelty	Readily available	Blends into the background
Deployment:	One-of-a-kind	Low production volumes	High production volumes	High production
Support:	Self-supported	Minimal, highly specialized	Extensive	Transparent to the user
Reliability:	Very low	Moderate	High	Very high
Cost per Unit:	Very high	High	Low, very affordable	User unaware of cost
User Base:	A few	Small	Very large	Very large
Commercialization:	Experimental, Limited	Moderate	High	Embedded
Example:	*UNIVAC I*	*Early mainframes*	*Laptop computers*	*Vehicle engine control modules*

PROCESS

1. Identify a technology area
2. Document historical trends
3. Establish current position along the technology evolution vector
4. Identify industry/technology drivers, enablers, and constraints
5. Apply the TRIZ 9 Laws, one at a time
6. Review and summarize

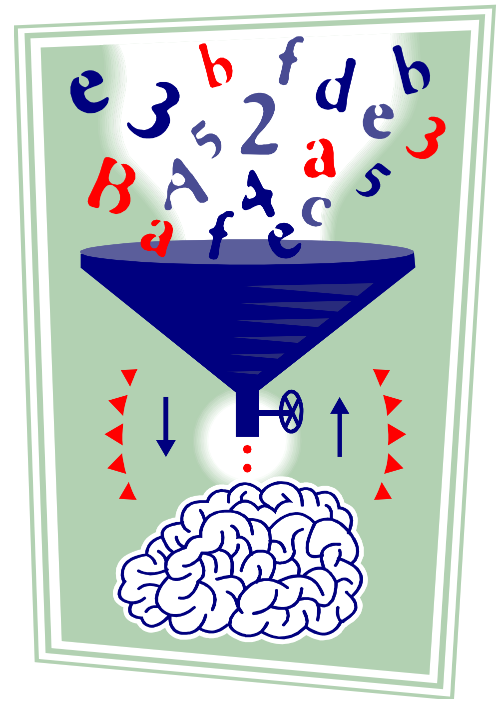
EXAMPLE: FUTURE OF BUSINESS INTELLIGENCE

Future of Business Intelligence

Problem	Where is Business Intelligence heading?
Type	Next Big Thing
TRIZ Tools	• Technology Evolution Vector • 9 Laws
Approach	1. Trends in Business Intelligence 2. Current position along the evolution vector 3. Possible future states of BI using the 9 Laws

Business Intelligence (BI)

- “Computer-based techniques used in spotting, digging-out, and analyzing business data, such as sales revenue by products and/or departments or associated costs and incomes” – *Wikipedia*

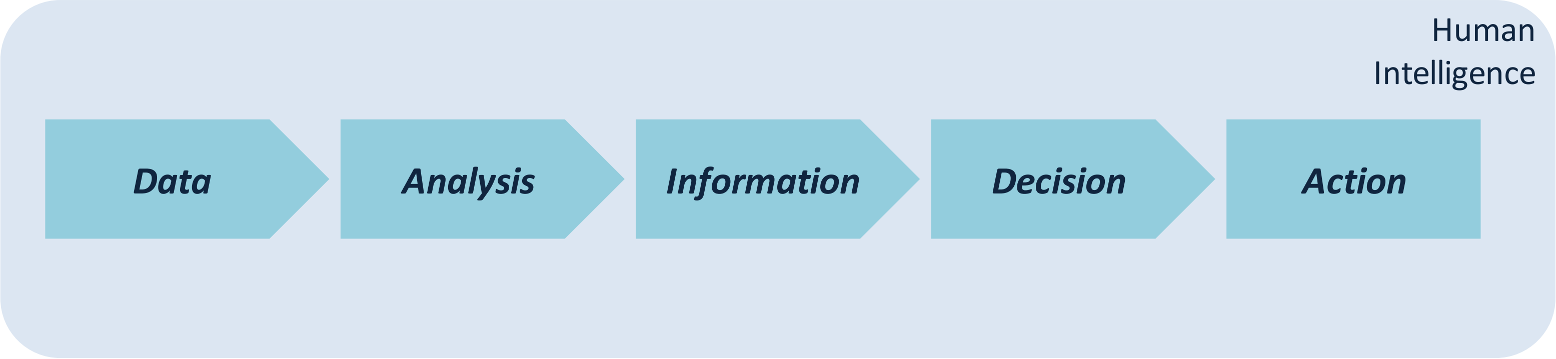


- Involved data warehousing, ETL, OLAP, data mining, text analysis, data visualization etc.

- Supports decision making via fact-based analysis of data

- BI has been gaining popularity over the years

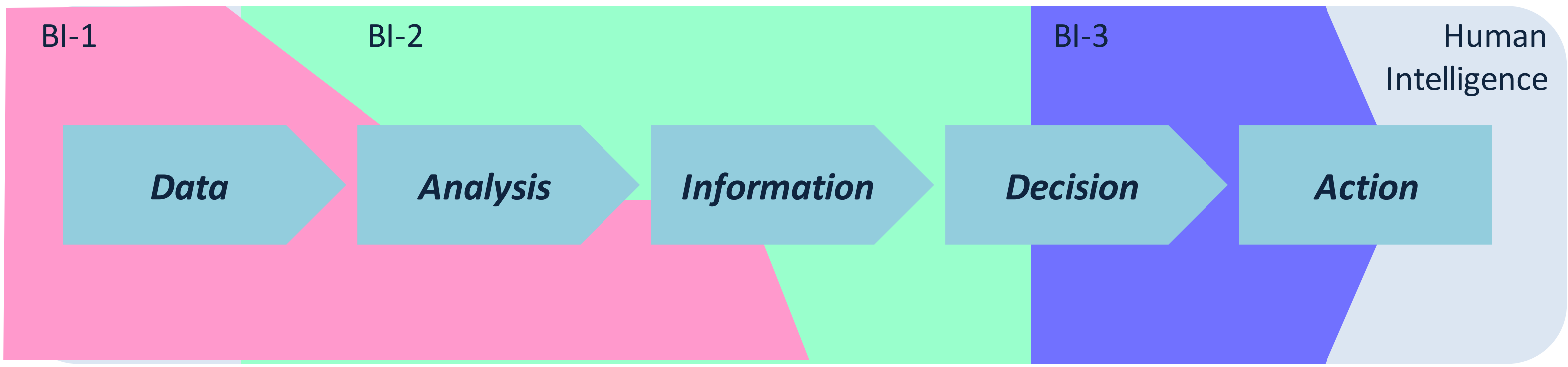
Stages of BI


Human
Intelligence
Data
Analysis
Information
Decision
Action

Phases of BI & Evolution Vector

Years	State	TRIZ Evolution Phase
Pre-1990	BI-0 • A manual process	**Rigid**
1990s	BI-1 • *Data warehouses* – a repository of organizational data • *Reporting* – usually after the fact • *OLAP* - slicing and dicing of data	**Modular**
2000s	BI-2 • *Enterprise data warehouses* – data repositories contain information at the enterprise level • *Unstructured data* – accessing the facts buried in textual data • *Real-time reports* – more actionable information than was available in BI-1 • *Advanced analytics* - correlations, forecasts, predictions, optimization • *Dashboards* – pulling together the results in an easy to digest form	**Early Programmable**
Future	BI-3 • Embedded BI • Mostly unknown	**Autonomous**

Pictorial



BI-1
BI-2
BI-3
Human Intelligence
Data
Analysis
Information
Decision
Action

Drivers & Enablers

Drivers

- The intense global competition, requiring enterprises to instantly tune to the rapidly changing market demands
- The need to make more fact-based decisions to avoid errors and lost opportunities
- Information explosion

Enablers

- Web 2.0 & social media
- Enterprise data warehouses
- Mashup data
- RFID, sensors, and location data
- Cloud computing

Constraints

- Cost of tools
- Complexity of analysis
- Infrequent use
- Data availability

Application of 9 Laws

No.	Law	Application to BI
1	Increasing degree of ideality	The future BI system will offer a broader range of analytical features, easier to use by non-technical knowledge workers, cost less, and has a broad range of business applications.
2	Non-uniform evolution of sub-systems	Different elements of BI solutions (e.g., data management, analysis, visualization, and application interfaces) will experience independent development by companies specializing in the respective fields. There will be continued start-up and acquisition cycles in the BI market.
3	Increasing dynamism (flexibility)	The future BI systems will grow from point solutions, to multi-discipline solutions, and to multi-industry solutions, spanning extended enterprises.
4	Transition to higher level systems	The future BI system will evolve from single applications, and begin to work in pairs (for more complex event analysis), then in groups with other BI applications (for even more complex event analysis).
5	Transition to micro level	BI solutions' footprints will shrink in size, requiring less computing resources, and accessible via mobile devices (yes, there will be an app for BI!)
6	Completeness	The future BI system will comprehend mashup data, include a full suite of analytical techniques, integration with all enterprise and personal applications, and driven by the users' context for decisions and actions.
7	Shortening of energy flow path	BI systems will be personalized - they will automatically locate the relevant data, eliminate data latency, analyze it in real-time, and serve the needs of the user via automated actions.
8	Increasing controllability	The user's context will control the operations of the BI system and the actions it takes.
9	Harmonization of rhythms	BI systems will be event-driven, based on a coordination among data availability, temporal and location factors, the characteristics of the user, and the needs of the user. BI systems will collaborate with each other.

Future of BI

– More transparent, and embedded within other applications and processes

– More pervasive, and used more broadly, by many more knowledge workers

– Less expensive

– Automated decision making and performing actions
  - Trust?
  - Liability?

INTERACTIVE WORKSHOP

PROCESS

1. Identify a technology area
2. Document historical trends
3. Establish current position along the technology evolution vector
4. Identify industry/technology drivers, enablers, and constraints
5. Apply the TRIZ 9 Laws, one at a time
6. Review and summarize

BIBLIOGRAPHY

BIBLIOGRAPHY

– Davenport, T.H., Harris, J.G., *Competing on Analytics*, Harvard Business School Press, 2007

– Fey, V., Rivin, E., *Innovation on Demand*, Cambridge University Press, 2005

– Kasravi, K., Risov, M., *Patent Similarity Detection*, Hawaii International Conference on System Sciences, 2009

– McCandless, D., *The Visual Miscellaneum*, 2009, Collins Design

– http://en.wikipedia.org/wiki/Charles_Joseph_Minard

– http://webvision.med.utah.edu

– http://www.health.com

APPENDIX
TRIZ 9 LAWS

1. INCREASING DEGREE OF IDEALITY

Definition	Systems evolve in the direction that increases degree of ideality.
Description	The primary law of evolution of technical systems. Over time, Systems will more efficiently perform their primary function, evolve new features, and deliver more value.
Example	Early digital computers required hardwiring the series of sockets and plugs. Over time this evolved to machine language, dramatically increasing the power and efficiency of the programmer. Machine language led to assembler, which led to Fortran, Cobol, various structured programming languages of the 70s, object-oriented programming languages, 4GLs, etc. With each step forward more could be done with less programming effort. Results have become more reliable and predictable. In addition, languages have kept in step with developments in hardware and infrastructure, so today's languages are capable of doing in a single line or two what could not be achieved at all in COBOL.

2. NON-UNIFORM EVOLUTION OF SUBSYSTEMS

Definition	The rate of evolution of subsystems is not uniform; the more complex the system, the more non-uniform the evolution of the subsystems.
Description	Driven by different market forces, different subsystems of a single technical system will evolve at different rates. This creates conflicts within the system, which open the door to opportunities for innovation.
Example	The evolution of PC subcomponents is constantly forcing development of others. Increased processor speed drove a need for more memory, driving memory prices down. Greater processing capacity drove the need for increased disk storage, which enabled the use of video and audio, again driving increases in processor, memory, and even I/O as parallel port and serial port gave way to USB. Improvements in each subsystem continually drove improvements in the others.

3. INCREASING DYNAMISM (FLEXIBILITY)

Definition	Systems evolve in the direction to increase adaptation and flexibility, changing environmental conditions, and of multi-functionality.
Description	Systems will initially be inflexible and devoted single-mindedly toward achieving their primary function. Over time they will evolve to perform well within a variety of conditions and regimes and may even add additional functionality.
Example	The default PC user interface was originally a printed report supplemented by data and code on punch cards or tape. This evolved into a simple black screen with green or amber phosphors that could display data and update it an instant later. Two colors became sixteen, permitting the more efficient display of data and rudimentary graphical data. More resolution and more colors led to high quality GUI interfaces. Applications were written for high color and supported by mouse, touchpad, or even touch screen. Today's users enjoy infinitely more degrees of visual freedom when working with their information.

4. TRANSITION TO A HIGHER-LEVEL SYSTEM

Definition	Systems evolve from mono-systems to bi-systems and poly-systems.
Description	Systems often start out simply with a single component, but grow by adding additional copies of the same component or copies that have different but complementary properties. Bi-systems feature two copies of the component (e.g. binoculars), where poly systems feature three or more (e.g. a disc harrow).
Example	Early PCs stood alone or connected to others via dial-up modem. Soon it became commonplace to connect computers over serial ports to share data. In time, peer-to-peer networks appeared to allow multiple PCs to connect within a local area. With the evolution of client/server LANs, small poly-systems became single mono systems supporting offices and small enterprises. Over a wide area network these began to share loads and become a poly system called a grid (e.g. SETI@Home). The latest evolution in this series is a poly-system which we currently called cloud computing.

5. TRANSITION TO MICRO-LEVEL

Definition	Systems evolve toward an increasing use of micro components.
Description	A system will evolve from using crystal lattices, to molecules, to atoms and ions, to elementary particles.
Example	The physical reduction of compute devices is easy to see. It has been said that today's musical greeting card has greater computing power than that which existed in all the world in 1960. What is less obvious is the reduction in the size of the impact that IT support has on our core business. Initially IT support services were expensive and consisted of teams of specialists that came to your desk as needed. With the advent of remote monitoring and support capabilities came outsourcing, off shoring, and finally automation. With each step we reduce cost and improve efficiency, reducing the cost and impact of IT support to far less visible dimension.

6. COMPLETENESS

Definition	An autonomous system consists of 1) working means, 2) engine, 3)transmission, and 4) control means.
Description	Technical systems will evolve into autonomous systems that do not require human intervention.
Example	Early in IT's history both data and programs might be stored on paper punch cards. Early systems required the physical storage and retrieval of trays filled with punch cards. As these were replaced with magnetic media the ability to store multiple files on one tape or disc was a huge leap forward. Still, the operator needed to tell the OS exactly where on the medium segments of the file could be stored. Later this process became transparent to the user as the system became capable of allocating data storage space without manual intervention. Today, data might be mirrored or striped and possibly distributed globally for safety, redundancy, and access.

7. SHORTENING OF ENERGY FLOW PATH

Definition	Systems evolve in the direction of shortening the energy flow passage through the system.
Description	Technological systems evolve to shorten the flow of energy between the source and the working means.
Example	Over time IT practitioners have developed systems and methods designed to apply energy directly to IT operations. Case tools are an example. These tools reduced time required to build many applications by smoothing the translation of ideas into working code. Other examples are Project Management and Lean Six Sigma methods and tools. By better structuring the application of resources to tasks, focusing on applying the right person at the right time in the right place, shortening the flow of energy to working means.

8. INCREASING SUBSTANCE-FIELD INTERACTIONS

Definition	Systems evolve in the direction of increasing controllability via more complex or complete tool-object-energy interactions.
Description	As the system evolves it will become more controllable through the application of energy to substances within the system.
Example	SLA agreements provide business/legal frameworks that provide control to IT service agreements in the same manner that a magnetic field might control a moving part. By measuring performance and applying a controlling force (generally in the form of money), SLA agreements drive IT performance ever closer to design specifications. In the IT industry, this effect is controlled in turn by Third-Party Intermediaries that exercise control over the formulation of the service level agreements. These forces continually drive IT service level closer to ideality.

9. HARMONIZATION OF RHYTHMS

Definition	Systems evolve in the direction of increasing coordination of the periodicity of the subsystems and components.
Description	To perform optimally, a system of control and coordinate the periodicity of the actions of its parts.
Example	As in any mechanical machine, the components of an IT system must work together in harmony. Early in its evolution, the IT industry struggled with the lack of standards that lead to inefficiency and high cost. One employee might not be able to use another's tools, or read another's data. One support team member alone could not support all of the tools extant in the workplace. By standardizing tools, flow of information was enhanced and the cost of maintaining the system reduced. Standard business processes, defined workflows, quality and efficiency methodologies (e.g. Lean Six Sigma) all work to harmonize the interactions of various parts of an IT system.

