# Data Visualization in the Real World

Tim Vlamis and Doug Schieder

February 26, 2020

www.vlamis.com

Vlamis Software Solutions

- Founded in 1992 in Kansas City, Missouri
- 400+ Enterprise Clients
- Consults in :
  - Enterprise Business Intelligence & Analytics
  - Analytic Warehousing
  - Machine Learning and Predictive Analytics
  - Data Visualization
  - ETL and data integration
- Vlamis consultants average 15+ years
- www.vlamis.com (blog, papers, newsletters, services)
- Co-authors of book "Data Visualization for OBI 11g"
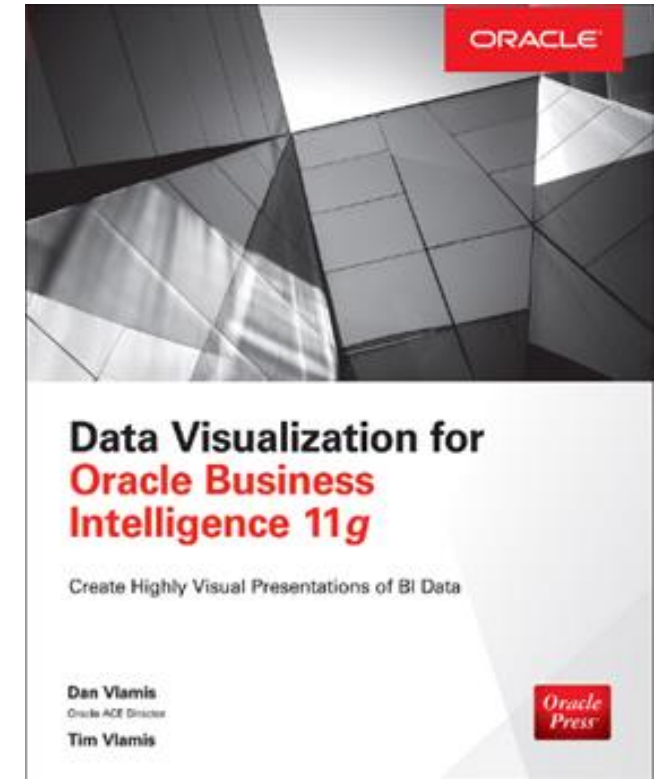
Presenter Background

Tim Vlamis – Vice President & Analytics Strategist
- 30+ years in business modeling and valuation, forecasting, and scenario analyses
- Instructor for Oracle University's Data Mining Techniques and Oracle R Enterprise Essentials Courses
- Professional Certified Marketer (PCM) from AMA
- MBA Kellogg School of Management (Northwestern University)
- BA Economics Yale University
- tvlamis@vlamis.com

Doug Schieder – Business Development Manager
- Works very closely with Dan Vlamis and Tim Vlamis
- Tracks all opportunities for Vlamis
- Background in Socio-Political Communication
- Worked on State and Federal campaigns

vlamis ™
SOFTWARE SOLUTIONS

Data is common, logic is rare

"*Crime is common. Logic is rare. Therefore it is upon the logic rather than upon the crime that you should dwell.*"

vlamis ™
SOFTWARE SOLUTIONS

Think Evidence, Not Analytics

Data Visualization is about supplying organizations with **evidence**.

vlamis ™
SOFTWARE SOLUTIONS

Simple Pivot Table with Easy Questions

2014 Monthly Sales by Company

D4 Company	2014 / 01 Sales	2014 / 02 Sales	2014 / 03 Sales	2014 / 04 Sales	2014 / 05 Sales	2014 / 06 Sales	2014 / 07 Sales	2014 / 08 Sales	2014 / 09 Sales	2014 / 10 Sales	2014 / 11 Sales	2014 / 12 Sales	Grand Total
Genmind Corp	$202,019	$296,178	$393,254	$401,352	$621,749	$921,152	$823,760	$576,288	$590,033	$477,079	$324,569	$326,255	$5,953,688
Stockplus Inc.	$317,533	$475,312	$650,825	$605,253	$868,347	$1,272,701	$1,076,425	$904,047	$947,674	$788,834	$515,927	$531,188	$8,954,066
Tescare Ltd.	$261,837	$422,774	$555,255	$550,912	$844,094	$1,222,869	$1,012,856	$810,286	$814,160	$691,479	$447,950	$457,773	$8,092,246
Grand Total	$781,389	$1,194,264	$1,599,334	$1,557,516	$2,334,190	$3,416,722	$2,913,041	$2,290,621	$2,351,868	$1,957,392	$1,288,446	$1,315,216	$23,000,000

- What was the highest monthly sales for a company?
- Which month had the largest drop in sales for a company?
- In which month(s) did Tescare have a greater gain than Stockplus?
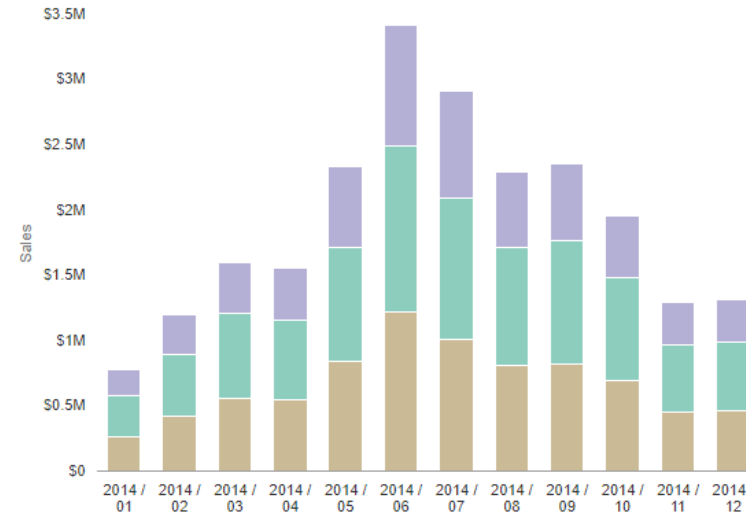- Is there a seasonal pattern to sales for all companies?

vlamis
SOFTWARE SOLUTIONS

Answers are Easy to Find in Graphs
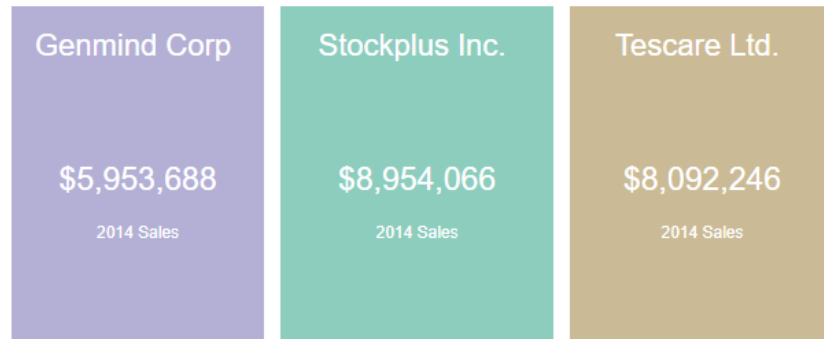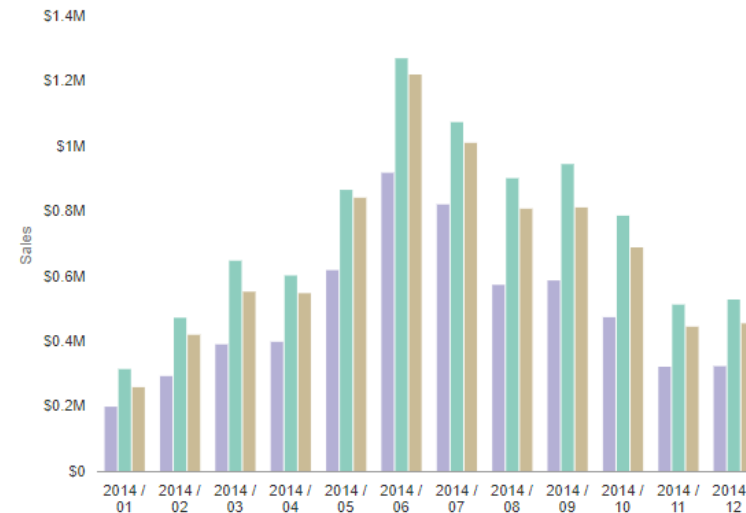
What was the highest monthly sales for a company?

Which month had the largest drop in sales for a company?

Total Sales 2014

$23,000,000

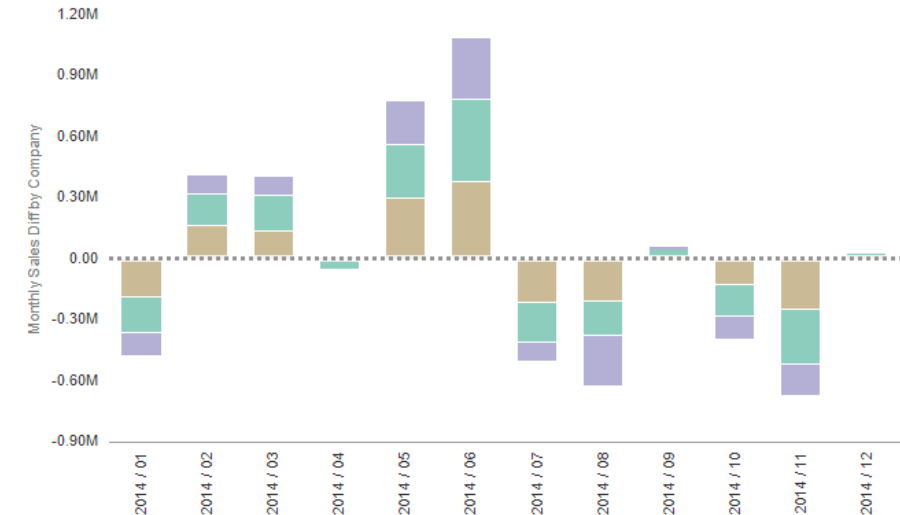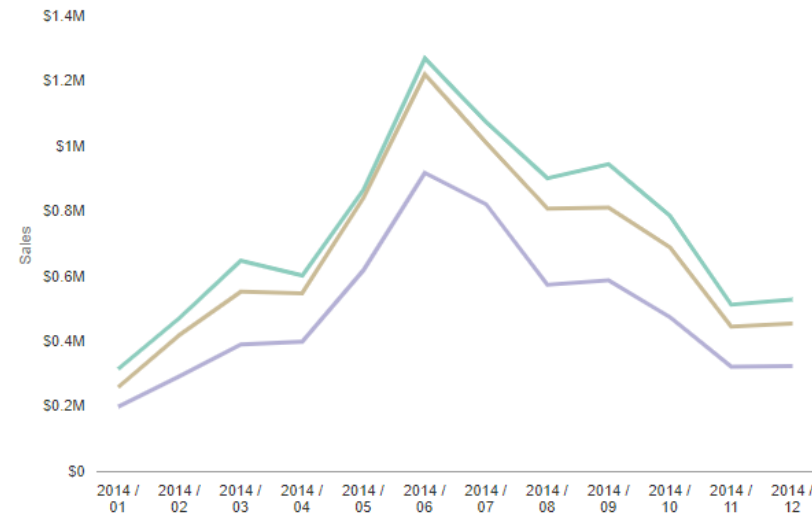| Genmind Corp	Stockplus Inc.	Tescare Ltd.
$5,953,688	$8,954,066	$8,092,246
2014 Sales	2014 Sales	2014 Sales

2014 Monthly Sales by Company
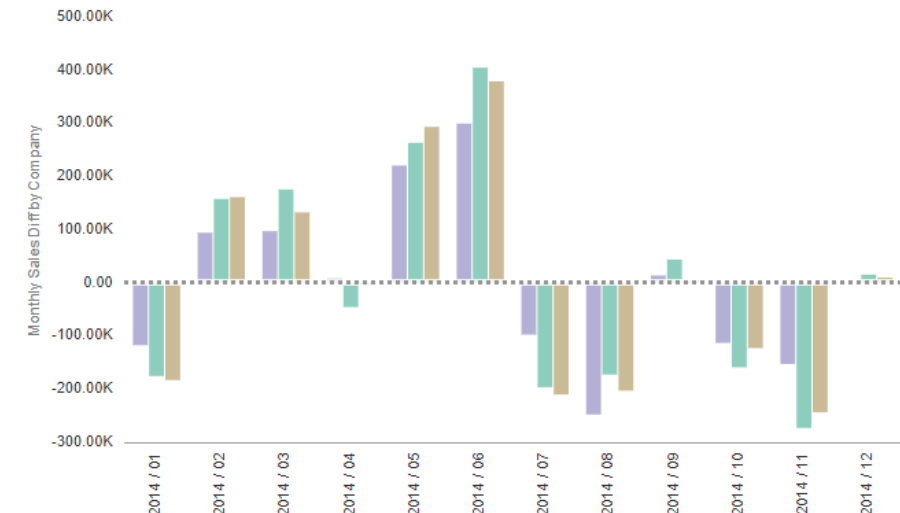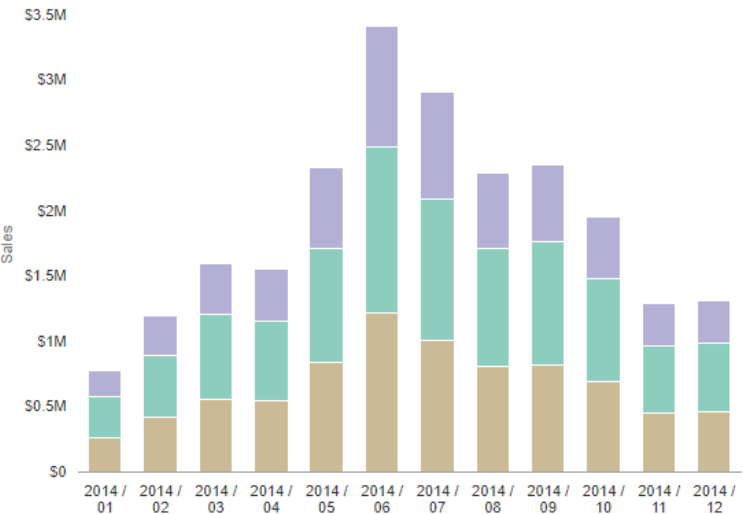

Difference from Previous Month Sales by Company

2014 Monthly Sales by Company
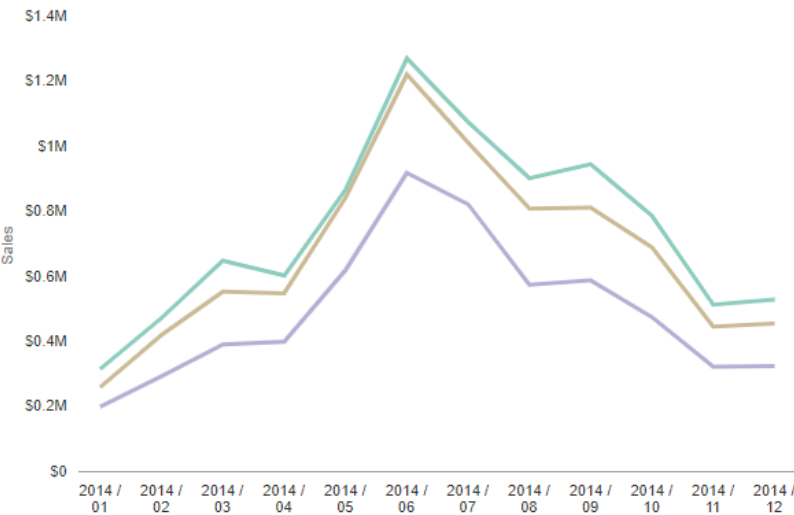
2014 Monthly Sales by Company
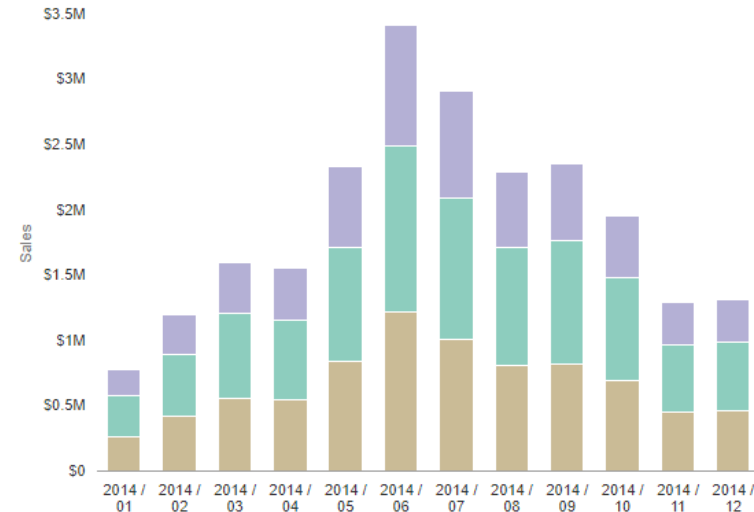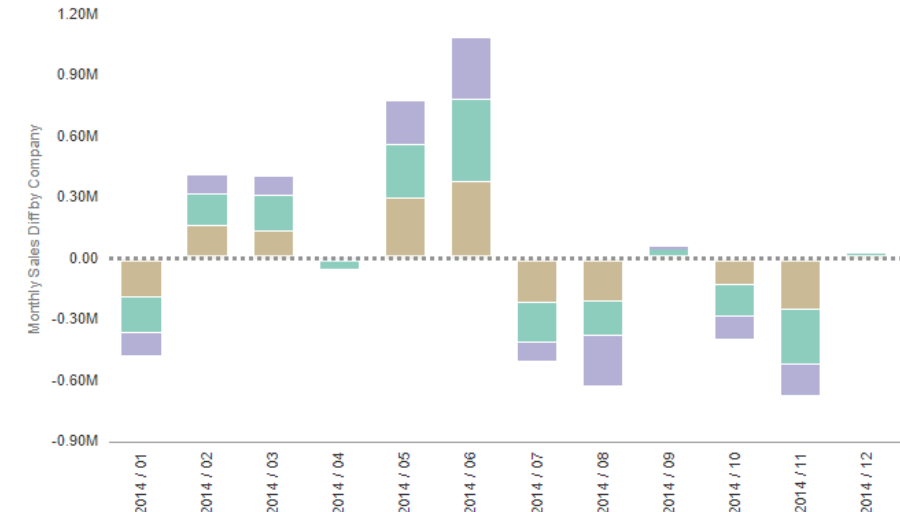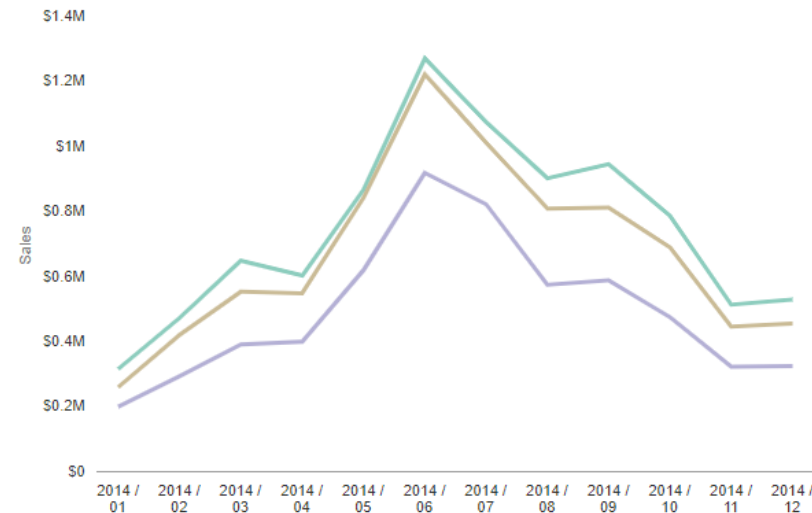

Difference from Previous Month Sales by Company
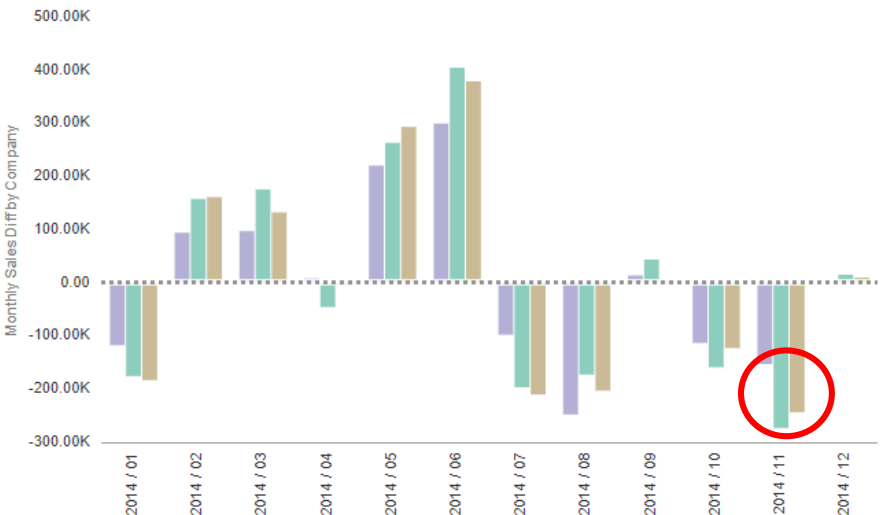
In which month(s) did Tescare have a greater gain than Stockplus?

Total Sales 2014

$23,000,000

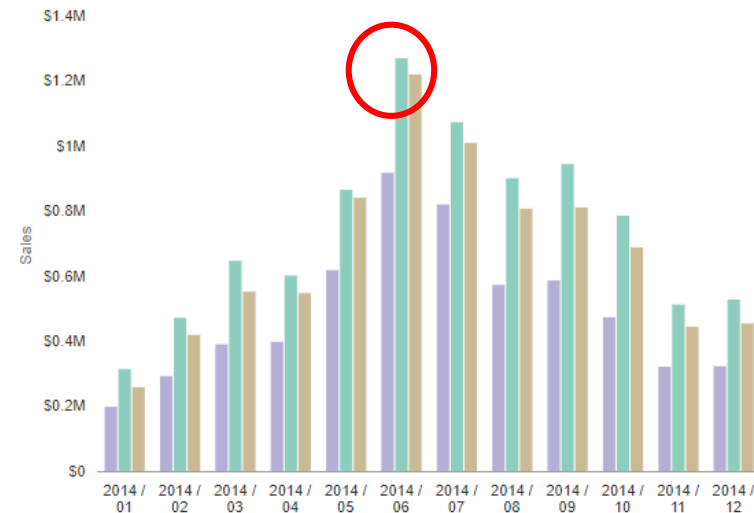| Genmind Corp	Stockplus Inc.	Tescare Ltd.
$5,953,688	$8,954,066	$8,092,246
2014 Sales	2014 Sales	2014 Sales

2014 Monthly Sales by Company
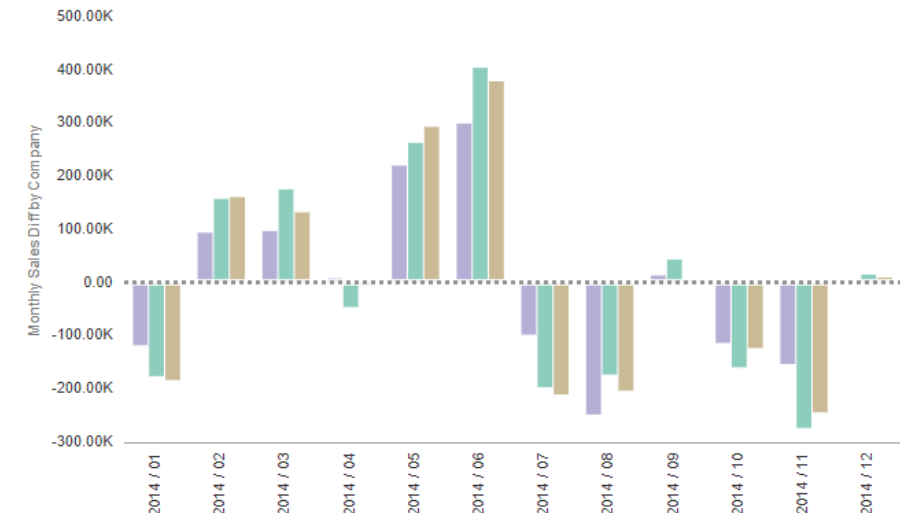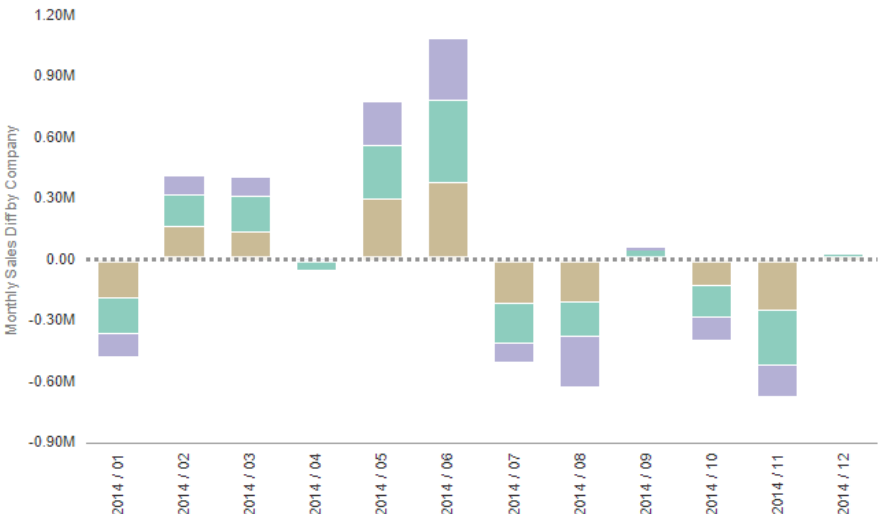
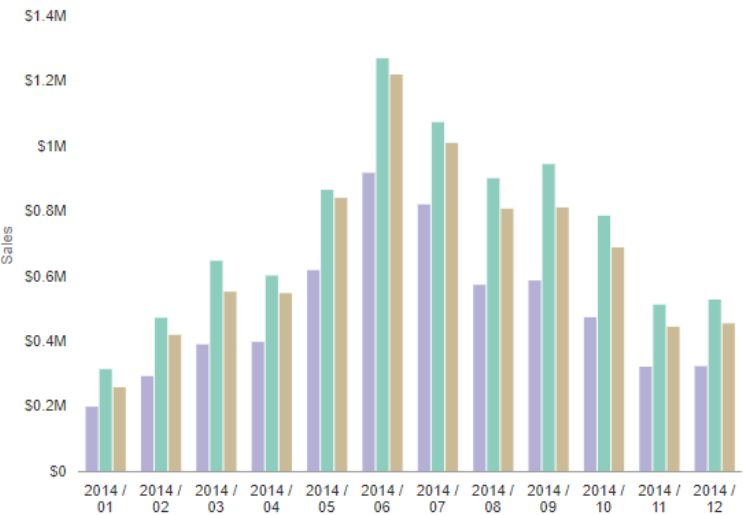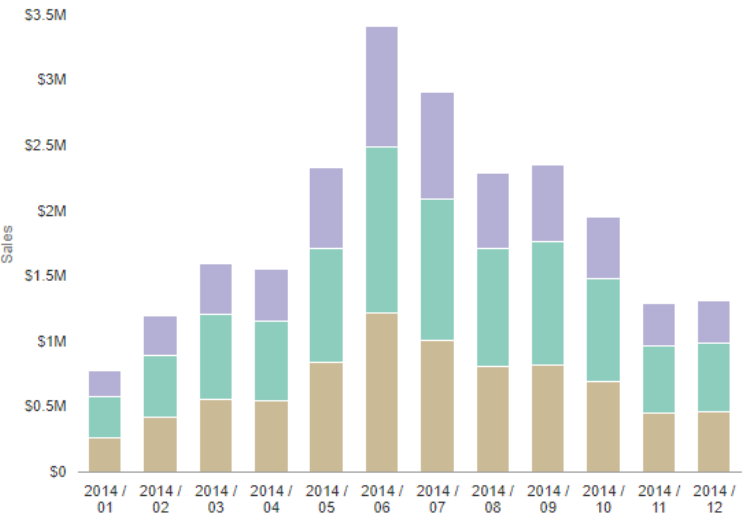Difference from Previous Month Sales by Company
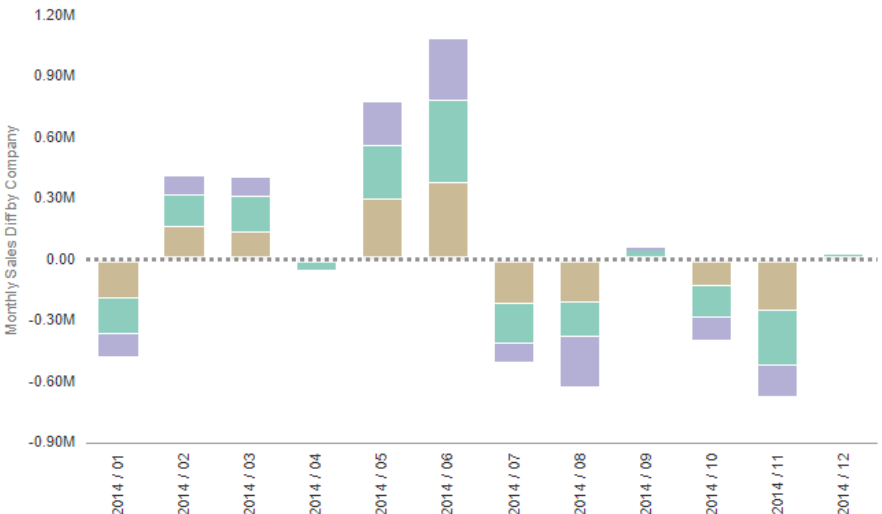
2014 Monthly Sales by Company
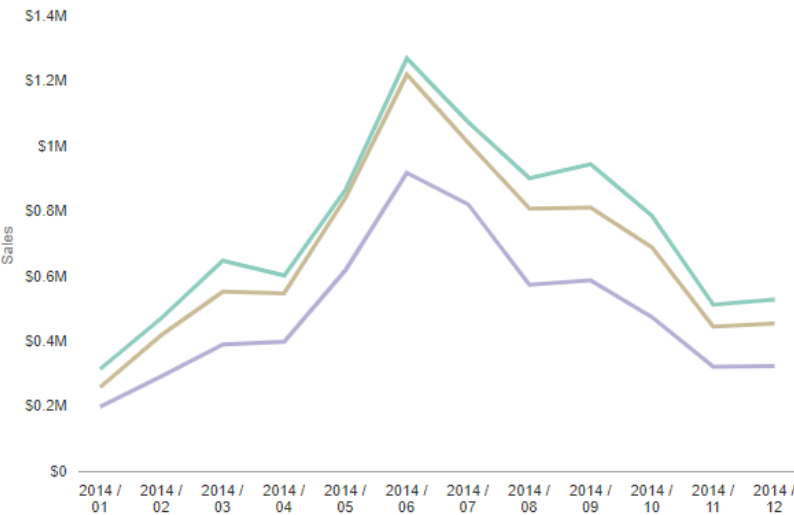
2014 Monthly Sales by Company
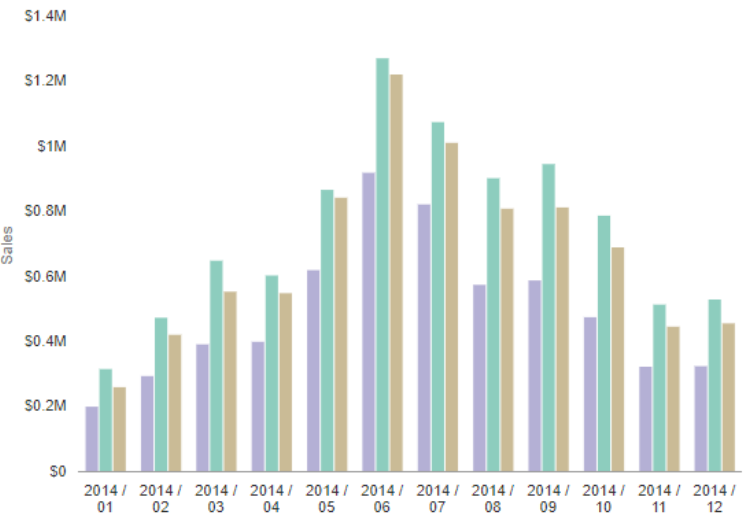
Difference from Previous Month Sales by Company

Is there a seasonal pattern to sales for all companies?
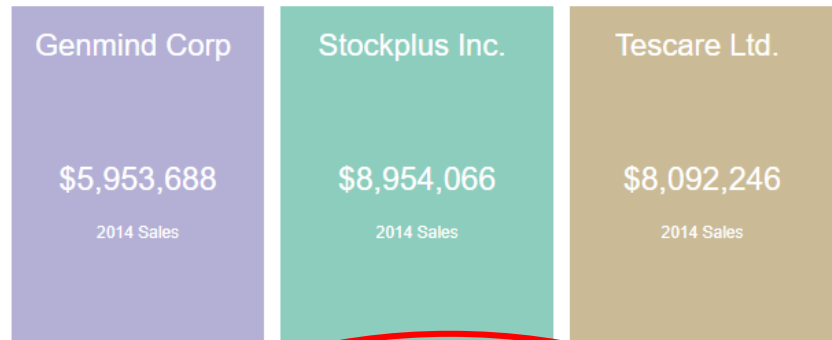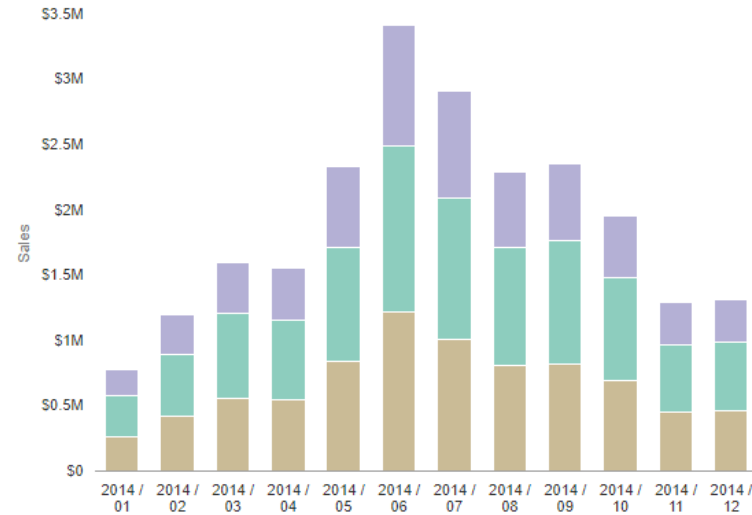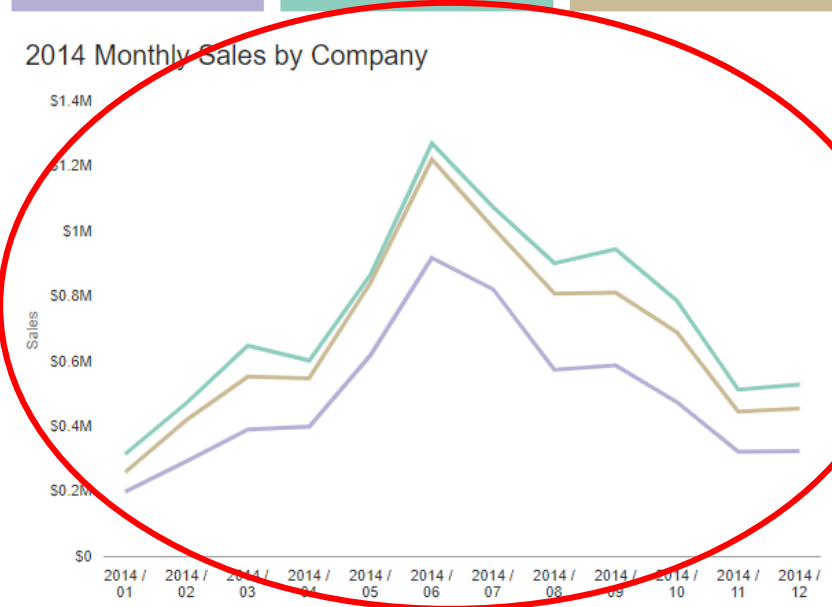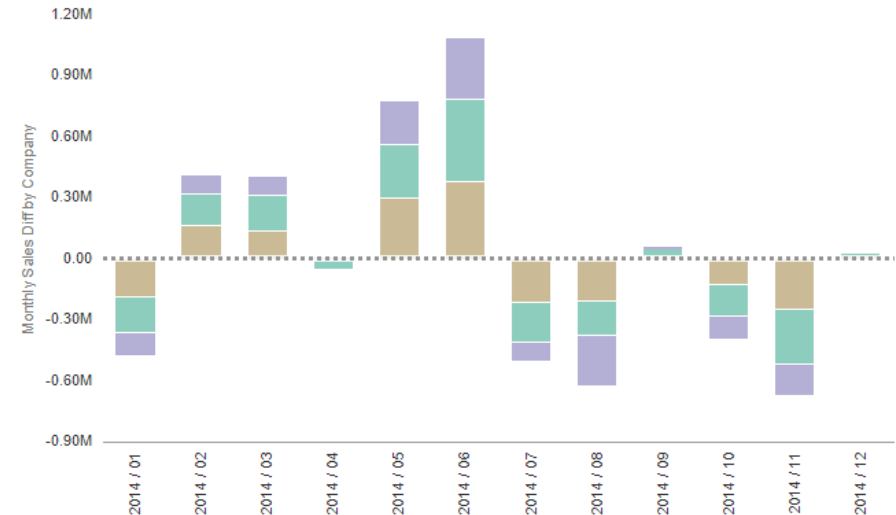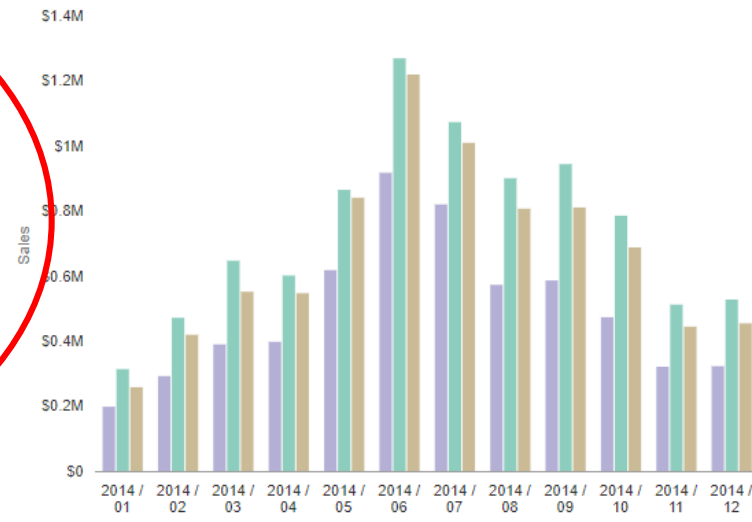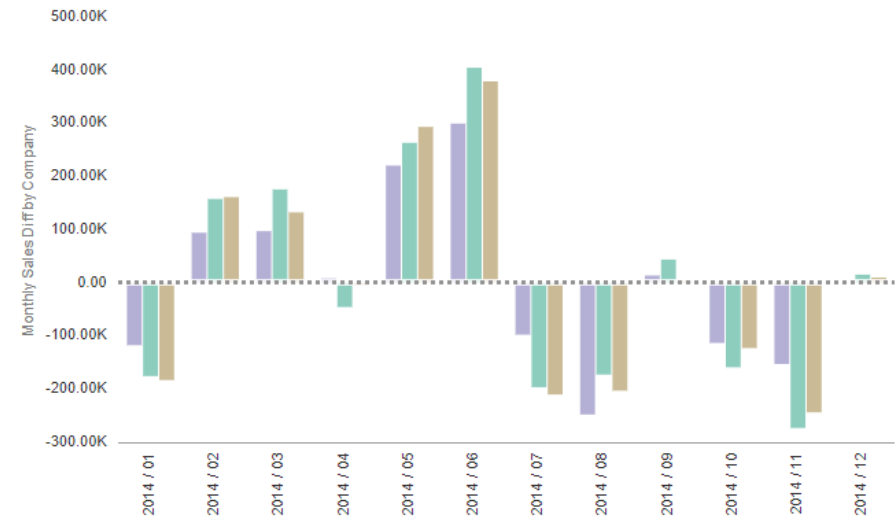
Shared Experiences!

Dashboards and displays drive coherence through a broad, shared understanding of organizational position, performance, flows, and influencers. Evidence!

https://xkcd.com/249/
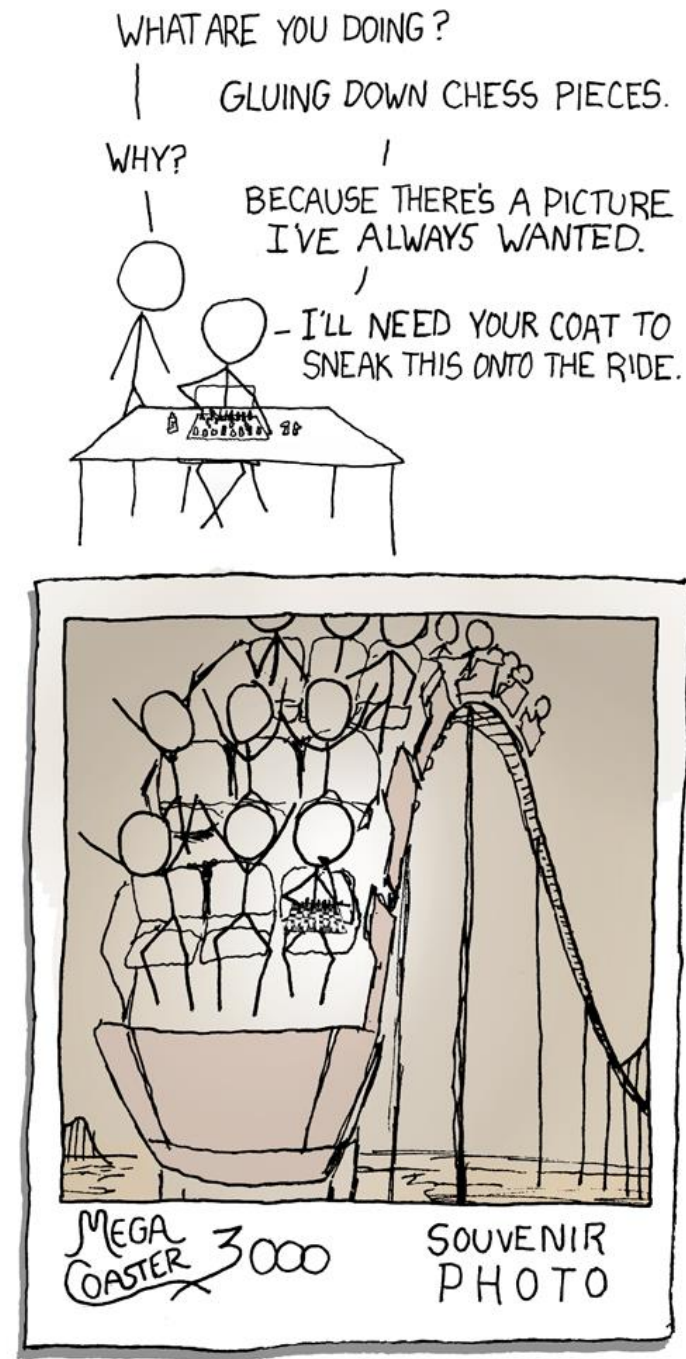
Data Visualization Scenarios

	Individual	Organizational
Deliberative Response	**Data Discovery**	**BI Dashboards**
Immediate Response	**Situational Awareness**	**Alerts Thresholds**

vlamis
SOFTWARE SOLUTIONS

Dashboard Definition

A Dashboard is a visual presentation of current summary information needed to manage and guide an organization or activity.

vlamis
SOFTWARE SOLUTIONS

Dashboard Definition

BI Dashboards should be designed to drive organizational coherence through a shared understanding of organizational position, performance, flows, and influencers.

vlamis
SOFTWARE SOLUTIONS

Great Dashboards Balance 6 Requirements

Manufacturing Summary | New Availability Analysis | New Injection Reject Analysis | Delivery Analysis

IQMS Manufacturing ERP
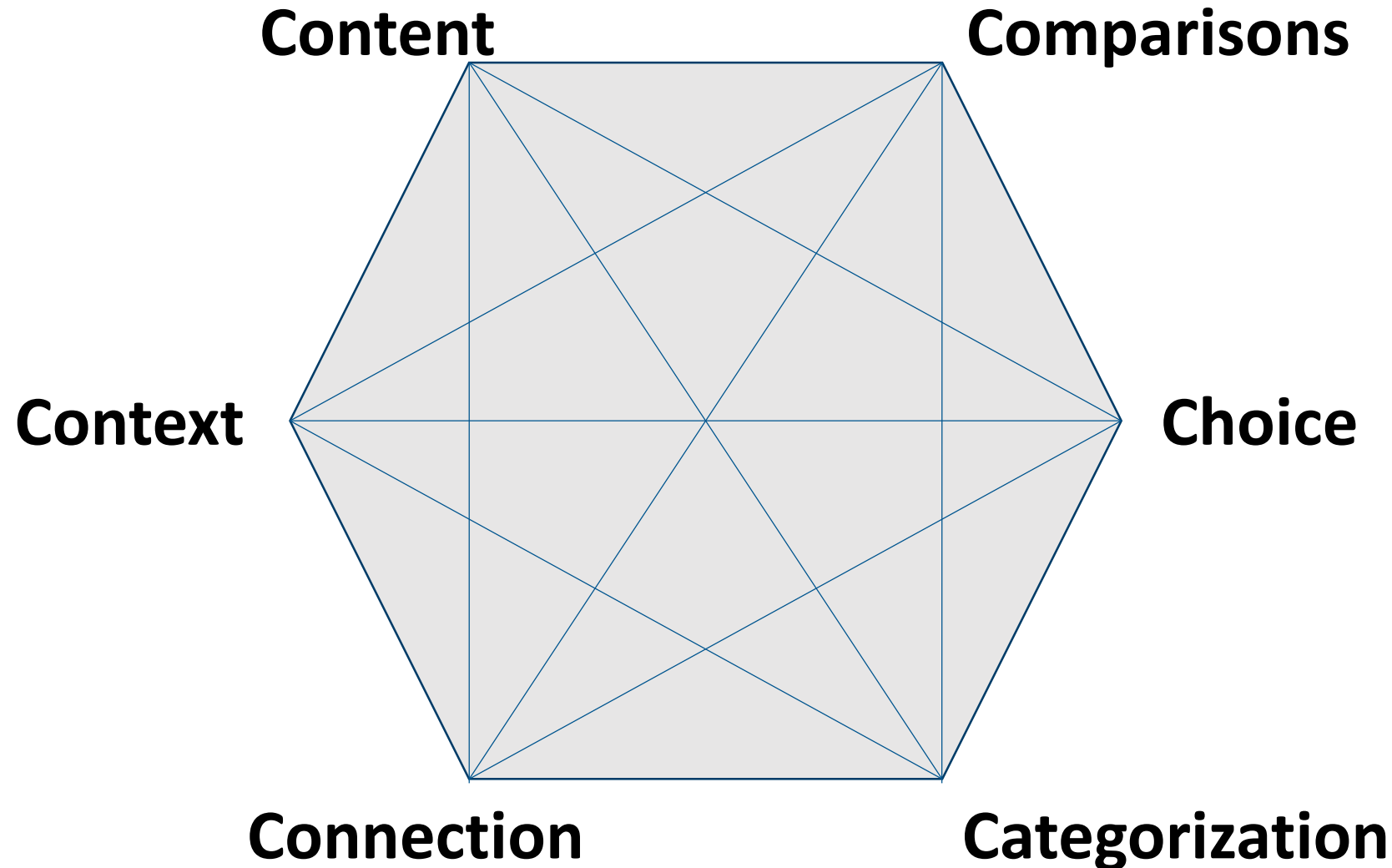
Manufacturing Summary Page

CALENDAR YEAR
2018

CALENDAR MONTH
(All Column Value)

EPLANT
--Select Value--

MFG TYPE
--Select Value--

MFG CELL
--Select Value--

RETURN TO MAIN MENU

38.0%
OEE % Actual

40.4%	94.9%	99.0%
Availability % Actual	Performance % Actual	Quality % Actual
45,308.6	99,089,453	1.04%
Production Hours	Total Good Parts	Scrap %

OEE Trend
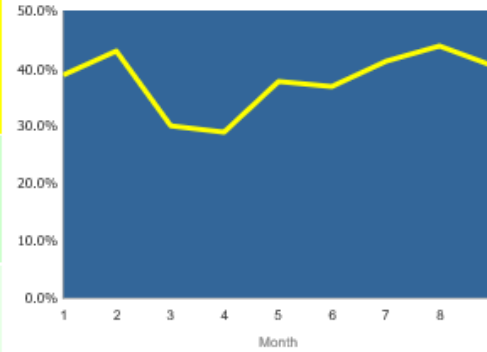
OEE by MFG Type
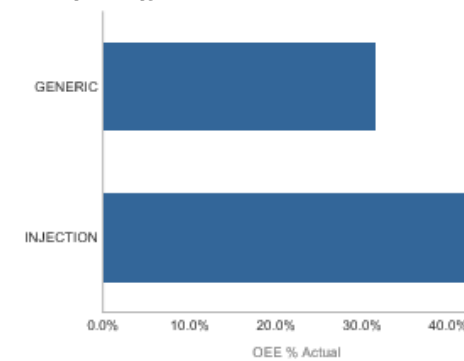
Production and Total Downtime Hours

Prod Hrs
Total DowntimeHrs
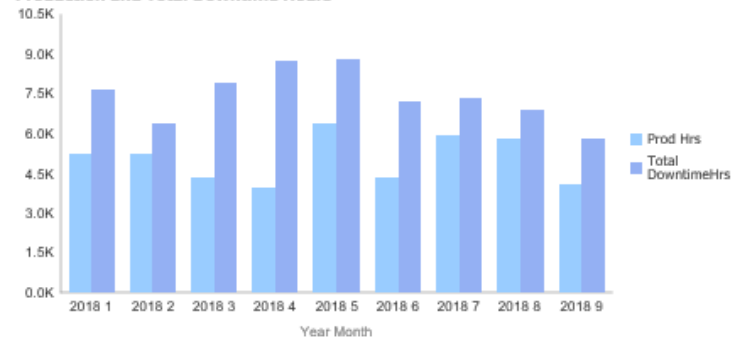
Standard and Actual Parts

Parts At Std
Total Parts
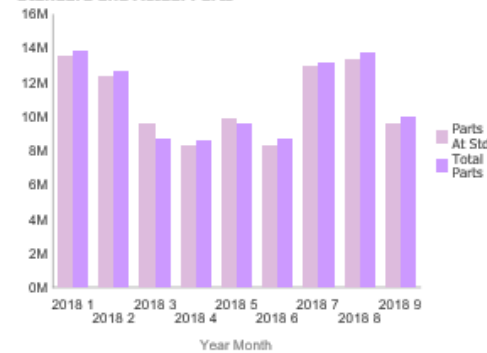
Scrap % by MFG Type
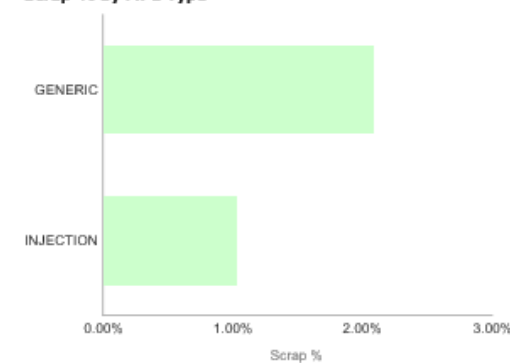
EPlant	Mfg Type	MFG Cell	Prod Hrs	Unplanned Downtime	Downtime %	Availability % Actual	Total Downtime Hrs	Hrs Available Actual
1	GENERIC	PASO GENERIC	14,491.1	28,684.4	66.35%	33.5%	28,684.4	43229.98
	INJECTION	PASO INJ	30,817.5	37,954.3	55.16%	44.8%	37,954.3	68811.90

EPlant	Mfg Type	MFG Cell	Parts At Std	Setup Hrs	Total Cycles	Line Utilization	OEE % Actual	Performance % Actual	Work Order ID	Prod Date	Total Parts	Prod Hrs
1	GENERIC	PASO GENERIC	2,096,975	54.5	1,747,778	166.23%	31.5%	95.9%			1,711,399	14,491.1
	INJECTION	PASO INJ	95,780,792	40.0	4,024,951	155.10%	42.1%	94.9%			97,422,136	30,817.5

EPlant	Mfg Type	MFG Cell	Good Parts	Quality % Actual	Scrap	Scrap %
1	GENERIC	PASO GENERIC	1,711,399	97.9%	2,227	2.08%
	INJECTION	PASO INJ	97,378,054	99.0%	15,322	1.03%

vlamis SOFTWARE SOLUTIONS

Manufacturing Summary | New Availability Analysis | New Injection Reject Analysis | Delivery Analysis

IQMS Manufacturing ERP

Availability Analysis by Work Center

This page shows a comparison of Production Hours and Downtime Hours

CALENDAR YEAR
2018

CALENDAR MONTH
(All Column Values)

DATE

WORK CENTER ID
--Select Value--

CLASS
--Select Value--

ITEM
--Select Value--

ITEM NO
--Select Value--

RETURN TO MAIN MENU

Availability % Actual by EQNO and Work Center

Availability % Actual
96.9%

> 98%
95% <> 98%
<95%

EQNO	Work Center	Availability % Actual
03	120 TON VAN DORN	99.9%
04	200 TON STOKES	99.7%
05	200 TON STOKES	98.8%
07	300 TON ENGLE	99.9%
08	300 TON NISSEI	98.5%
09	300 TON KRAUSSMAFFEI	91.0%
10	050 TON BATTENFELD	96.4%
11	050 TON BATTENFELD	92.2%
14	500 TON JSW	99.5%
16	2.5" EXTRUDER	100.0%
17	2.5 EXTRUDER	94.3%
ASSY-01	ASSEMBLY TABLE	97.2%
ASSY-02	ASSEMBLY TABLE	18.1%

Export

Production Hours
54,320.7

Downtime Hours
1,475.5

Production Hours by EQNO and Work Center

Downtime Hours by EQNO and Work Center

EQNO and Work Center Bubble - Production Hours vs Downtime Hours

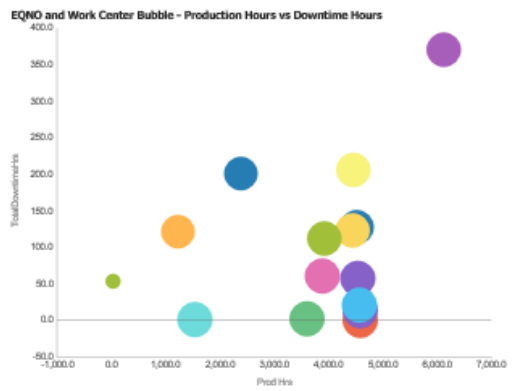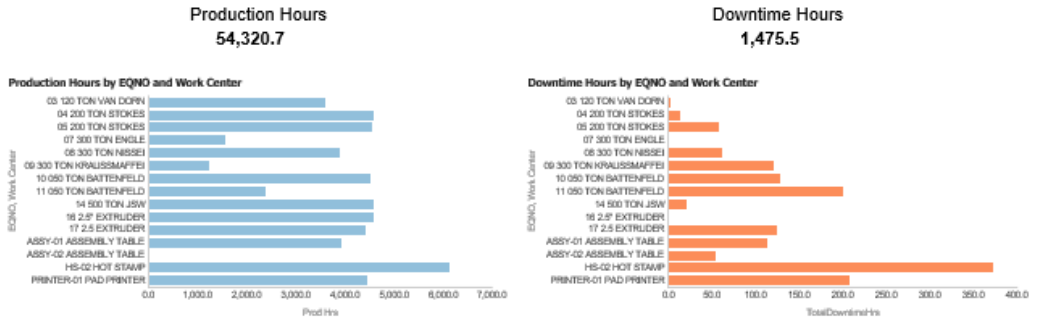| EQNO	Work Center	Item No	Item	Class	Total Hours	Line Utilization	Prod Hrs	Setup Hrs	Total Downtime Hrs	Downtime %	Availability % Actual
03	120 TON VAN DORN	A-200-H-BY	HANDLE, BATTERY	WP	3,607.1	1.00%	3,604.6	0.0	2.5	0.07%	99.9%
04	200 TON STOKES	05543-101	VOLUME CUP	FG	4,592.0	1.00%	4,578.5		13.5	0.29%	99.7%
05	200 TON STOKES	SH-020999-CLR	SHELF BRACKET 2 " CLEAR	FG	4,591.0	1.01%	4,533.6		57.4	1.25%	98.8%
07	300 TON ENGLE	9988-300	DOOR SEAL PLUG, SERIES 300	FG	1,537.8	1.00%	1,536.7	0.0	1.0	0.07%	99.9%
08	300 TON NISSEI	05092-102	RETAINER CLIP	FG	3,944.4	1.02%	3,883.7	0.0	60.7	1.54%	98.5%
09	300 TON KRAUSSMAFFEI	9988-300	DOOR SEAL PLUG, SERIES 300	FG	1,343.1	1.09%	1,222.1		121.0	9.01%	91.0%
10	050 TON BATTENFELD	HBUT-RD	1/2" BUTTON - RED	FG	4,680.0	1.02%	4,511.8	40.0	128.2	2.74%	96.4%
11	050 TON BATTENFELD	HBUT-BK	1/2" BUTTON - BLACK	FG	2,579.1	1.08%	2,378.3	0.0	200.8	7.79%	92.2%

vlamis SOFTWARE SOLUTIONS ™

Reject Analysis

IQMS® Manufacturing ERP

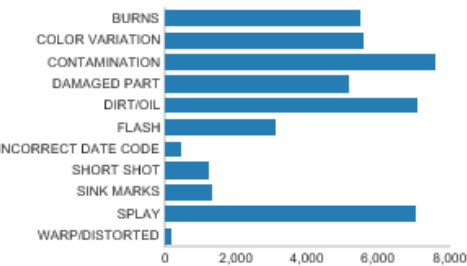CALENDAR YEAR
2018

CALENDAR MONTH
(All Column Value)

DATE

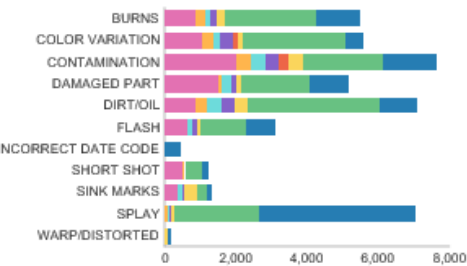EPLANT
--Select Value--

MFG CELL
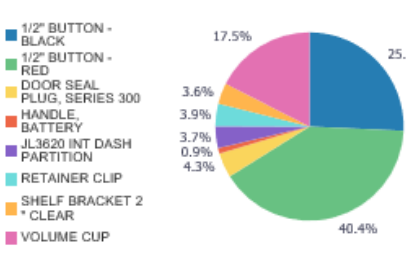--Select Value--

ITEM
(All Column Value)

RETURN TO MAIN MENU

Reject Qty by Reject Description
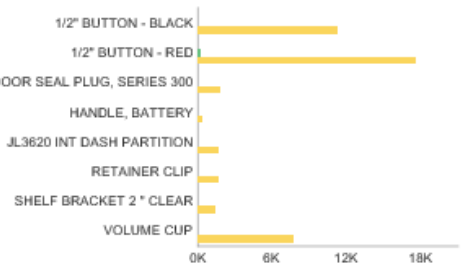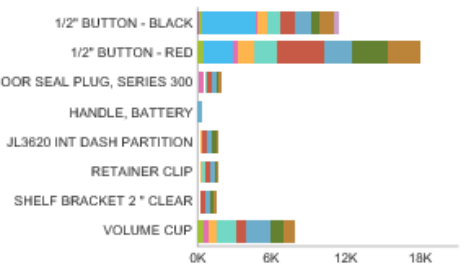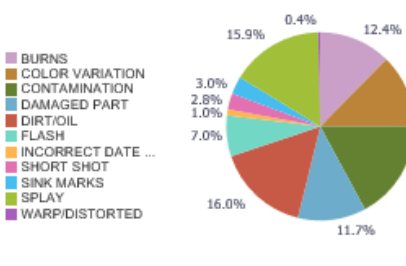### Reject Qty by Reject Reason and Product
Reject QTY by Product
Reject Qty by Product
Reject Qty by Product and Reject Reason
Reject QTY by Reject Reason

	BURNS	COLOR VARIA...	CONTAMINATI...	DAMAGED PART	DIRT/OIL	FLASH	INCORRECT D...	SHORT SHOT	SINK MARKS	SPLAY	WARP/DISTOR...
1/2" BUTTON - BLACK	1,237	521	1,446	1,103	1,038	811	462	154	131	4,411	99
1/2" BUTTON - RED	2,576	2,870	2,287	1,916	3,738	1,317		500	330	2,381	
DOOR SEAL PLUG, SI	246	131	420	166	373	85			321	76	68
HANDLE, BATTERY		139	277								
JL3620 INT DASH PAF	176	381	342	164	328	123			95	51	
RETAINER CLIP	138	174	437	262	409	114		27	89	69	
SHELF BRACKET 2 " (	266	313	409	82	369	35		49		67	
VOLUME CUP	850	1,063	1,990	1,506	843	632		495	382		

Color   Reject Qty   Low   High

Item	Reject Qty										
	BURNS	COLOR VARIATION	CONTAMINATION	DAMAGED PART	DIRT/OIL	FLASH	INCORRECT DATE CODE	SHORT SHOT	SINK MARKS	SPLAY	WARP/DISTORTED
1/2" BUTTON - BLACK	1,237	521	1,446	1,103	1,038	811	462	154	131	4,411	99
1/2" BUTTON - RED	2,576	2,870	2,287	1,916	3,738	1,317		500	330	2,381	
DOOR SEAL PLUG, SERIES 300	246	131	420	166	373	85			321	76	68
HANDLE, BATTERY		139	277								
JL3620 INT DASH PARTITION	176	381	342	164	328	123			95	51	
RETAINER CLIP	138	174	437	262	409	114		27	89	69	
SHELF BRACKET 2 " CLEAR	266	313	409	82	369	35		49		67	

Small Tables are Better

Consistency of interpretation

Size of table

vlamis™
SOFTWARE SOLUTIONS

Table View Use Cases

of
Columns

Direct Comparisons	**Exploration**
Explanation	**Sorted Views**

# of Rows
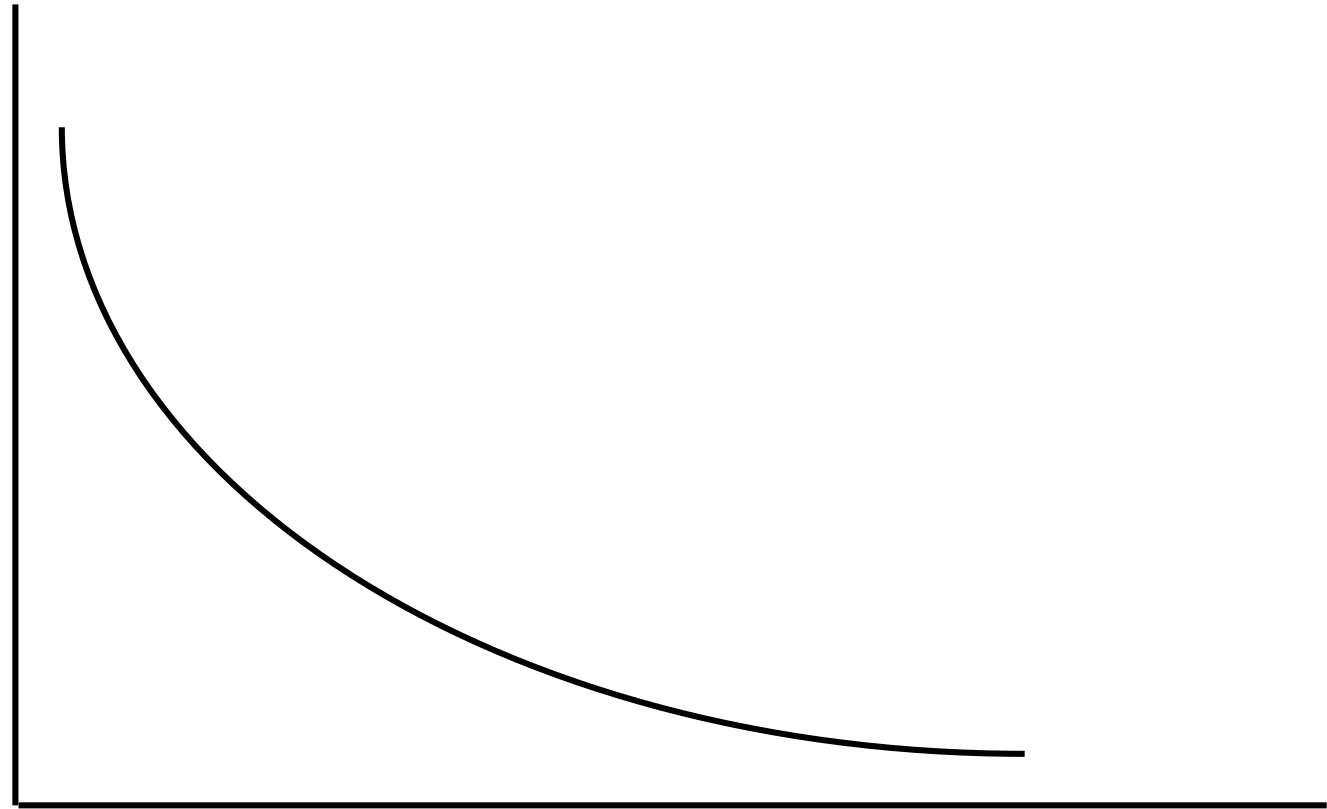
vlamis
SOFTWARE SOLUTIONS
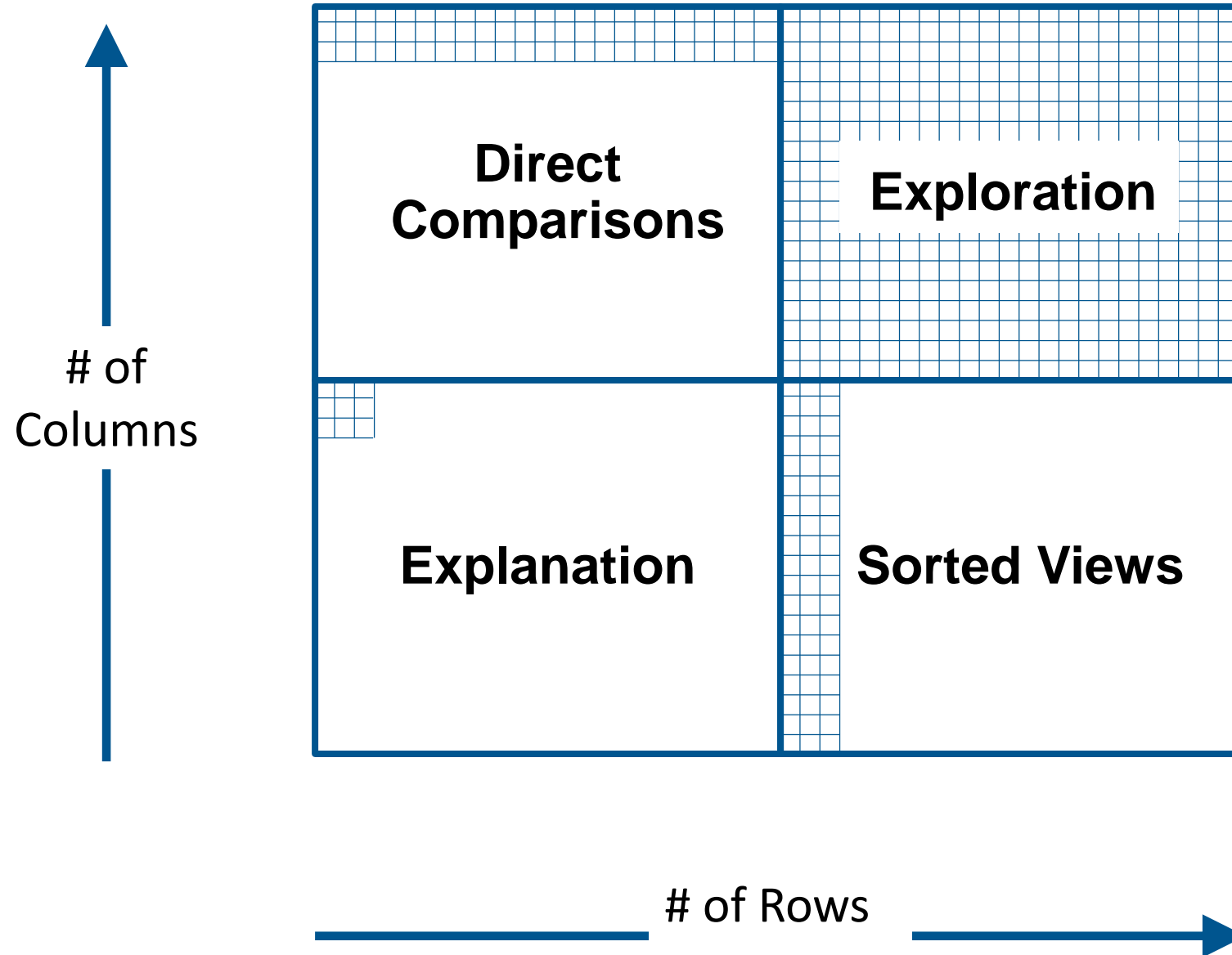
of Orders and Gross Profit Contribution for Customers 2014

Quadrant 1

Quadrant 2

Which quadrant should marketing target?

# of orders

Quadrant 4

Quadrant 3

Gross Profit Contribution

··· Average Gross Profit Contribution   ··· Average # of orders

ColorBrewer2.org

i want hue

W3Schools Color Picker

Four Realms of Decision Making

	Short Term	Long Term
Hi Impact	**Crisis**	**Strategy**
Low Impact	**Responsiveness**	**Planning**

Duration of Decision ↑

Latency of Decision →

vlamis™
SOFTWARE SOLUTIONS
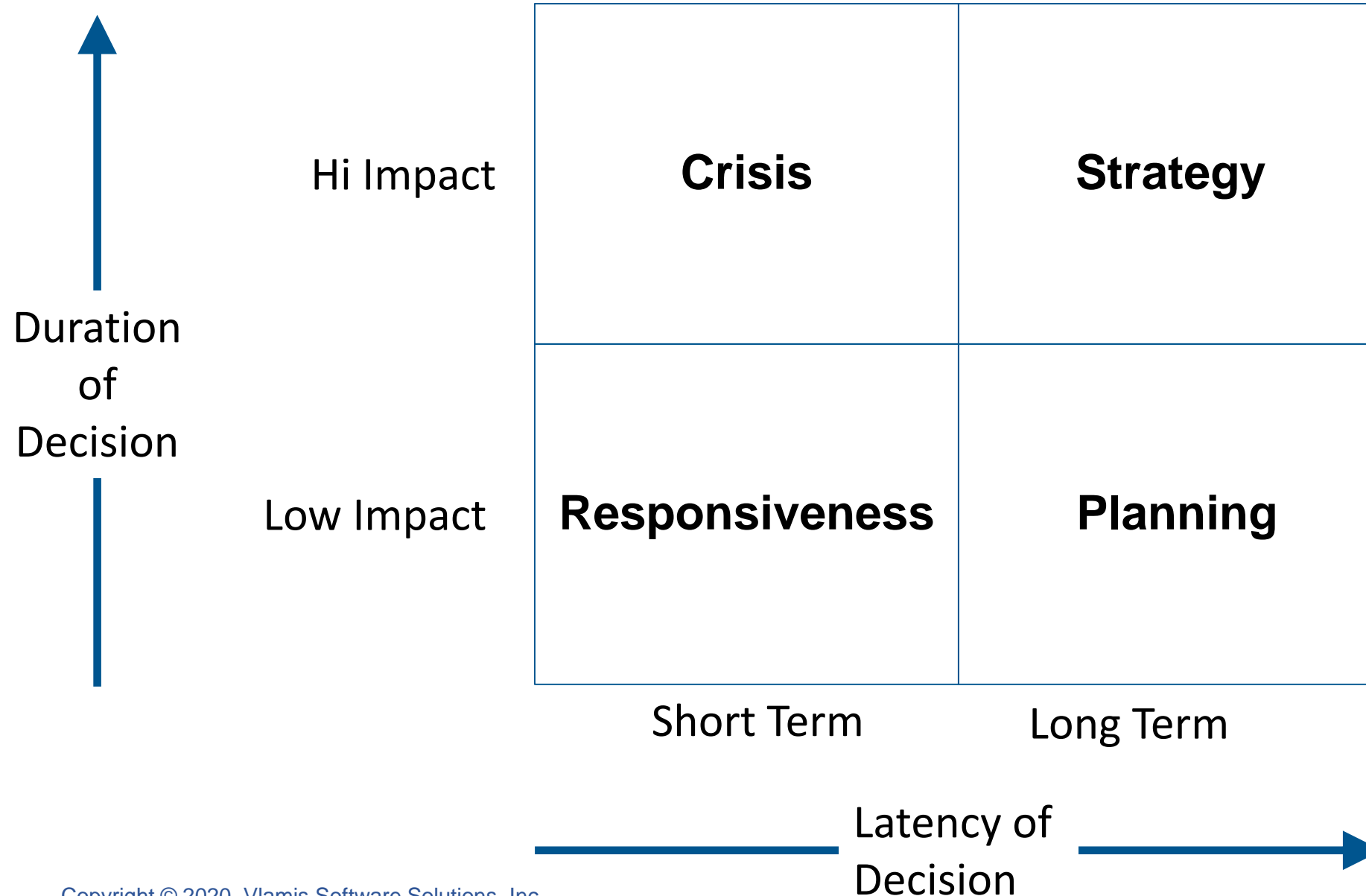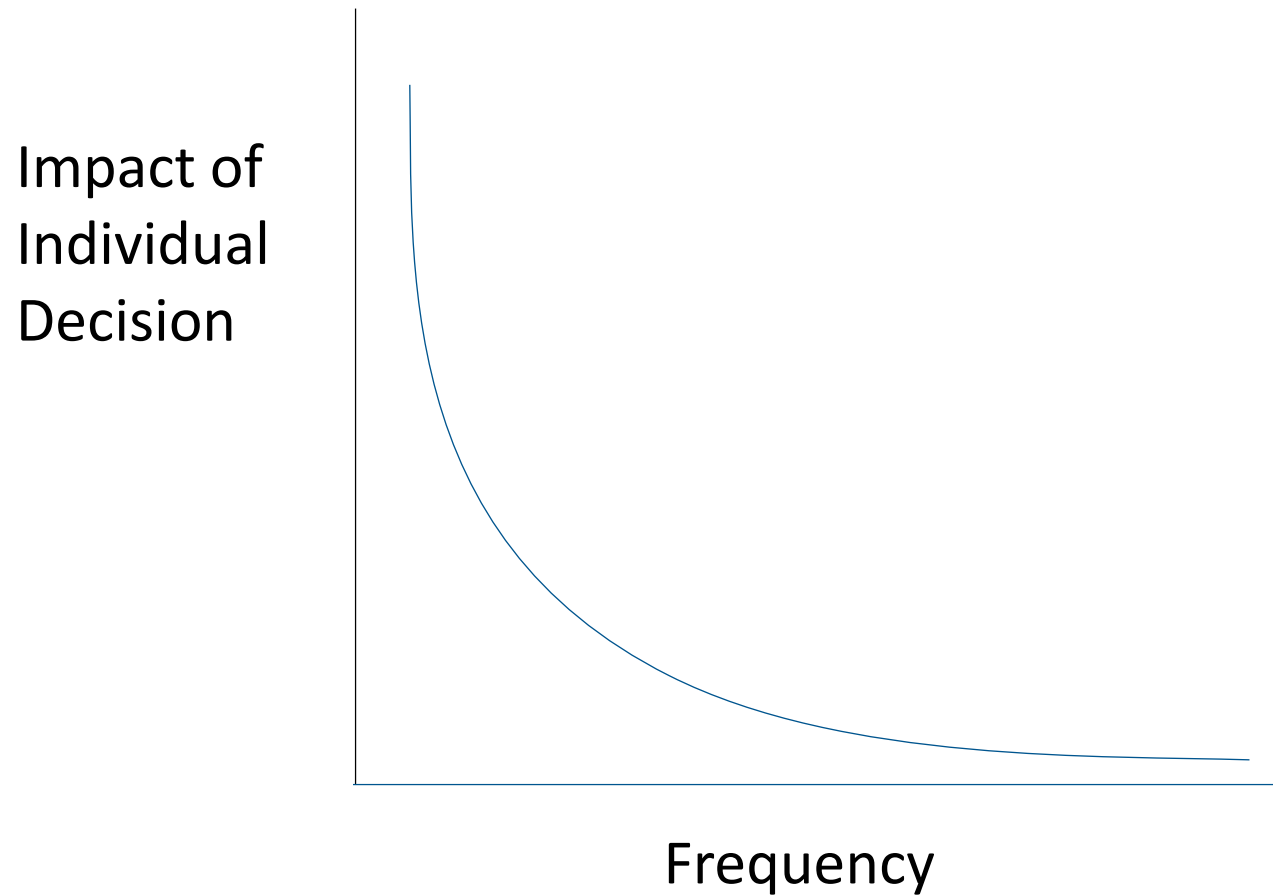
Organizational Decision Making

Impact of
Individual
Decision
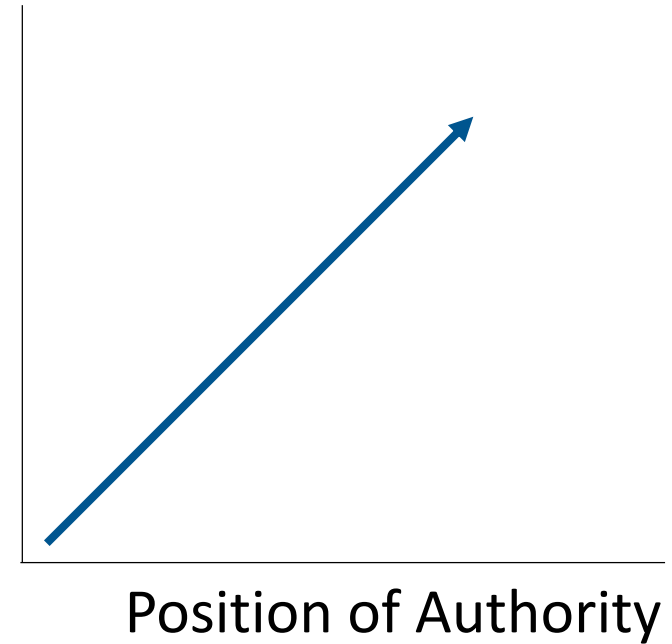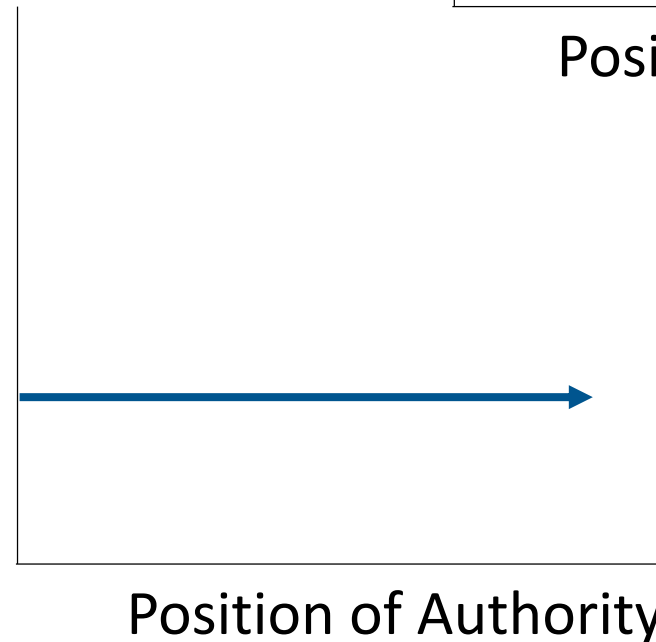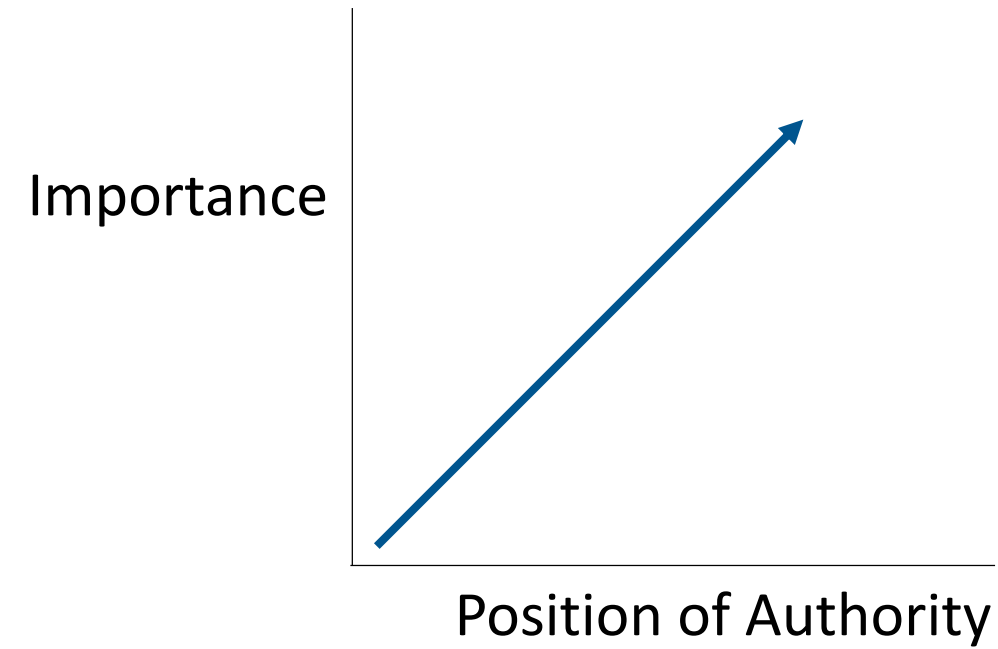
Frequency

vlamis ™
SOFTWARE SOLUTIONS

Organizational Decision Making

Dashboard Importance Score

Dashboard Importance Score

Importance

Position of Authority

Importance

Organizational Coordination Req.

vlamis™
SOFTWARE SOLUTIONS

Dashboard Importance Score

Importance

Position of Authority

Importance

# of Users

Importance

Organizational
Coordination Required

vlamis™
SOFTWARE SOLUTIONS
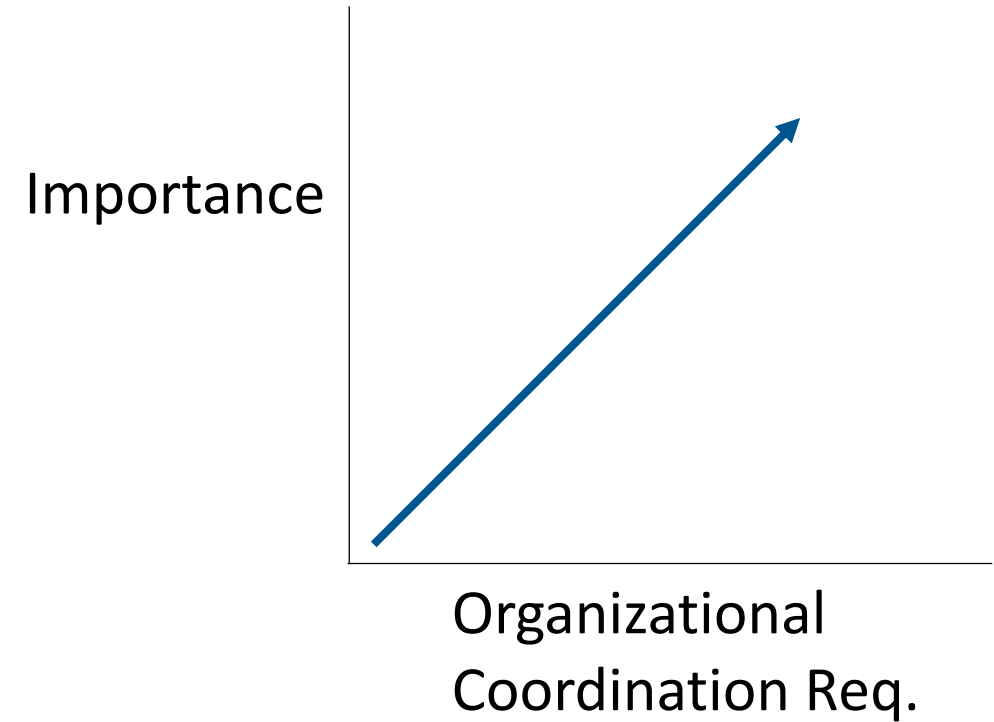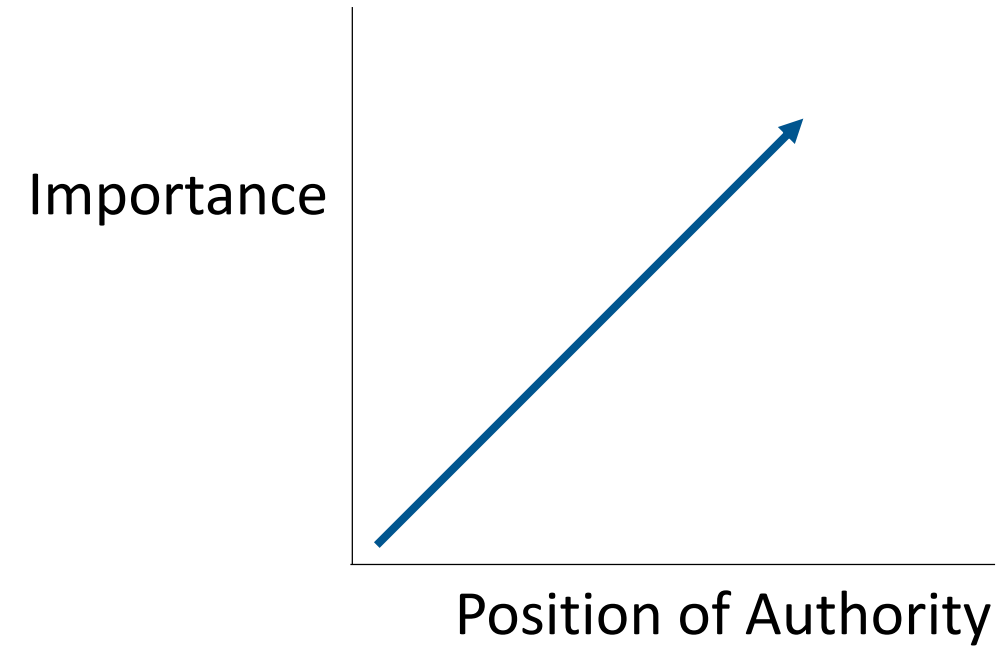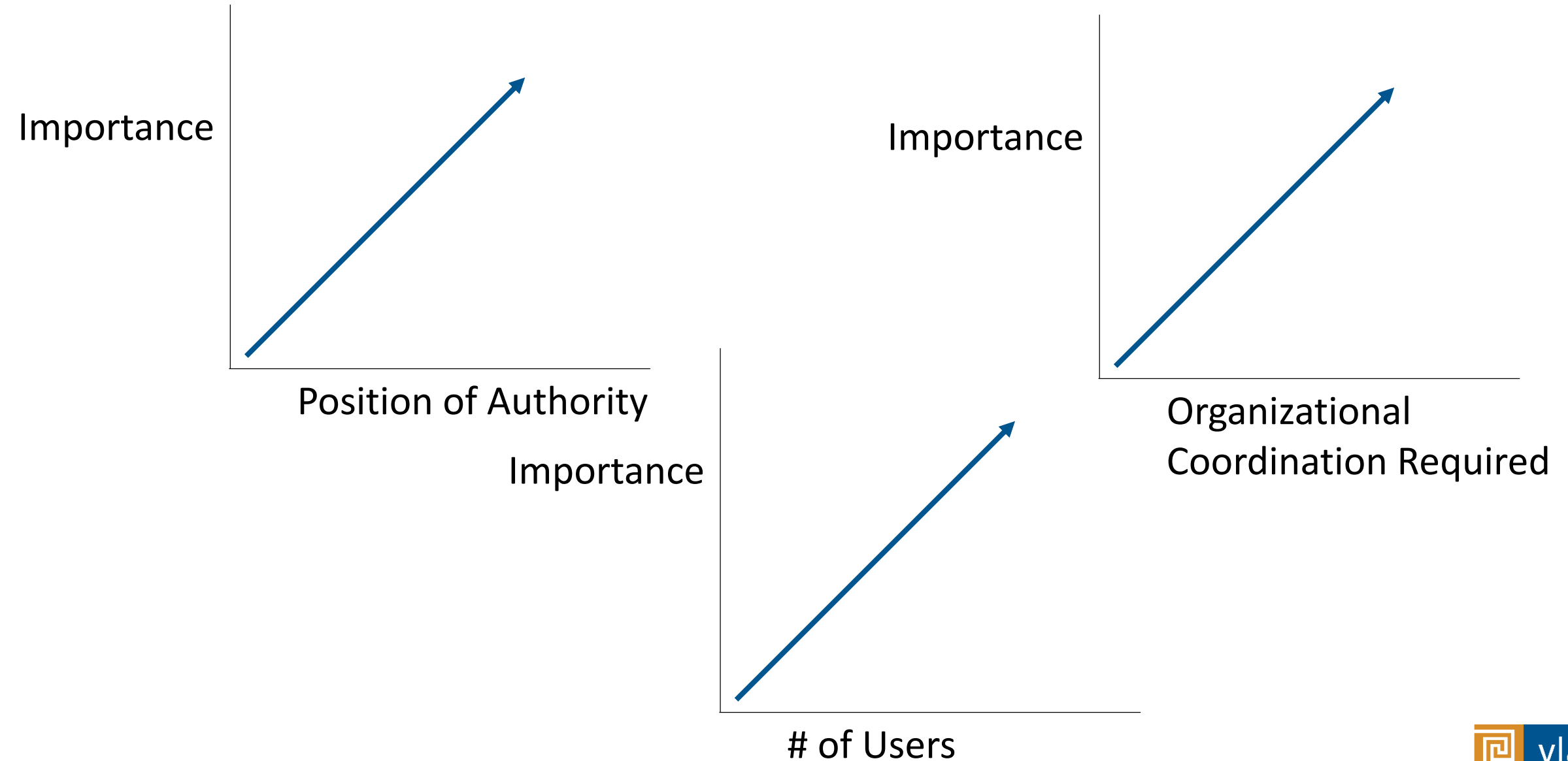
Dashboard Importance Score

$$\text{Dashboard Importance} = \sum_{1}^{n} \text{Position X Coordination X Users}$$

vlamis ™
SOFTWARE SOLUTIONS

Example Dashboard Importance Rubric

Division Operational Expenses YTD Dashboard

Role	Authority	Org Coord	Users	Product	Notes
Director	8	3	3.5	84	
Fin Analyst	2	3	4	24	
Managers	4	1	9	36	
Dashboard Importance				**144**	

Media Manager Monthly Dashboard

Role	Authority	Org Coord	Users	Product	Notes
Marketing Analyst	1	1	2	2	
Managers	4	2	2	16	
Dashboard Importance				**18**	

vlamis ™
SOFTWARE SOLUTIONS

Balance Choices with Shared Views

Shared Understanding

Data story/ visual narrative

Scorecard style executive dashboard

Traditional role-based dashboard

Prompt heavy dashboard

Individual Choice

vlamis
SOFTWARE SOLUTIONS
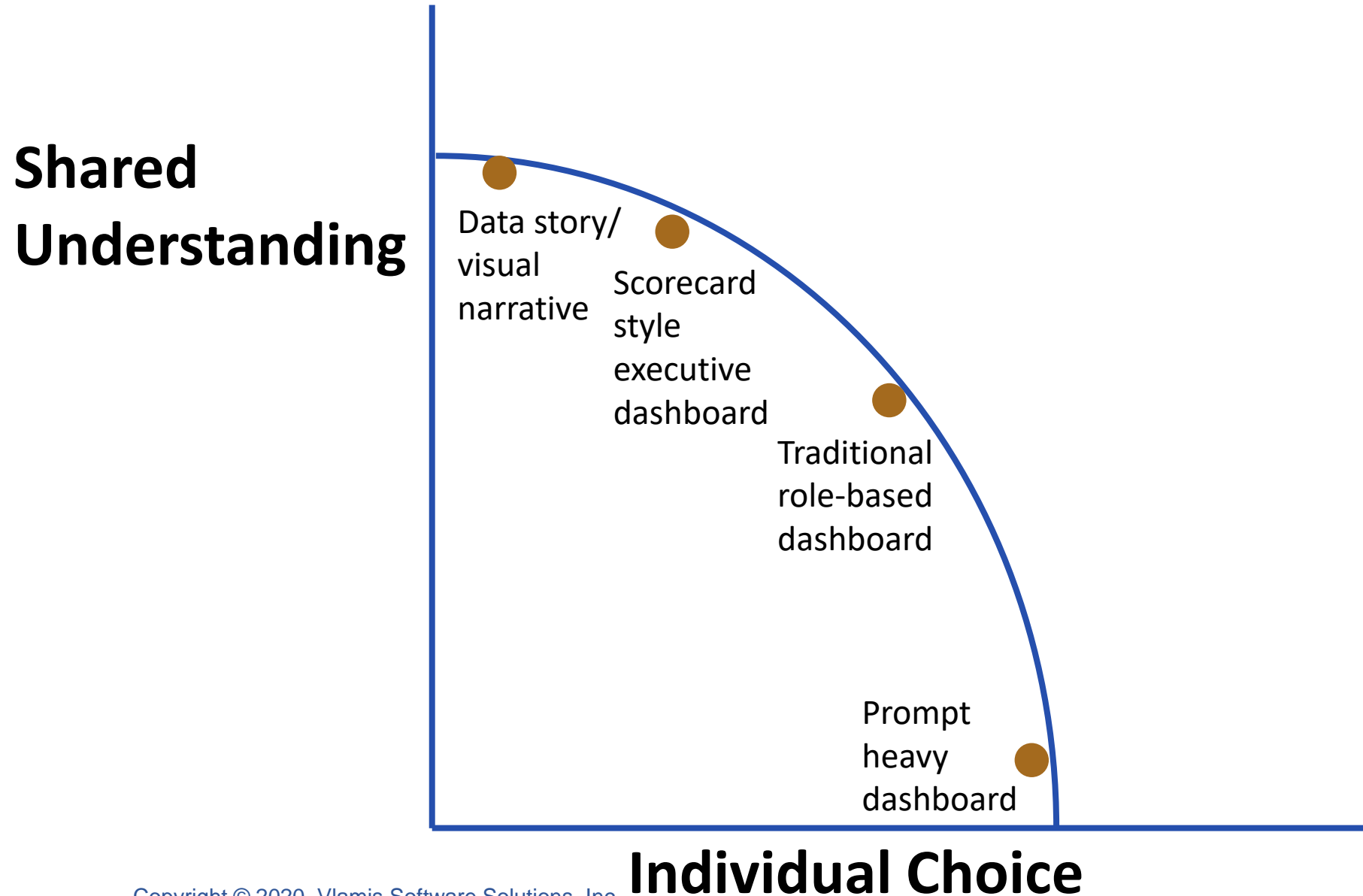
Strategic vs Tactical Dashboards

Required Organizational Coordination

General use case, strategic

Specific use case, tactical
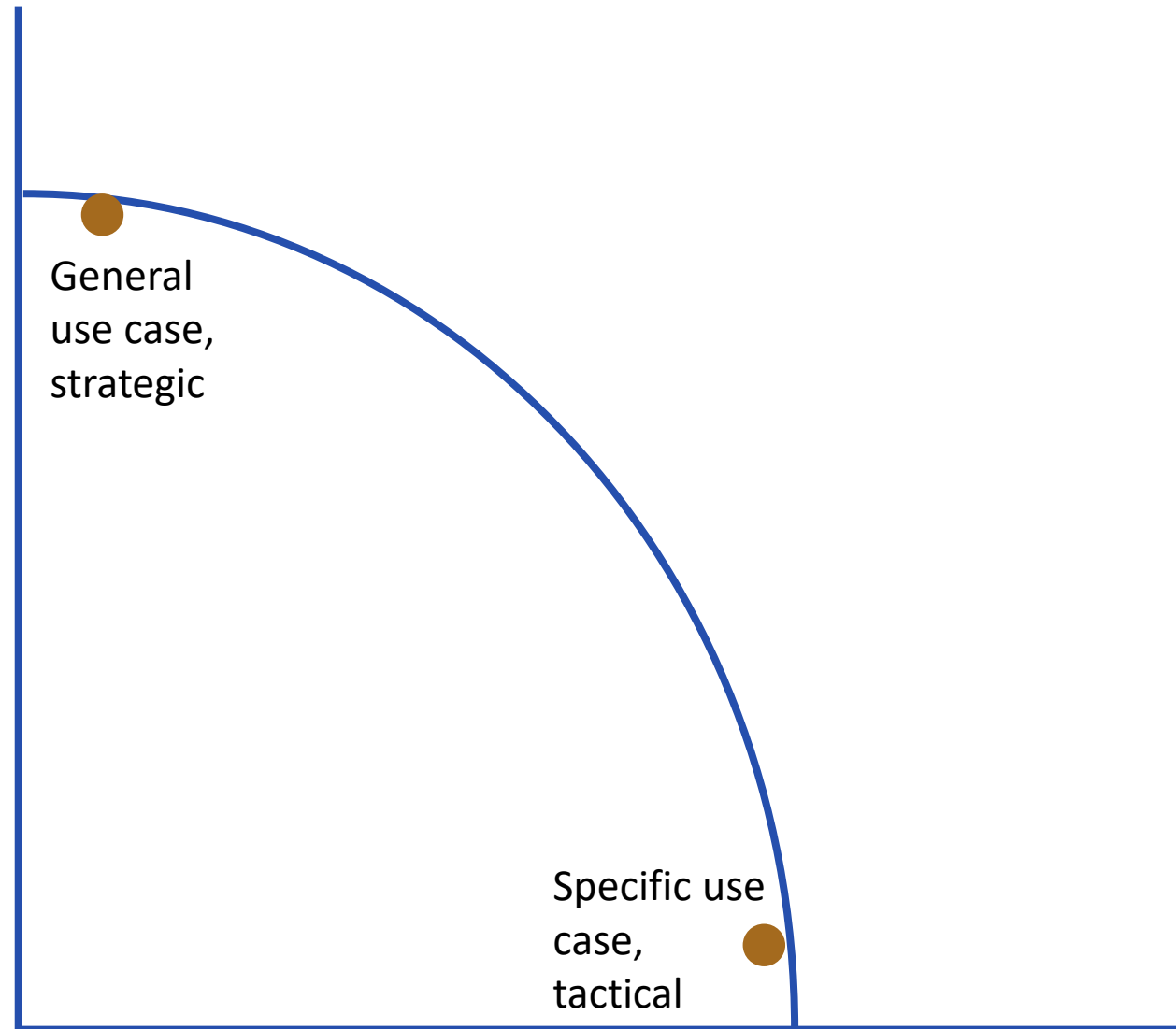
Individual Decision/Action

vlamis
SOFTWARE SOLUTIONS

Dimensional Analysis

- Enterprise systems have highly complex data
  - Facts/measures
    - raw data
    - calculated
  - Dimensions/attributes
    - Flat
    - Levels
    - Hierarchies
- Lots of strategies

vlamis ™
SOFTWARE SOLUTIONS

Hierarchies are Common in Enterprise Data

- Natural hierarchies
  - Levels
    - Often implied in share-type of calculations—e.g. Brand share (of category? Total?)
    - Sometimes level not important or identified—e.g. employee hierarchy
    - Drill order of levels e.g. Brand within Type or Type within Brand?
      - Define for organizational consistency and ease of use
      - Allow any combination for ad-hoc flexibility
    - How handle value that skip through parent?—e.g. Washington, D.C within US State
  - Names/keys—Are values at a level unique across parents?
    - Transparency in displays
    - Affect how values are labeled and identified

vlamis ™
SOFTWARE SOLUTIONS

Dimensional Analysis

- Use brushing and selection with multiple graph layouts.
  - Build four or five graphs with related attributes or measures.
  - Too many graphs or several highly dense graphs exceed limitations
- Consider alternative graph types
  - Scatter plots
  - Trellis charts
  - Sankey graphs
  - Parallel coordinates
  - Grid heat maps

vlamis™
SOFTWARE SOLUTIONS

Dimensional Analysis

- Order of importance for Scatter Plots
  1. Y Axis typically has the "response variable", i.e. highest interest
  2. X axis has the "independent variable".

  3. Color (can be categorical or numeric)
  4. Size
  5. Trellis by category
  6. Shape
  7. Filters

- Use logarithmic scale for "long tail" distributions or break into two or more graphs.

vlamis ™
SOFTWARE SOLUTIONS
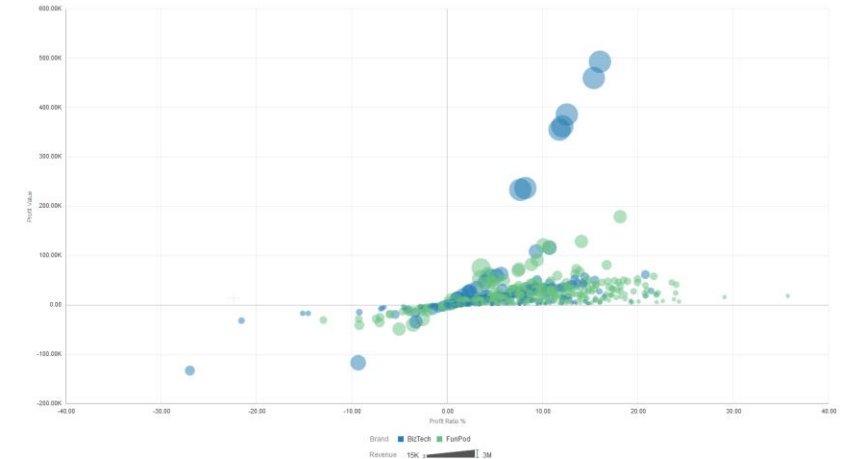
Trellis Charts

- Make sure that the major axis of interest is aligned with Trellis chart choice.
  - Vertical when X axis is important
    - Example: compare patterns over time
    - Compare length of horizontal bar graph
  - Horizontal when Y axis important
    - Compare lengths of vertical bar graphs
- Use horizontal for long, scrolling trellis charts with many members
- Use both to create a grid of graphs

vlamis ™
SOFTWARE SOLUTIONS

Sankey Graphs

- Used in "flow" analyses and comparative analyses
- Used to show relative strengths of relationships between attributes
- Line weight and size are proportional to flow/relational measure
- Hover and click on lines to show relationships
- Sort order is very important
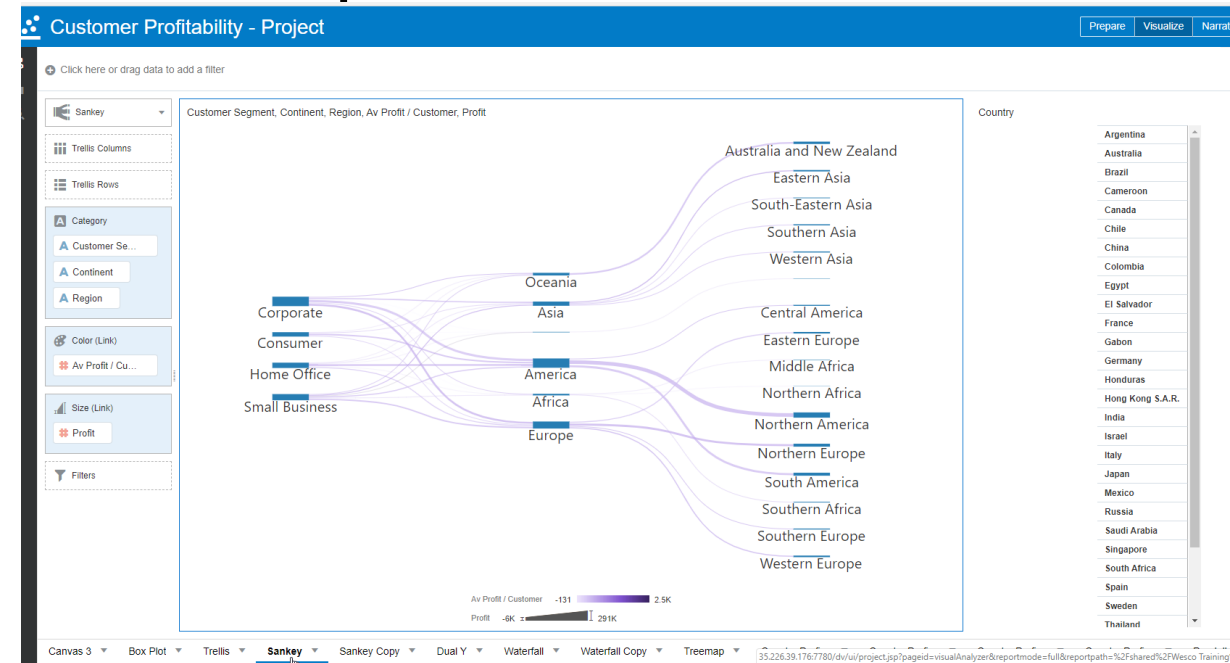
Parallel Coordinates Graphs

- Used to show otherwise disparate relationships
- "Custom join graph"
- Each line represents a record in the **active** data set
- Sort order is extremely important
- Highly interactive
- Not recommended for general users
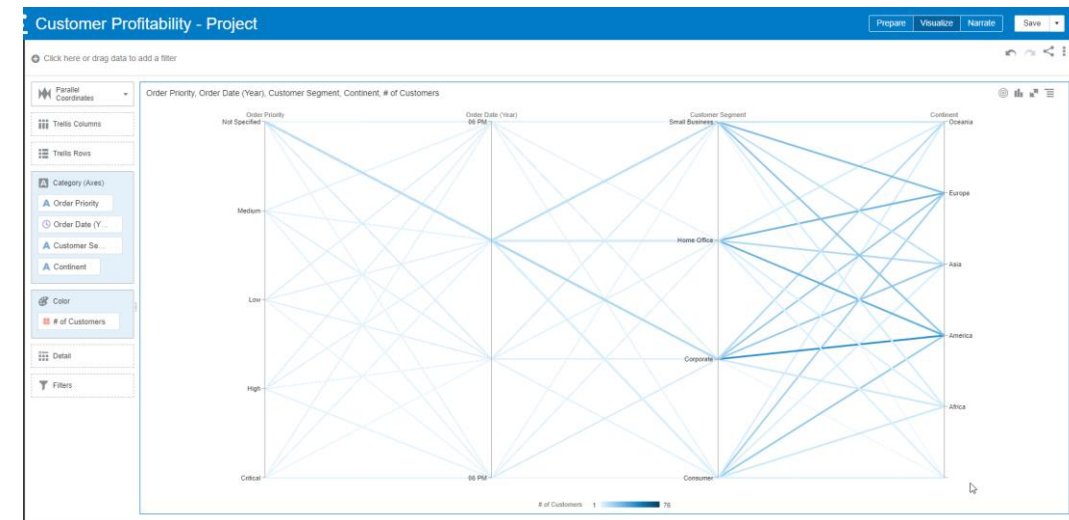
Bin or Bucket data to facilitate insights

- Equal width bins
  - Read number of members in bins
- Ntile bins
  - Read level of and relationship between bin values
- Custom range bins
  - Evaluate member sets within bins

vlamis ™
SOFTWARE SOLUTIONS

Keys to Data Discovery

- Identify your main topic of interest with a performance tile
- Summary
- Evaluating a fact or a dimension?
  - Sales analysis
  - Customer or product analysis
- Fact analysis
  - Find lowest grain
  - Flat low distribution
  - Event or transaction
- Look for clustered distribution
  - Scatter with points as event in fact table
    - Set fact on X axis and response variable on Y axis

Fact	Dimension 1
Dimension 3	Dimension 2

vlamis ™
SOFTWARE SOLUTIONS

Keys to Data Discovery

- Identify your main topic of interest with a performance tile
- Summary
- Evaluating a fact or a dimension?
  - Sales analysis
  - Customer or product analysis
- Fact analysis
  - Find lowest grain
  - Flat low distribution
  - Event or transaction
- Look for clustered distribution
  - Scatter with points as event in fact table
    - Set fact on X axis and response variable on Y axis
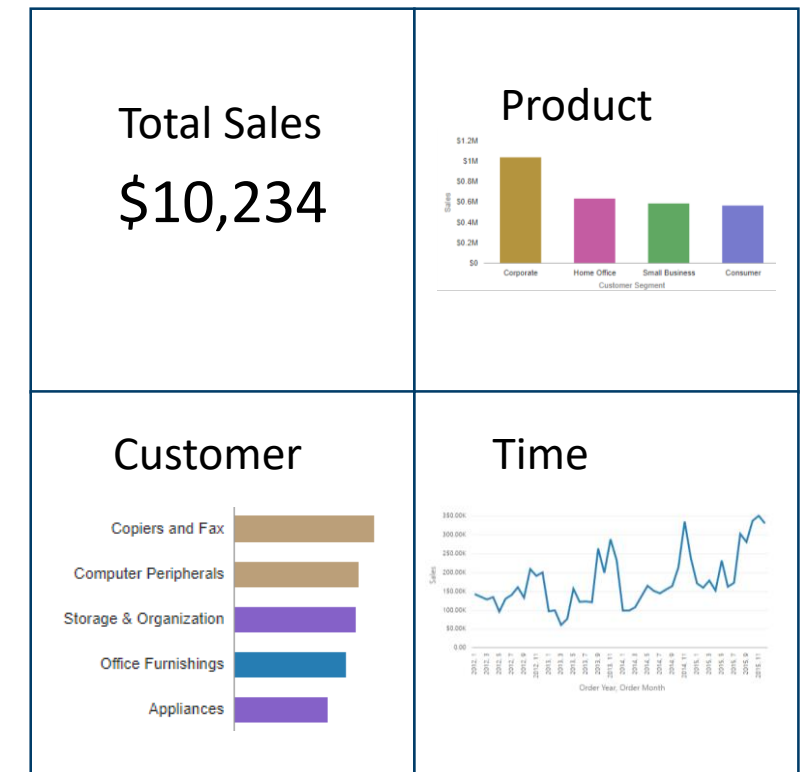
Visual Discovery and Analytic Techniques

- Graph distributions of data

- Seek outliers

- Graph differences directly

- Normalize data to facilitate comparisons

- Bin or Bucket data to facilitate insights

- Use high density graphs to uncover potentially meaningful attributes

- Choose a meaningful sort order for every visualization

- Determine the importance of different measures and attributes and place them in the appropriate place for every visualization.

Progression of Data Explanations

1. True exploration (new data set, unknown insights)
2. Ad hoc discovery (known data set, seeking new insights)
3. Guided navigation
4. Selections and reading
5. Summary dashboard
6. Narration and storytelling

vlamis
SOFTWARE SOLUTIONS

Starting with Data Discovery

- Begin either with a specific question or a framework

- Avoid "wandering around"

- Most of your visualizations will not produce new insights

- Move quickly through visualizations

- Be prepared to open a lot of browser tabs

vlamis™
SOFTWARE SOLUTIONS

Finding is not Explaining

- Process of interaction has a huge impact on the contextual understanding of an insight

- When someone discovers something, they believe it more

- Human Cognition Biases

vlamis ™
SOFTWARE SOLUTIONS

Six Steps of Persuasion

- Introduction (capture attention, establish credibility)
- Narration (summary of facts)
- Division (organize and subset facts)
- Proof (logically derive your conclusions)
- Refutation (address opposing arguments)
- Conclusion (satisfy emotional appeal)

vlamis
SOFTWARE SOLUTIONS

Bad Exec Dashboards Cause Grind

- People see different things
- Assumed context
- Data is distorted

vlamis ™
SOFTWARE SOLUTIONS

Great Exec Dashboards Reduce Friction

- Common data
- Shared context
- Established prioritization

BI Standards Drive Value

- Best Practice Driven
- Work best when documented and reinforced with examples
- Dashboard layout and style
  - Prompt placement
  - Prompt styles
  - Navigation
- Naming and Titles
- Graph design
- Table design
- Color palettes and assignments

vlamis™
SOFTWARE SOLUTIONS

Keys to Effective Data Story Telling

- Have a main idea or key point every visualization/layout
- Give your key point the most visual weight
- Provide supporting context and data for your key point
- Address potential objections and justify choices/assumptions
- Summarize your main point

vlamis ™
SOFTWARE SOLUTIONS

Good Questions Guide Priorities

- What is the key message or insight?

- Which is more important, this or that?

- What comparison do you want? Difference? Percentage difference? Etc.

- This version emphasizes this, this other version emphasizes that, which do you prefer?

- What do you want everyone else to understand from this?

- Would you rather we spend time on this or that?

- Who else in your team will use this? Will others outside your team use this version?

- Currently we have x hours estimated to work on this, which is (high, low, average), sound right?
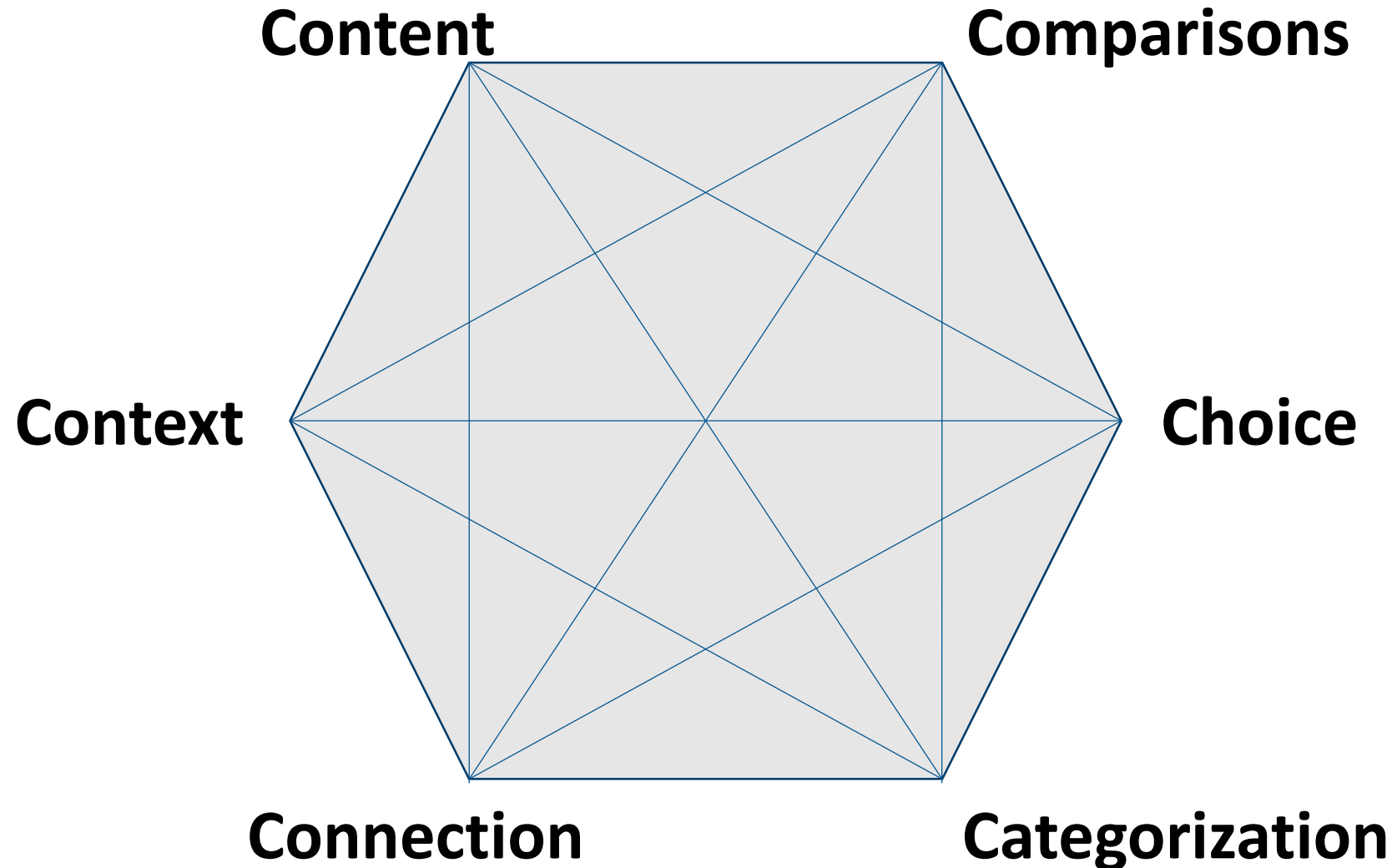
vlamis ™
SOFTWARE SOLUTIONS

Bad Questions Lead Executives Astray

- What do you want?

- How do you want me to lay this out?

- What kind of graphs do you like?

- How much data do you want?

- Do you like the colors?

- How much time do you want me to spend on this?

vlamis ™
SOFTWARE SOLUTIONS

Three-Way Role Playing

- Developer, client, coach
- Developer
  - Practice suggestions and techniques learned during training
  - Lead by asking good questions
  - Thank the coach for observations and client for challenge
- Client Role:
  - Be a typical, demanding client
  - Don't be unreasonable or purposefully difficult
  - Don't be a pushover
- Coach Role
  - Observe developer and client interaction without interference
  - Offer positive feedback and things to consider
- Rotate roles every round

Great Dashboards Balance 6 Requirements

Different Strategies

- If it's worth doing, it's worth doing right.

- The perfect is the enemy of the good.

vlamis
SOFTWARE SOLUTIONS

Questions?

Thank You!!

- Tim Vlamis   tvlamis@vlamis.com

- Dan Vlamis   dvlamis@vlamis.com

vlamis
SOFTWARE SOLUTIONS

Resources

- ColorBrewer website. Color ramps. http://colorbrewer2.org/

- Iwanthue. color ramps and scripts for data scientists https://medialab.github.io/iwanthue/

- W3 Schools Color Picker. Adjust color intensity

https://www.w3schools.com/colors/colors_picker.asp

vlamis
SOFTWARE SOLUTIONS