



SOFTWARE SOLUTIONS

Strategies and Layouts for Data Visualization Success

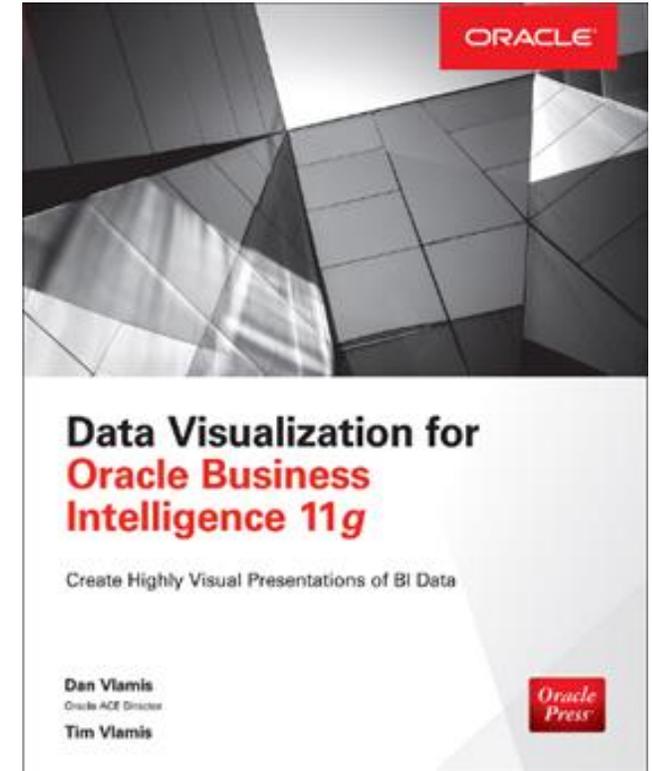
Tim and Dan Vlamis

May 24, 2022

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

Dan Vlamis – President

- Founded Vlamis Software Solutions in 1992
- 30+ years in business intelligence, dimensional modeling
- Oracle ACE Director  
- Developer for IRI (expert in Oracle OLAP and related)
- BIWA Board Member since 2008
- BA Computer Science Brown University
- @dvlamis

Tim Vlamis – Vice President & Analytics Strategist

- 30+ years in business modeling and valuation, forecasting, and scenario analyses
- Oracle ACE Director  
- 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
- @TimVlamis

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.”*

Sir Arthur Conan Doyle

The Adventure of the Copper Beeches

Copyright © 2022, Vlamis Software Solutions, Inc.



Think Evidence, Not Analytics

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



Simple Pivot Table with Easy Questions

2014 Monthly Sales by Company

D4 Company	2014 / 01	2014 / 02	2014 / 03	2014 / 04	2014 / 05	2014 / 06	2014 / 07	2014 / 08	2014 / 09	2014 / 10	2014 / 11	2014 / 12	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?

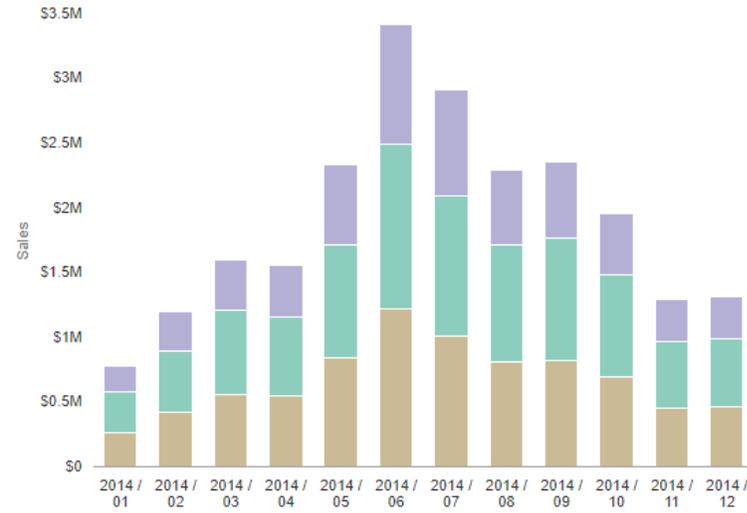
Answers are Easy to Find in Graphs

Total Sales 2014

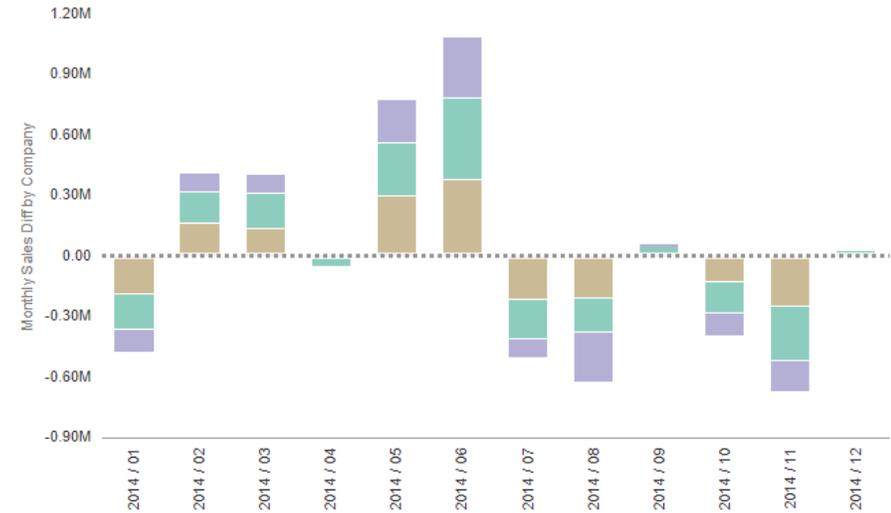
\$23,000,000



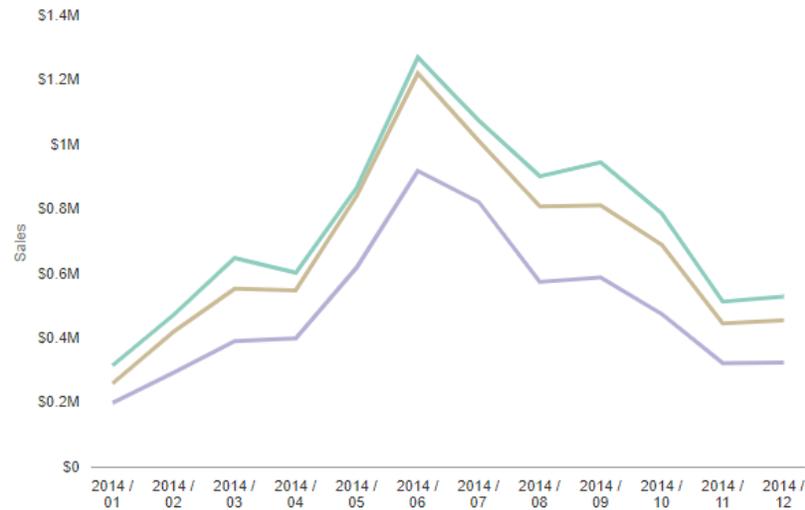
2014 Monthly Sales by Company



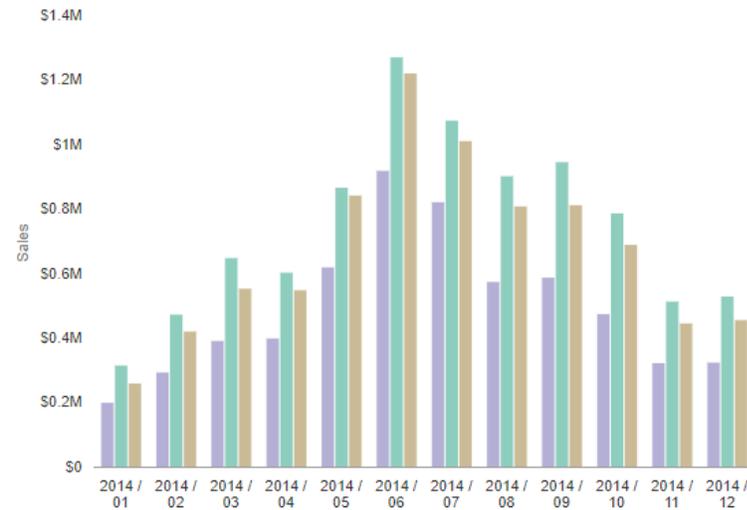
Difference from Previous Month Sales by Company



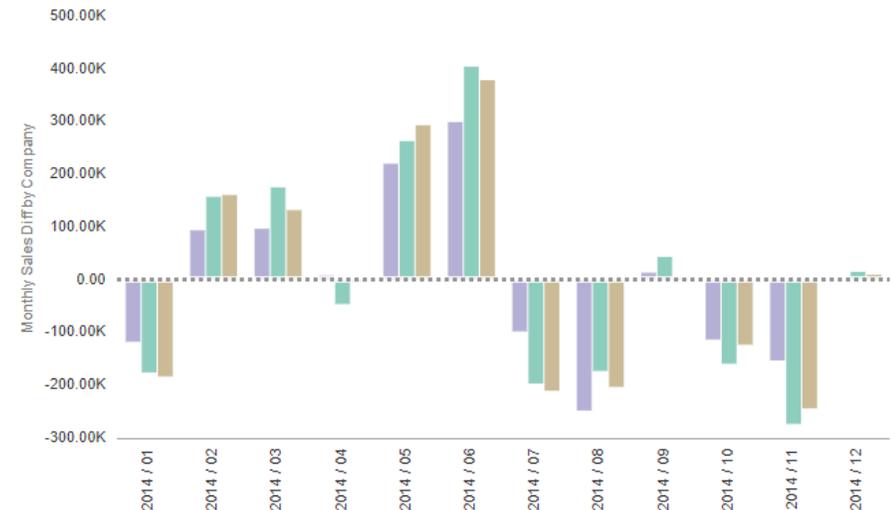
2014 Monthly Sales by Company



2014 Monthly Sales by Company



Difference from Previous Month Sales by Company



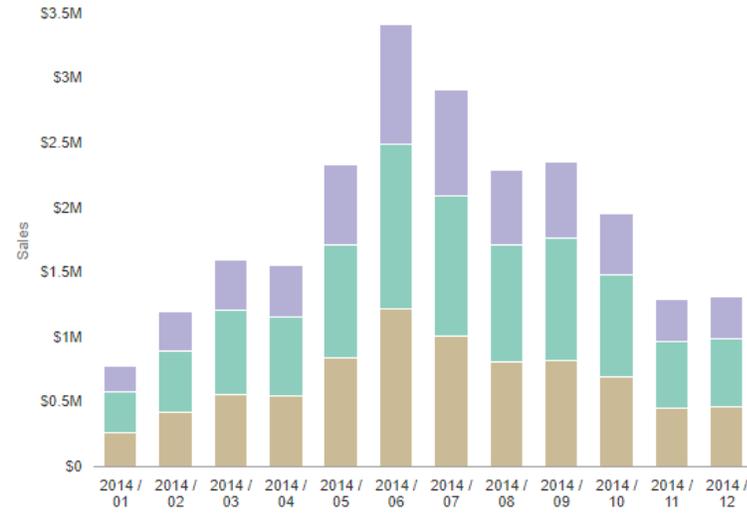
What was the highest monthly sales for a company?

Total Sales 2014

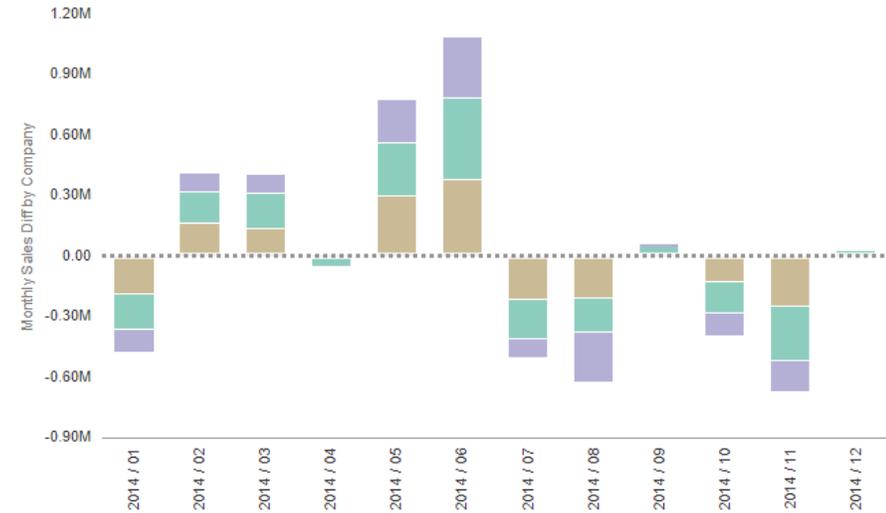
\$23,000,000



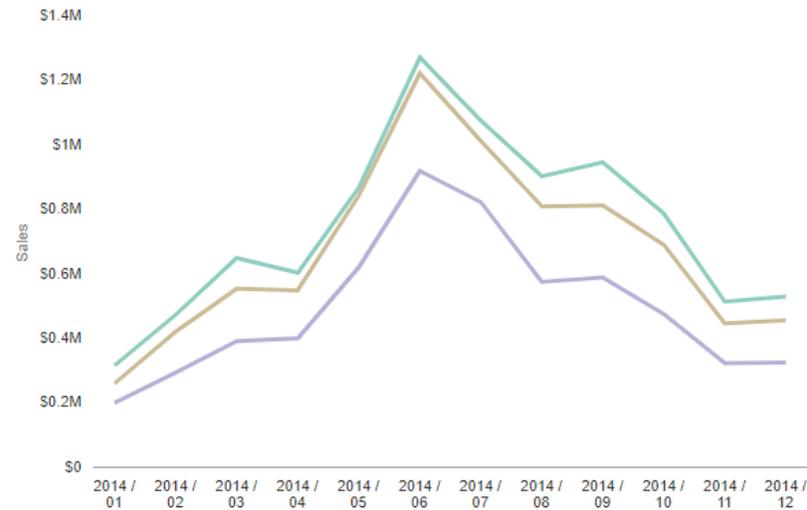
2014 Monthly Sales by Company



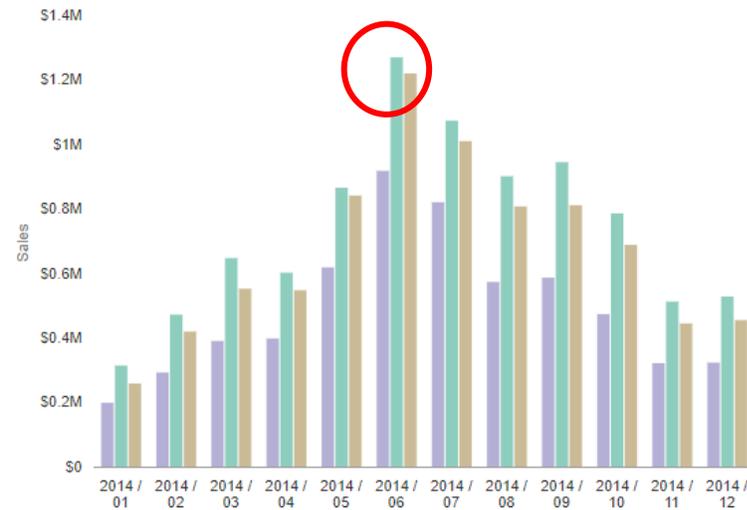
Difference from Previous Month Sales by Company



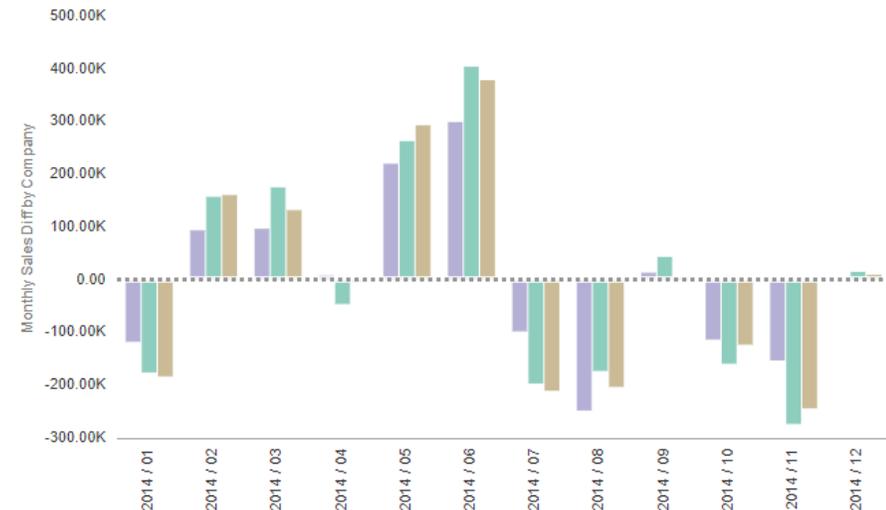
2014 Monthly Sales by Company



2014 Monthly Sales by Company



Difference from Previous Month Sales by Company



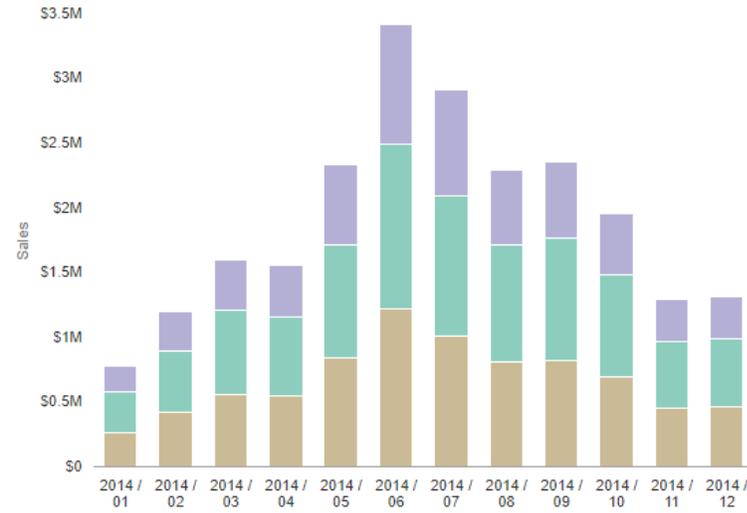
Which month had the largest drop in sales for a company?

Total Sales 2014

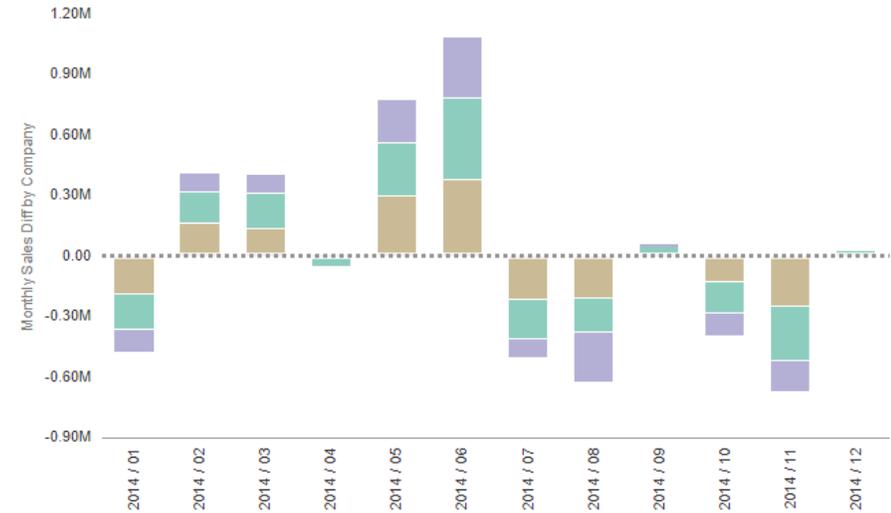
\$23,000,000



2014 Monthly Sales by Company



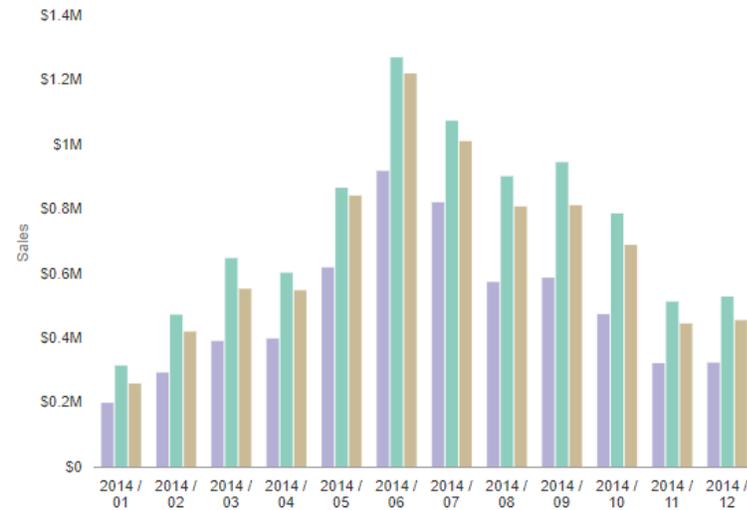
Difference from Previous Month Sales by Company



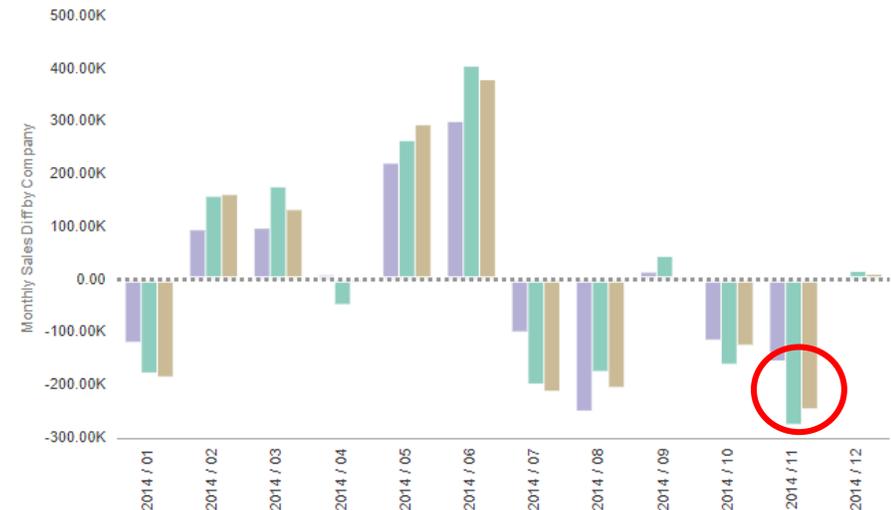
2014 Monthly Sales by Company



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

\$5,953,688

2014 Sales

Stockplus Inc.

\$8,954,066

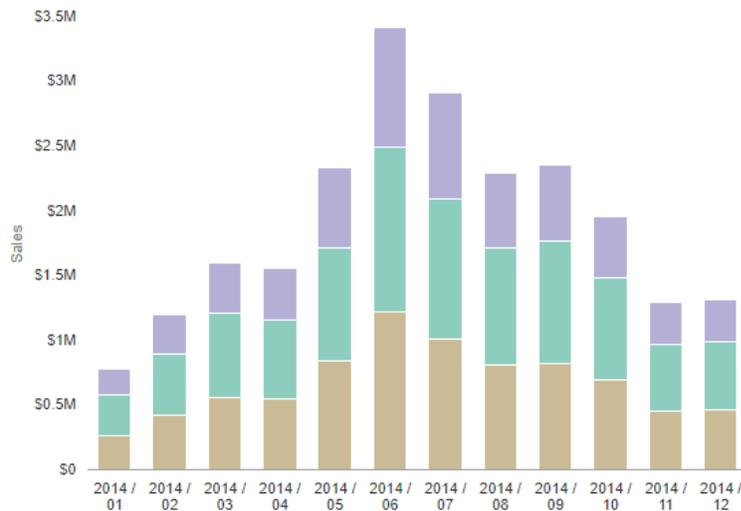
2014 Sales

Tescare Ltd.

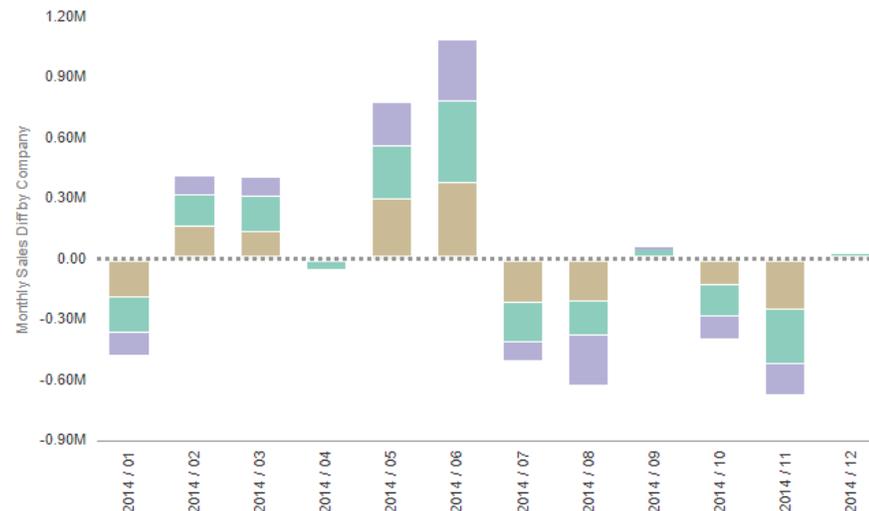
\$8,092,246

2014 Sales

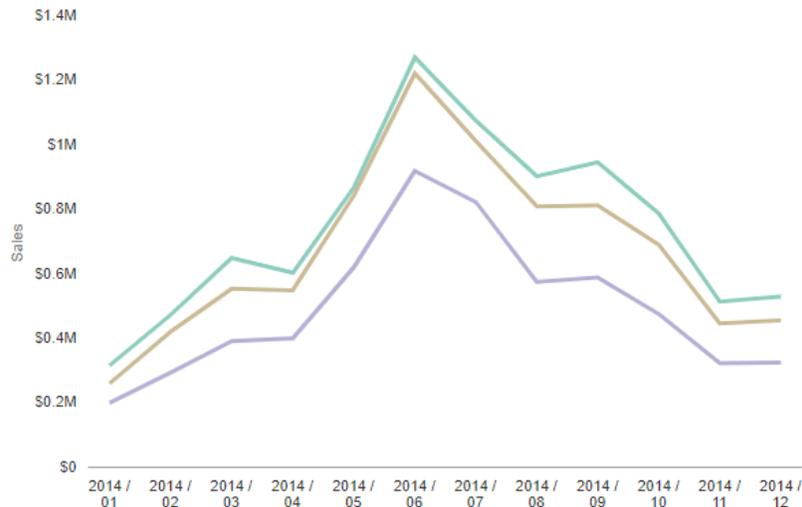
2014 Monthly Sales by Company



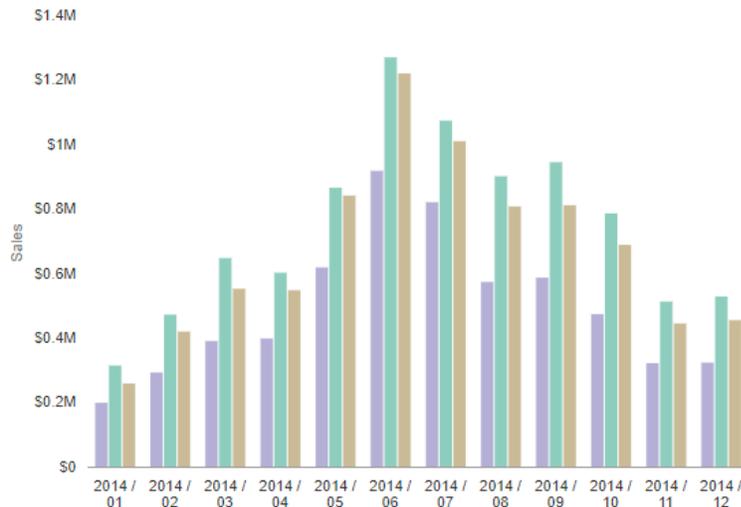
Difference from Previous Month Sales by Company



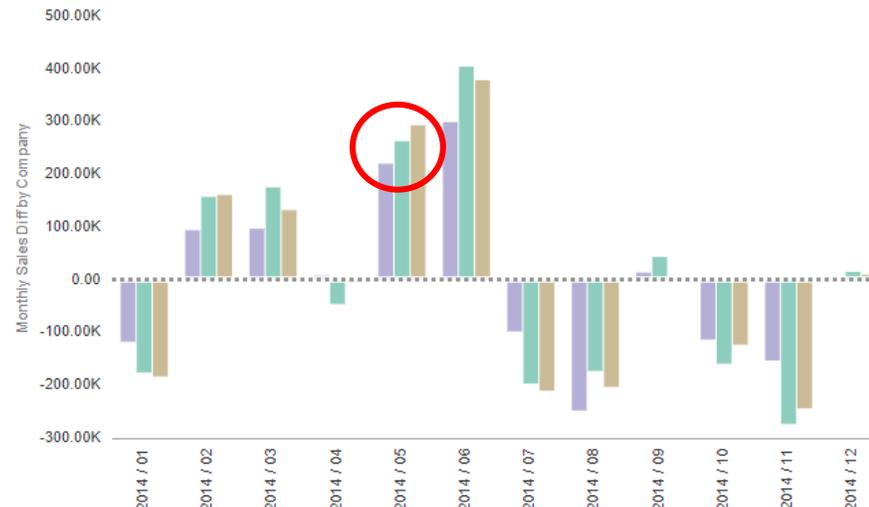
2014 Monthly Sales by Company



2014 Monthly Sales by Company



Difference from Previous Month Sales by Company



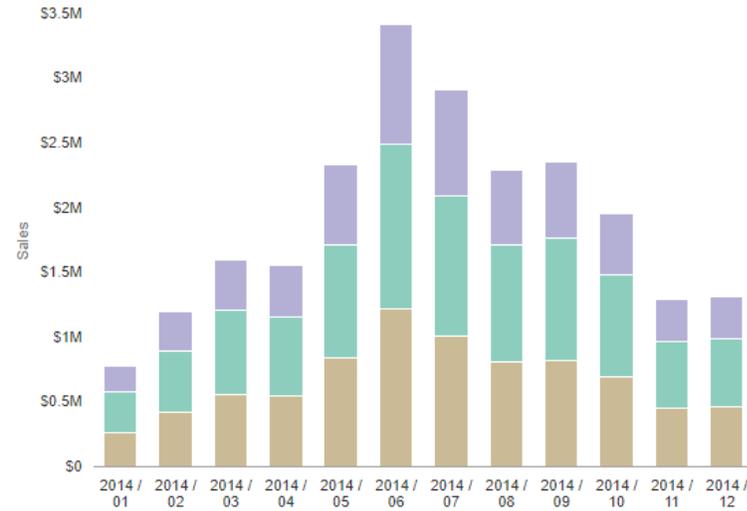
Is there a seasonal pattern to sales for all companies?

Total Sales 2014

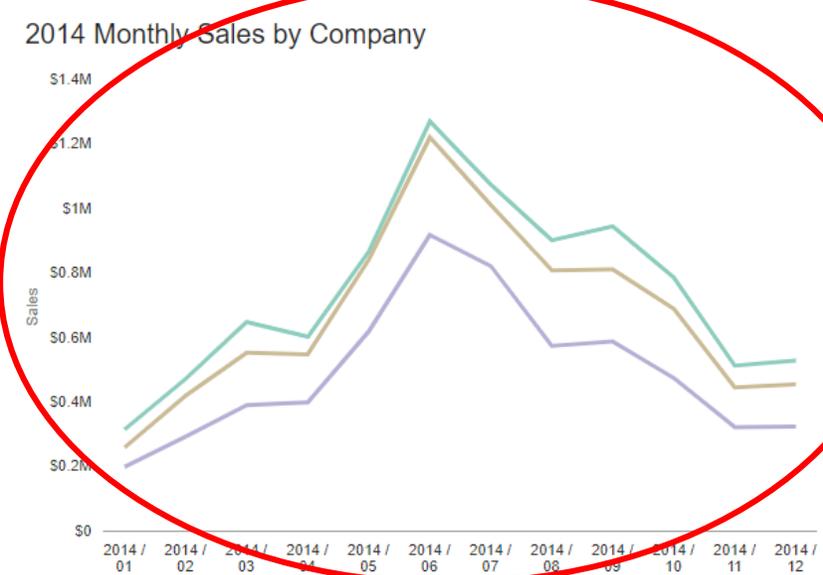
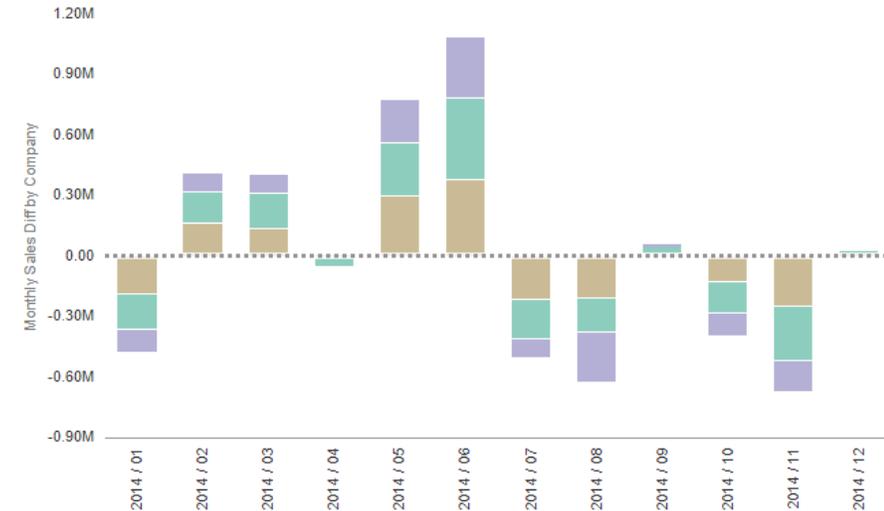
\$23,000,000



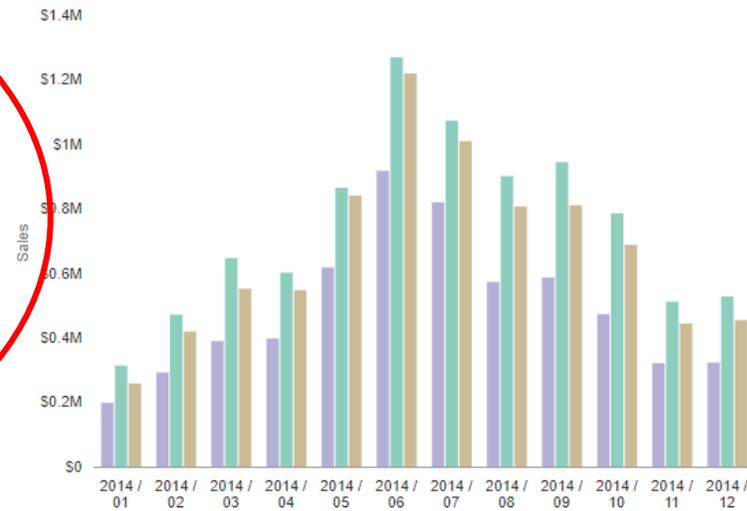
2014 Monthly Sales by Company



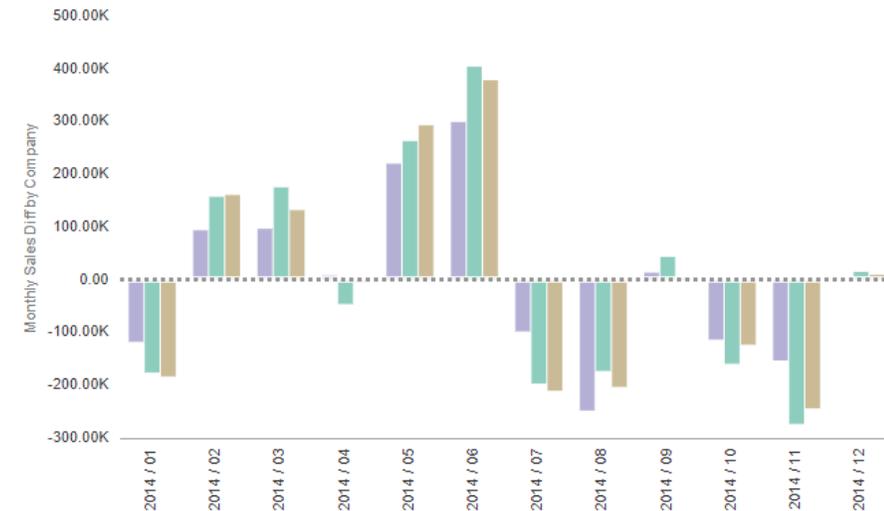
Difference from Previous Month Sales by Company



2014 Monthly Sales by Company



Difference from Previous Month Sales by Company



It's OK to Repeat Data in Multiple Views



Layouts

- Four Quadrant layout
- Summary KPI layout
- Multi-dimension layout
- Driving graph layout

Four Quad Summary Layout

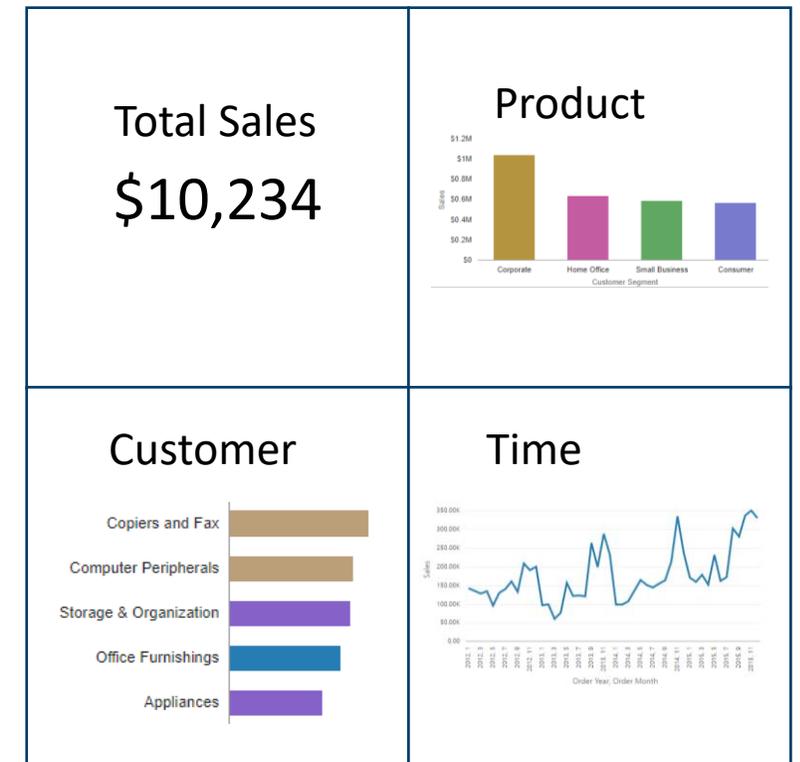
Keys to Data Discovery

- 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

Summary 4 Quadrant Layout

- Evaluating a fact or a dimension?
 - Sales analysis
 - Customer or product analysis
- Identify your main topic of interest with a performance tile
- Summary view of three key dimensions for fact topics
- Summary view of dimension measures and key comparison for dimensions
 - Versus last period/forecast/budget
 - Top N (magnitude, change, %, new)
 - RFM (recency, frequency, magnitude)

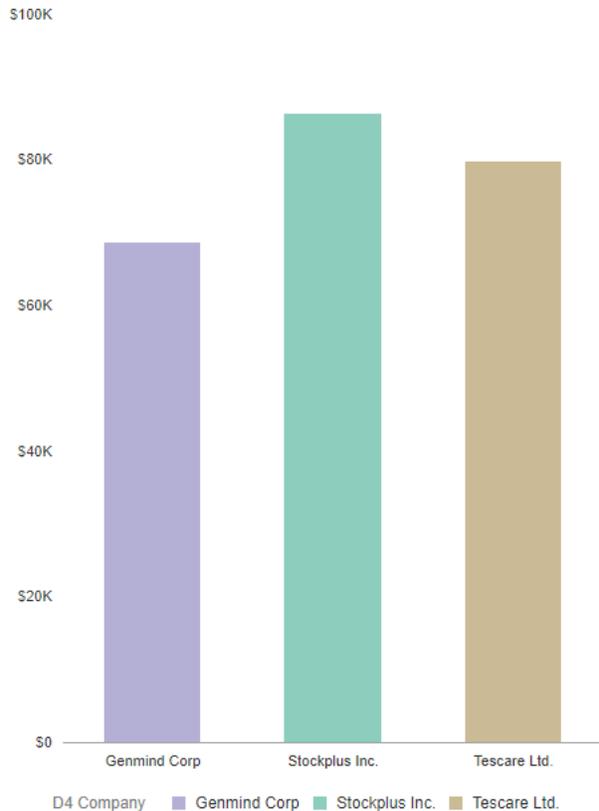


Summary KPI Layout

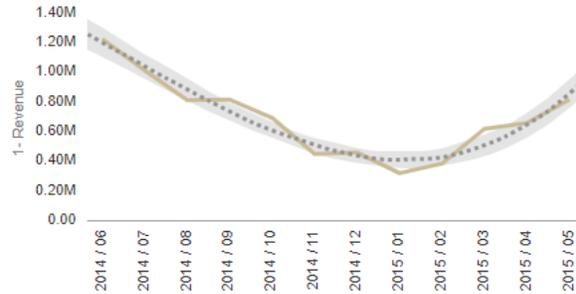
Selected Company

Tescare Ltd.

EBITDA by Company for May 2015



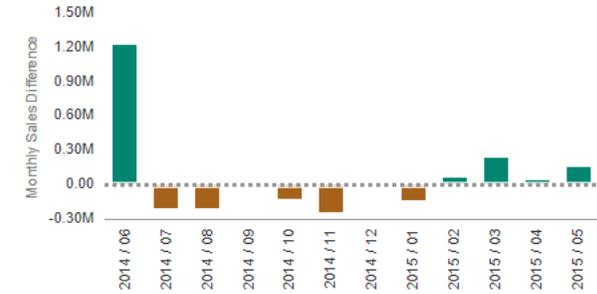
Revenue Rolling 12 Months



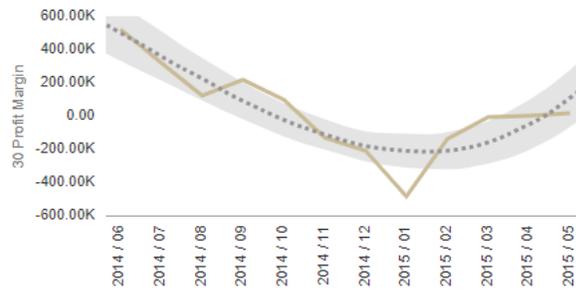
Revenue - May 2015

\$814,948

Revenue Difference Rolling 12 Months



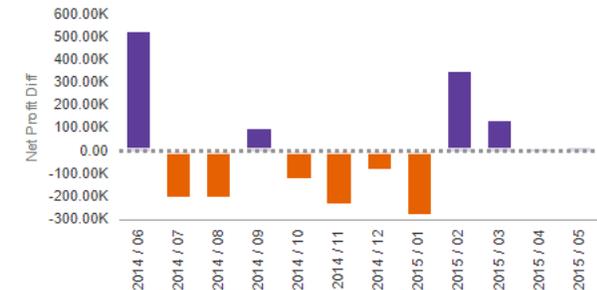
Net Profit Rolling 12 Months



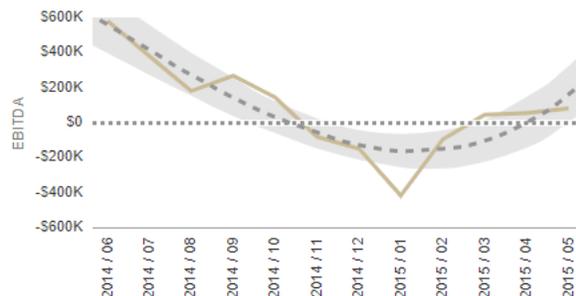
Net Profit May 2015

\$16,951

Net Profit Difference Rolling 12 Months



EBITDA Rolling 12 Months

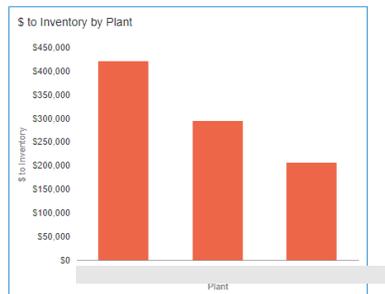


EBITDA - May 2015

\$80,002

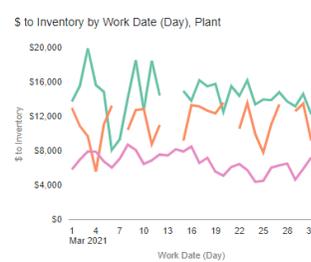
EBITDA Difference Rolling 12 Months



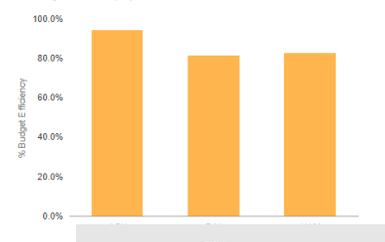


\$ to Inventory

\$926,383



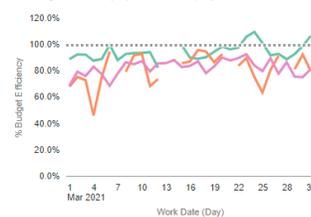
% Budget Efficiency by Plant



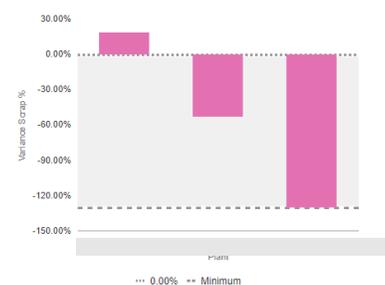
% Budget Efficiency

84.3%

% Budget Efficiency by Work Date (Day), Plant



Variance Scrap % by Plant



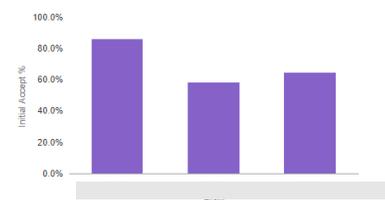
Variance Scrap %

-69.5%

Variance Scrap % by Work Date (Day), Plant



Initial Accept % by Plant



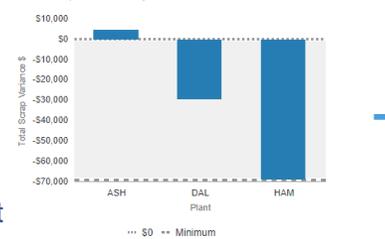
Initial Accept %

66.9%

Initial Accept % by Work Date (Day), Plant



Total Scrap Variance \$ by Plant



Total Scrap Variance \$

-\$93,628

Total Scrap Variance \$ by Work Date (Day), Plant



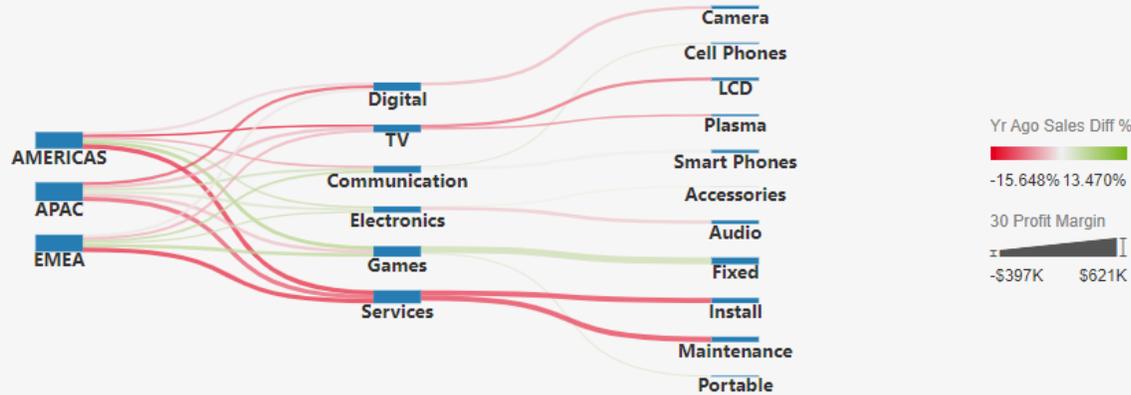
Summary KPI Layout

- Organize KPIs in a prominent column
- Place views in alignment with KPI measures
- Meant to read across
- Place “selector” visualization on far left

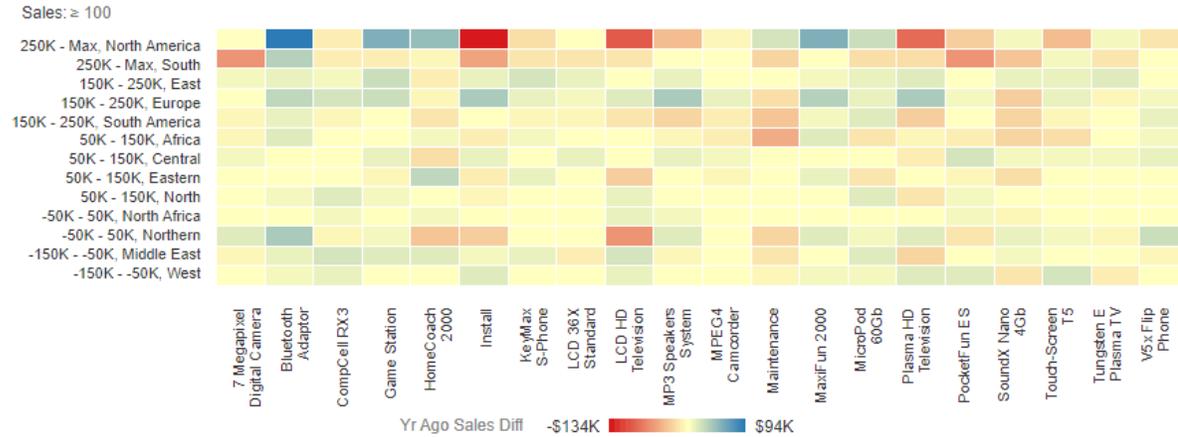
Multi-Dimension Layout

T05 Per Name Year 2014 C2 Customer Status All C3 Customer Type (5) Gold, Non Member, Platinum, Referrals, Silver C5 Income Level All C4 Customer Segment All C20 Years as Customer All

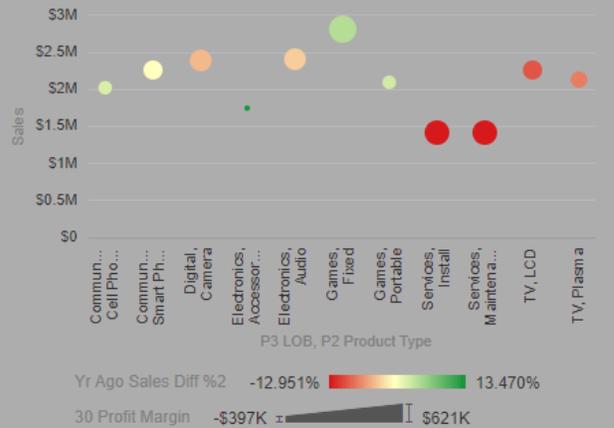
Sales Difference from a Year Ago by Region LOB Product Type



Sales Difference from Last Year for Geo Area and Products Sorted by Profitability



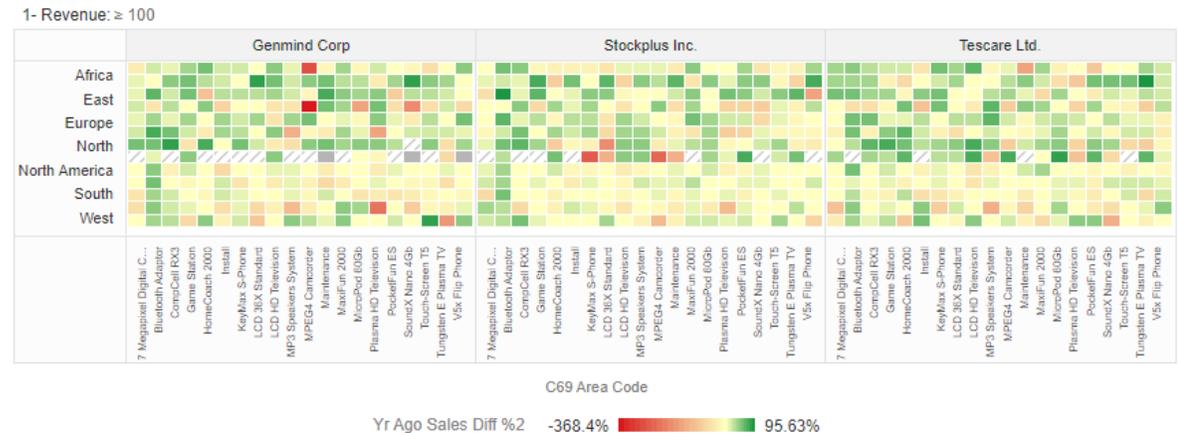
Sales for LOB and Product Type



Sales Diff % from a Year Ago for LOB and Product Type



Sales Difference % from Last Year for Area, Company and Products

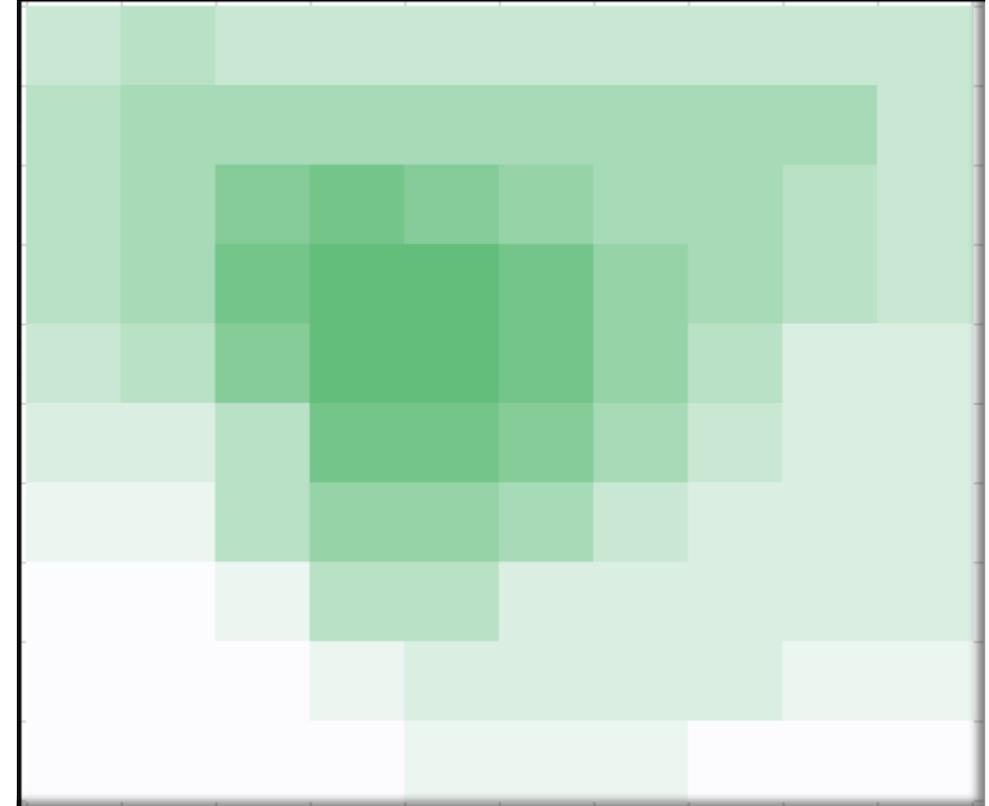


Key Questions to Ask About Dashboards

- Who is the audience?
- What is the subject or topic?
- What is the major insight?
 - What element has the most visual weight?
 - What element has the most prominent position?
- What context is provided and what comparisons are made?
- What are the decision scope and the data scope?
- Are navigation options understood?
- Are prompts and selectors understood?
- Are dashboard pages ordered logically?

Dashboard Layout Strategies

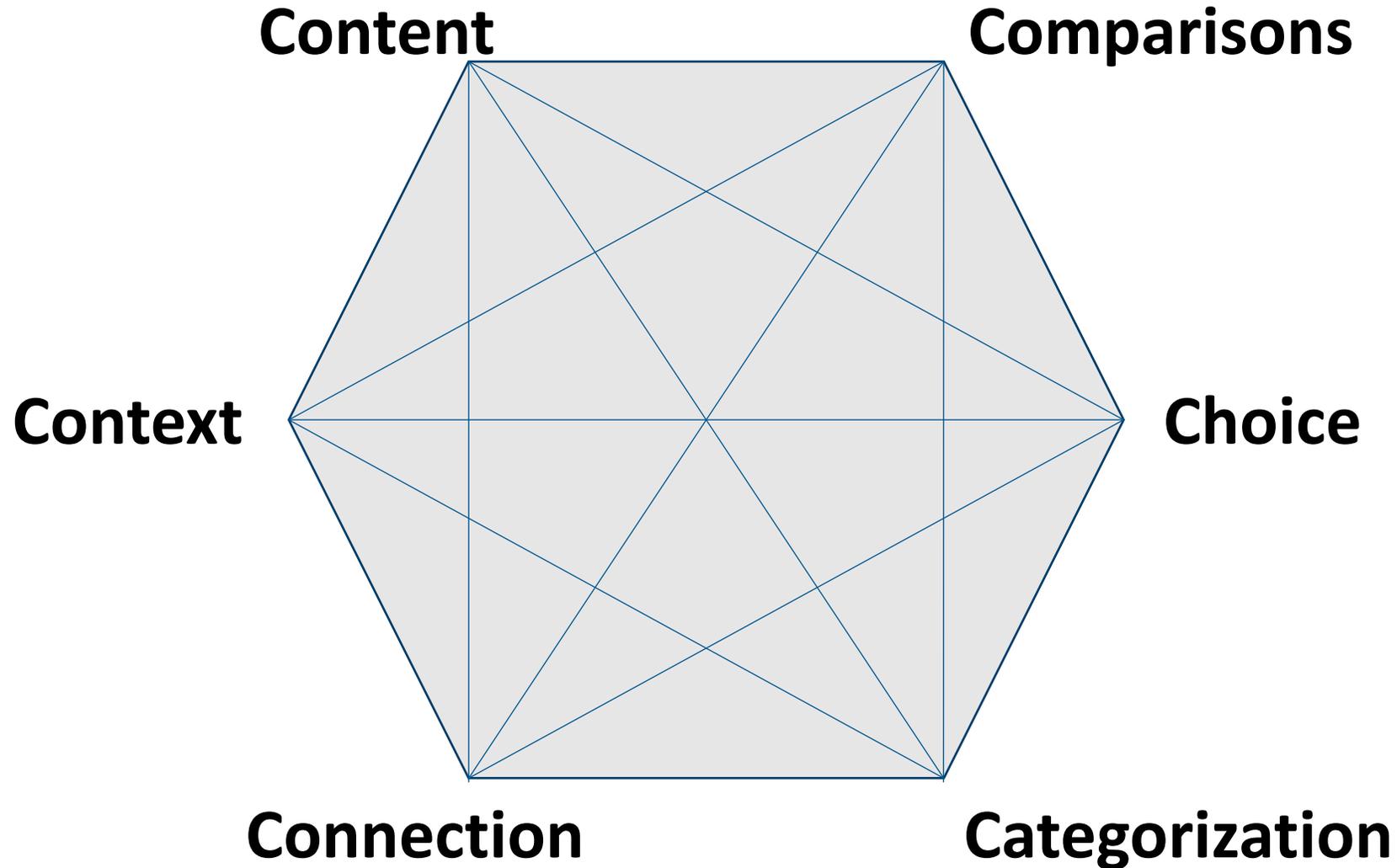
- Center and Top are most visible.
- Place prompts and navigation links in a consistent location.
 - Leftmost column is best for prompts
 - Navigation can be less prominent
- More important info gets more visual weight and better location.
- Place detailed look up information at the bottom of dashboards.
- Alignment
- Proximity



BI Dashboards

- Role-based.
- Data selection and filtering are extremely important.
- Dashboards support evidenced-based decision making.
- Shared understanding of business situation is a key benefit.
- Content may be individualized.
- Design should be standardized.

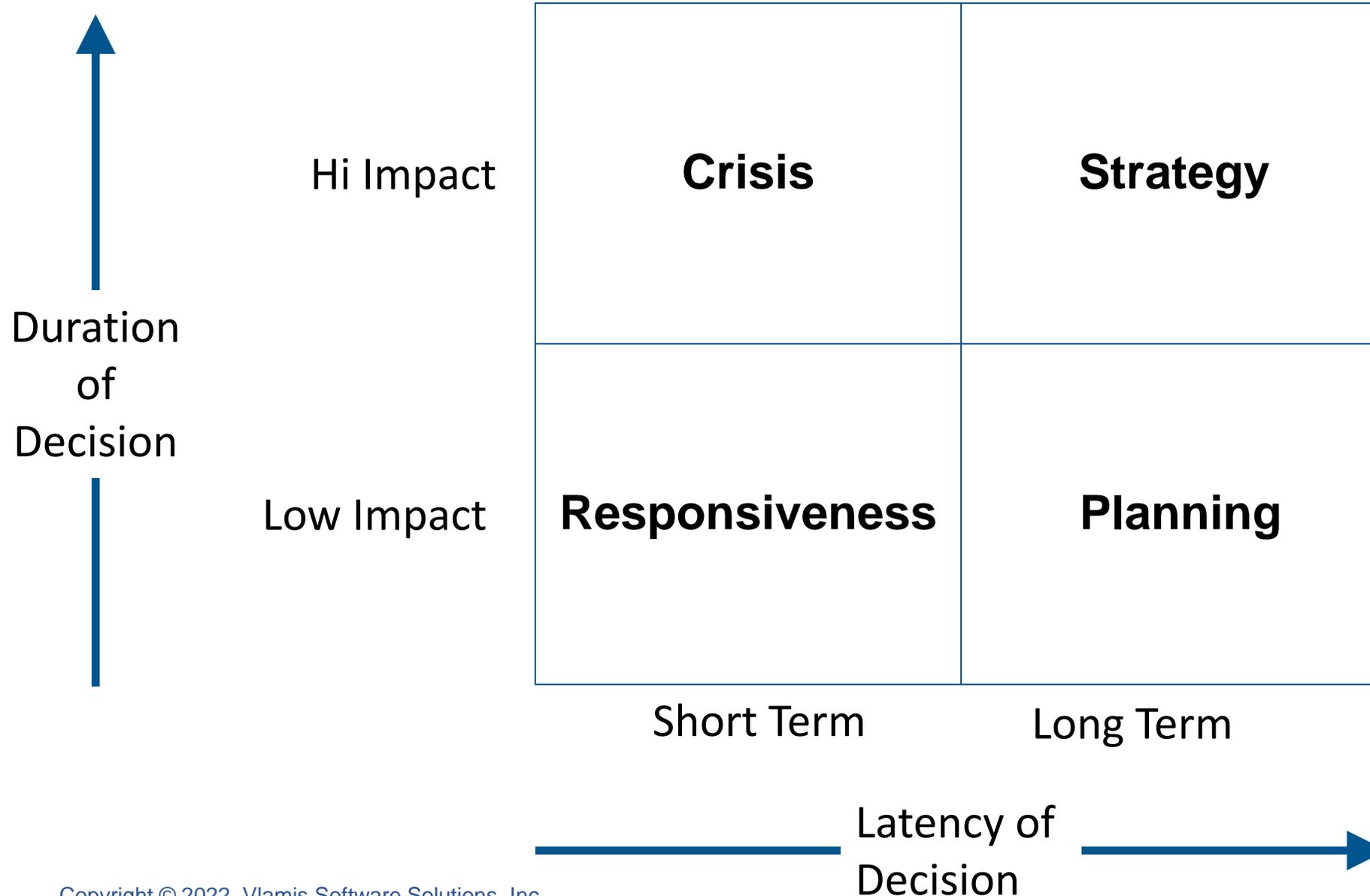
Great Dashboards Balance 6 Requirements



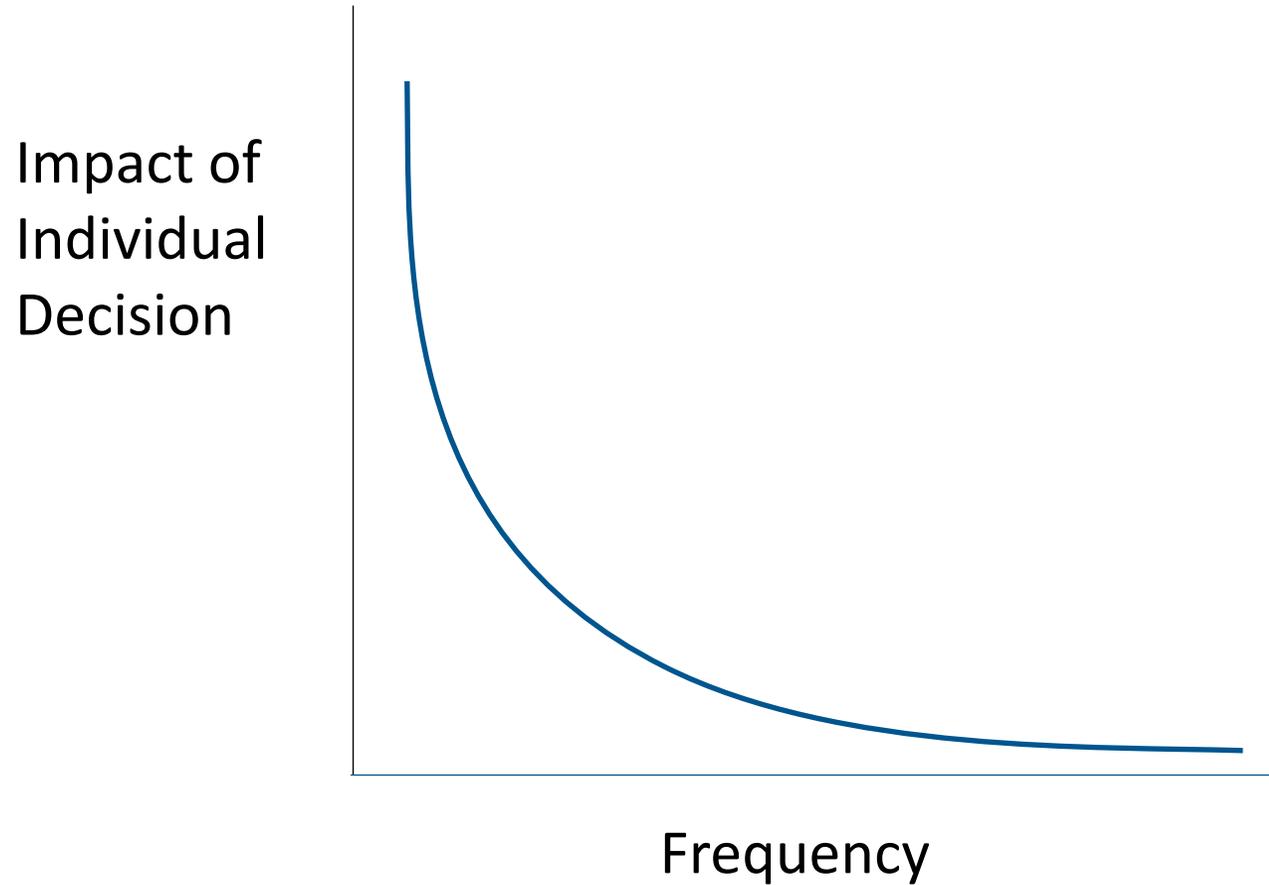
Great Dashboards Balance:

- Content (what's displayed)
- Comparisons (relative importance)
- Choice (user-invoked changes to the display like prompts)
- Connectivity (navigation, drill-to-detail, cross-drills, linked displays)
- Context (how displayed, i.e. design)
- Categorization (how organized and logically collected)

Four Realms of Decision Making

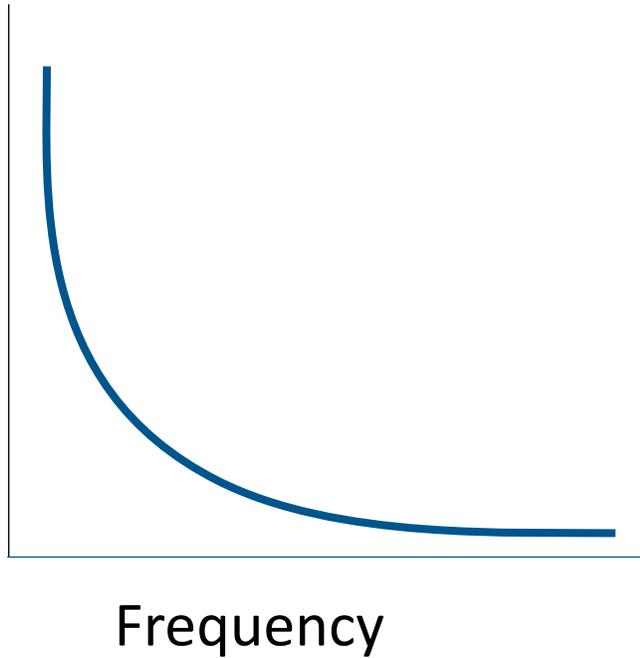


Organizational Decision Making

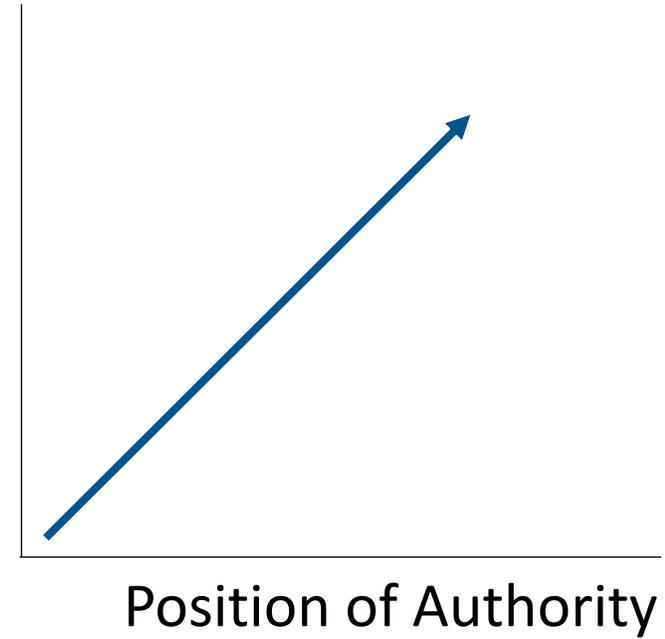


Organizational Decision Making

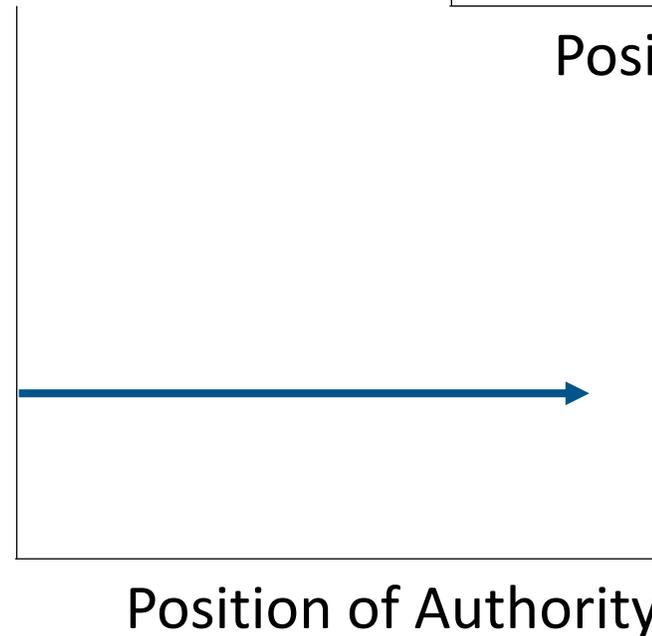
Impact of Individual Decision



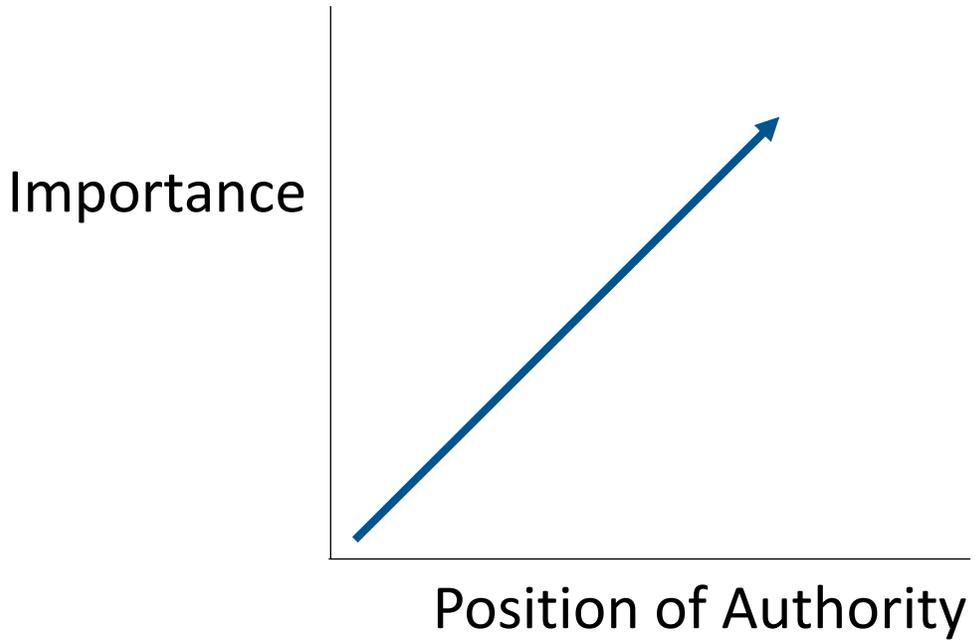
Impact of Individual Decision



of Decisions

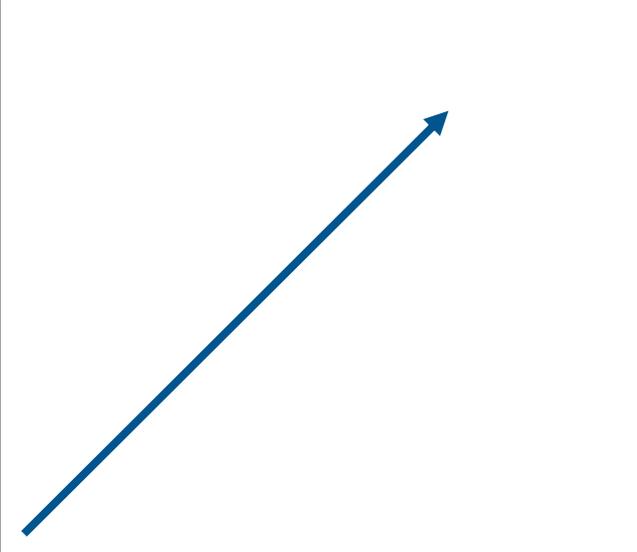


Dashboard Importance Score



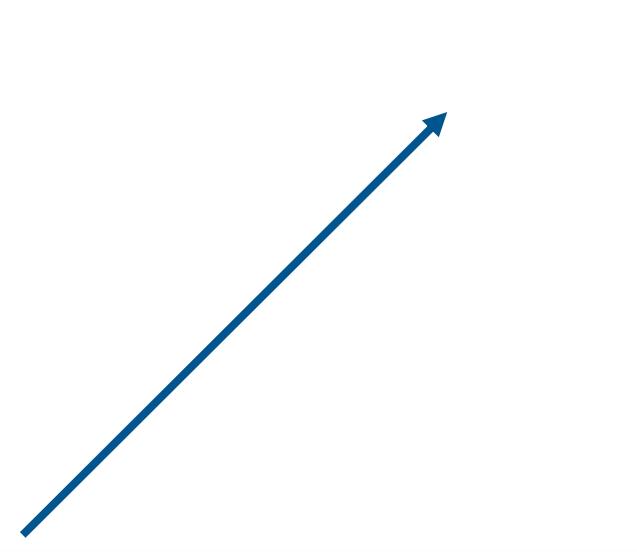
Dashboard Importance Score

Importance



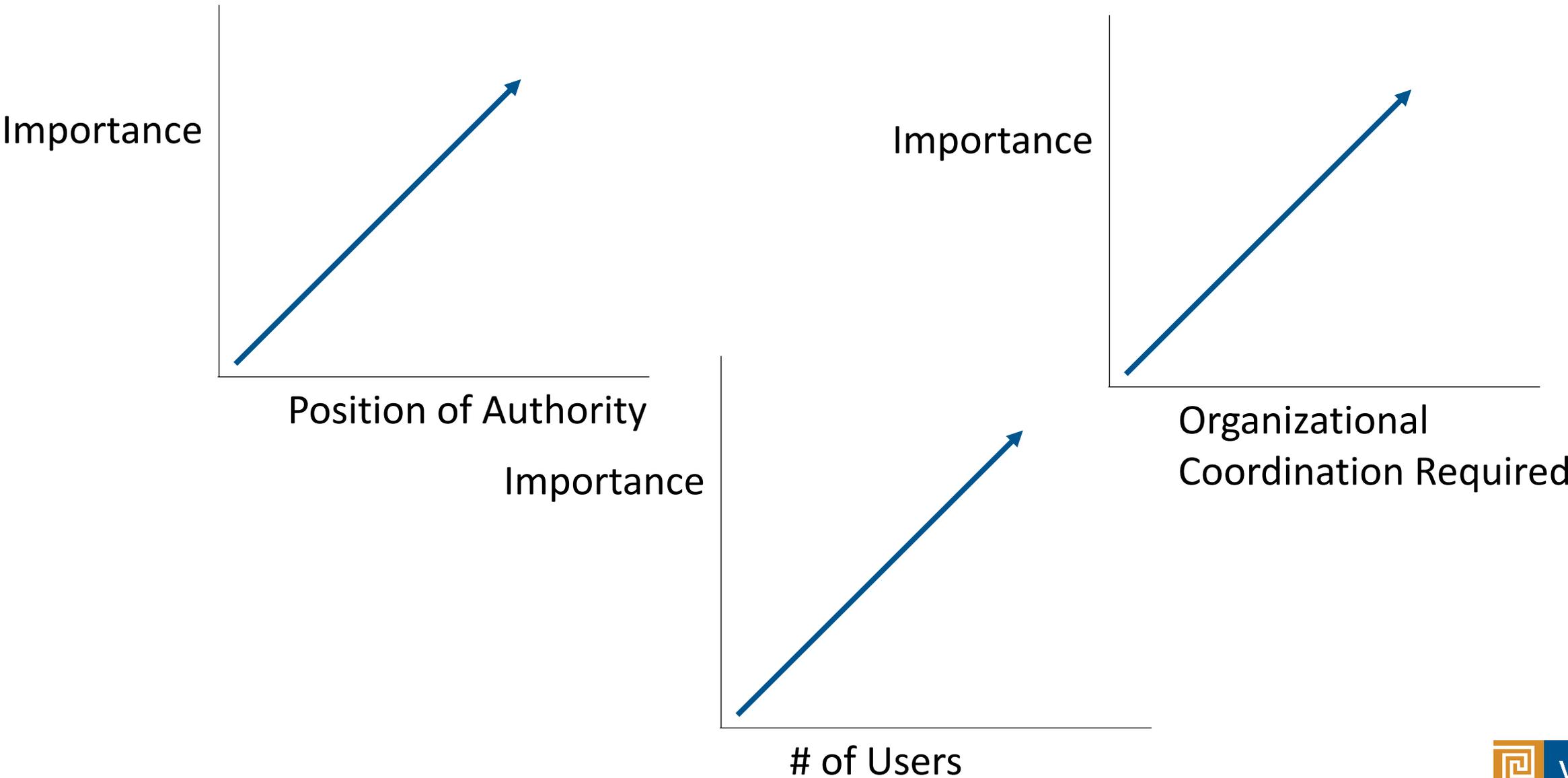
Position of Authority

Importance



Organizational
Coordination Req.

Dashboard Importance Score



Dashboard Importance Score

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

Example Dashboard Importance Rubric

Division Operational Expenses YTD Dashboard

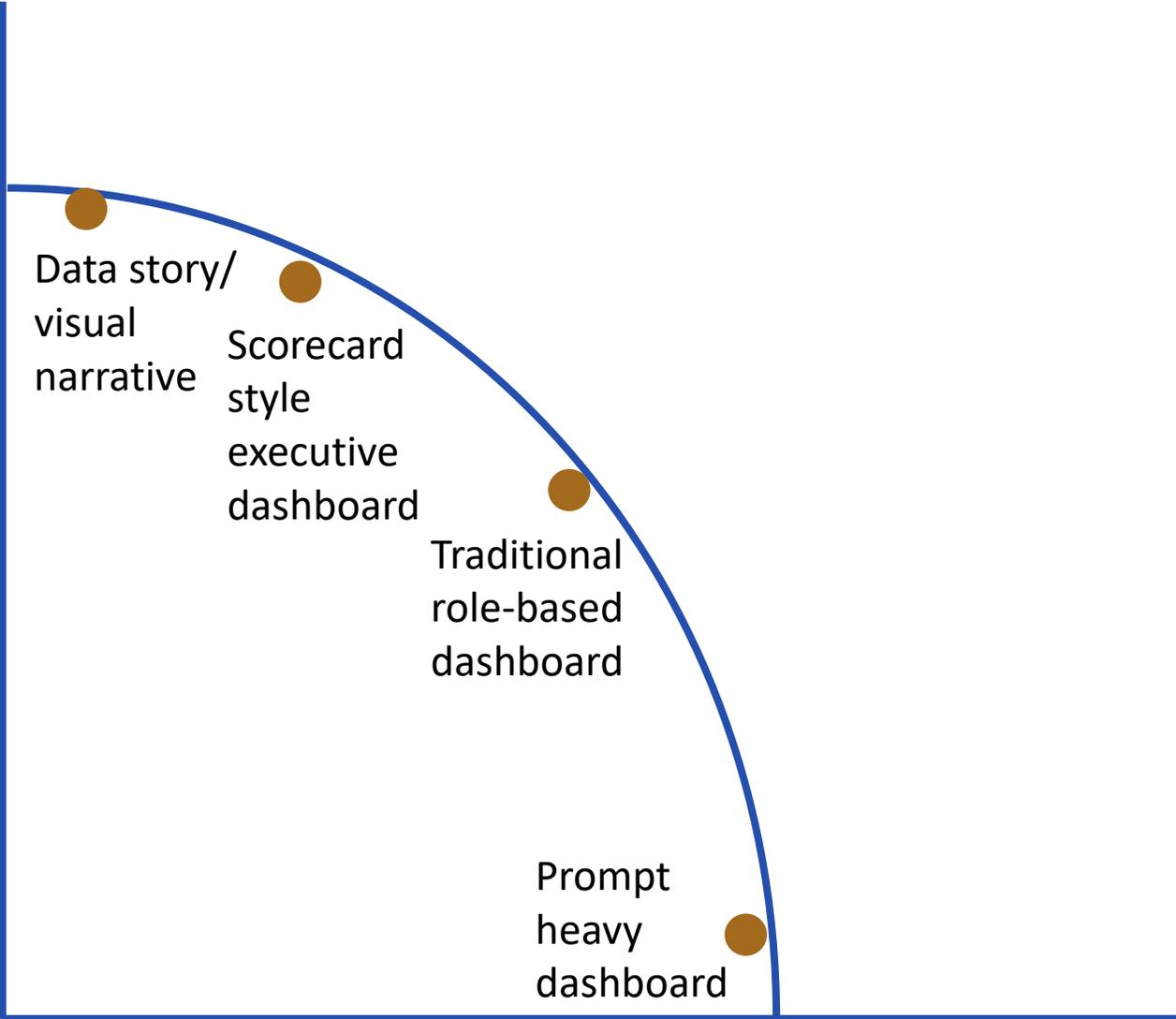
Role	Authority	Org Coord	Users	Product	Notes
Director	8	3	3.5	84	
Fin Analyst	2	3	4	24	
Managers	4	1	9	36	
Dashboard Importance				144	

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	

Balance Choices with Shared Views

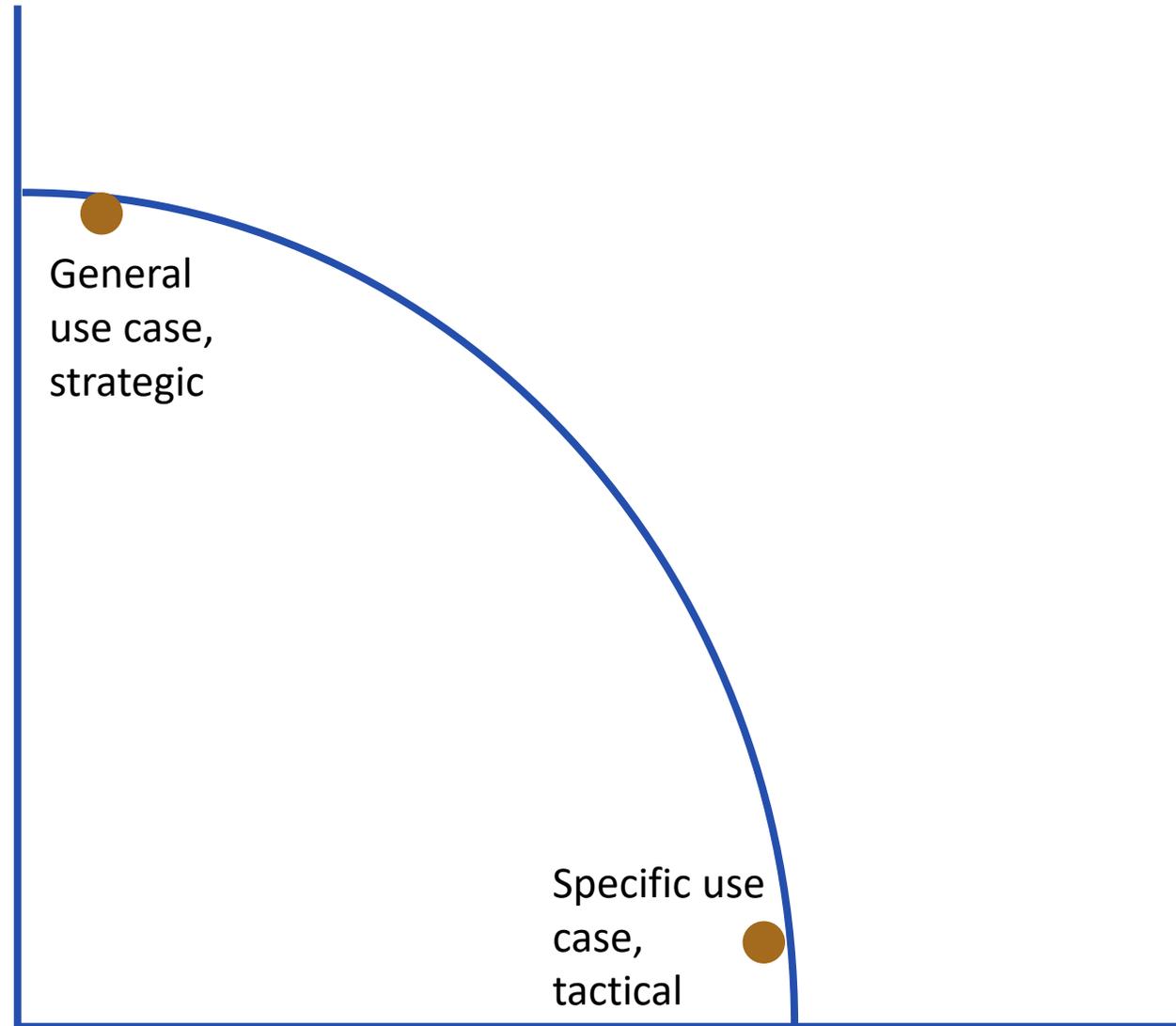
**Shared
Understanding**



Individual Choice

Strategic vs Tactical Dashboards

**Required
Organizational
Coordination**



Individual Decision/Action

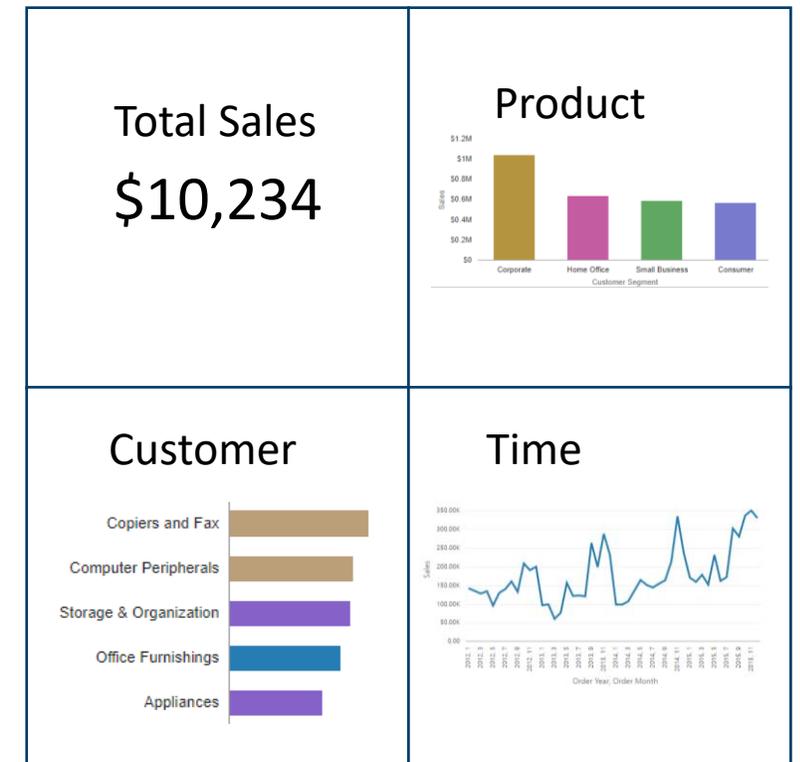
Keys to Data Discovery

- 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

4 Quadrant Layout - 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



2 Dimensions 1 Fact

- Summary bar graphs for single attributes at top of page
- Use stacked bar graphs with attributes switched below
- Grid heat map with fact as color below that
- Summary pivot table at bottom



2 Major Facts 1 Major Dimension

- Use scatter plot
 - Placement of data is ordered.
 - Use multiple scatter plot graphs to reveal insights when working with more than two facts.
 - Only add grid lines with meaning (quadrants)
- Consider using box plot for large data sets with large number of data points and several attributes.

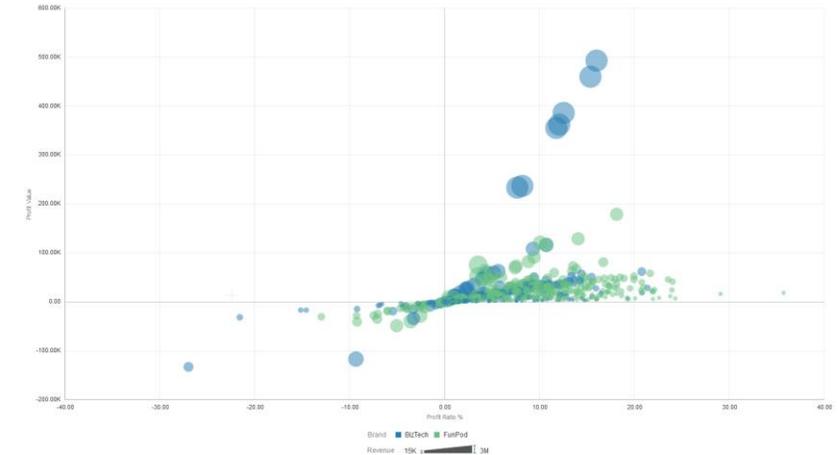
Summary 4 Quadrant Layout

- Evaluating a fact or a dimension?
 - Sales analysis
 - Customer or product analysis
- Identify your main topic of interest with a performance tile
- Summary view of three key dimensions for fact topics
- Summary view of dimension measures and key comparison for dimensions
 - Versus last period/forecast/budget
 - Top N (magnitude, change, %, new)
 - RFM (recency, frequency, magnitude)



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.



Dimensional Analysis

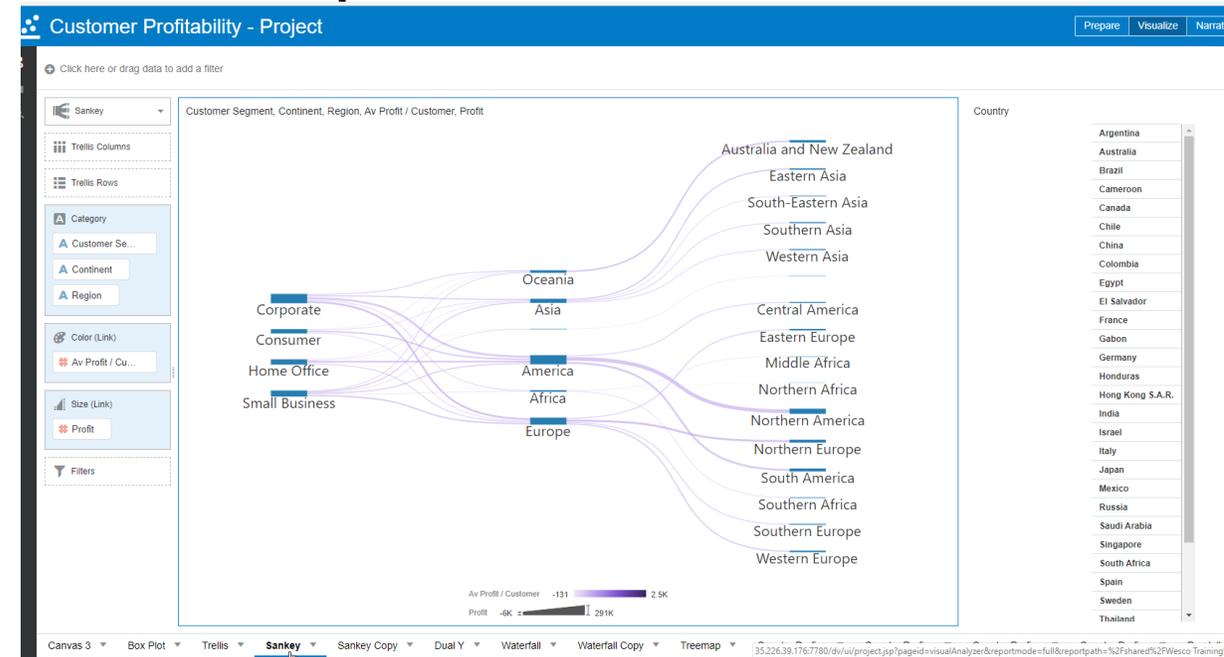
- Use brushing and selection with multiple graph layouts.
 - Build four or five graphs with related attributes or measures.
 - Too many graphs or several highly dense graphs exceed limitations
- Consider alternative graph types
 - Scatter plots
 - Trellis charts
 - Sankey graphs
 - Parallel coordinates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

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

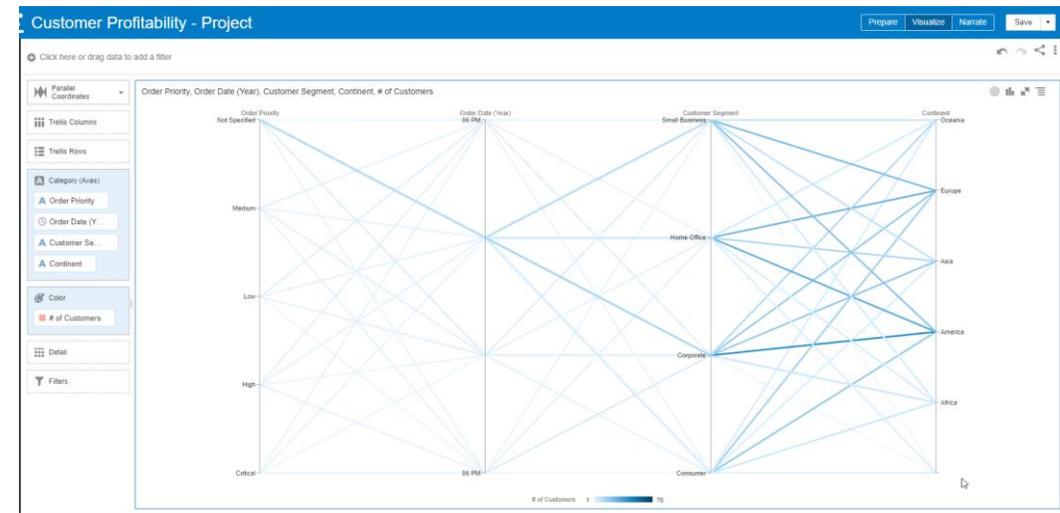
Sankey Graphs

- 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



Dimensional Columns

High number of factors
/
cardinality

Low number of factors
/
cardinality

Lowest Grain	Trend/cycle Correlation Outlier
Trellis	Comparative Correlation

Flat

Shaped

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

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

Different Strategies

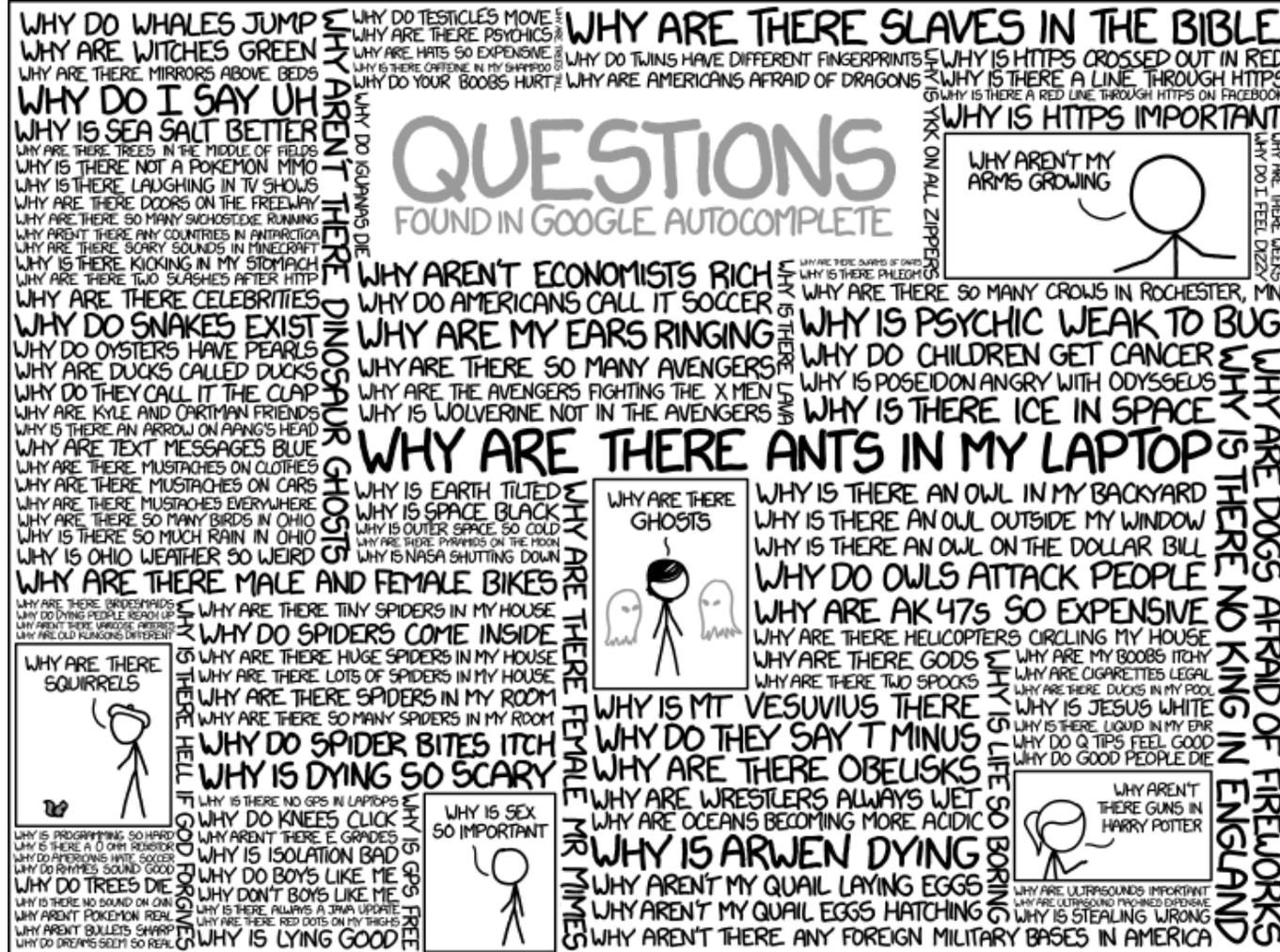
- If it's worth doing, it's worth doing right.



- The perfect is the enemy of the good.



Questions?



Thank You!!

- Tim Vlamis
- Dan Vlamis

tvlamis@vlamis.com

dvlamis@vlamis.com

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