Data Visualization Workshop

Mitch Campbell
Oracle
Analytic Journey with Data Visualization

- Foundation and Current Trends
- Telling a Story with Data
- Analytic Design and Best Practices
- Selecting the Right Chart type
Data Visualization helps people understand data using visual analysis.

Patterns, trends, and correlations are much easier understood visually.

People – anyone can use it.
Data – any data.
Visual – rich, easy to use visual experience!
Data Visualization Demo

Mitch Campbell, Oracle
@Smitchc26
Analytic Journey – Data Visualization

- Foundation and Current Trends
- Telling a Story with Data
- Analytic Design and Best Practices
- Selecting the Right Chart type
The World Has Gone Mobile

We look at our phones more than we look at our loved ones.

150
the average number of times a person look as their mobile each day.
My Phone is Alive!

**IN YOUR MIND:**
- Fetch my emails!
- Navigate to John's house!
- Show me the news!
- Send this photo to Lynn!

**IN REALITY:**
- Charge me!
- Gimme some wifi! Now!
- New email! Read!
- Answer this call!
- A restaurant! Check in!

Elite Readers
JUNE 22 AT 18:04
Strong Tailwind Of Technology Trends

MOBILE

CLOUD

BIG DATA
“Unless data is communicated across an organization, it becomes worthless.”
5 Steps To Building A Data-Driven Culture

1. **Data-oriented mindsets and infrastructure support metrics.** Data-driven companies define processes that support key performance indicators fundamental to their businesses, and they communicate these metrics to staff. Seventy percent of these organizations have a senior executive who champions analytics, Aberdeen reports.

2. **Data is centralized and organized.** To ensure data is recent and relevant, data-driven cultures gather and organize information from across the organization. Centralizing data allows for constant updates, keeping data fresh and up-to-date throughout the business.

3. **Policies govern data access.** Within data-driven organizations, formal policies control user access. They establish a traceable chain of custody to mitigate risk, specify levels of access, and determine how and by whom information may be altered.

4. **Data access is layered.** Data-driven organizations have a sophisticated, layered approach to managing key indicators in which each metric is clearly linked to organizational function and desired outcome.

5. **Analytics are integrated into tools.** Analytic tools used at data-driven organizations tend to be among the most innovative tools available. They’re also typically embedded into existing tools, making them intuitive and more likely to be used.
Business Analytics: Major Disruption Under Way

• Broader adoption of Analytics by business users
• Transfer of Buying Power from IT to line of business
• Data Gravity shifting to the Cloud
• Growth in Embedded Analytics
• Innovation powered by Big Data
Business user demands are changing

Speed over consistency

Visual data discovery

Embedded analytics
Google Search “Data Visualization” and Infographics

https://www.google.fr/search?q=Data+visualization&es_sm=93&tbm=isch&tbo=u&source=univ&sa=X&ei=kfTSVOfvlSHnaqPhgpAO&ved=0CD4QsAQ&biw=1600&bih=799

http://www.datavizcatalogue.com/

http://d3js.org

http://flowingdata.com/2015/08/11/real-chart-rules-to-follow/

http://i.imgur.com/BxX1Gex.jpg
Analytic Journey – Data Visualization

- Foundation and Current Trends
- **Telling a Story with Data**
- Analytic Design and Best Practices
- Selecting the Right Chart type
telling a story or showing data
# Q3 Highlights

<table>
<thead>
<tr>
<th>Source</th>
<th>Revenue</th>
<th>Gross Profit</th>
<th>Net Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>$20,348M</td>
<td>$15,445M</td>
<td>$2,569M</td>
<td></td>
</tr>
</tbody>
</table>

### Pharmaceuticals

- **30%** Revenue growth for Q3 2012 from the same quarter in 2011. $15M sale to Pharma Source Distributors contributed 5% to this growth.

### Cosmetics

- **9%** Revenue growth for Q3 2012. Revenue was $891M, an increase of 9% compared with $819M in the same quarter a year ago.

### New Drugs Approved

- **$109M** Revenue for Zylar.

### Featured Product

- **Advantax**
  - Firming and anti-aging cream.
  - Approved: July 2012
  - Introduced: September 2009
  - Today, it is one of the top.

### Revenue Trend

- **Series**: Cosmetics Group: 2011-Q1 Value: 744,274

### Revenue vs. Budget
don’t let your audience interpret data
right data

• valid
• reliable
• complete
• relevant

visualizations

• type
• color, graphics
• elements
• placement
Number of Dimensions Drives Complexity

Poorly designed charts obfuscate the important information.

Designing Data Visualizations, Iliinskey & Steele, O’Reilly Media, 2011.
Infographic

Manually Drawn, Specific to Data at Hand, Aesthetically Rich

Designing Data Visualizations, Illinskey & Steele, O’Reilly Media, 2011.
US Smartphone Marketshare 2008

- 39.0% RIM
- 21.2% Other
- 19.5% Apple
- 9.8% Palm
- 7.4% Motorola
- 3.1% Nokia

Source: http://www.elartedepresentar.com/
2012 PRESIDENTIAL RUN

GOP CANDIDATES

70% BACK PALIN
63% BACK HUCKABEE
60% BACK ROMNEY

SOURCE: OPINIONS
DYNAMIC
turn it upside down

learn to distill

think infographic
turn it upside down

• ask (= conclusion)
  – 3 key messages
    • data / evidence for each
• conclusion

• dashboard ≠ storyboard

• don’t let audience interpret data
Learn to distill

- noise vs. support for messages
- avoid “infoxication”
- balance validity, reliability, completeness
Think Infographic

- clean & simple
- handout vs. presentation
- say it without speaking
Analytic Journey – Data Visualization

• Foundation and Current Trends
• Telling a Story with Data
• **Analytic Design and Best Practices**
• Selecting the Right Chart type
Two Categories of Data Visualization

1) Data Exploration
   – Lots of data, not sure what’s in it
   – Best done at high level of granularity
   – Find the story the data is telling

2) Data Explanation
   – Already know what data has to say & need to tell the story to someone else.
   – Design to highlight that story
   – Make editorial decisions as to what data to keep, what to omit
Two Categories of Data Visualization

1) Data Exploration (Data Discovery)
   – Lots of data, not sure what’s in it
   – Best done at high level of granularity
   – Find the story the data is telling

2) Data Explanation (Business Intelligence)
   – Already know what data has to say & need to tell the story to someone else.
   – Design to highlight that story
   – Make editorial decisions as to what data to keep, what to omit
Data Visualizations are the Great innovation of our Time

- Quantitative communication through graphical representation of data and analytical concepts is essential to survive amid the deluge of data flowing throughout our world.

Telling Stories with Data
More Than Numbers!

- Data is a representation of real-live; more than a bucket of numbers
  - Tell Story:
    - Convince you of something.
    - Compel you to action.
    - Enlighten you with new info.
    - Force you to question your preconceived notions of reality.
  - See what the data has to say.

Nathan Yau, Visualize This: The FlowingData Guide to Design, Visualization, and Statistics
Seven Questions to Know Your Audience

- What are they like?
- Why are they here?
- What keeps them up at night?
- How can I solve their problem?
- What do you want them to do?
- How might they resist?
- How can I best reach them?

Nancy Duarte, Resonate: Present Visual Stories that Transform Audiences
SUCCESS Rules – Hichert Partners

Prof. Rolf. Hichert

- Developed a concept for Management Information Design in companies on the basis of E. Tufte’s principles.
- SUCCESS:
  - Say
  - Unify
  - Condense
  - Check
  - Enable
  - Simplify
  - Structure
Technology Enables Non-Technical Users
Deliver More Advanced Insights About Quantitative Information

• Internet-based Standards for Rich Analytic Applications
• Graphical User Interfaces
• New Charting Engines
• New Visualization Available Graphics Libraries
Increasing Expectations for Visual Excellence

• **Graphical interaction** with data is fast becoming expected norm for many business users (Executives to front-line personnel).

• **Visualization** is a key concern for BI, Analytics & Discovery Professionals.

• **Good data visualization** is critical to making smarter decisions and improving productivity.

• **Poorly designed visualizations** can mislead business users and more difficult to overcome the daily digital onslaught

• **Users lose confidence** in their BI Systems if they don’t understand the data in the system.

Traditional Dashboards Are Not Data Visualization

- TDWI Research finds that only 7% of the 453 respondents to the survey for this report are “very satisfied” with their ability to view and interact with data visually so that they can communicate information effectively and make decisions based on information through graphical means.
- Almost a quarter (23%) said they are “not satisfied,” and about one-third each are somewhat satisfied or somewhat dissatisfied.

Most Popular Visualization Types

TDWI Recommendation:

• Organizations should consider technology architecture that allows their users to expand the variety of visualization types they use, rather than restrict selection.

• Because technology progress is enabling ever broader and deeper data analysis, we expect no slowdown in innovation with data visualization.

Analytic Journey – Data Visualization

• Foundation and Current Trends
• Telling a Story with Data
• Analytic Design and Best Practices
• Selecting the Right Chart type
Chart Type Selection Framework

What would you like to Show?

- Comparison
- Distribution
- Composition
- Relationship

Adapted from ExtremePresentation.com
Showing Relationship

Analysis

Variables

Chart

Relationship

2

Scatter Chart

3

Bubble Chart
Showing Distribution

Analysis

Variables

Chart

Distribution

1

Column Histogram

Line Histogram

Box Plot

2

Scatter

3

3D Area
Showing Composition

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Time Period</th>
<th>Composition</th>
<th>Chart</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Simple Share of Total</td>
<td>Pie Chart</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accumulation or Subtraction to total</td>
<td>Waterfall Chart</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Components of Components</td>
<td>100% Stacked Column or Treemap</td>
</tr>
<tr>
<td></td>
<td>Few periods</td>
<td>Only Relative Difference Matter</td>
<td>100% Stacked Column Chart</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Relative &amp; Absolute Difference Matter</td>
<td>Stacked Column Chart</td>
</tr>
<tr>
<td></td>
<td>Many periods</td>
<td>Only Relative Difference Matter</td>
<td>Stacked Column Chart</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Relative &amp; Absolute Difference Matter</td>
<td>Stacked Area Chart</td>
</tr>
</tbody>
</table>

Composition

- Few periods: Only Relative Difference Matter
- Many periods: Only Relative Difference Matter
- Static: Simple Share of Total, Accumulation or Subtraction to total, Components of Components

Chart

- Pie Chart
- Waterfall Chart
- 100% Stacked Column Chart
- Stacked Column Chart
- Stacked Area Chart
Showing Comparison

- Analysis
- Time Period
- Comparison
- Data
- Chart

Comparison Over Time
- Few Periods
  - Single or Few Categories
  - Many Categories
  - Single Non-cyclical Data
  - Single Cyclical Data

Comparison Among Items
- Many Categories
  - Single Item
  - Many Items
  - Few Items

Comparison Few Categories
- Few Items
  - Column or Bullet Chart

Comparison Many Categories
- Single Item
  - Trellis Chart

Comparison Single Item
- Many Periods
  - Line Chart

Comparison Many Periods
- Single Non-cyclical Data
  - Line Chart

Comparison Single Non-cyclical Data
- Single Cyclical Data
  - Circular Area Chart

Comparison Single Cyclical Data
- Many Categories
  - Line Chart

Comparison Many Categories
- Column Chart
# Chart Type: Treemap

**What is it? When to use it?**

- To display up to 2 quantitative variables at different level of a hierarchy
- Display three dimensions simultaneously:
  - Size
  - Color
  - Relationships using grouping or containment

<table>
<thead>
<tr>
<th>Advantage</th>
<th>Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Well suited to find patterns, clusters, gaps, or outliers in large data sets.</td>
<td></td>
</tr>
<tr>
<td>- Particularly useful when the data can be grouped in hierarchies.</td>
<td></td>
</tr>
<tr>
<td>- Able to point out important values in a matter of seconds</td>
<td></td>
</tr>
<tr>
<td>- Data sets with millions of small equally sized nodes</td>
<td></td>
</tr>
<tr>
<td>- Unable to support time series data.</td>
<td></td>
</tr>
<tr>
<td>- Data sets with negative values.</td>
<td></td>
</tr>
</tbody>
</table>
Chart Type: Trellis

What is it? When to use it?

- A Trellis Chart is a layout of smaller charts in a grid with consistent scales. A “Chart of Charts”
- It is a basic chart type repeated many times.

Advantage

- Having many small charts enables you to view complex multi-dimensional data in a flat 2D layout; avoiding the need for confusing 3D charts.
- The grid layout combined with consistent scales makes data comparison simple.

Limitation

- Could support only 1 variable.
Chart Type: Waterfall

What is it? When to use it?

• A waterfall graph lets you visualize how a value increases or decreases sequentially and cumulatively.
• The columns are color-coded for distinguishing between positive and negative values.

Advantage

• Focus the user’s attention on how each measure contributes to the overall total.
• Communicate through simple formatting by using color.
• Great for visually showing the contribution of parts to a whole

Limitation

• Could not support time series data.
Chart Type: 100% Stacked Bar

What is it? When to use it?

• This chart type displays multiple sets (i.e. series) of data as stacked columns, and the cumulative proportion of each stacked element always totals 100%.

Advantage

• Useful when you have three or more data series and want to compare distributions within categories, and at the same time display the differences between categories.

Limitation

• Could not support time series data.
Chart Type: Bullet

What is it? When to use it?

• Bullet graphs feature a single quantitative measure along with complementary measures to enrich the meaning of the featured measure.
• Relate the featured measure to defined quantitative ranges that declare its qualitative state.
• A variant of the bar graph developed by Stephen Few.

Advantage

• Need only a small footprint,
• Supports more efficient reading than radial meters.

Limitation

• Does not support multiple time periods.
Chart Type: Box Plot

What is it? When to use it?

• Box plot graphically depict groups of numerical data through their quartiles.
• Useful to show the distribution of a dataset
• Also known as the box-and-whisker plot.

Advantage

• Handle large dataset easily.
• Provide some indication of the data’s symmetry & skewness.

Limitation

• Limited usage and comprehension outside statistical community.
• Exact value not retained.
Chart Type: Map

What is it? When to use it?

• Map geographies and business information together to see results.

Advantage

• Maps bring data to life.
• Make it easier for business users to identify the significant trends and issues contained in most reports and dashboards.

Limitation

• Unable to show time series data
Bird’s Eye: Geographic View
Customers, RMs, Branches
Understand Demographics of Customers

Grouped Charts

In this multi-series bar chart, we group bars together rather than stack them. A grouped chart allows accurate comparison of individual values thanks to an aligned baseline: a position, rather than length, judgment is used.

An ordinal scale positions the groups vertically; the bars are then replicated inside a panel, a technique that is also used for small multiples.
Understand Customer Relationship Over Time

Area Charts

This simple area chart is constructed using an area mark, with an added line for emphasis on the top edge. Next, rules and labels are added for reference values.

Although this example is basic, it provides a good starting point for adding more complex features. For instance, mouseover interaction can be added to allow precise reading of data values. Or multiple series of data can be added to produce a stacked area chart.
**Product Cross-Sell/Up-Sell Likelihood**

**Scatterplots**

Scatterplots can facilitate visual analysis along multiple dimensions, though care should be taken to avoid interference. In this example, we encode three dimensions: two are encoded using position, while the third is redundantly encoded as both area and color. Mouseover any of the dots to read the exact values.

A useful extension of scatterplots is the scatterplot matrix, as shown in the Anderson's Flowers example.
This example demonstrates loading of CSV data, which is then quantized into a diverging color scale. The values are visualized as colored cells per day. Days are arranged into columns by week, then grouped by month and years.
Analytic Journey – Data Visualization

• Foundation and Current Trends
• Telling a Story with Data
• Analytic Design and Best Practices
• Selecting the Right Chart type

• Empowering End Users…
See the Signals

In Sales

In Marketing

In Procurement

In Human Resources

In Finance

In Supply Chain
Visual Analytics for Business Leaders

- **Blend More Data**: Easily upload and analyze data yourself
- **Get More Answers**: Quickly uncover hidden patterns and previously unknown trends
- **Scale More Insights**: Craft and share a story supported by your data
Visualization Space
## Industry Focus...What’s Important

<table>
<thead>
<tr>
<th>Healthcare</th>
<th>Retail</th>
<th>Higher Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tracking revenue cycles</td>
<td>Driving a highly-personal, real-time shopping experience.</td>
<td>Successfully attracting and retaining students</td>
</tr>
<tr>
<td>managing the practice from anywhere</td>
<td>Customers are shopping, sharing and perusing at the same time. Understanding omni-channel retail sales model.</td>
<td>Improving rankings</td>
</tr>
<tr>
<td>Tracking patient care quality</td>
<td>campaign profitability, CSAT and retention scores</td>
<td>Managing endowments allocation and usage</td>
</tr>
<tr>
<td>Decreasing cost per outcome</td>
<td>Maximizing brand and campaign profitability</td>
<td>Scrutiny from government, media and alumni.</td>
</tr>
<tr>
<td>Increasing operational efficiencies</td>
<td>Driving customer acquisition and loyalty</td>
<td>Monitoring marketing and student acquisition</td>
</tr>
<tr>
<td>Monitoring the effectiveness of new treatments</td>
<td>Comparing rep and manager sales metrics</td>
<td>Identify and track at-risk students</td>
</tr>
<tr>
<td>Tracking Gross bookings, gross churn, CMRR</td>
<td>Tracking social media reach and engagement</td>
<td>Improve job and grad school placement</td>
</tr>
<tr>
<td>Measuring manufacturing and production efficiency</td>
<td>Monitoring bad debt</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

@LORASANALYTICS #LCBIGDATA16 2016 BUSINESS ANALYTICS SYMPOSIUM
## Industry Focus...What’s Important

### Public Sector
- Modernizing the workforce of diverse user populations
- Fulfilling complex integration requirements
- Streamlining service delivery
- Prioritizing maintenance efforts
- Increasing procurement efficiency and results
- Providing transparency
- Maintaining security & access

### Hotels & Restaurants
- Following customer from acquisition through retention
- Following customer trends, customer satisfaction staying ahead of the competition
- Tracking specialty resourcing availability
- Following mobile purchasing trends
- Managing customer satisfaction scores
- Comparing capacity tracking and planning
- Monitoring sales to costs by category

### Hospitality & Transportation
- Driving customer loyalty
- Managing Costs
- Viewing campaign performance at-a-glance
- Tracking web analytics
DVCS Trials are LIVE!
THANK YOU