KPMG

cutting through complexity

Data Visualization

You've got to see it to believe it

Fall 2015 Casualty Actuaries of the Northwest Meeting

Speakers:
Jenny Andrzejewski, ACAS, MAAA
Nate Loughin

Antitrust Notice

The Casualty Actuarial Society (CAS) is committed to adhering strictly to the letter and spirit of the antitrust laws. Seminars conducted under the auspices of the CAS are designed solely to provide a forum for the expression of various points of view on topics described in the programs or agendas for such meetings.

Under no circumstances shall CAS seminars be used as a means for competing companies or firms to reach any understanding—expressed or implied—that restricts competition or in any way impairs the ability of members to exercise independent business judgment regarding matters affecting competition.

It is the responsibility of all seminar participants to be aware of antitrust regulations, to prevent any written or verbal discussions that appear to violate these laws, and to adhere in every respect to the CAS antitrust compliance policy.

Agenda

Our objective today: A survey designed to give attendees a good understanding of the breadth of possibilities of data visualization, and of the many choices in approaching and creating high quality data visualization.

1. **What is data visualization?**
2. **Visualization software**
3. **Data**
4. **Interface/Using the Web**
5. **Examples**

1. What is data visualization?

Why?

- Aid users in speedy and deep understanding of complex data
- Answer common questions, but also give insights and answers to questions no one thought to ask
- Socialize information with broad communities of users with diverse backgrounds and skills

How?

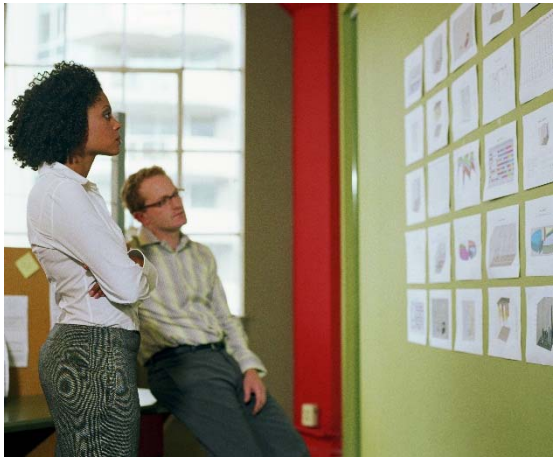
- Data visualization (DV) is a form of displaying significant amounts of data in visual formats so that it is easier to understand, gain insights, and observe trends.
- DV software gives the user flexibility to slice and dice the data by its various segments (line of business, company, channel, region/state, loss vs. expense) in order to analyze its various dimensions (premium, losses, exposures) and key performance indicators (KPIs) (loss ratios, rates, calendar year development, expense ratios).

1. What is data visualization? (continued)

Without DV:

- Experiences with the pain of massive reports and digesting them

- Issues with creating and using visuals in Excel® or PowerPoint® (for example) that are static and not interactive



1. What is data visualization? (continued)

With DV:

- Enhance actuarial analyses
- Convey observations and results to non-actuaries using various dashboards
- Tailor the dashboards depending on the audience (e.g., claims, underwriting) and their level (e.g., manager, C-suite)
- Allow users to filter and drill to share insights between various segments within the company

1. What is data visualization? (continued)

Benefits:

- Incorporates various data source types—allows you to bring data together from different channels to get a more complete picture versus looking at information in silos
- Leads to insights and trends that are sometimes not apparent when analyzing the numbers
- Facilitates scalability—quickly refresh data so it is up to date or add another set of data chronologically

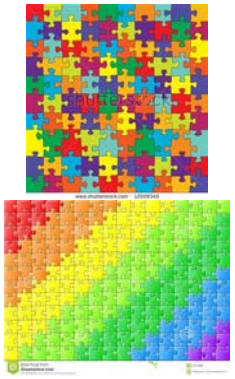
1. What is data visualization? (continued)

Building your visuals:

- Think through the questions you want answered in order to determine the dimensions and measures you need
- Keep flexibility to drill; each answer can raise new questions
- Consider the arrangement of the charts, navigation, and filters so that users can find the information they want in a clear manner

Designing your visuals:

- Consider the use of graphics
- Grab the user's attention
- Avoid too much information (cluttered)
- Contrast to draw attention; use conformity to avoid confusion



1. What is data visualization? (continued)

Some Chart Examples:

- Bar charts – vertical vs. horizontal; pareto (largest first); waterfall; butterfly
- Line charts – standard or stacked; also in radar format to show seasonality
- Pie chart – percentage of total
- Treemap (farm plot) - percentage of total as well as subtotals
- Box and whisker chart – depicts data by their quartiles
- Bubble charts – four dimensions (two axes, size and color of bubble)
- Scatter plots
- Speedometer – shows how a KPI falls within a range of acceptable standards
- Traffic light – shows when a KPI achieves a certain outcome
- Geographic maps – country, state, county
- Word clouds
- Opportunity for creative innovation, new types of visuals, visuals in motion
- Standard visuals can be quick and easy, but new types are being created as computing power increases

1. What is data visualization? (continued)

Chart Options:

- Tooltips – hover over a chart item (e.g., point, bar, bubble) to get useful information about the item
- Click on a chart item to drill down
- Add filters in order to view or exclude particular segments
- Trellis charts – grid of smaller charts
- Mini charts (sparklines) in tables

2. Visualization software

- Many software programs are available, e.g., IDEA®, QlikView, Tableau, D3, InfoSys, BOARD, Sisense, Pentaho, Birst, phocas, Halo, Spotfire. Many others.
- Considerations when selecting a visualization software: purpose, cost, support, functionality, data

2. Visualization software (continued)

- **Purpose**
  - *Users* – company-wide or one unit/division
  - *Objective* – only for the creation of visuals/dashboards or to complement the real purpose of the software (e.g., predictive analytics, reserving, modeling, claims)
  - *Flexibility* – quick and easy use of canned functionality or more creative innovative visuals that might take more time to create
- **Cost**
  - *Price* – licensing and sharing of software within the company; free software available but little support to help ensure software works as it should
  - *Other costs* – time and money spent on training and deployment
  - *Needed skillset* – consider if programming skills are required and what they will cost to acquire
- **Support**
  - *Training and support* – availability, quality, and cost of training and support; availability of an Internet community for collaborative development support
  - *Prior work* – availability of existing, completed dashboards adaptable to your needs, in a public or private repository

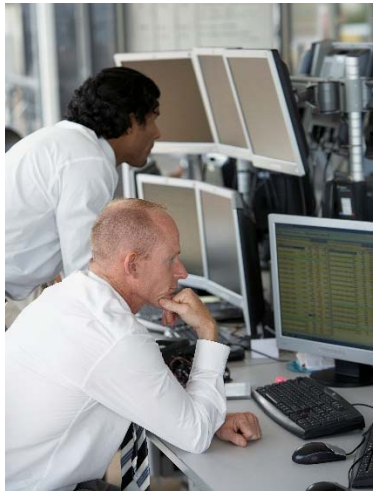
2. Visualization software (continued)

- **Functionality**
  - *Ease of use* – drag-and-drop technology and quality/quantity of pre-made visuals
  - *Web-ready format* – ease of uploading the dashboards online
  - *Bookmark/notes* – ability to bookmark particular selections and attach explanations or provide information about the data or observation on particular visuals to help guide the users
  - *Graphics* – specific visualization needs (e.g., geospatial mapping, specialty charts, etc.); look for software that provides the right native capabilities, avoid workarounds where possible
  - *Look and feel* – each tool provides a different set of style options; choose the one that matches your existing collateral and preference
- **Data**
  - *Capabilities of querying and exploring the data* – consider the volume of data that can be uploaded and the speed at which this occurs and visuals can change
  - *Data volume* – find a tool that performs well at your expected volume of data; do you expect megabytes, gigabytes, terabytes, etc.?
  - *Data format* – find a tool with known performance characteristics accessing your data structure and hardware configuration

3. Data

- *Inventory of data sources* – can be in various formats and from multiple sources:

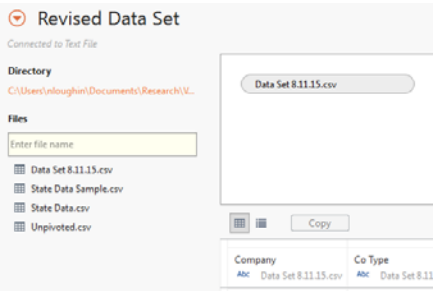
Data Source	Description	Examples
Traditional servers	Visualization software offers direct interfacing with ODBC and OLEDB servers	SQL Server, Oracle, Hadoop
Standard table files	Traditional table and data file formats are supported as well	Excel, CSV, text, Access
Custom sources	Non-traditional cloud-based connections are supported and can be accessed through add-in or direct connections supported by third-party vendors	Salesforce, Google™ Analytics, Twitter

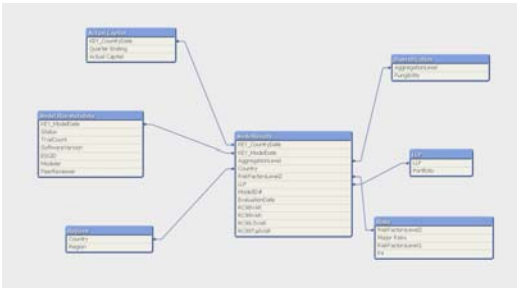


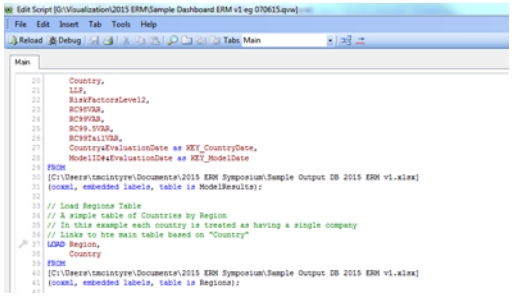
3. Data (continued)

- *Software interfaces* – common data upload features:

Data Source	Description	Examples
Upload templates	Most common data sources have pre-built upload interfaces specific to the data source	SQL Server, Oracle, Hadoop
Query/Code view	Solve visualization programs allow you to edit the code supporting the data upload directly, which can provide greater control over the data processing and access to data that is otherwise unsupported through the upload templates	Edit Script View (QlikView)
Pre-processing	Software provides varying degrees of pre-processing, such as automatic association of fields with the same names between various data sources, or manual override of data formats for specific fields	Data Interpreter (Tableau), Associative Engine (QlikView)







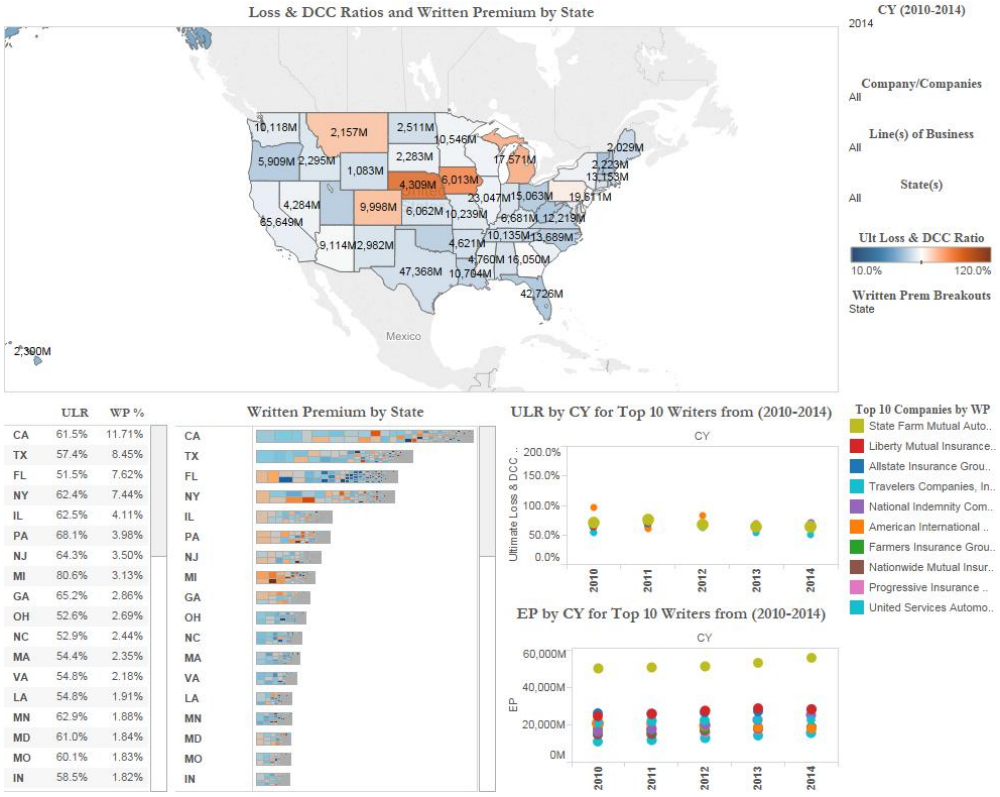
4. Interface/Using the Web

Visualization is for sharing and collaboration

- In order for the visualization dashboards to be effective on mobile devices, a larger screen will be needed. Check that your selected software provides this functionality and support.
- Confirm that the Web interface provides a similar breadth of functionality to that of the desktop software; note that performance may differ substantially between the two.
- Compare the look-and-feel of the Internet deployment with the desktop version.
- Assess the availability and reliability of Internet access.

5. Examples

Dashboard examples



Discussion/Questions

Presenters

KPMG LLP

www.kpmg.com

Jenny Andrzejewski
Director
314-244-4165
jandrzejewski@kpmg.com

Nate Loughin
Senior Associate
610-230-2068
nloughin@kpmg.com

cutting through complexity