MILLIMAN BRIEFING NOTE

Data Visualisation

An introduction to the uses and types of data visualisation

Gillian Tucker, FSAI
Donal McGinley, FSAI

"A picture paints a thousand words"

Data visualisation is the graphical representation of data. In this briefing note we will explore who may be interested in data visualisation, why it is so useful, and what it is, including examples of some of the many types of data visualisation techniques.

Who may be interested in Data Visualisation?

Data visualisation is useful for both data science practitioners and management/key decision makers of companies.

- It may be used by data science practitioners to help to understand the data they are working with, to build and validate models, and to communicate the outcome of data analysis undertaken to management.
- It may be used by management and key decision makers to aid their understanding of any analysis which has been carried out. This may enable key decisions to be made. These users would include management of sales, risk, and actuarial teams, amongst others.

Why is Data Visualisation useful?

Ease of communication

Data visualisation enables large amounts of information to be reduced to a graph, chart or picture, which can be much more easily consumed and understood by the intended audience.

Easy identification of issues

When data is displayed graphically, trends, patterns and outliers are often easier to identify.

Efficiency

The ease of identification of issues and ease of communication makes data visualisation an efficient tool. It is much more efficient to use charts or graphs to visualise large amounts of data than poring over databases.

Enriched analysis

Data visualisation tools can often combine data from multiple sources to provide an enriched analysis taking all relevant data into consideration. More advanced data visualisation can also produce interactive or animated charts, which add an additional dimension to the information being presented. Further information on these options is outlined in the following section on "Complexity in Data Visualisation".

Additional insights

Data visualisation can yield many additional and useful insights for companies. For example, it can:

- Improve understanding of customer behaviour, such as lapse or claim behaviour. This information can be used, for example, in customer sales and retention activities, and also in the actuarial assumption setting process and modelling.
- Identify areas in the business that need attention or improvement, for example it can identify weaker agents within a given sales channel.
- Detect fraud, by identifying out of pattern claims behaviour.

Complexity in Data Visualisation

Data visualisations can range in complexity. The simplest forms are simple static visualisations such as charts and graphs where the information displayed does not change.

Visualisations can also be developed to include interactive features so that users can examine the information in them in more detail. One particularly useful feature is the ability to hover over a particular data point. Users can hover their mouse over a data point and further information about that data point is displayed. Charts can also include drop-down menus and sliders so that certain information can be chosen to be displayed at any one time.

Visualisations can also be developed into animations. Animations are visualisations which move. This can be very effective especially for data showing change across multiple groups or time periods.

Types of Data

Data being used can come in many forms, for example, it can be:

- At a point in time or time series; time series data measures data over time rather than at one fixed point in time.
- Structured or unstructured; structured data is highly organised, while unstructured data has no pre-defined

format or organisation, making it much more difficult to process and analyse.

- Big data or small data; big data comprises large volumes of data, a variety of types of data and takes a long time to process, making it more difficult to manage compared to small data.
- Raw data or analysed data; the data being visualised can sometimes be raw data, or sometimes it can be data which has already been analysed and relationships with other data sources or between different variables within the same data source identified.

If working with raw data, data visualisation will primarily be used as a tool to aid analysis. If working with previously analysed data, data visualisation will primarily be used as a tool for communicating the outcome of the analysis.

Data visualisation can be particularly useful for unstructured data and big data as these types of data can often be very difficult to analyse and communicate results from otherwise.

Types of Data Visualisation

The type of visualisation chosen will need to consider the type of data available but should also consider the type of graphic which demonstrates the expected patterns in the data best or the story that needs to be told by the data. For example, does the data visualisation need to show the relationship between different data points, the distribution of data, comparisons of different data points or comparisons over time?

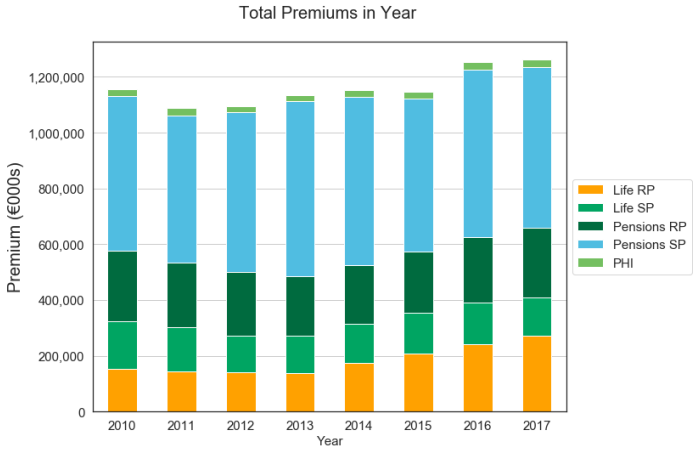
In this section we will look at just a few different types of data presentation for different types of data. However, there are many more types available other than those mentioned in this note.

Basic Charts

Basic charts are the simplest type of data presentation. These can be used, for example, to compare different data sets, to see the composition of data sets, or to analyse trends. Examples of basic visualisation charts include:

- Bar charts
- Scatter plots
- Line graphs

FIGURE 1: BAR CHART OF SALES BY YEAR AND PRODUCT TYPE FOR A SAMPLE COMPANY



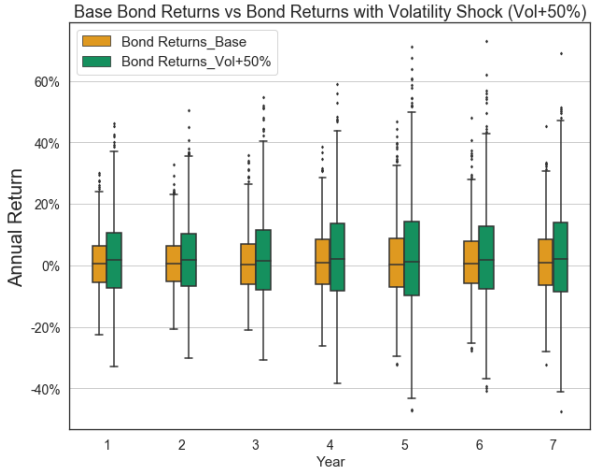
This bar chart shows the composition of premiums in each year, and the trend over the years.

Statistical Charts

Statistical charts can display statistical distributions of data. These can help to identify outliers and averages. Examples include:

- Box plots
- Histograms
- Density Plots

FIGURE 2: BOX PLOT OF PROJECTED BOND RETURNS OVER 1000 SCENARIOS, COMPARING THE BASE SCENARIOS TO SCENARIOS WITH SHOCKED VOLATILITY

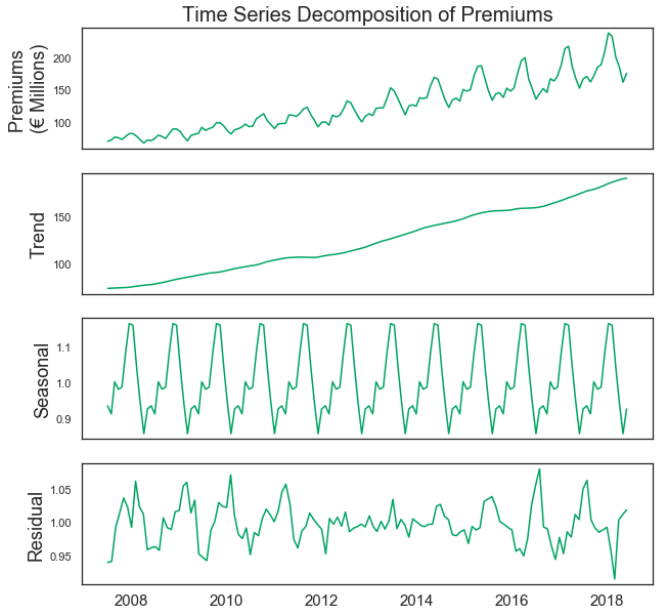


The interquartile range (the range of values between the 25th and 75th percentile) is illustrated by each box. The median of each distribution is illustrated by the line dividing each box. The outliers are any points shown beyond the whiskers of the boxes (beyond the lines coming from the boxes). The end points of the whiskers are the highest and lowest observations which are not outliers.

Time Series

Time series charts can be used to visualise data values over time. This can be useful for spotting trends and seasonal changes in data. The following time series example decomposes the history of premiums received so that the general trend, the impact of seasonality and any residual movements over time can be seen.

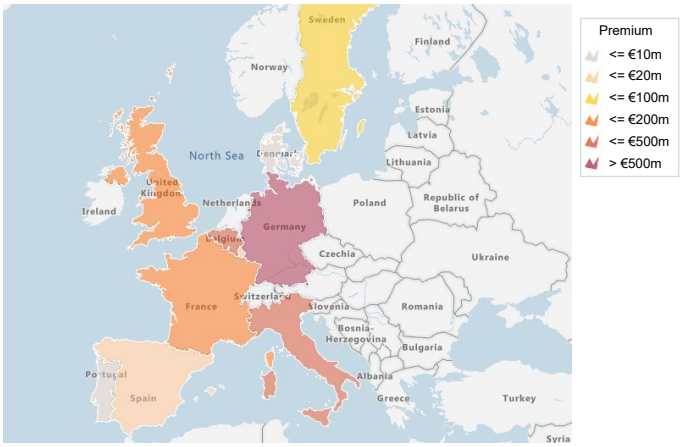
FIGURE 3: TIME SERIES DECOMPOSITION OF PREMIUMS



Geographical Maps

Data can also be displayed visually on a geographical map using scatter plots, bubbles, lines or colour, for example, to display data points in a specific region. In the chloropleth map below, the intensity of the colours indicates the relative level of sales in each country.

FIGURE 4: CHLOROPLETH MAP OF SALES BY COUNTRY



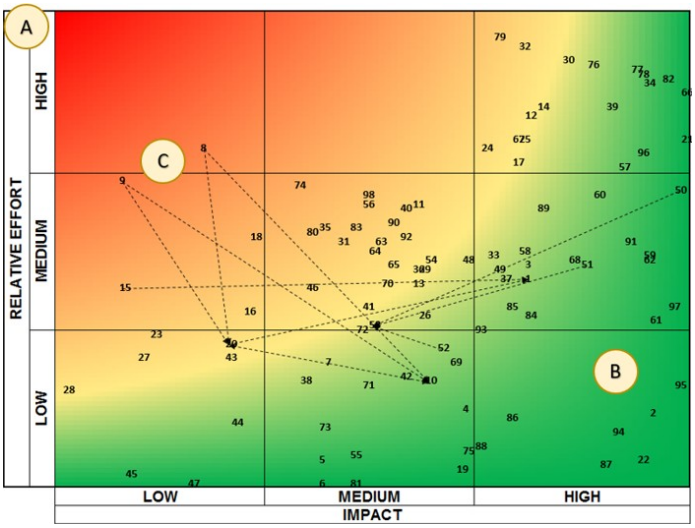
Heat Maps

Heat maps are used to display data using colour.

Correlation matrix heat maps show the correlation coefficients between sets of variables and use colour to indicate the most/least correlated variables in the plot.

The following heat map from a presentation by the Institute and Faculty of Actuaries ("IFoA") Data Visualisation Working Party uses colour to identify the relative importance of client recommendations (from a consultancy report), i.e. those with high impact and low relative effort are in the most green area of the chart, and those with low impact and high relative effort are in the most red area of the chart. This chart also displays dependencies between the recommendations.

FIGURE 5: HEAT MAP OF CLIENT RECOMMENDATIONS FROM IFOA DATA VISUALISATION WORKING PARTY PRESENTATION



A – Overall heat map plotting numbered recommendations by impact and importance, split into High/Medium/Low sections

B – Coloured background to ease identification of relative importance, particularly enabling a user to rapidly identify "quick wins" – recommendations that are low effort and high impact

C – Dependencies between recommendations plotted graphically by linked dotted lines

3D Charts

An extra dimension can be added to data visualisations using 3D charts, for example 3D surface plots, 3D scatter plots, 3D bubble charts . However, care must be taken as 3D charts are notoriously hard to read.

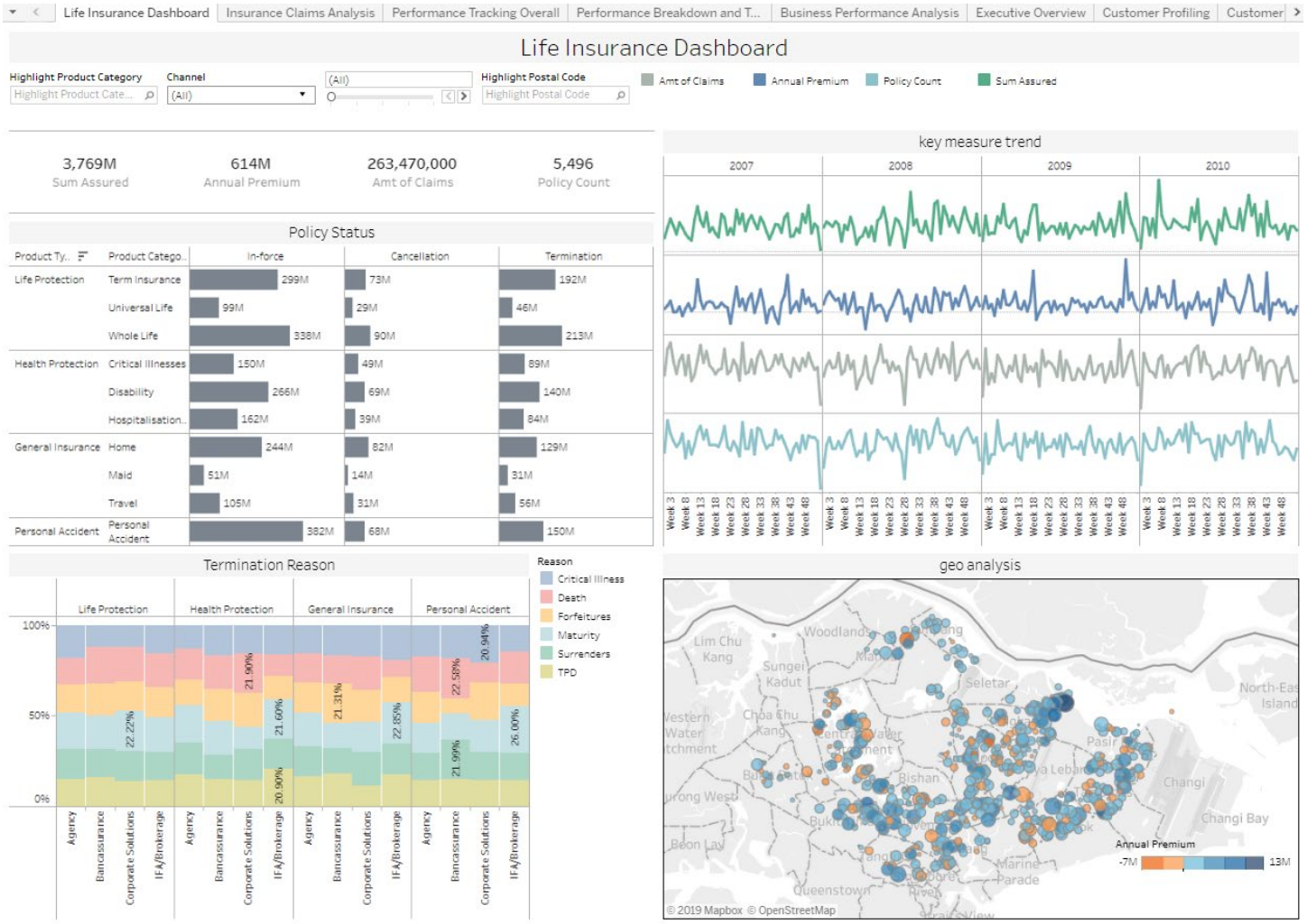
An alternative option is to use a 2D chart which can display three variables, for example a contour chart. In a contour chart the third variable is plotted as contour levels in the chart.

Dashboards

Data dashboards are used to display a number of data visualisations together to provide an “at a glance” snapshot of the business. Data dashboards work in a similar way to car dashboards – data is updated regularly on the dashboard for ongoing monitoring and rapid decision-making.

Dashboards can be customised to show information of particular importance to the company. They provide a central location for businesses to monitor and analyse key performance indicators, metrics and key data points. They can be interactive and allow users to apply filters and drill down into specific information as required.

FIGURE 6: EXAMPLE LIFE INSURANCE DASHBOARD FROM TABLEAU



Dos and Don'ts of Data Visualisation

Data visualisation is not simply a matter of representing information but rather of conveying important messages that will be understood and retained. Particular attention must therefore be paid to the choice of graphic type, colour palette, scale etc.

There are many dos and don'ts for good data visualisations, but some important things to remember are:

- Keep it simple – good visualisations should simplify messages and be as easy to understand as possible.
- Pay attention to the use of colour – use colour as an aid to help understanding but do not over-complicate things with colour. Be mindful of colours with strong connotations or cultural associations.
- Use "honest" data axes – ensure the data axes do not conceal or mislead the true message in the data.
- Use clear and direct labelling.
- Ensure the data visualisation can be easily understood.
- Ensure the point you want to convey in the data visualisation is being made clearly.

How Milliman can help

At Milliman, we have been actively working with our clients for many years to effectively harness the power of data science in order to help meet their business needs.

We can assist you with all aspects of your data science initiatives including providing advice on:

- Data visualisation techniques
- Best practice frameworks for data science processes
- Model interpretation
- Model validation
- Collection and processing of data
- Identifying applications for data science techniques
- Identifying suitable tools and techniques for particular circumstances
- Implementing solutions
- Understanding the implications of results
- Constraints and practical challenges

For further information please contact any of the consultants below or your usual Milliman consultant.

Milliman

CONTACT
Gillian Tucker
Gillian.Tucker@milliman.com

Donal McGinley
Donal.McGinley@milliman.com

Milliman is among the world's largest providers of actuarial and related products and services. The firm has consulting practices in life insurance and financial services, property & casualty insurance, healthcare, and employee benefits. Founded in 1947, Milliman is an independent firm with offices in major cities around the globe.

milliman.com

LinkedIn
Follow our 'Milliman Ireland' page:
https://www.linkedin.com/company/milliman-ireland