Creating effective visualisations

for descriptive and predictive analytics

Dr. Julia Perl, Chessman Wekwete

hannover re®

Note

This document gives an overview of some of the key points discussed during a presentation held at the Australian Actuaries Institute Virtual Insights Session in April 2021. If you are interested in receiving the full presentation including various vivid examples, please get in touch with us.

Introduction

- Traditional visualisations responded to various questions
  - Generally one variable at a time
- Expansion due to increases in
  - Amount/volumes of data and
  - Number and granularity of variables available
  - Expanded analytical methods
  - Audience groups and medium of dissemination
- Elevated importance of data visualisations

Agenda

1 Why should I care?

2 General principles

3 Visualisations to interpret a predictive model

4 Best practices

5 Tools and resources

somewhat different

1 Why should I care?

hannover re®

Communication risks

Dangers

- Audience fails to get message
- Poor/partial message
- Misleading message
  - Defective
  - Biased
- Inaccurate/wrong message

Benefits

- Good visuals tell clearer stories
- Help understanding
- Good recollection

somewhat different

2 General principles

hannover re®

Data visualisation in context

- Visuals
  - Data driven or concept driven
  - Exploration or communication purpose
- Data driven visuals
  - Audience
  - Targeted audience response/actions
  - Medium – visuals package

https://hbr.org/2016/06/visualizations-that-really-work

The reaction and the audience

Typical audience	Target reaction for audience	Objects of audience reaction
• Self-exploratory analysis • Niche audience (pricing team, valuation team, operation teams, Exco, Board, client, industry conference) • Mass audience (in particular non-chart people, end-consumer)	• Compare • Observe • Discover	• Quantities • Trends and projections • Relationships • Distributions • Compositions • Variations

Visuals that achieve communication objective

Follow the science

- Studies on how audiences consume visuals
- Adopt visual aspects to maximise impact

Follow the art

- Select visuals for
  - intended response/actions
  - type of data/message
- Utilise ‘best practice’ approach to creation of visuals

Customise for target audience

Psychology of reading visuals

- How are visuals recognised and recalled?
  - Michelle Borkin et al (2015)
  - Analysis of eye movements
- Title and text are key
- Redundancy helps recall and understanding
- Ranking of channels
  - Cleveland and McGill (1984)

- Studied most accurately decoded
  - Position along a common scale
  - Position on identical but nonaligned scales
  - Length
  - Direction, angle
  - Area
  - Volume
  - Curvature
  - Shading
  - Colour saturation

http://faculty.washington.edu/aragon/classes/hcde511/s12/readings/cleveland84.pdf

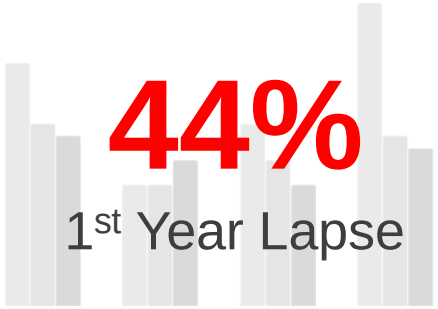
somewhat different

3

Visualisations to interpret a predictive model

hannover re®

Visualisations to interpret a predictive model



XGBoost
using 27 variables

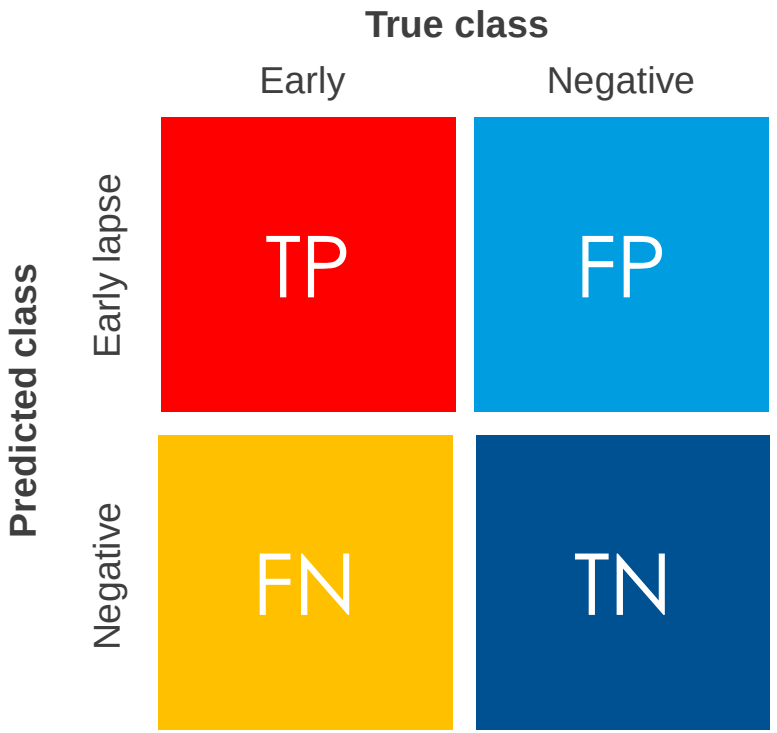
AUC value of 0.84

Do you care only **how well** the model predicts
or also
want to **understand** the model
to **convince** others?

hannover re®

1. Confusion matrix

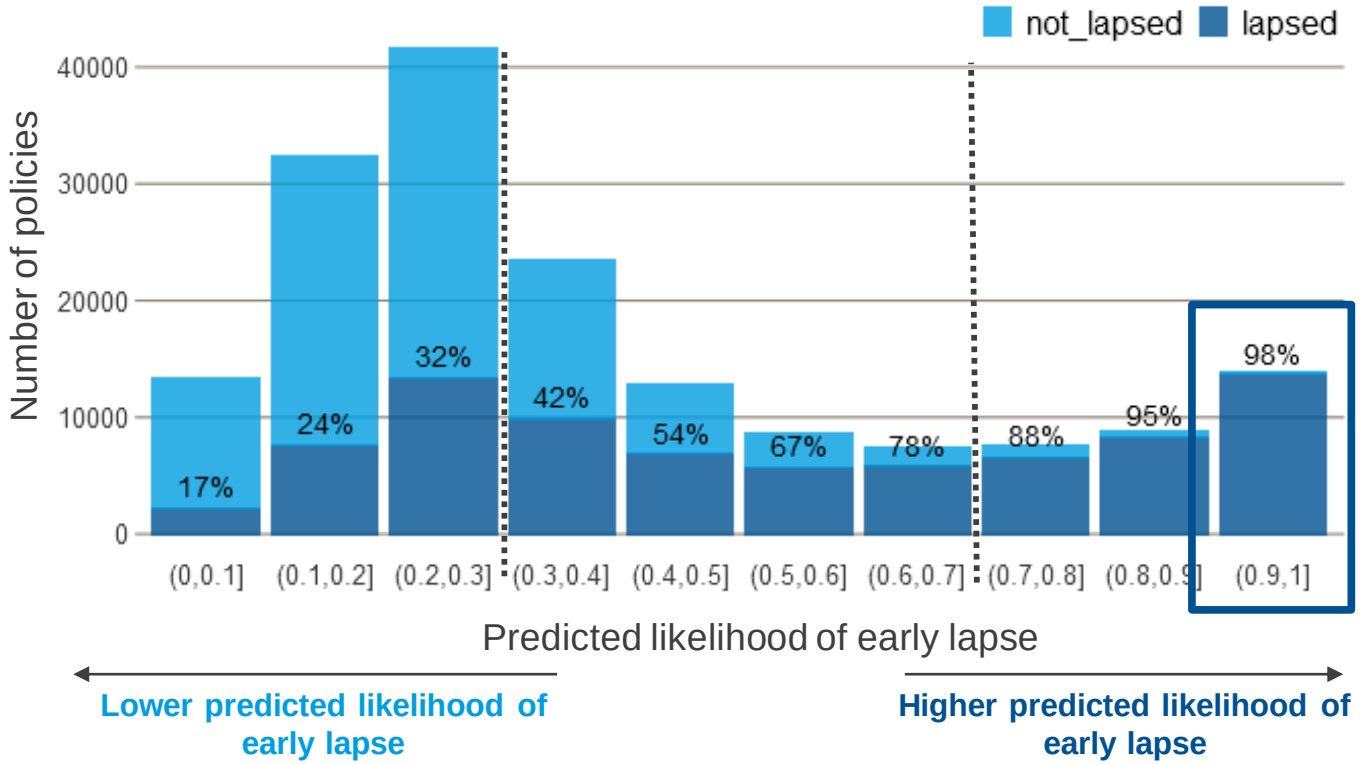
How well does the model predict lapses?



- The confusion matrix shows how well predicted and actual values match
- How often does my model correctly predict actual early lapses as early lapses (TP) or actual negatives as negatives (TN)?
- How often does my model falsely predict either actual early lapses as negative (FN) or actual negatives as early lapses (FP)?

2. Propensity score plot

How well does the model predict lapses?

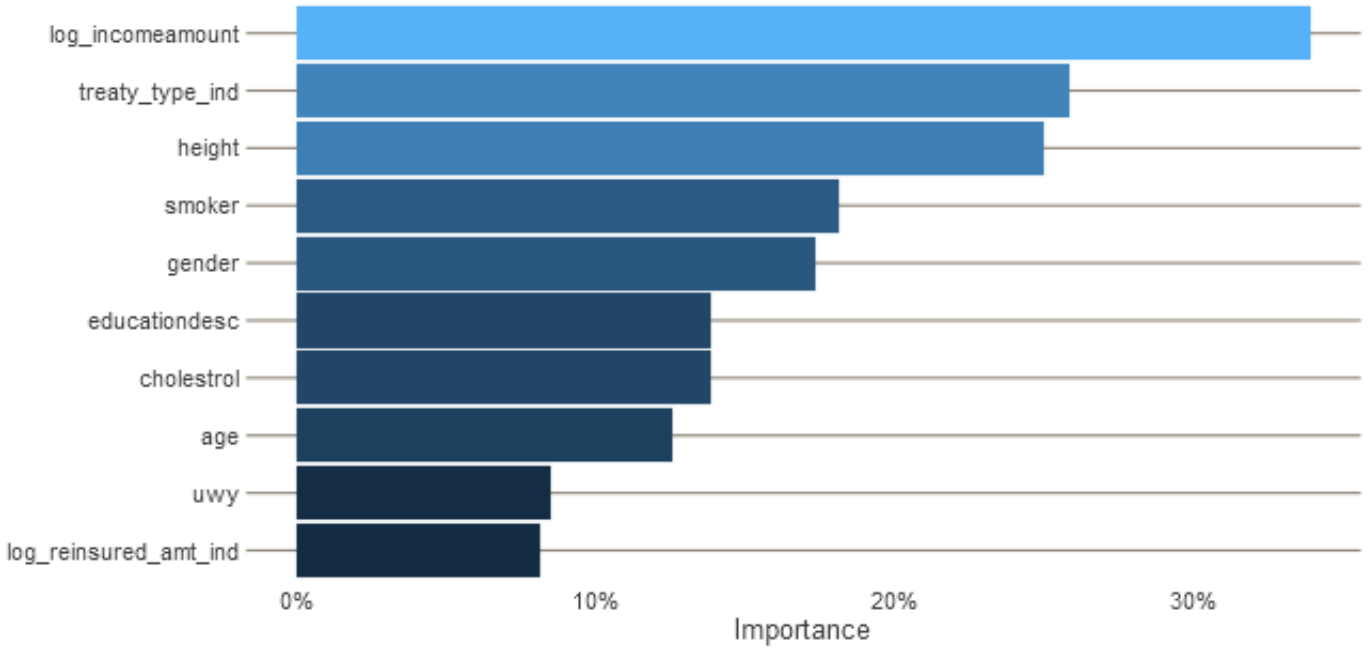


- Distribution of the predicted likelihood of the model versus the actual likelihood
- Out of the 15K policies scored as highly likely lapses ((0.9, 1.0] likely to lapse) 98% actually lapsed.
- Selected propensity range defined individually based on client's strategy

hannover re®

3. Feature importance plot

Which features have the highest impact on the prediction?

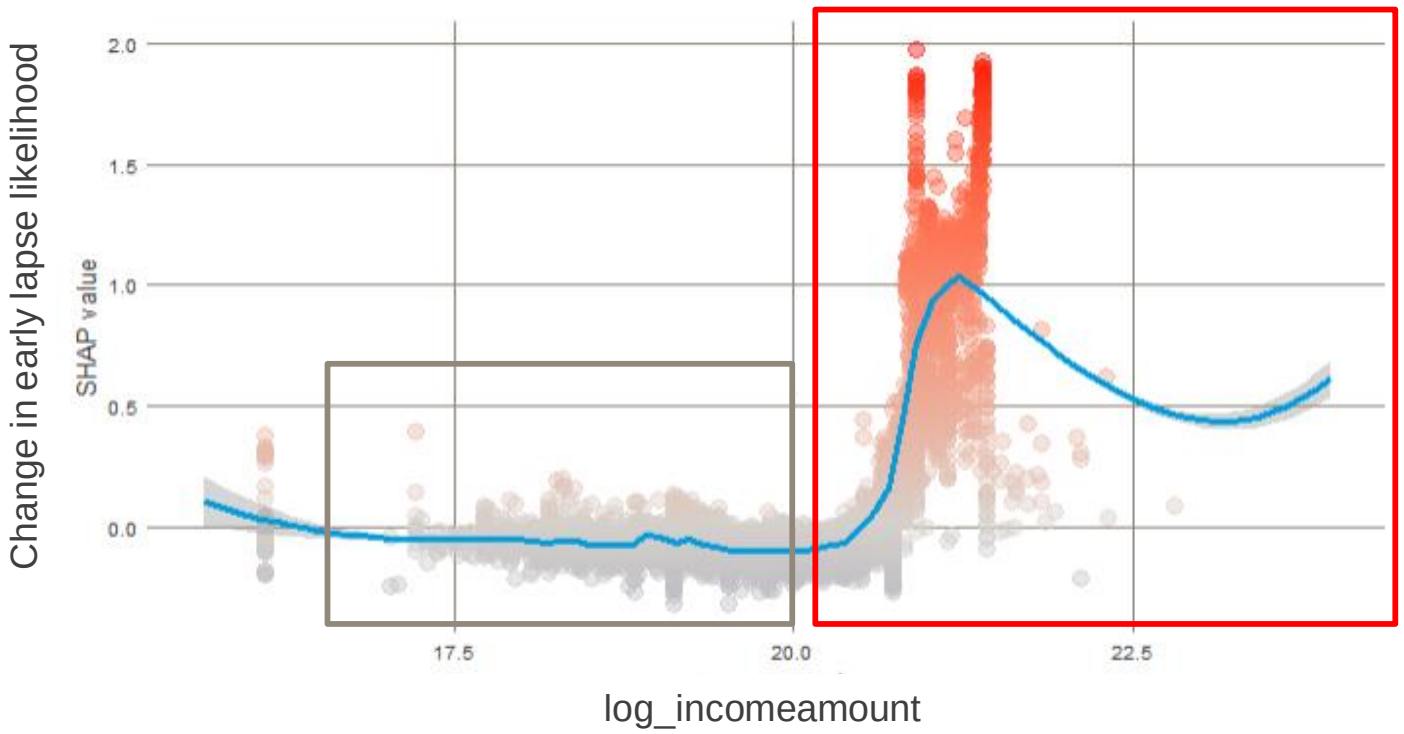


- The ten variables with the highest influence (SHAP feature importance) on the predictions are depicted.
- A value of 30% indicates that on average this variable changes the early lapse likelihood by +30%.

hannover re®

4. Feature dependence plot

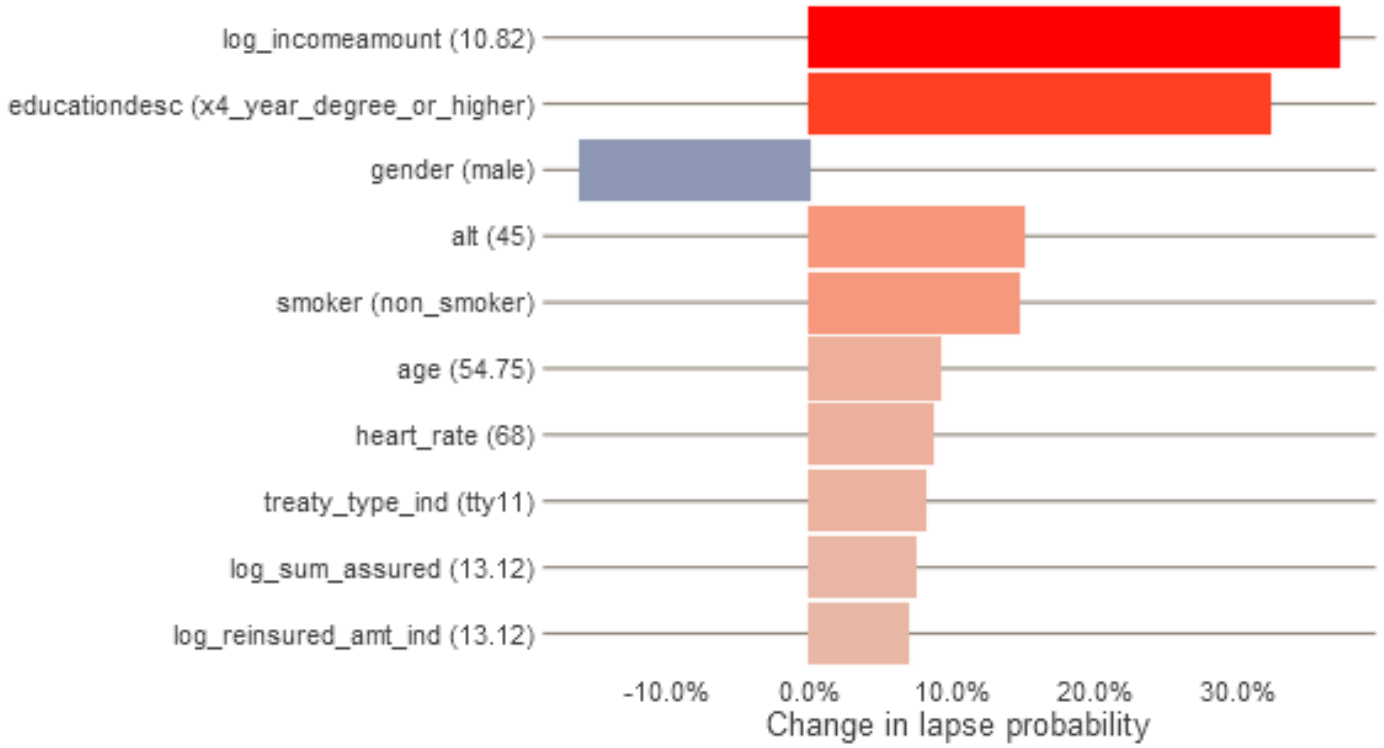
How does a feature impact the prediction?



- Each individual point corresponds to a policy. The higher the policy is being placed, the likelier it is to lapse.
- Median income amounts (grey box), don't have an impact on the lapse prediction.
- The higher the income amount goes, the likelihood of a policy to lapse early tends to increase. (red box)

5. Feature contribution plot

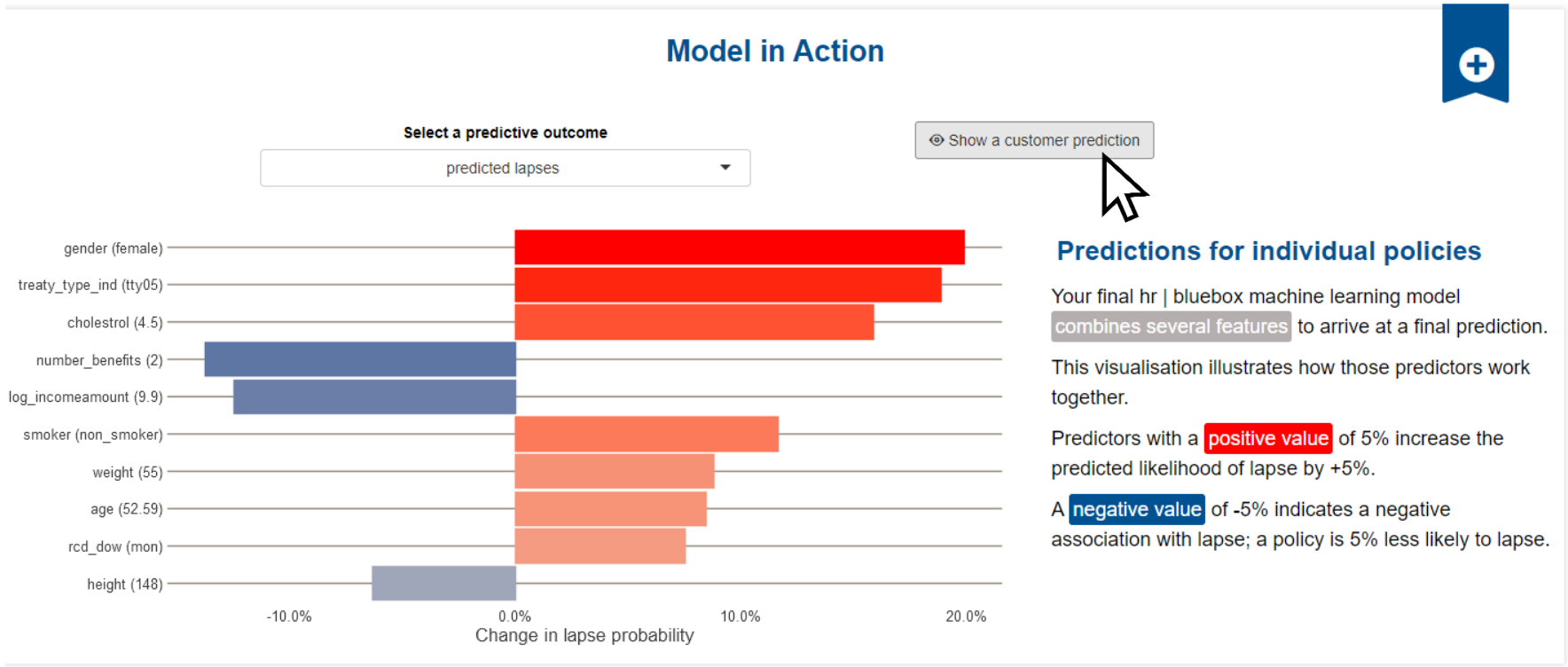
How does a model make individual predictions?



- The model combines several features into a final prediction. This visualisation illustrates how those features work together.
- Predictors with a positive value (e.g. 30%) increase the predicted likelihood of lapse (by +30%).
- A negative value (of e.g. -10%) indicates a negative association with lapse; a policy is (-10%) less likely to lapse.

6. Dashboard

How does a model make individual predictions?
Allow decision makers to interact with your model to make it more real!



hannover re®

somewhat different

4 Best practices

hannover re®

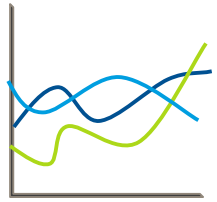
Choosing the right chart type

What would you like to show?

Comparison

How does one set compare to another?

Bar Chart



Line Chart

Relationship

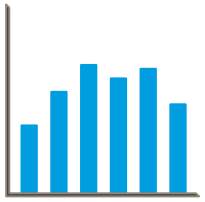
What's the relationship between sets?

Scatter Plot

Bubble Chart

Distribution

How is the data distributed?



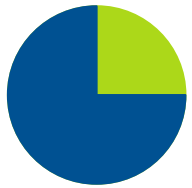
Histogram

Scatter Plot

Composition

Which parts does the data consist of?

Stacked Columns

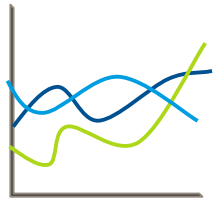


Pie Chart

Trend

How does the data change over time?

Stacked Columns

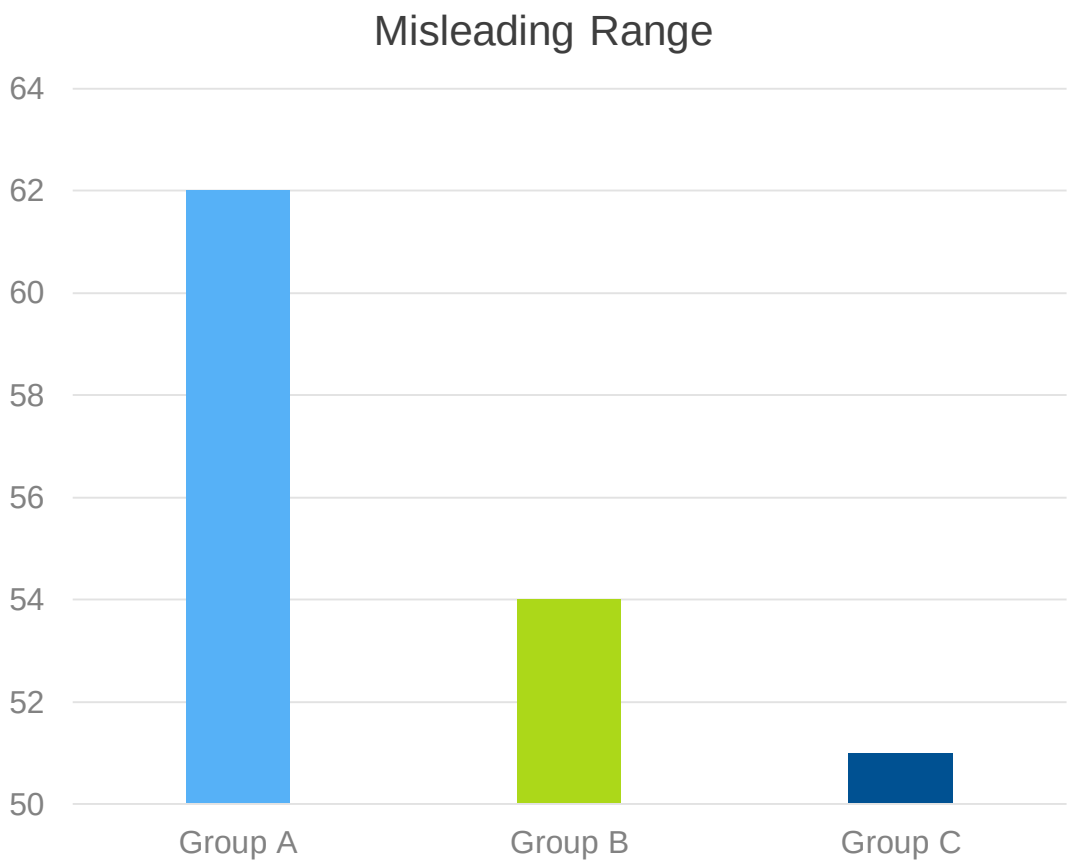


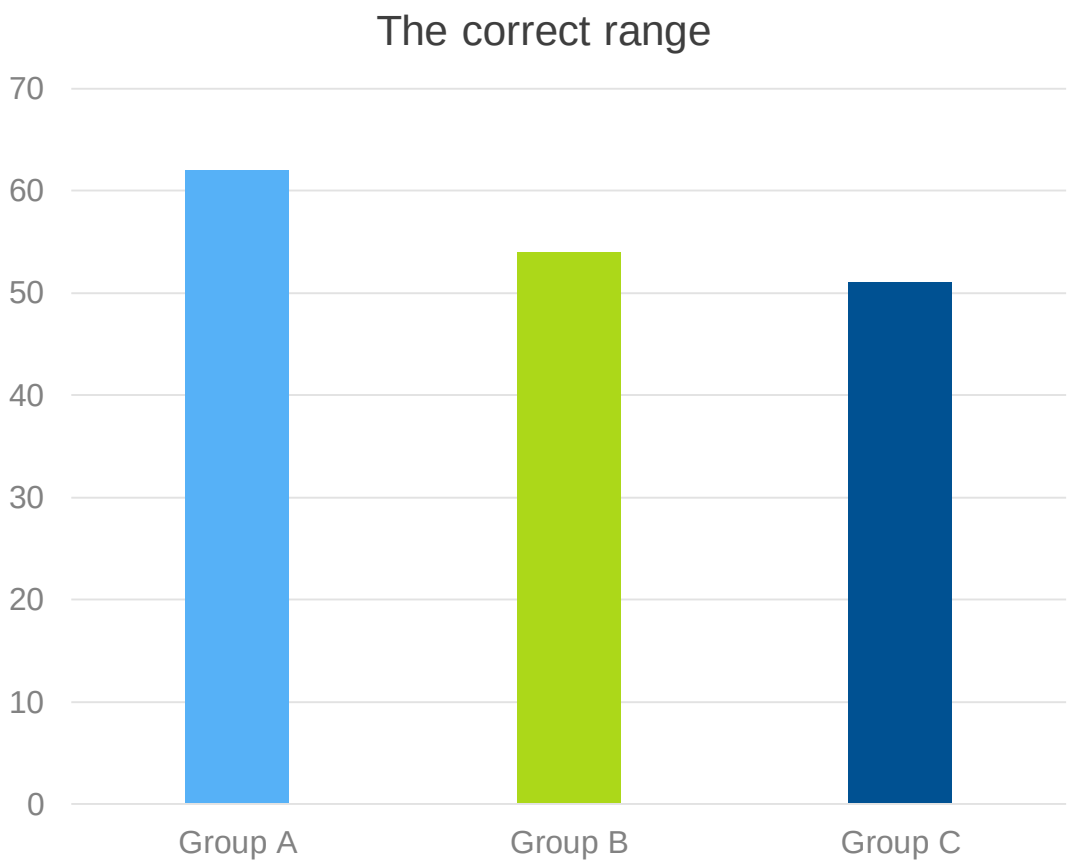
Line Chart

In depth: http://ft.com/vocabulary or https://www.data-to-viz.com/

Scales & ranges

Select the correct and expected range.





hannover re®

Colours, colour scales & contrasts

Choose a colour which fits the expectations of the audience*
*Different cultures/regions have different expectations towards colours

Good

Bad

hannover re®

Overview

1. Choosing the right chart type

- Which objective do you want to achieve?
  Then select the matching chart type.
- Don’t be afraid to experiment with different chart types

2. Scales & ranges

- Label axis properly and clearly
- What is the best scale of your axis? (Logarithmic, percentage, absolute)
- What is the expected range for your data?

4. Text: Title, description & annotations

- Title: What is the main message of your viz?
- Description: What is depicted?
- Use annotations to highlight important points in your viz
- Use colour to relate to coloured elements in the viz

3. Colours, colour scales & contrasts

- Choose distinctive colours
- Choose the colour in line with audience’s expectations
- Be careful with contrasts

somewhat different

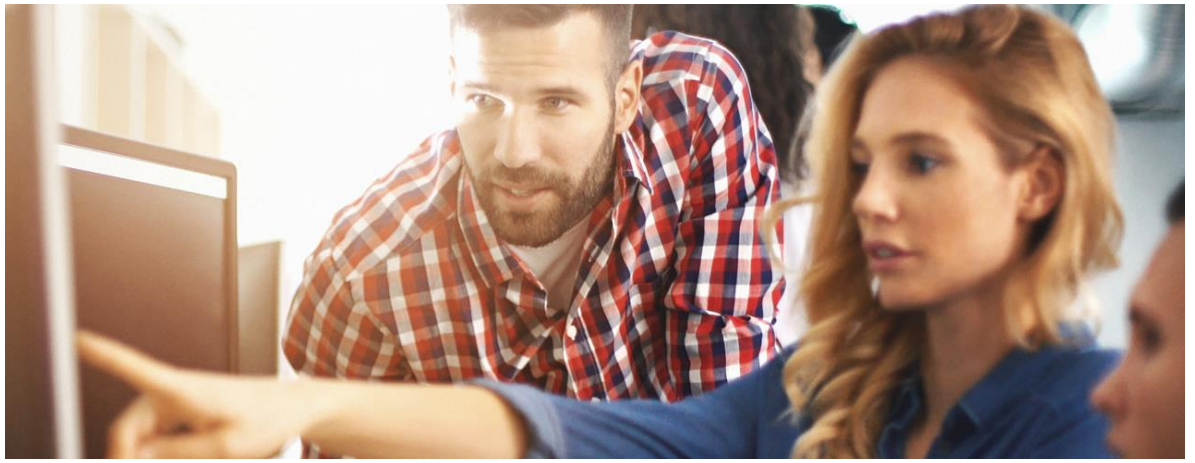
5 Tools and resources

hannover re®

Software, further resources and books

Software tools

- Excel
- Tableau
- Power BI
- R
- Python



Further resources

- How to select the right chart type:
  - Interactive tool
  - Cheat sheet for charts
- Checklist for data viz
- Interpretable Machine Learning

Books

- Healy, K. (2018). *Data visualization: A practical introduction*. Princeton University Press.
- Wilke, C. O. (2019). *Fundamentals of data visualization: A primer on making informative and compelling figures*. O'Reilly Media.
- Munzner, T. (2014). *Visualization Analysis and Design*. A K Peters/CRC Press.

Summary

Visualisations have increased in importance

- greater data volumes and dimensions
- to explain complex processes/relationships
- greater need to convince people
- viz are particularly important for predictive analytics

There's little excuse for getting it wrong

- there are many great resources (theory and practical)
- lots of tools and templates
- seek audience feedback

Contacts

Dr. Julia Perl
Senior Data Scientist

Tel: +49 511 5604-2949
julia.perl@hannover-re.com

Chessman Wekwete
Head of Data Insights

Tel: +61 282462-648
chessman.wekwete@hlra.com.au

Disclaimer

The information provided in this presentation does in no way whatsoever constitute legal, accounting, tax or other professional advice.

While Hannover Rück SE has endeavoured to include in this presentation information it believes to be reliable, complete and up-to-date, the company does not make any representation or warranty, express or implied, as to the accuracy, completeness or updated status of such information.

Therefore, in no case whatsoever will Hannover Rück SE and its affiliated companies or directors, officers or employees be liable to anyone for any decision made or action taken in conjunction with the information in this presentation or for any related damages.

hannover re®