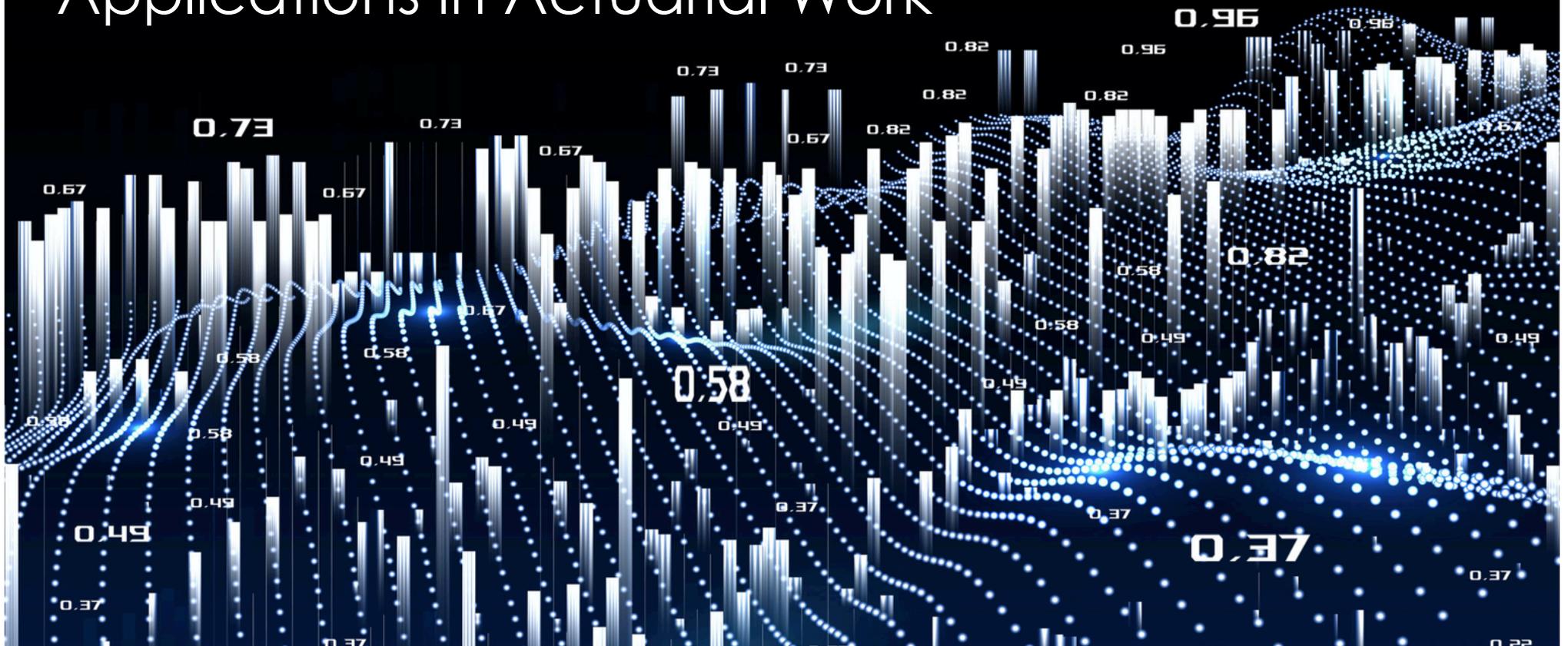


Data Science Applications in Actuarial Work



actuarTech



Agenda

- Introductions
- Opportunities of data science
- Challenges of data science
- Practical Case Study
- Where to start: learning data science hands on

Opportunities of Data Science

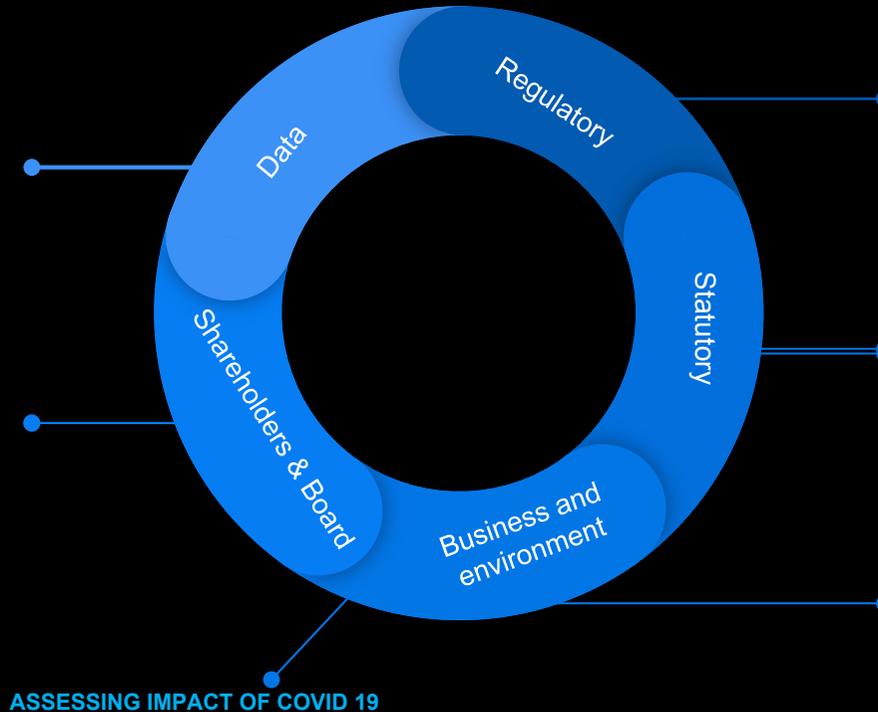
What are some of the pressures on insurers?

ACCESS TO THE RIGHT DATA, AT THE RIGHT TIME

A huge increase in data generation, data capture and data storage combined with significantly increased computing power is providing insurers with a unique opportunity to re-evaluate the value that their data can provide; and the technologies available to do that.

SPEED, GRANULARITY, MORE MI, REDUCED COSTS

Supporting initiatives around strategic direction, value add, insight, intelligence and advanced risk management



ASSESSING IMPACT OF COVID 19

ADVANCED ANALYTICAL AND CONDUCT REVIEWS

BOE Transforming data collection from the UK financial sector. "The FCA's refreshed Data Strategy sets out a transformation plan to become a highly data-driven regulator."

IFRS 17 AS A STRATEGIC TRANSFORMATIONAL INITIATIVE?

Assess data infrastructure as well as analytical, modelling, data governance, storage and reporting capabilities.

PRESSURE TO OPERATE AS FUNCTION OF THE FUTURE

More review, more analysis, more challenge. Less handle turning, less processing and less manual data transfer activities.



actuartech

THE DATA OPPORTUNITY

NEW FORMS
OF DATA

NEW TYPES
OF MODELS

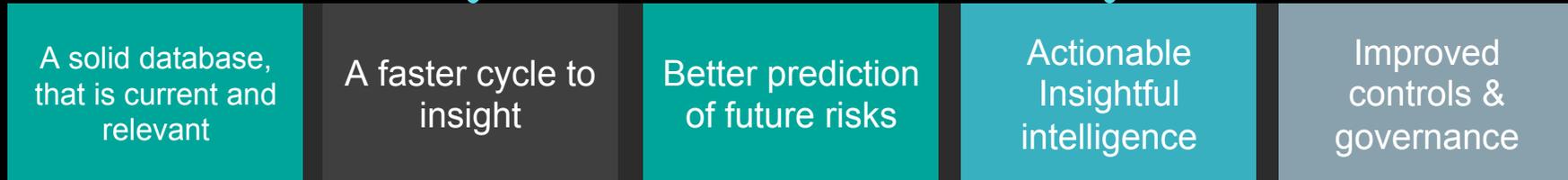
NEW FORMS
OF ANALYSIS

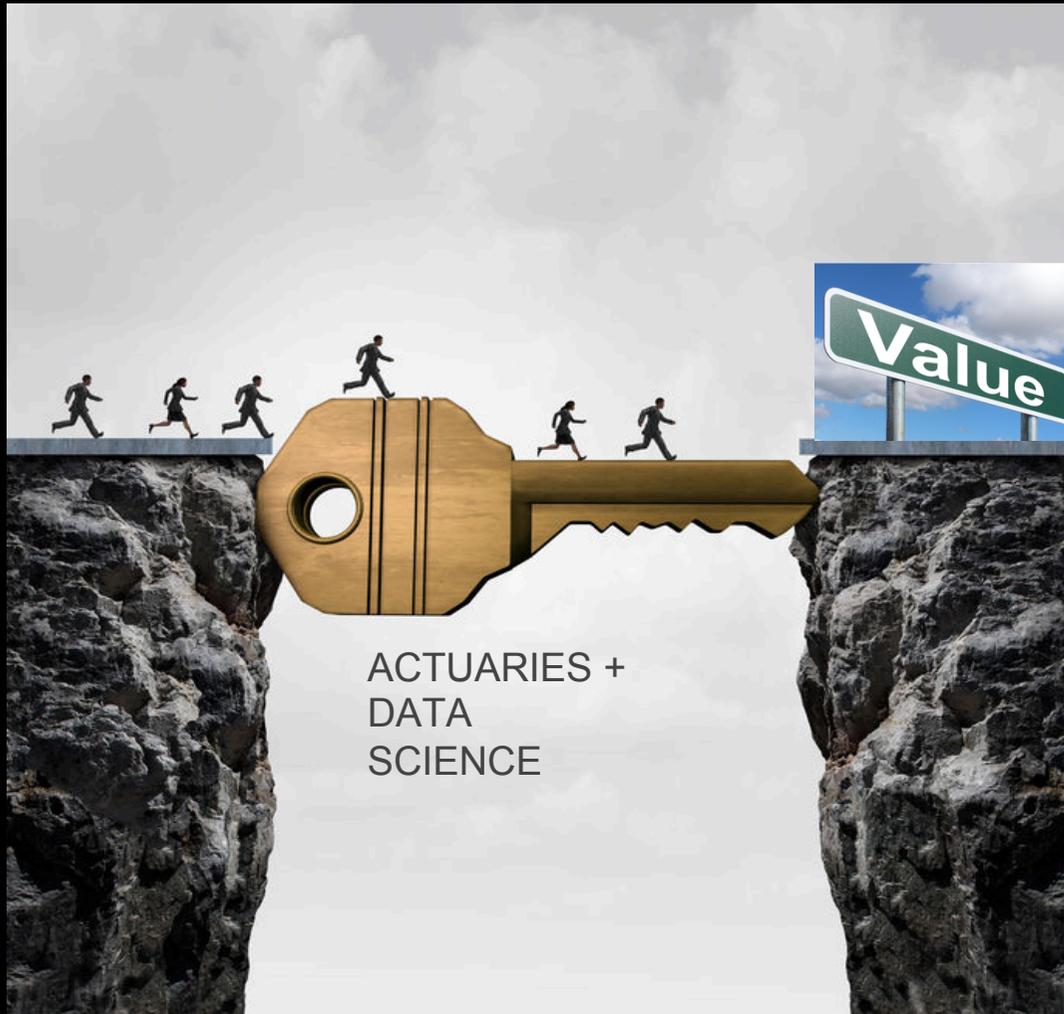


ENHANCED CYCLE OF ANALYSIS

The actuary's wish list?

- Mindset
- Skillset
- Toolkit





EMERGING TRENDS IN THE FIELD OF DATA SCIENCE

Data Science Related Trends

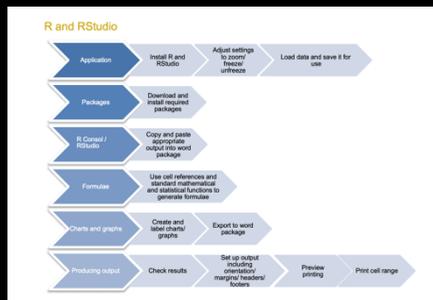
- Availability of Data increasing: more data, new sources
- Better & faster computing power
- More research in statistical fields and applications combined with computer science
- More pressure to use business knowledge to add value

Institute and Faculty of Actuaries

- Modelling Analytics & Insights Data working party 2016
 - Practical Applications Papers
 - Collaboration between Life & Non-Life Actuaries
- Data Science Member Interest Group 2018
 - Virtual conference: Data Science and Opportunities for Actuaries
 - Optimise use of data science techniques
- Publishes 'A guide for ethical data science' 2019 (with Royal Statistical Society)
- Launches Data Science Certificate Course

Other relevant trends

- R included in actuarial syllabus
- Need for Data Science skills growing
- Growing presence of online communities and availability of training courses
- Trends from actuartech.com: Interest from market: Interpretable Machine Learning
- Data & trend analyses and prediction of impact of COVID-19 pandemic



<https://www.actuaries.org.uk/system/files/field/document/CS1%20and%20CS2%20Guide.pdf>

https://southamptondata.science/ugc-1/1/2/0/candidate_info_pack_and_policies_ifo.pdf

www.actuartech.com



Challenges of Data Science

Some Risks of Data Science

- Using open-source programming languages which could pose risk from a governance and security perspective
- Building models which are wrong, not validated or poorly understood
- Using incorrect, inappropriate or otherwise flawed data or drawing conclusions from data which may not be statistically significant
- Coded models being reviewed by inexperienced, unskilled staff
- Models used and appraised out of context
- New Staff unfamiliar with context
- Not understanding data, data quality issues, validation
- Not enough governance in place when models are deployed

Please note the Risks and practical considerations outlined here are examples of what needs to be considered and is not an exhaustive list. In addition these are largely relevant to all areas of data science, not just in respect of the use of R

Important Considerations

Normal professional standards and ethical conduct apply

Data source, checks and controls need to be put in place

Assumptions need to be verified and document

Model approach and testing to be validated and documented

Results with limitations and uncertainty communicated

Tailor communications to audience, i.e.:

Avoid jargon

High level or detailed results as appropriate / possible

Ensure clear audit trail with version control, independent peer review and model validation and interpretation governance in place

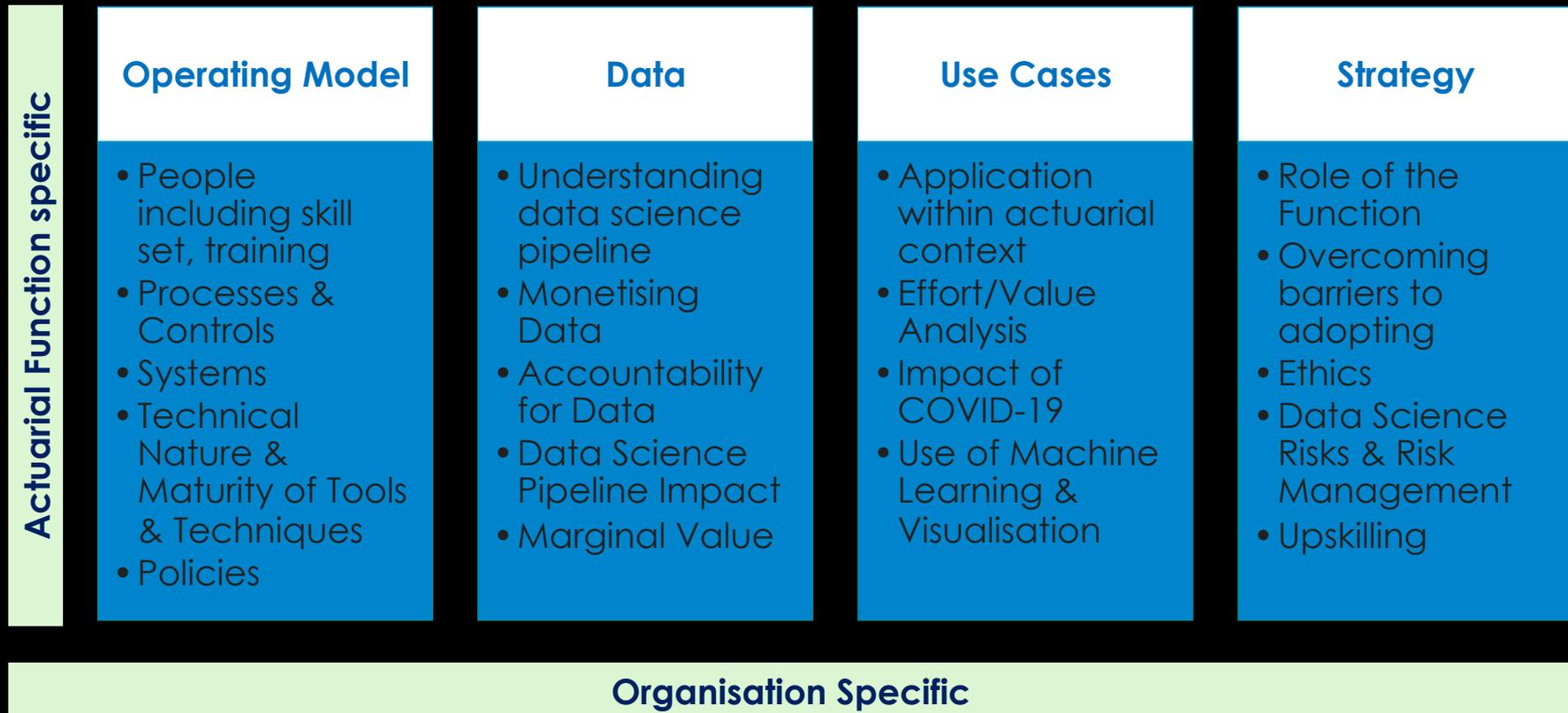


actuartech

The use of data science

Benchmarking: key takeaways

Key Focus Areas of benchmarking exercise



Data Science Example Use Cases in Insurance

Pricing

Predictive Underwriting

Telematics

Price Optimisation

Risk and Capital Management

Assumption Setting

Fraud and Anomaly Detection

Audit and Review Tools



Customer Insights

Predictive Persistency
Management

Recommender Systems

Segmentation

Management Information

Dashboarding

Automated Reporting Analysis

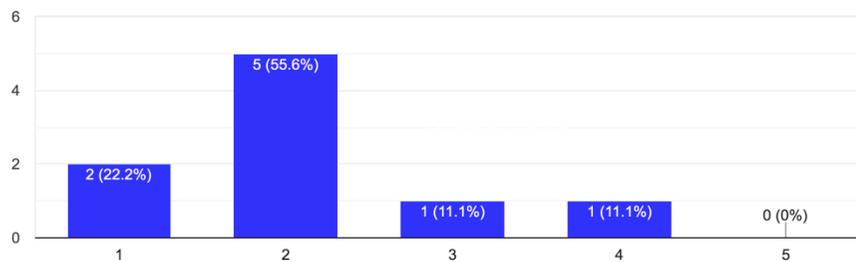


actuarTech

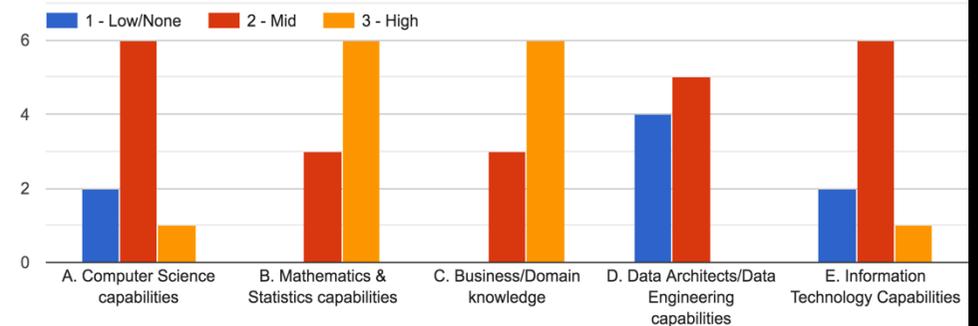
Trends in the use of data science

2.1.1 Please select the function's overall competency with data science.

9 responses



2.1.5 How would you rank the following aspects of your talent within the function:



The following areas of up-skilling required have been identified by our respondents

- High amount of up-skilling required in the business application of data science and validation & interpretation of data science results;
- Medium to High amount of up-skilling required in the use of programming languages and data science algorithms
- Medium to High amount of up-skilling required in data science risk management and communication

Actuaritech analysis

What are the barriers to increasing the application of data science in your function?

Lack of infrastructure or computing software

Low quality of internal data

Difficulties accessing internal data

Privacy issues related to 3rd-party data

Challenges in obtaining and aggregating data from multiple internal data sources

Difficulties in finding and collecting relevant external data

Not a high priority for us

Lack of executive support

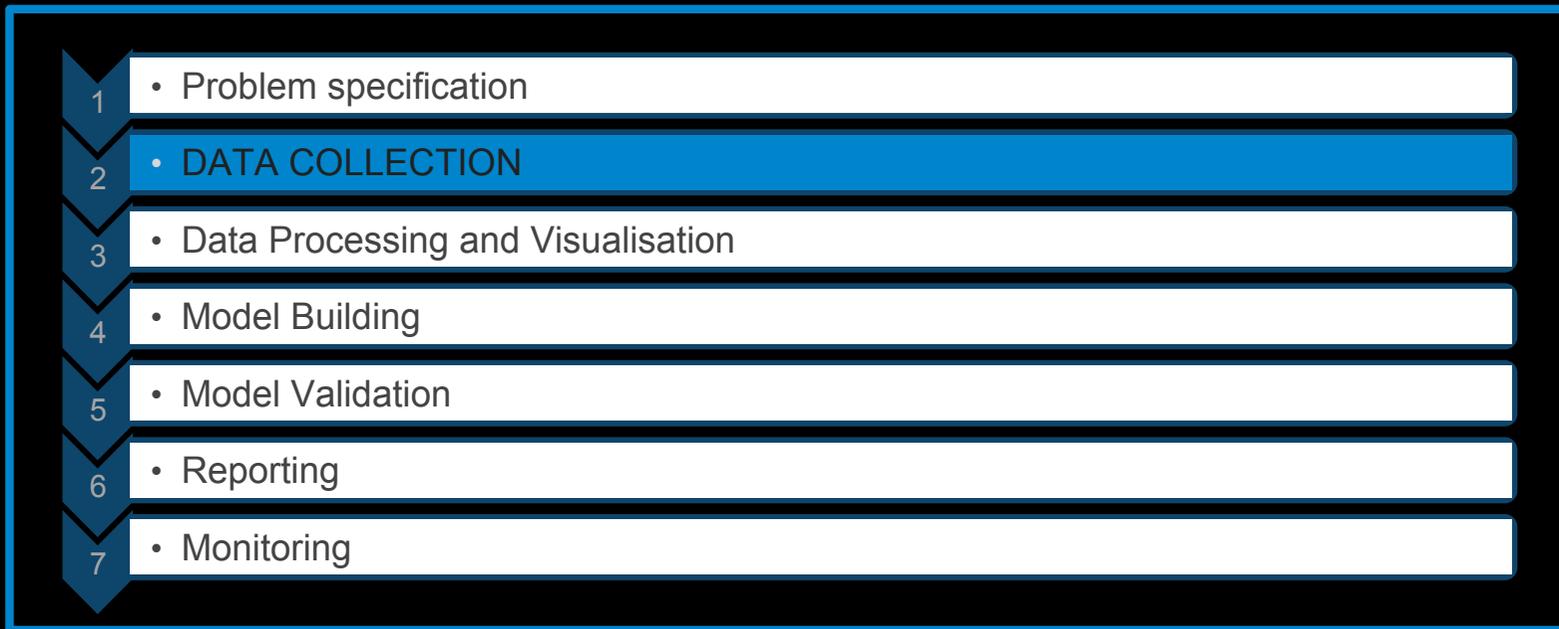
Lack of internal talent & expertise

No or few relevant use cases identified

The reality is that the data, tools, skills and resources are not readily accessible for use case design & development, decision making, risk management or to derive insight.

Case Study

Case Study – Data Science Process



LAPSE CASE STUDY: Hands On

The screenshot shows a web application interface for a course. On the left is a dark sidebar with the ActuarTech logo at the top. Below the logo is a link to 'Go to my Dashboard', the course title 'Case Study: Lapse Rates', and a progress indicator '0% complete'. A search bar is present. A list of lessons follows: '1 - Introduction' (0/3), '1.1 - Introduction' (LESSON), '1.2 - Overview of Lapses' (LESSON), '1.3 - Overview of the Lapse Study' (LESSON), and '2 - Problem Specification' (0/4). The main content area is titled '1.1 - Introduction' and contains the following text:

Introduction

1.1 - Introduction

This course will serve as an end-to-end walk-through of an investigation into lapse rate data. We will guide you through importing, cleaning, investigating, model fitting, visualising, and interpreting the data.

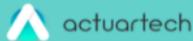
We will use the open source programming software, R, as the main programming language of instruction, and it will be presented in a similar fashion to the *Foundations in R for Actuaries* course (see [here](#) for those unfamiliar with the fundamentals of using R) by making use of Jupyter notebooks with the code and explanations embedded. The code is interactive so you can run it and make tweaks to see how it effects output. We encourage you to work through the course in your own time and make of the code provided in a personal capacity (i.e. run the same lapse analysis using a slightly different dataset in RStudio).

This end-to-end case study is presented as a course and aims to assist users in answering how life insurers can address low persistency and how data and analytics can help.

At the bottom right of the main content area is a button that says 'MOVE TO NEXT LESSON →'. The ActuarTech logo and name are also visible in the bottom right corner of the interface.

Interactive coding: understanding lapse rates

You are previewing all course lessons



< Go to my Dashboard

Case Study: Lapse Rates

0% complete

Search by lesson title

- 1 - Introduction 0/3
- 2 - Problem Specification 0/4
- 3 - Data Collection and Management 0/5

- 3.1 - Introduction LESSON
- 3.2 - Importing and Transforming the Data MULTIMEDIA

3.3 - Preliminary Analysis

jupyter 3.3 - Preliminary Analysis (autosaved)

File Edit View Insert Cell Kernel Widgets Help Not Trusted | R O

3.3.6 - Lapse Rates by AGE_BAND

Now by age band:

```
In [8]: #LAPSE RATES BY AGE_BAND
plot <- summarize(group_by(data,AGE_BAND ), EXPOSURE_CNT=sum(EXPOSURE_CNT),
                  LAPSE_CNT=sum(LAPSE_CNT),EXPOSURE_AMT=sum(EXPOSURE_AMT)) %>%
mutate(lapse_rate=LAPSE_CNT/EXPOSURE_CNT)

ggplot(plot, aes(AGE_BAND, lapse_rate, fill=AGE_BAND))+
  geom_bar(stat = "identity")
```

``summarise()` ungrouping output (override with `.groups` argument)`



MOVE TO NEXT LESSON →

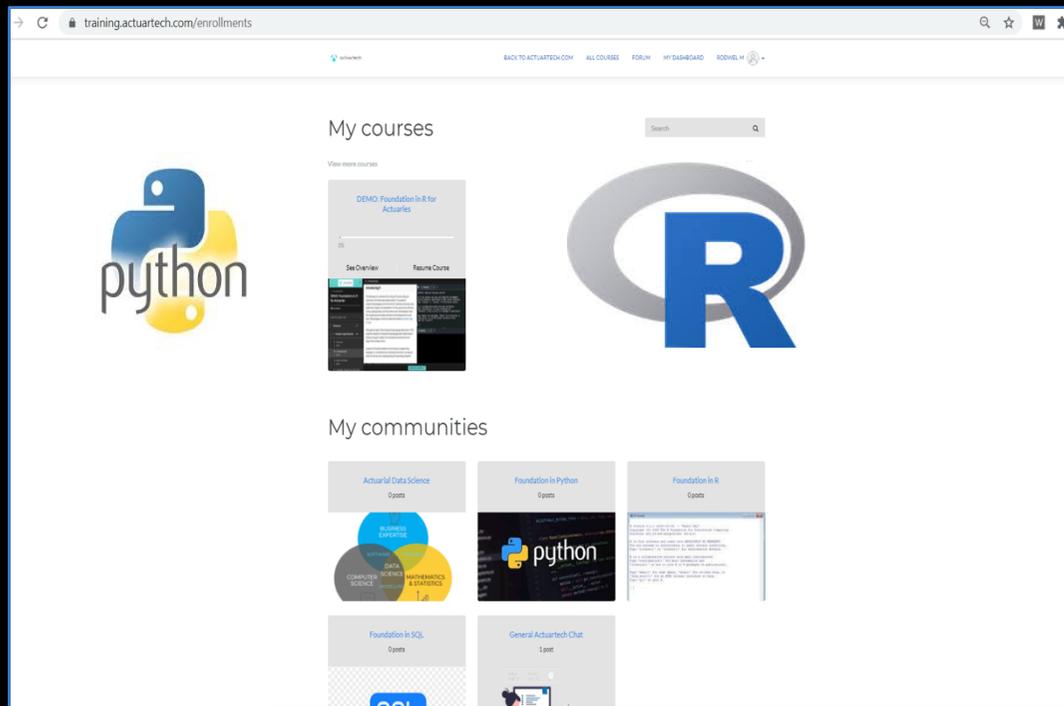
 actuartech

Comparison To Traditional Approach

	Traditional Lapse	Predictive Modelling
Data	<ul style="list-style-type: none">• Internal data• Lower volume	<ul style="list-style-type: none">• Internal and external data• Suited to large datasets
Feature Selection	<ul style="list-style-type: none">• Limited by internal data sources• Incorporates judgment that could lead to bias	<ul style="list-style-type: none">• Model based• Explores wider feature set
Modelling	<ul style="list-style-type: none">• Grouped data analysis• Average experience data cell	<ul style="list-style-type: none">• Policy level modelling• Wider model toolbox• Multivariate analysis
Communication	<ul style="list-style-type: none">• Limited	<ul style="list-style-type: none">• Richer and transparent analysis of drivers• Incorporates visualisations

Where do you start?

TRAINING PLATFORM



- Training Platform designed by actuaries for actuaries
- Web based interface supporting multiple languages-no need to install software
- Case studies and Assignments relevant to actuarial work, based on relevant datasets provided
- Track your progress
- Learn in your own time
- Option to interact and network with peers



actuartech

ACTUARTECH TRAINING PLATFORM

The screenshot displays the ActuarTech training platform interface. At the top, a browser address bar shows the URL: `training.actuartech.com/courses/take/demo-foundations-in-r-for-actuaries/multimedia/15323309-4-5-creating-histogram-and-density-plots`. Below the browser, a teal header contains the ActuarTech logo and a navigation link: `Go to my Dashboard`.

The main content area is a Jupyter notebook titled "4.5 - Creating Histogram and Density Plots". The notebook interface includes a menu bar (File, Edit, View, Insert, Cell, Kernel, Widgets, Help) and a toolbar with icons for navigation and execution. The notebook content consists of a text cell, a code cell, and a plot cell.

The code cell contains the following R code:

```
claim_size, probability = TRUE, xlab = "Claim Size",  
main = "Histogram of claim size", ylim = c(0, 0.0002), col = "light blue")  
lines(density(claim_size, from = 200, na.rm = TRUE),  
col = "black", lwd = 2)  
abline(v = mean(claim_size, na.rm = TRUE),  
col = "red", lty = 2)
```

The plot cell displays a histogram titled "Histogram of claim size" with a density curve overlaid. The x-axis is labeled "Claim Size" and ranges from 0 to 60,000. The y-axis is labeled "Density" and ranges from 0 to 0.00020. A vertical red line indicates the mean claim size.

Annotations on the screenshot include:

- Web browser interface**: Points to the browser address bar.
- Case Studies Examples Assignments**: Points to the left sidebar navigation menu.
- Interactive coding**: Points to the code cell in the Jupyter notebook.

At the bottom of the notebook, there is a text cell: "If we want to try identify outliers, we can start by viewing the top 10 highest claims:" and a button labeled "MOVE TO NEXT LESSON →".

Web browser interface

Case Studies Examples Assignments

Interactive coding



VIEWING YOUR COURSES ON THE PLATFORM

My courses

Search

All ▾

View more courses

DEMO: Foundation in R for Actuaries

1%

See Overview | Resume Course

Resource Library

0%

See Overview | Start Course

Q&A

Contact us on info@actuaritech.com for any questions.

