Society of Actuaries in Ireland

Analytics & Big Data
What, Why and How

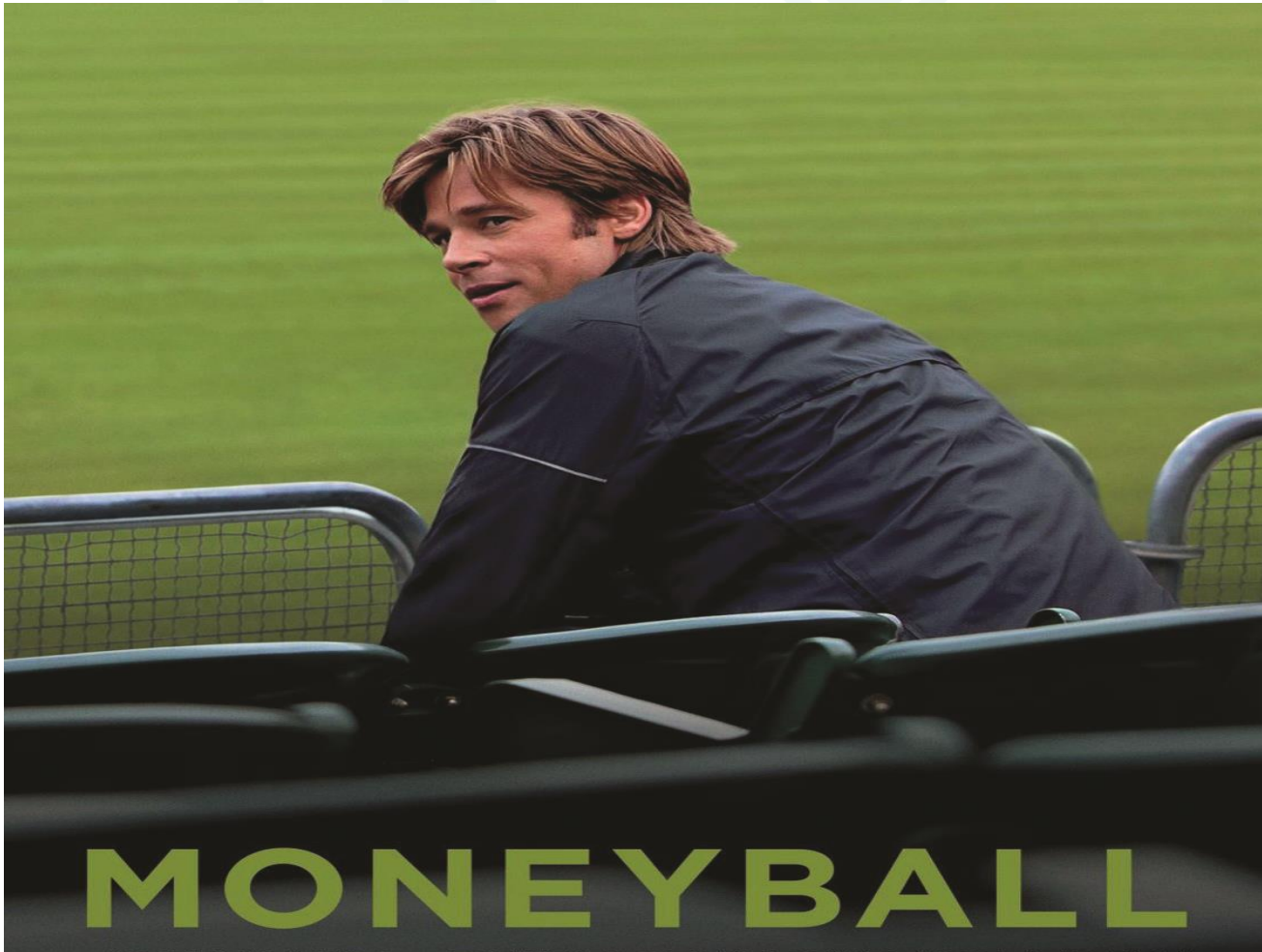Colin Murphy FSAI
Dr. Richard Southern
Sinead Kiernan FSAI

07.04.2014

Agenda

- Introduction

- What is Analytics and Big Data?

- Growth of Analytics and Big Data

- What does analytics mean for the insurance sector? What is the role of actuaries?

- Analytics applied to the insurance sector

- Tools used in data analytics

- Questions

MONEYBALL

What is Analytics and Big Data?

Understanding Analytics

Foresight

Understand the signals generated across being generated across your ecosystem to shape the future

Insight

Use data from within the organisation to drive changes here and now

Hindsight

Conduct "rear-view mirror" assessments based on data generated by operations

Foresight	Predictive	Simulation & optimisation
		Advanced data mining
		Predictive modelling
		Advanced forecasting
Insight	Descriptive	Segmentation insights
		Data trend and pattern analysis
Hindsight		Slice and dice queries and drill-downs
		Management & KPI reporting
		Enterprise data management

Increasing business advantage and sophistication

"Analytics" is a discipline which focuses on the conversion of data to information, and information to actionable insight, to drive decisioning across the entire spectrum

Why Business Analytics?

Powerful trends are driving the adoption of a new approach to business analytics. An unforgiving demand for performance, a wake-up call for better risk management in the face of tougher regulatory enforcement, and exponentially increasing amounts of data to process, comprehend, and react to.

Profitability – The ability to compete profitably. Businesses are increasingly competing on their ability to exploit the data they are holding and address the following;
- Revenue generation
- Cost reductions
- Reduce exposures to risks

Rapid response to new and changing business signals: changes happen increasingly quickly in today's markets and to prosper businesses need to be able to identify the earliest signals, plan, and react immediately

Hidden insight: buried in the flood of data, businesses collect are the facts about what is really happening - decision makers need the expertise and capability to uncover hidden patterns that would otherwise go undetected

Data volumes & technology capacity: global data volumes continue to grow exponentially – businesses have to be able to control, store and publish this data in numerous forms
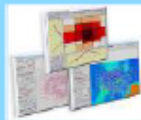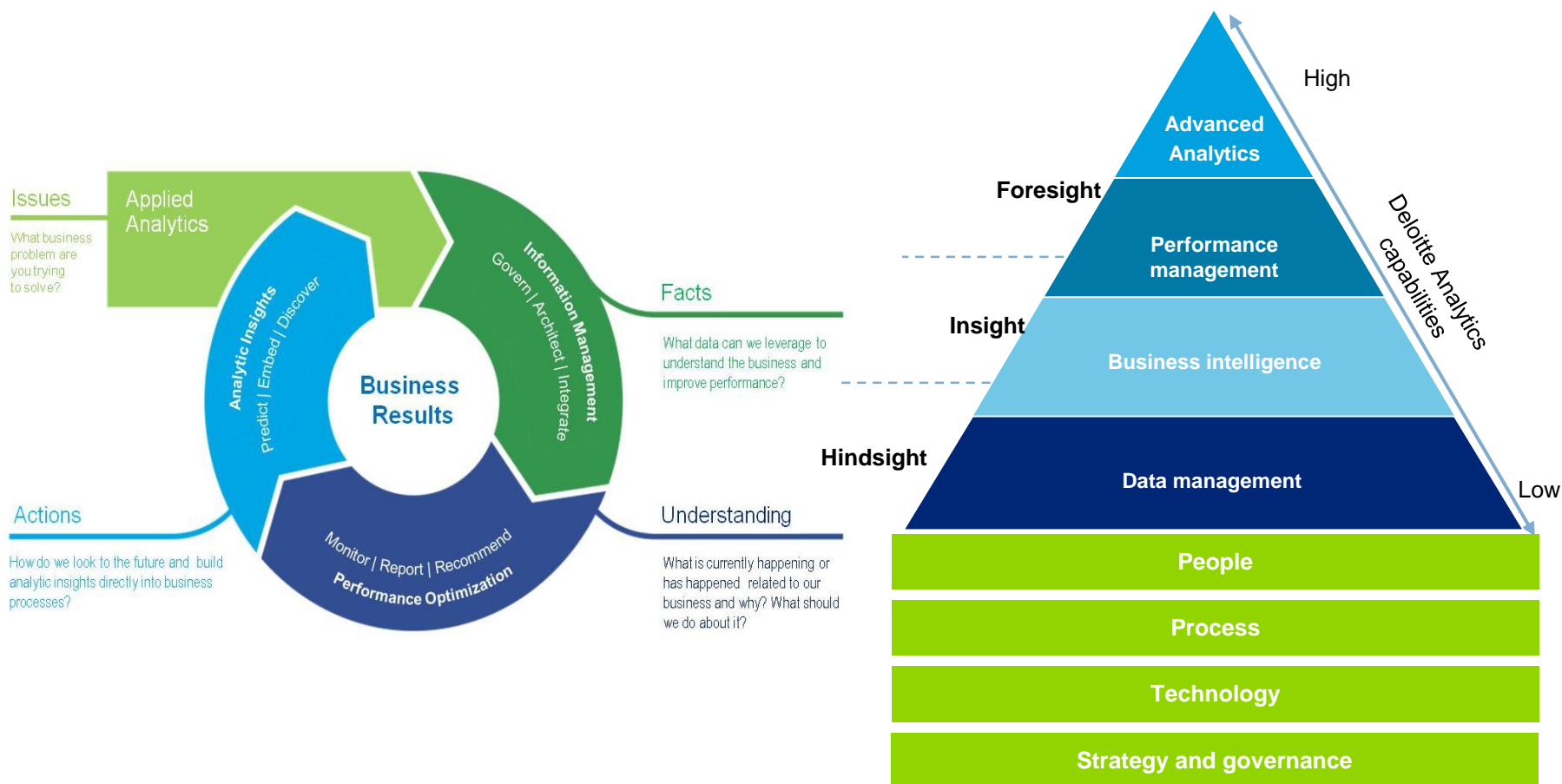

Regulations: regulators are demanding deeper insight into risk, exposure, and public responsiveness from financial, public and many other sectors requiring integrated data across the enterprise
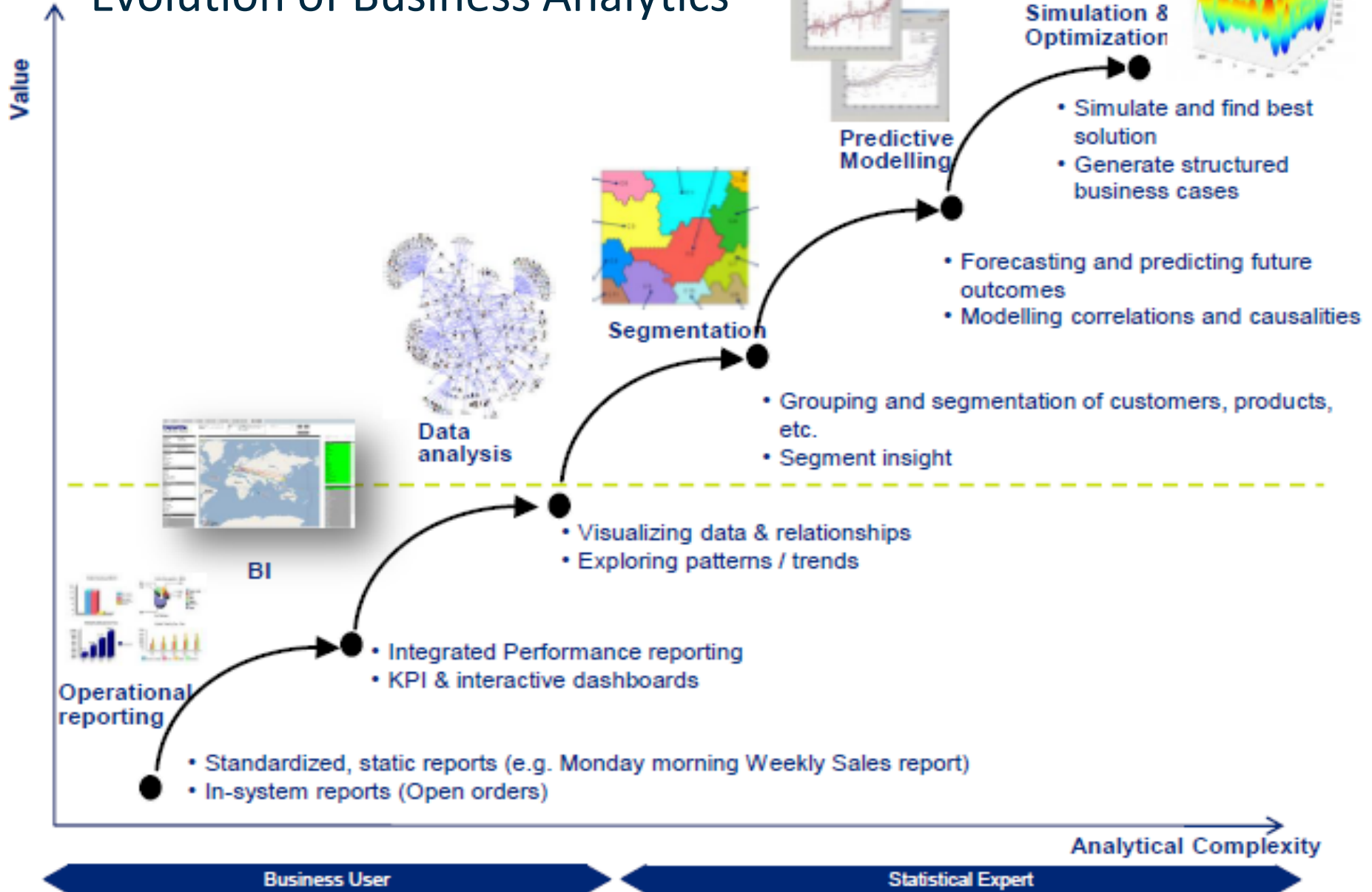
Analytics defined

Business analytics is the practice of using data to drive business strategy and performance. It includes a range of capabilities – from looking backward to evaluate what happened in the past, to forward-looking approaches like scenario planning and predictive modeling. It spans the capabilities stack from data management and business intelligence up through performance management and advanced analytics.
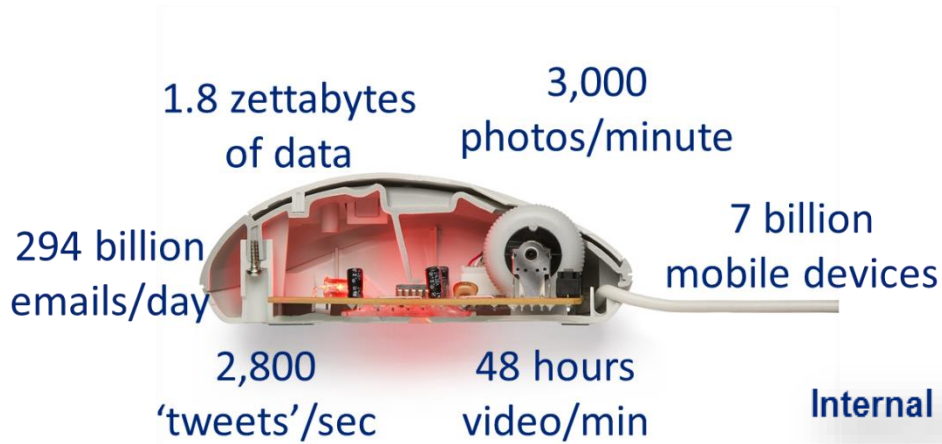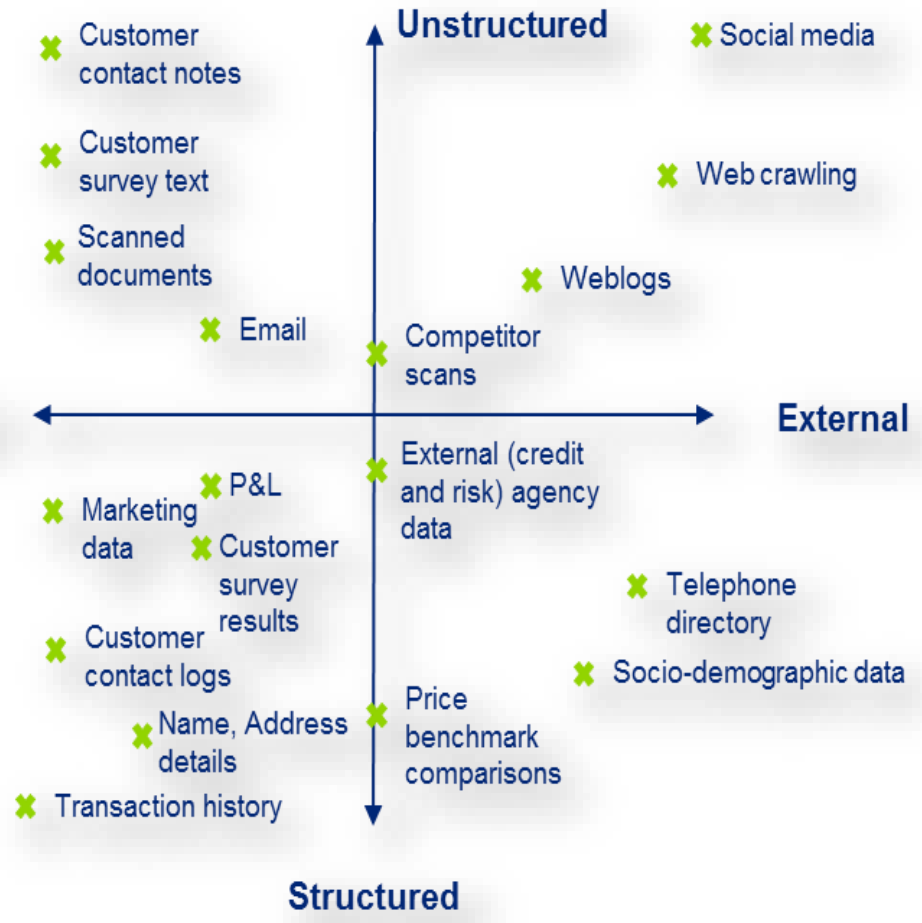
Evolution of Business Analytics

Value

Simulation & Optimization
- Simulate and find best solution
- Generate structured business cases

Predictive Modelling
- Forecasting and predicting future outcomes
- Modelling correlations and causalities

Segmentation
- Grouping and segmentation of customers, products, etc.
- Segment insight

Data analysis
- Visualizing data & relationships
- Exploring patterns / trends

BI
- Integrated Performance reporting
- KPI & interactive dashboards

Operational reporting
- Standardized, static reports (e.g. Monday morning Weekly Sales report)
- In-system reports (Open orders)

Analytical Complexity

Business User | **Statistical Expert**

Growth of Analytics and Big Data

1.8 zettabytes of data

3,000 photos/minute

294 billion emails/day

7 billion mobile devices

2,800 'tweets'/sec

48 hours video/min

1 kilobyte	**1,000**,000,000,000,000,000,000	
1 megabyte	**1,000,000**,000,000,000,000,000	
1 gigabyte	**1,000,000,000**,000,000,000,000	
1 terabyte	**1,000,000,000,000**,000,000,000	
1 petabyte	**1,000,000,000,000,000**,000,000	
1 exabyte	**1,000,000,000,000,000,000**,000	
1 zettabyte	**1,000,000,000,000,000,000,000**	

Unstructured

Customer contact notes

Customer survey text

Scanned documents

Email

Social media

Web crawling

Weblogs

Competitor scans

Internal

External

P&L

Marketing data

Customer survey results

Customer contact logs

Name, Address details

Transaction history

External (credit and risk) agency data

Telephone directory

Socio-demographic data

Price benchmark comparisons

Structured

Analytics Survey

Basically, analytics is about making good business decisions. Just giving reports with numbers doesn't help. We must provide information in a way that best suits our decision-makers.
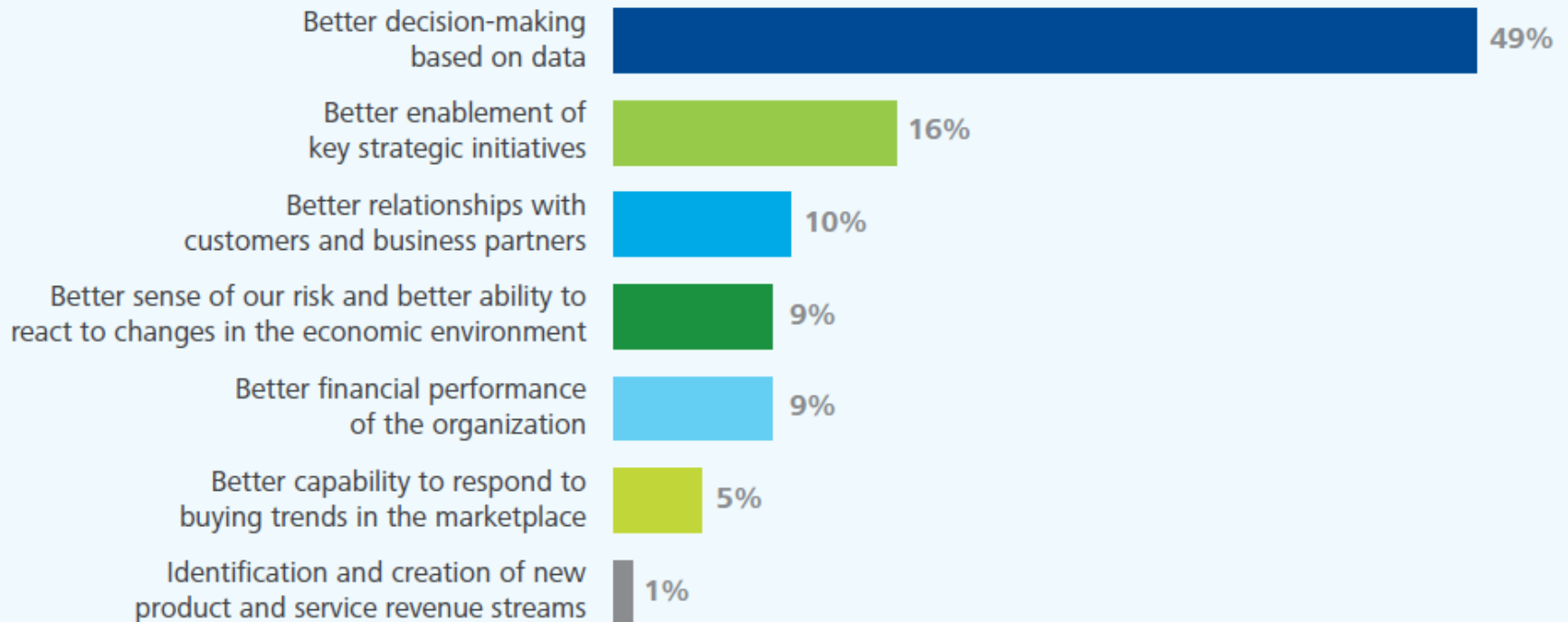
Two areas where analytics drives ROI

Income production   **Cost reduction**
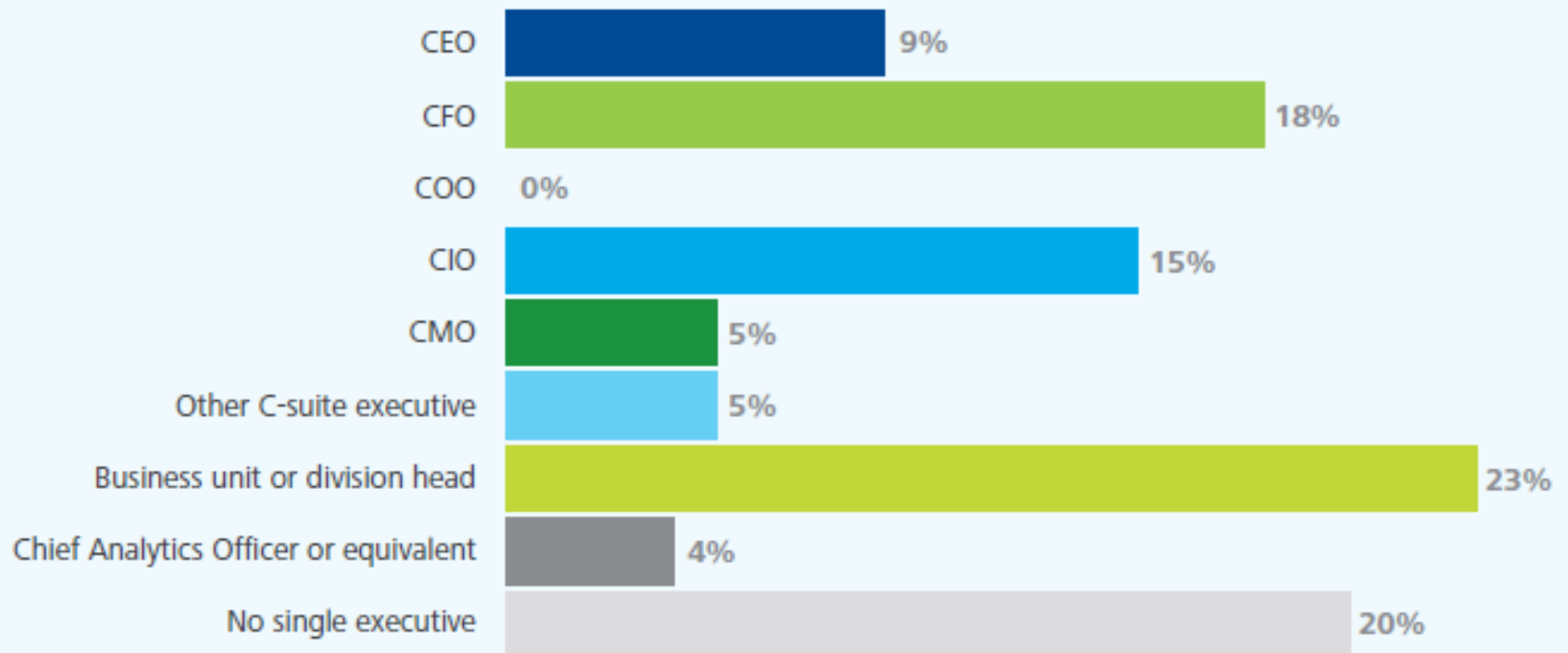
ANALYTICS IS KEY FACTOR IN DECISION-MAKING

Better decision-making based on data	49%
Better enablement of key strategic initiatives	16%
Better relationships with customers and business partners	10%
Better sense of our risk and better ability to react to changes in the economic environment	9%
Better financial performance of the organization	9%
Better capability to respond to buying trends in the marketplace	5%
Identification and creation of new product and service revenue streams	1%

Analytics Survey

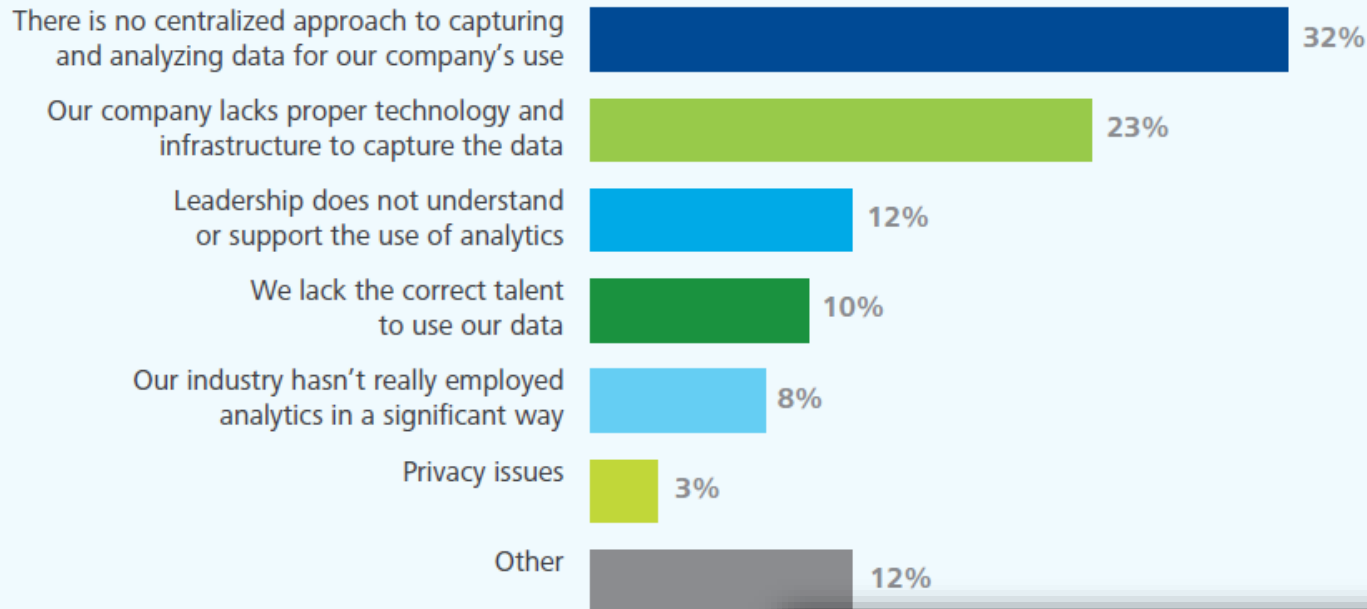"We're really not spending money on data analytics. We're using it to find better alternatives for making money."
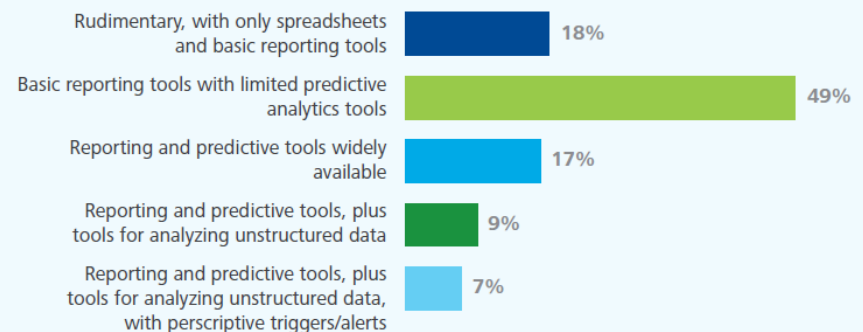
WHO OVERSEES ANALYTICS INITIATIVES?

Role	Percentage
CEO	9%
CFO	18%
COO	0%
CIO	15%
CMO	5%
Other C-suite executive	5%
Business unit or division head	23%
Chief Analytics Officer or equivalent	4%
No single executive	20%

Analytics Survey

LEADING BARRIERS TO ANALYTICS USE

Barrier	Percentage
There is no centralized approach to capturing and analyzing data for our company's use	32%
Our company lacks proper technology and infrastructure to capture the data	23%
Leadership does not understand or support the use of analytics	12%
We lack the correct talent to use our data	10%
Our industry hasn't really employed analytics in a significant way	8%
Privacy issues	3%
Other	12%

MAJORITY OF ORGANIZATIONS LACK TECHNOLOGY TO SUPPORT ANALYTICS

Category	Percentage
Rudimentary, with only spreadsheets and basic reporting tools	18%
Basic reporting tools with limited predictive analytics tools	49%
Reporting and predictive tools widely available	17%
Reporting and predictive tools, plus tools for analyzing unstructured data	9%
Reporting and predictive tools, plus tools for analyzing unstructured data, with perscriptive triggers/alerts	7%

Analytics Survey

"It's just a huge amount of data to deal with. Data management is becoming a bigger and bigger part of the puzzle, and a bigger and bigger challenge for us to overcome."

DATA QUALITY CHALLENGES MAY IMPEDE ANALYTICS ADOPTION

Poor. Inferior quality and organization of data that makes it difficult for substantial analysis. No groups with strong data orientation. — 4%

Un-integrated. Data is usable but in functional or process silos. Our senior executives do not discuss data management. — 31%

Adequate. We have identified key data domains and created central data repositories. — 31%

Good. Integrated, accurate, and common data maintained in a central warehouse. Data is still mainly an IT matter. We have little unique data. — 21%

Excellent. We have a relentless search for new data and metrics. A team separate from IT oversees information, and data is viewed as a strategic asset. — 13%

Analytic Trends

The talent crunch that wasn't

When you're recruiting analytical people, be clear what your needs are. Just saying you need a data scientist is like advertising a slot for a *smart person who's good with numbers*.

- Data Science is an emerging and growing area

- Growing demand for talent across all industries

- Significant growth in availability of courses: 3rd level, MOOC etc

- There are shortages of skilled people but…

- Good data scientists take time to develop and rely on past experience

- The effective embedding of analytics across Industries is evolving all the time

Analytic Trends

A picture is worth a thousand numbers

- Data visualisation is real, and valuable

- Helps to give a better understanding to the data

- Well used hypothesis generation method to understand patterns, trends and behaviours

- However, there is a right way and a wrong way

- The overuse of visualisation can get in the way of clarity -
    - Making a complex story too simple or
    - Making simple facts seem too complex

Analytic Trends

Machine learning finds a big data niche

- Not new – In theoretical existence since 1960's and academic use since 1970/80's

- Assists in task automation & separating "signal from the noise"

- Several success stories of application to real world business problems

- Don't exclude expert intervention - Bring back the Humans!!

- "Garbage in – garbage out" holds for any modelling scenario

- Overly complex estimators are badly understood and accepted in business - transparency

Analytic Trends

The rise of the Chief Analytics (and Big Data) Officer

- Emerging as a new C-level position - sitting at the Exec table

- Chief Analytics Officer (CAO), Big Data Officer (BDO), Chief Data Officer (CDO). . .

- Helps support, sponsor and drive the analytics initiative

- Other CXO level execs have vested interests – CIO, CFO, CMO…

- Is there a measurable benefit to the organisation in creating this function – time will tell!

Data Scientists

Perception, Skills and shortage

"The sexy job in the next ten years will be statisticians… The ability to take data—to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it—that's going to be a hugely important skill."

Hal Varian, Google's Chief Economist

DATA MINING
Collecting and formatting the information

HACKING AND COMPUTER SCIENCE
Coding, programming, and theory behind computational science

STATISTICS
The collection, organization, analysis, interpretation and presentation of data

CREATIVITY AND INSIGHT
Business insight and understanding, and the ability to draw novel conclusions from the data

COMMUNICATION
Representation or visualization in the form of presentations and infographics

PREDICTIVE ANALYTICS
Forming predictions based on the data

Case Study – Business Intelligence
Creating valuable insight from multiple data sources

Business Problem	Data	Analytics	Insight

Dynamic Dashboarding

To allow quick and easy visualisation of unemployment data:

- Nationwide
- By areas with and without a welfare office
- By town or city

And plotting of unemployment rates:

- Against the level of education in a town
- Against the occupation held

Data Sources

- Welfare office locations (www.welfare.ie)
- Irish Census 2011 data (www.cso.ie):
  - Economic status of citizens
  - Occupation of citizens
  - Social class of citizens
  - Level of education of citizens

Analysis

- Unemployment rates around the welfare office
- Unemployment rates for each town and city
- Unemployment versus average education level for each town and city
- Unemployment versus occupation

Insight

- Areas of high unemployment with no welfare office
- Areas of low unemployment with a welfare office
- Distribution of principal economic status, occupation, social class and highest level of education completed for each town and city

Principal Economic Status (Population aged 15 years and over)

Deloitte Analytics. — Education & Occupation Da
Unemployment v Average Education Level
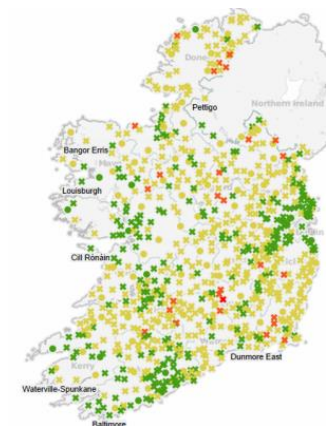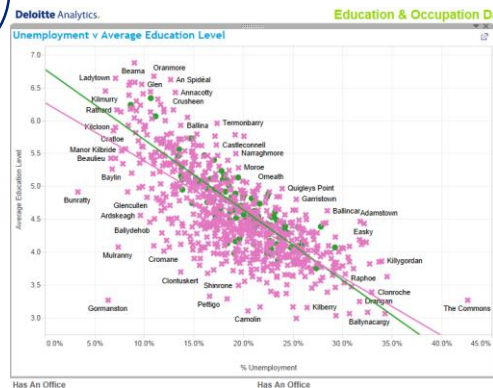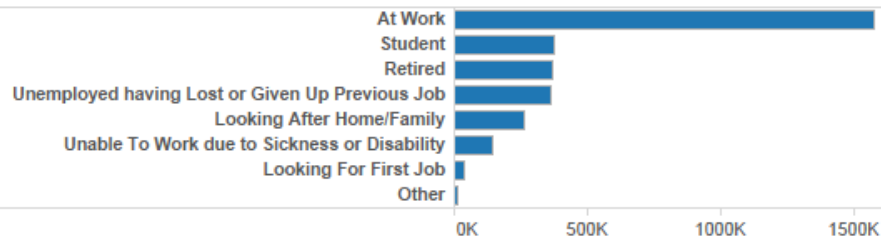
What does analytics mean for the insurance sector?

What are the opportunities for actuaries in analytics?

Challenges Facing the Insurance Industry

- External Business Pressures
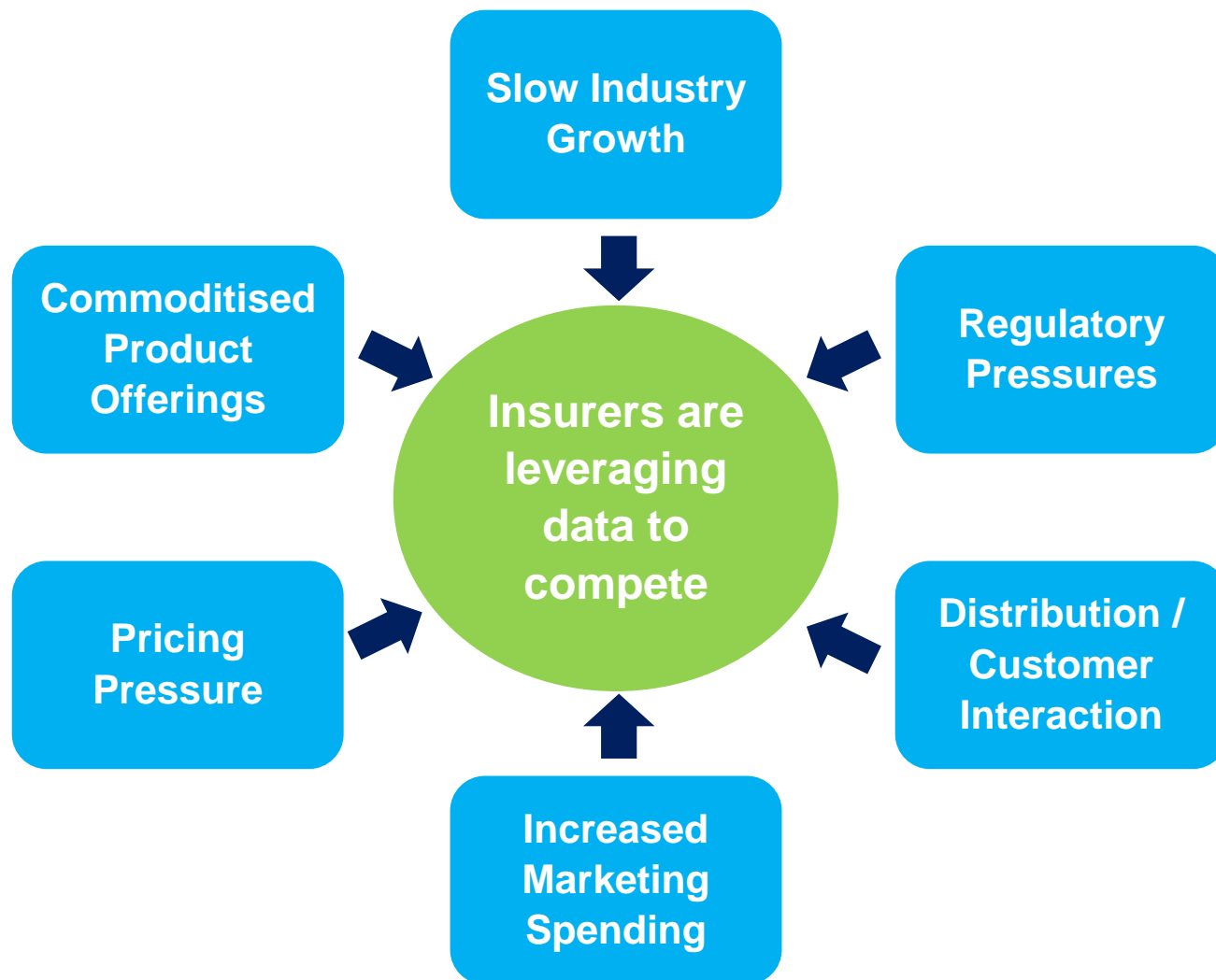
Challenges Facing the Insurance Industry
- Internal Business Challenges

- Multiple versions of metrics regarding customers, policies, claims

- Multiple owners of data across the enterprise

- Legacy or siloed systems that are not integrated

- A partial view of the "real" profitability or state of the organization

- Incomplete information available for decision making

- Lack of consistent fact based decision making, due to incomplete information

Observed Trends

- Organisations are looking for ways to harness internal and external information to gain a competitive edge to facilitate timely and fact based tactical and strategic decisions – seeing data as a strategic asset

- Companies are shifting their operations from product-centric to customer-focused business models requiring more sophisticated information management and analytical capabilities

Barriers to Realising Data Benefits

- Many insurers fail to leverage the business value of data

Implementation Challenges

- Insurers maintain data in separate functional or product silos with multiple owners and no overriding Master Data Management or Enterprise Data Strategy.

- Multiple versions of customer information and disparate data sources need to converge to create a coherent single view of the customer.

- Implementing information management or analytics capabilities without fully understanding industry leading capabilities often leads to failure.

- Implementing data and analytical investment programmes without understanding what business challenge you are attempting to be solve.

At what stage is your organisation's data strategy?

12%
17%
36%
35%
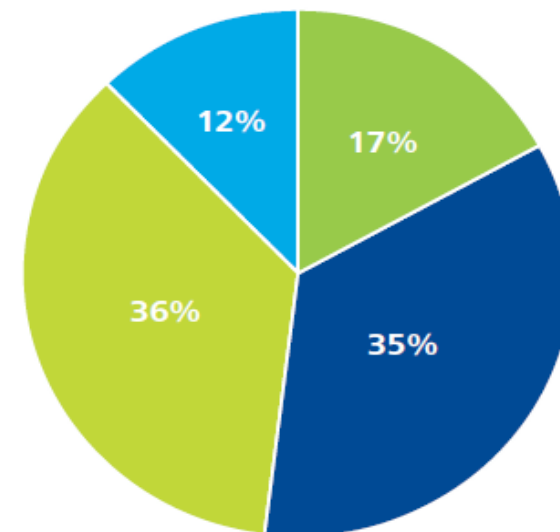
Complete
Currently under development
Currently a piecemeal approach
There is no information strategy

Source: Deloitte Survey

Analytics impacts the entire insurance value chain
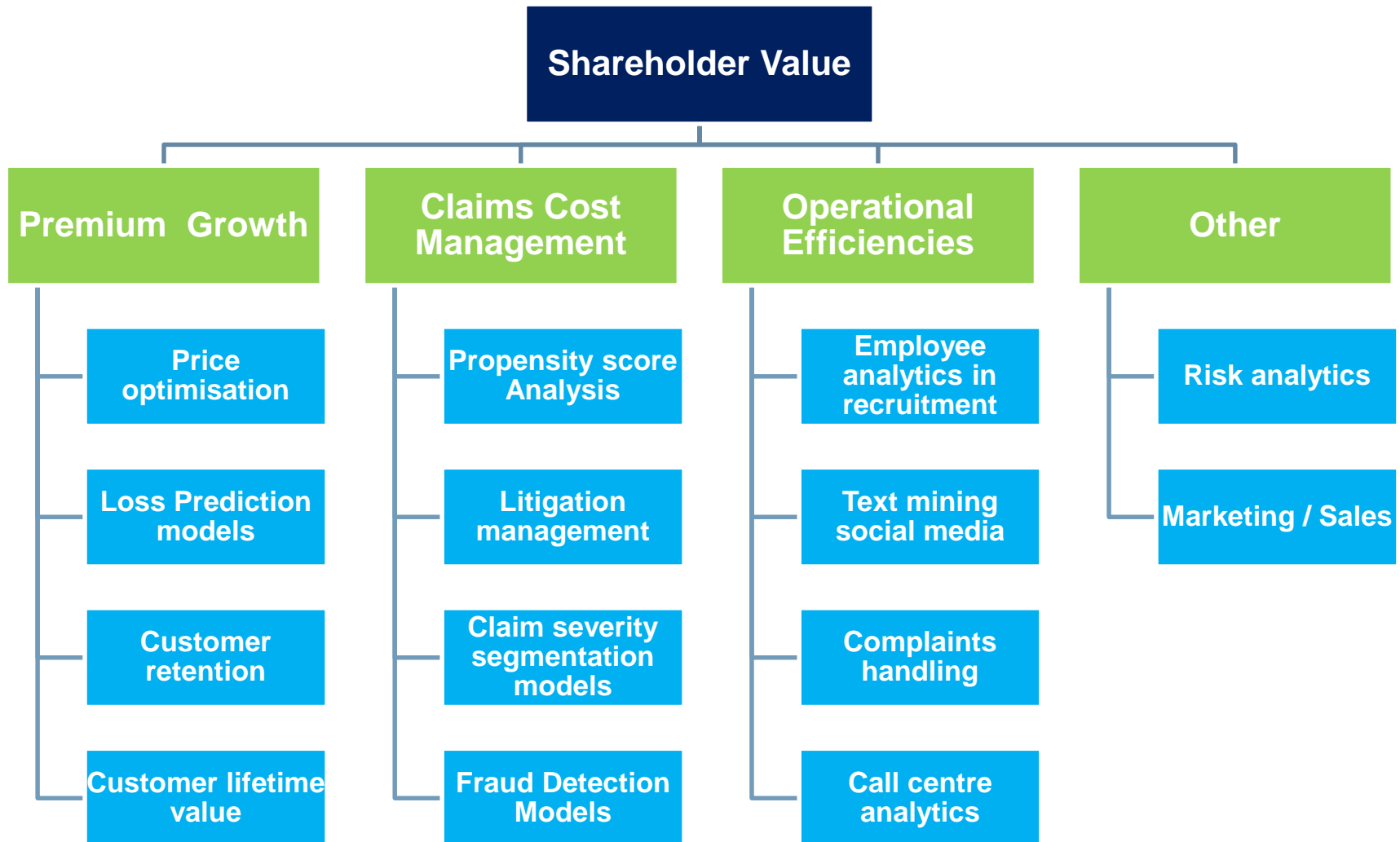
Implementing Analytics

5 Adopt in operational processes

4 Implement pilots and prototypes

3 Build data and analytical competencies

2 Assess current capabilities

1 Strategic questions

Opportunities for Actuaries

Professionalism

Data

Actuary

Strategic / Business Insight
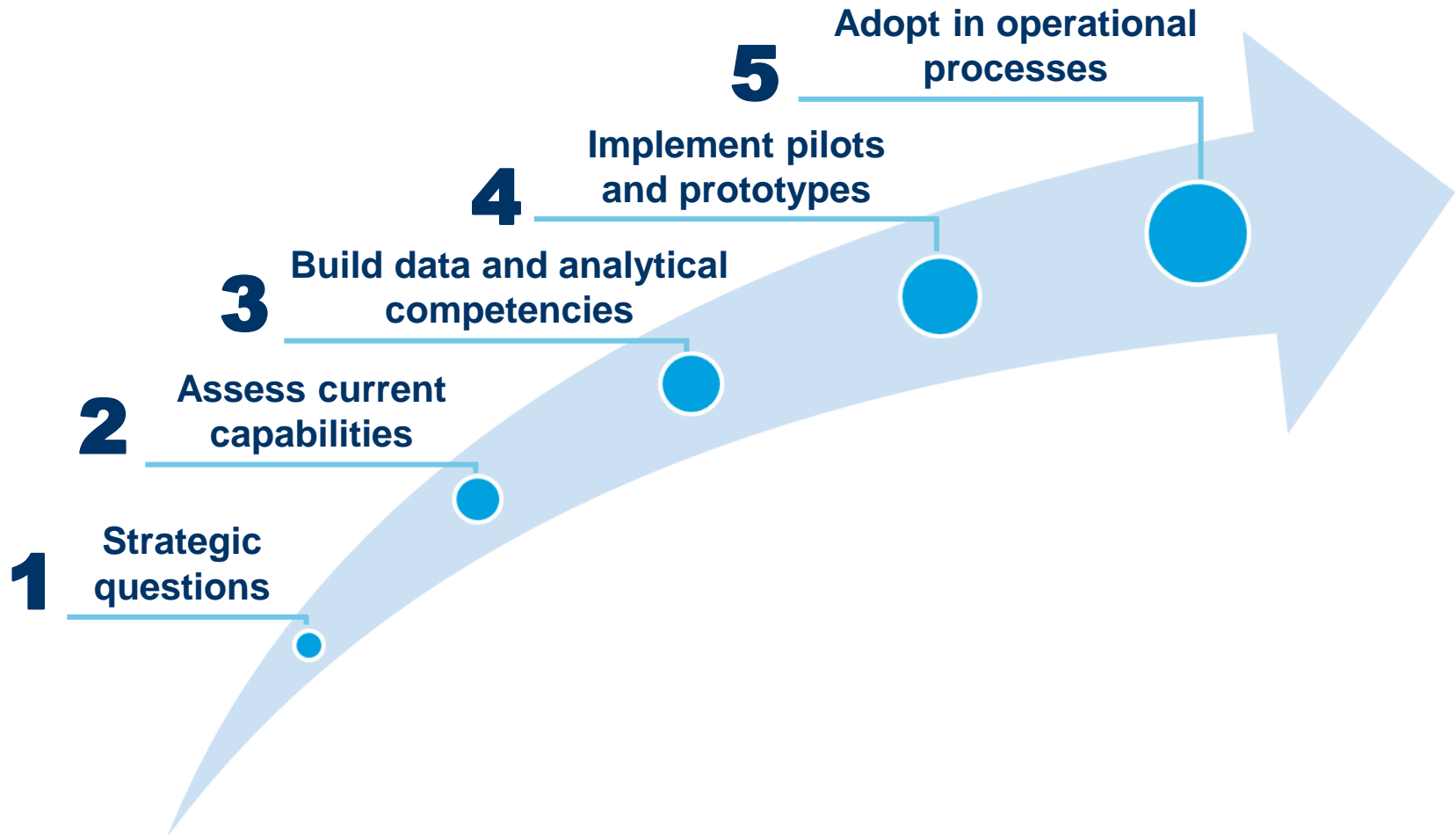
Modelling / Analysis
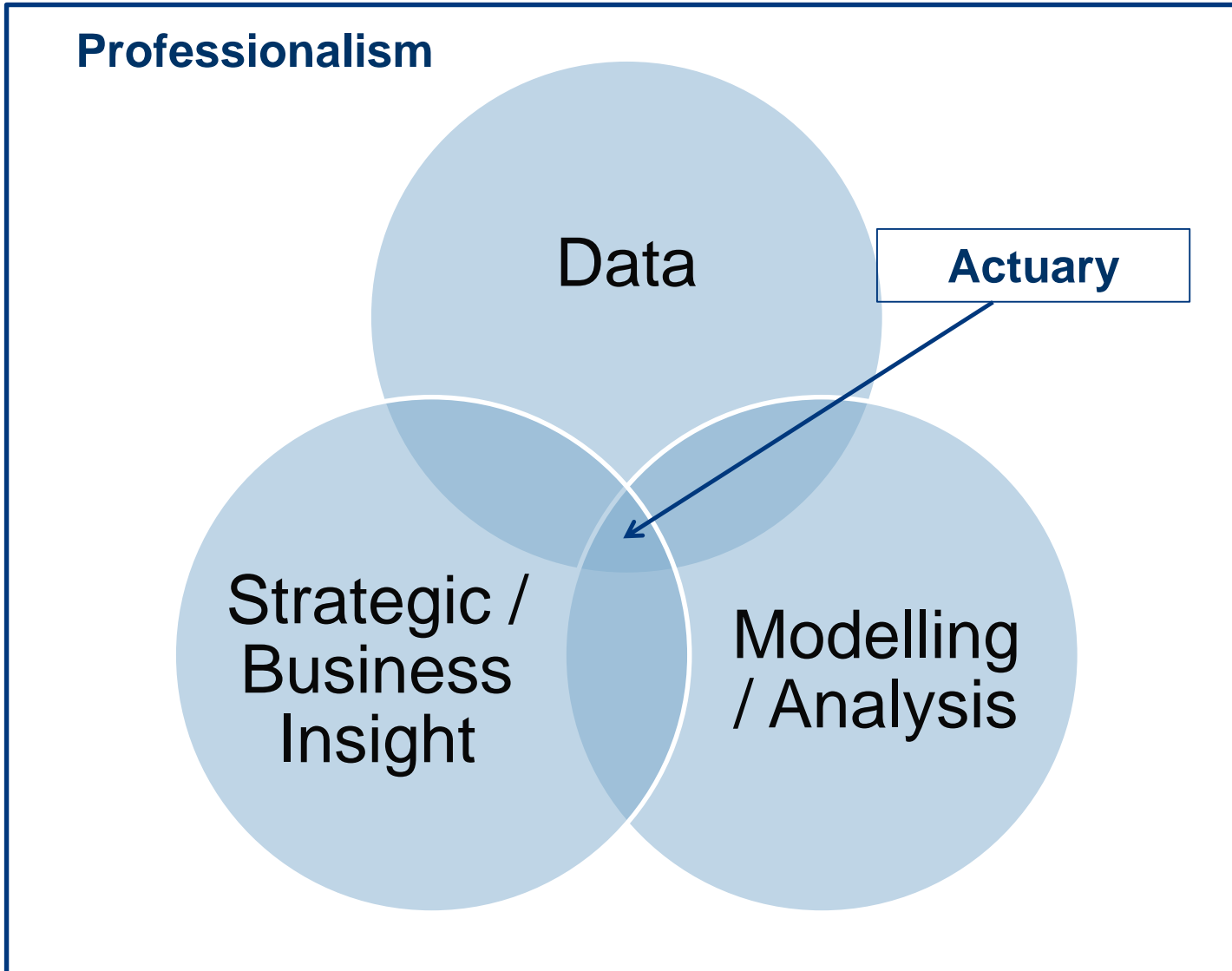
Examples of analytics applied to the insurance sector

Retention Analytics

- Approach

1 Identify required data to generate insights

2 Combine data into a single view of the customer

3 Create customer segments

4 Build propensity model

5 Calculate customer value metric

6 Prioritise customers for inclusion in the initiative based on probability and value

7 Design benefit based on customer segment requirements

8 Integrate customer contact list into campaign management tools

Fraud Detection

- Companies are increasingly turning to data analysis and predictive analytics for fraud mitigation

- Data analytical techniques can be utilised to help detect fraud at both point of sale and point of claim (including suppliers)

- Techniques can aid identification of known fraudsters at point of application

- Identification of applications whose features in combination are indicative of higher fraud likelihood i.e. a relative probabilistic finding

- Given a claim has happened, what is the probability that it is fraudulent?

Fraud Detection

- Data preparation - building a suitable database for modelling

- Applying data analytical and modelling techniques to establish statistically verifiable fraud signals and fraud scoring engines - requires access to skilled modelling competencies

- Apply developed models in operational processes – for example establishing referral rules in underwriting and claims processes

- Establish measures to monitor effectiveness

Telematics

1 Data and Data Collection

2 Analytics and Predictive Modeling

How do I capture data? How much data do I need?

How do I use this data to drive credible risk selection and pricing?

Do I want to use telematics data for more than pricing?

What can I offer my policyholders to increase engagement and loyalty?
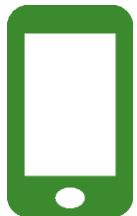
4 Customer Engagement

3 Leveraging the Data

What is process mining?

The discovery of process models based on available event log data

What do you need for process mining?

- ID track to track the process

- The Activity

- A time stamp to order and analyse performance
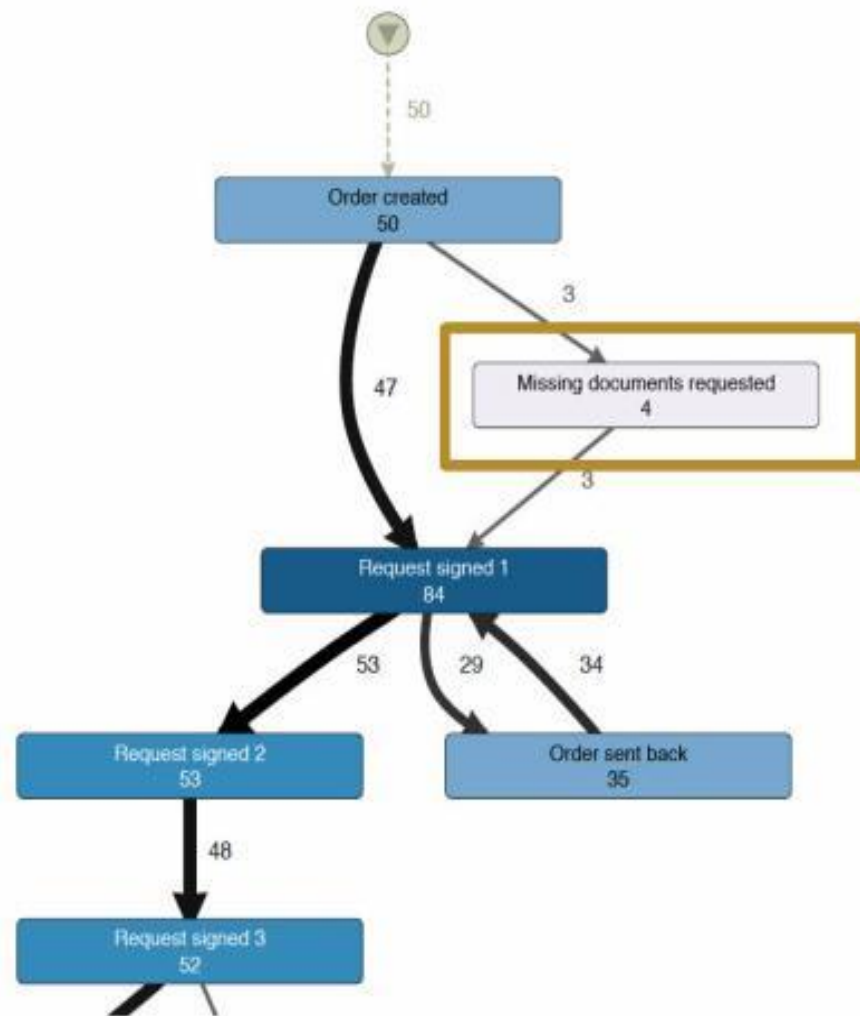
Core Idea:
Use data to
automatically discover
and deeply analyze a
Business Process

What is sentiment analysis?

Measuring a persons attitude towards a brand or product.

How does it work?

Through a set of methods, sentiment analysis classifies the polarity (positive or negative) of comments or statements which determine whether sentiment around a topic is positive negative or neutral.
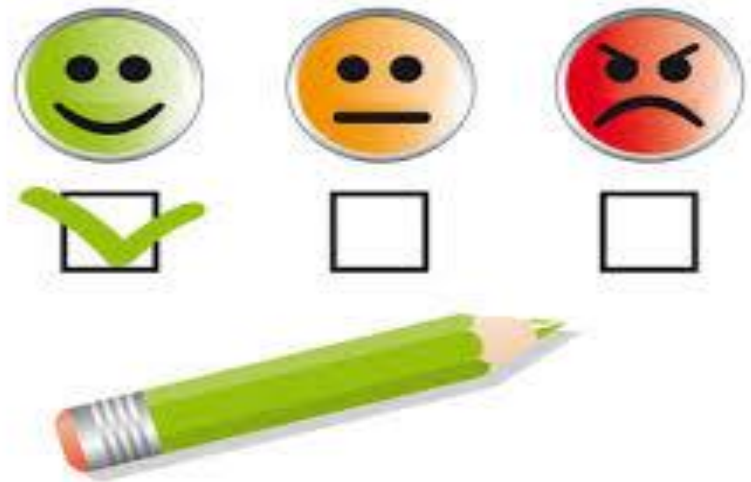
Sentiment Analysis – An Example

Visualisation

Dashboards
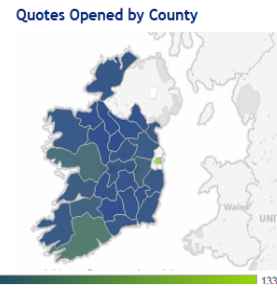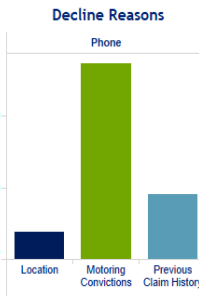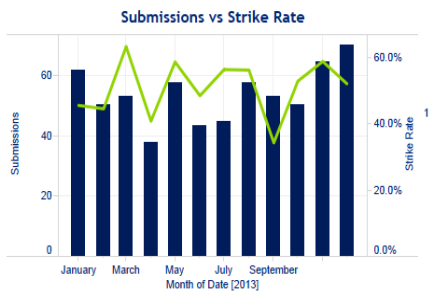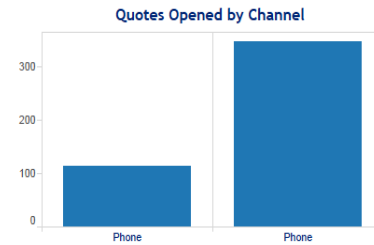
Geospatial analysis
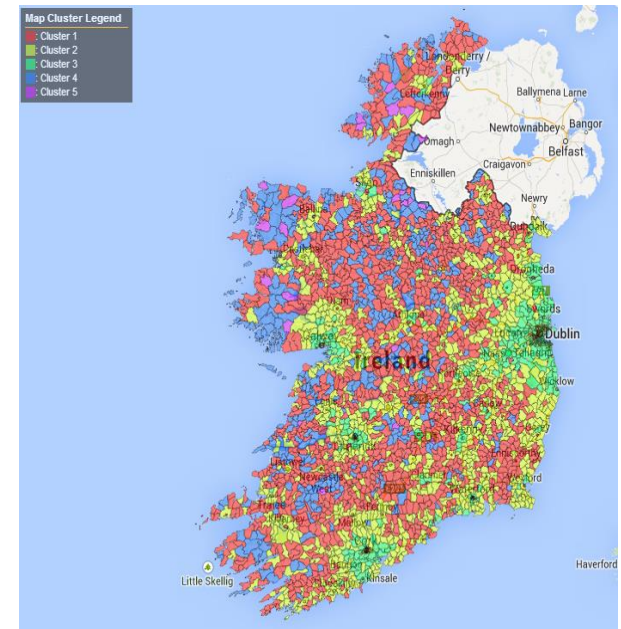
Tools used in Data Analytics

Example of Tools used

Open Source

Commercial

Questions