



The Impact of Big Data in the Role of Future P&C Actuaries

Carl Lambert & Frédérick Guillot

CAS In Focus Seminar

October 27-28, 2016

“We always overestimate the change that will occur in the next 2 years and underestimate the change that will occur in the next 10 years.”

- Bill Gates

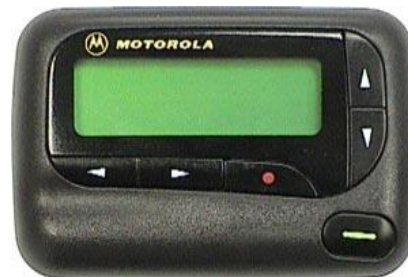


We always overestimate the change that will occur in the next 2 years and underestimate the change that will occur in the next 10 years.

- Bill Gates



COMPACT
disc
DIGITAL AUDIO



We are a Co-op Owned by 42 Co-op



Ajouter Edge
Benefits

We are the Canadian Analytic Carrier of the Year 2016



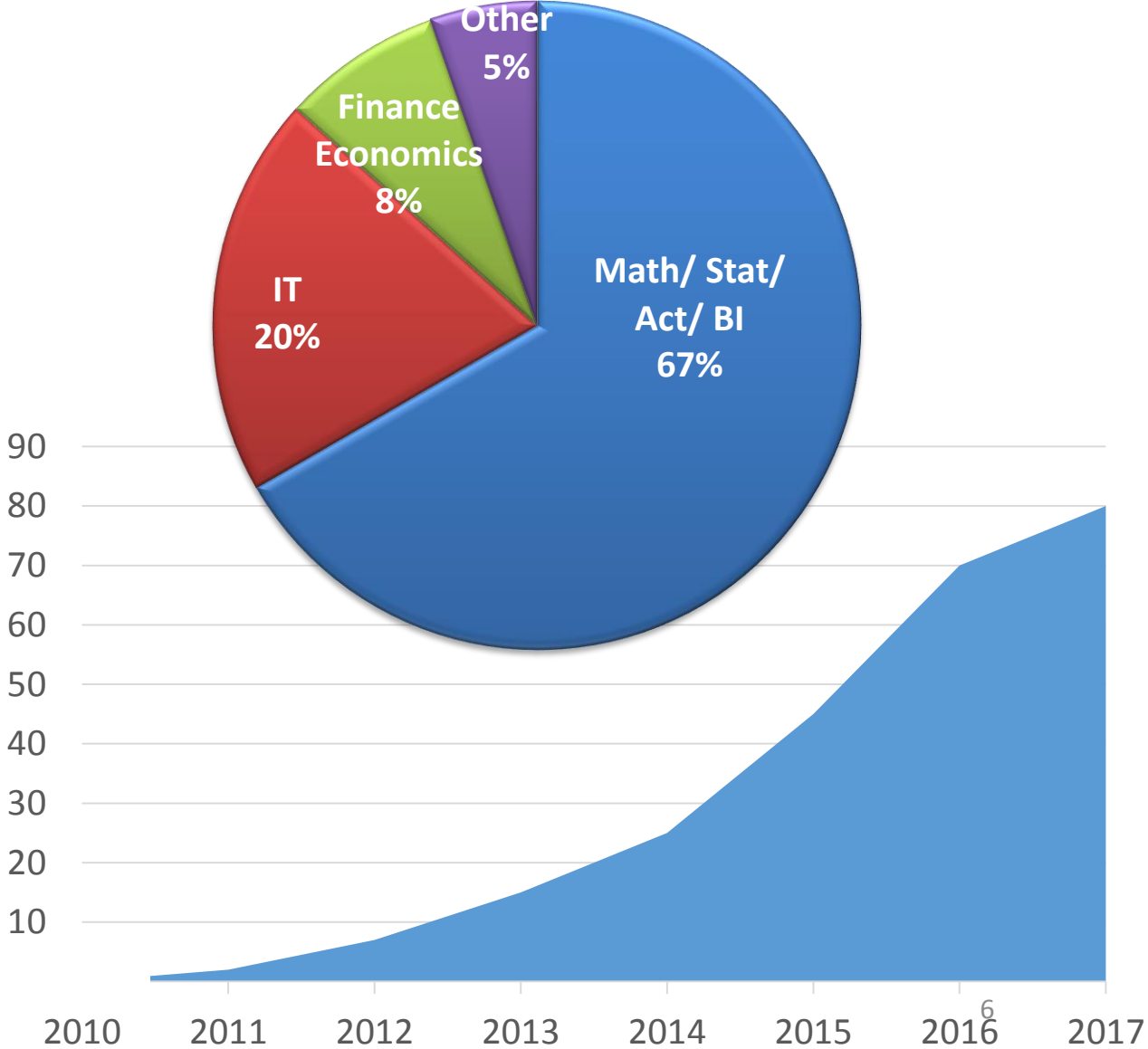
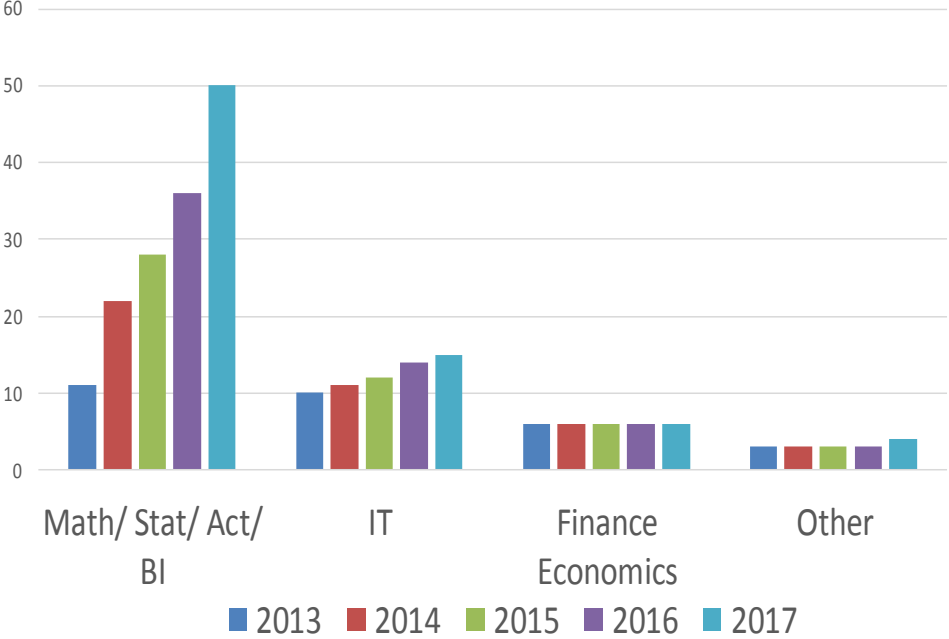
**Canadian analytics
carrier of the year 2016**

And more ...











JD Powers 2016
Best Canadian insurer, auto and home
Alberta, Ontario, Atlantic provinces

Business Intelligence at The Co-operators



TOP SECRET

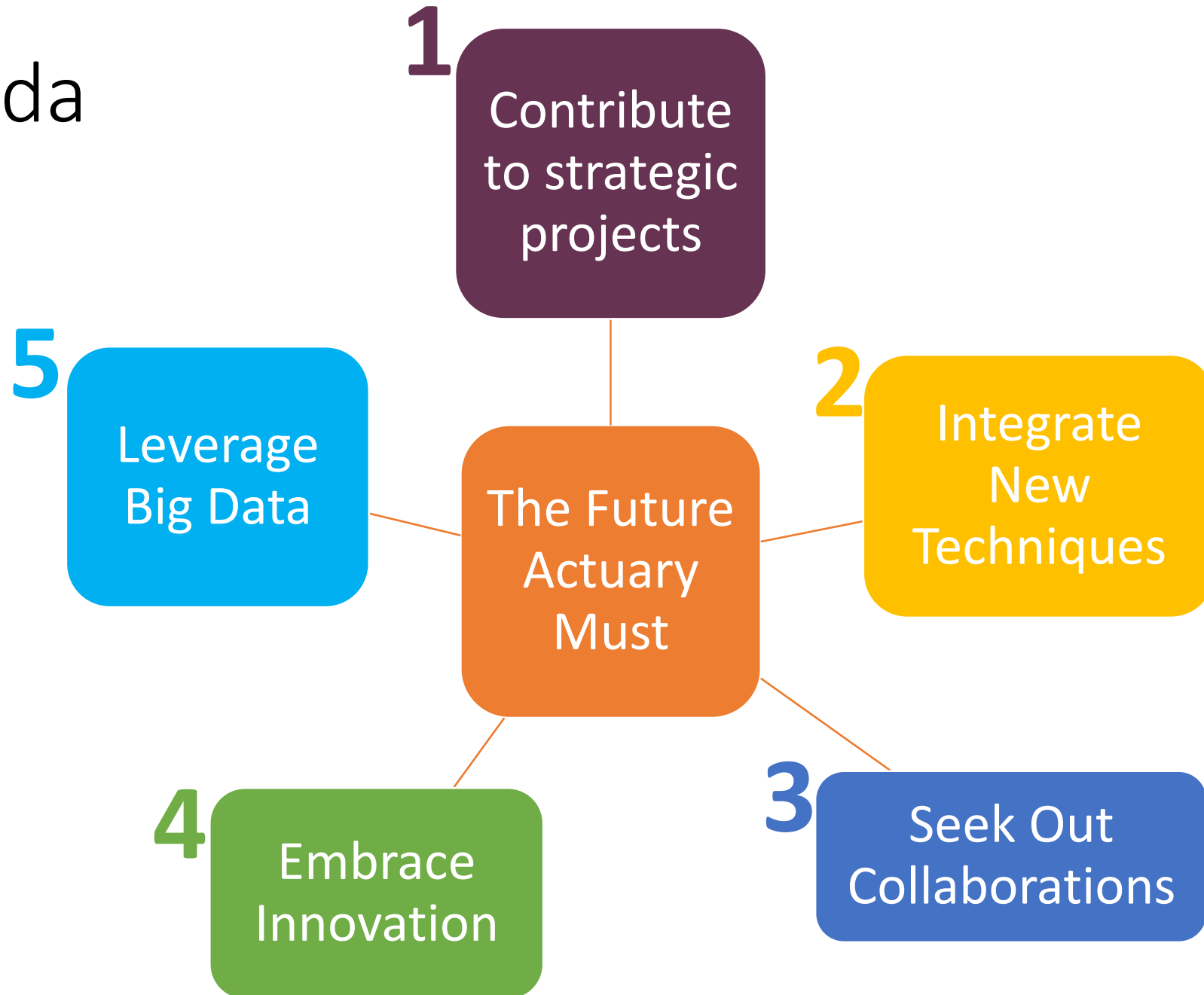
Not like this

			
			
1	2	3	4

Like this!

				
				
1	2	3	4	5

Agenda







Example 1.1 – Flood Insurance

1

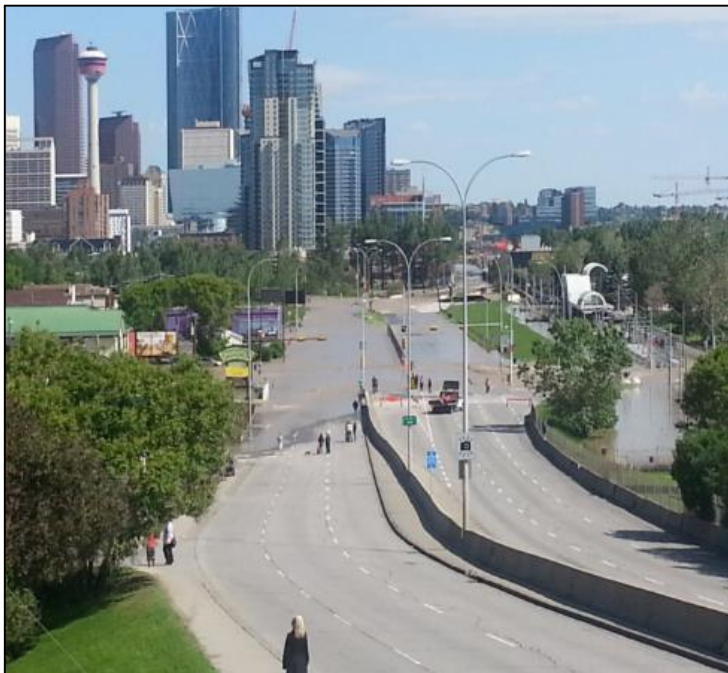
2013

=

2 floods in 2 weeks

+

Canada was the only G8 country without flood insurance



Challenge #1- Misunderstandings

Coverages



“70% of insureds believed they are covered for flood”

Confusion with Sewer Backup, DFAA, ...

Risk



“Insureds are underestimating their flood risk potential”

1 in 100/500 years events don't occur in a lifetime ...

1

Challenge #2- New & Complex Phenomenon

Causes

Fluvial, Pluvial, Coastal

Hydrology

Flow, depth, floodplains

Computations

Geocoding, Point-in-polygons



1 

Challenge #3- No Claim Experience Data

No Coverage



No internal data



Can you imagine a traditional
actuary without data ?

As useless as a skier without snow ...

**Fortunately, our
actuaries had creativity!**



Solution #1- Involve Academics

Find specialists + Research opportunities



1 

Solution #2- Involve Experts

Validate findings + Create partnerships



Solution #3- Use External Data

Elevation

- Resolution
- Canadawide (?)

Watercourses

- Geolocation
- Flow

Rainfalls

- Averages
- Observations

Floodplains


- Internal vs external model
- Canadawide (?)

Other Sources

- Soil type – Past events – Defenses
- Other partners



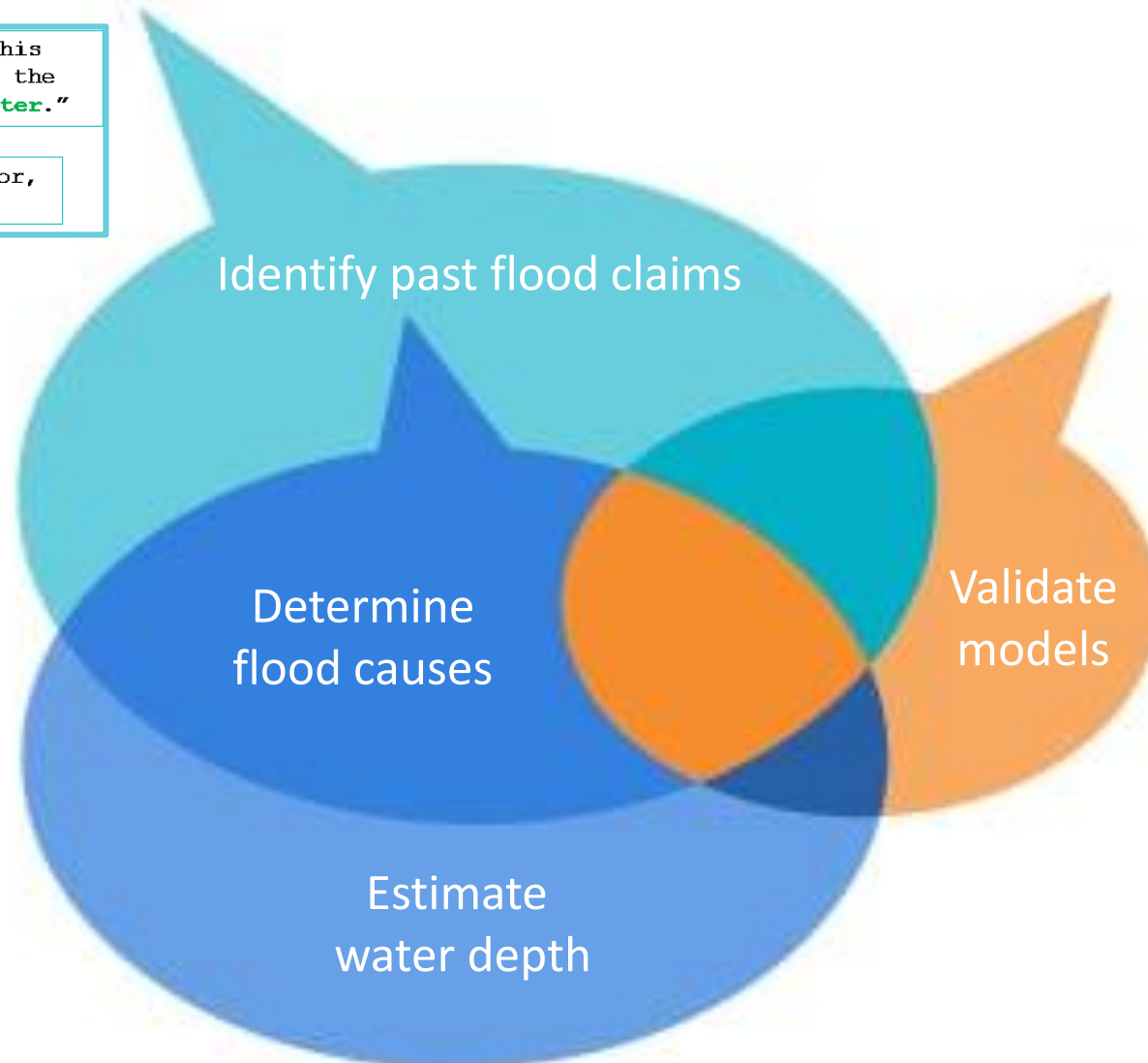
Solution #4- Use text mining



enterprise technology infrastructure operations
scorecards objectives
analytics text mining capitaliz
manage
metrics applications connection technical
solution stakeholder

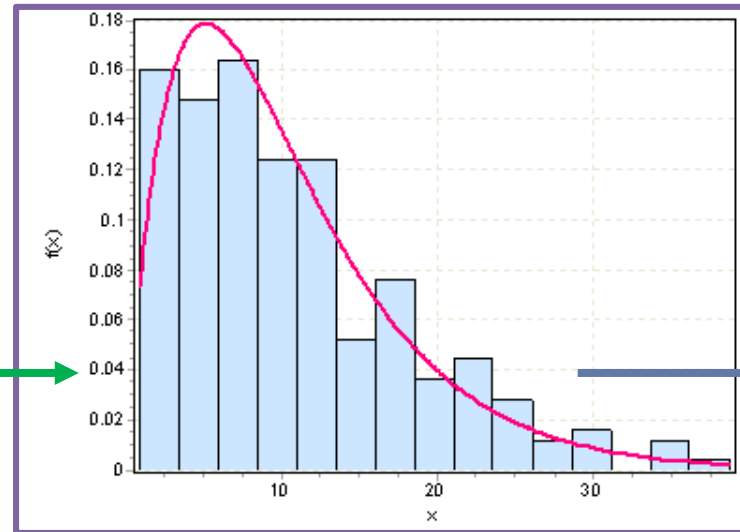
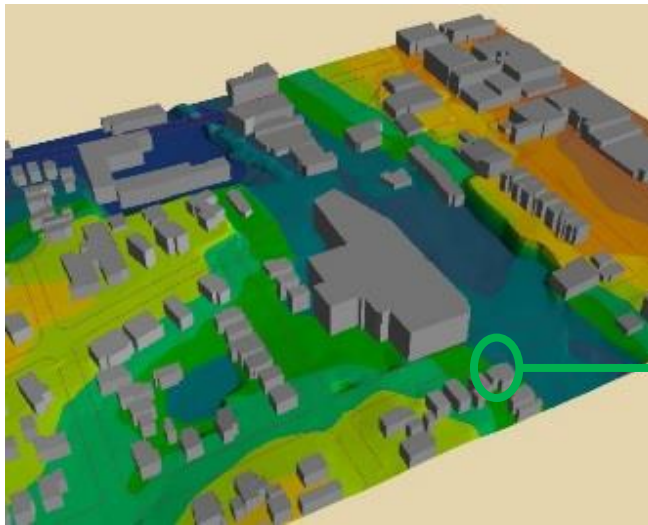
"Due to the **severe flooding** this time the water almost reached the furnace - approx **3 feet of water.**"

"**6 inches of water** & raised subfloor, was not sewage but **clear water.**"



Solution #5- Involve Hydrologists

Model water / Assess pure risk !



Risk \$

1

Solution #6- Involve Civil Engineers

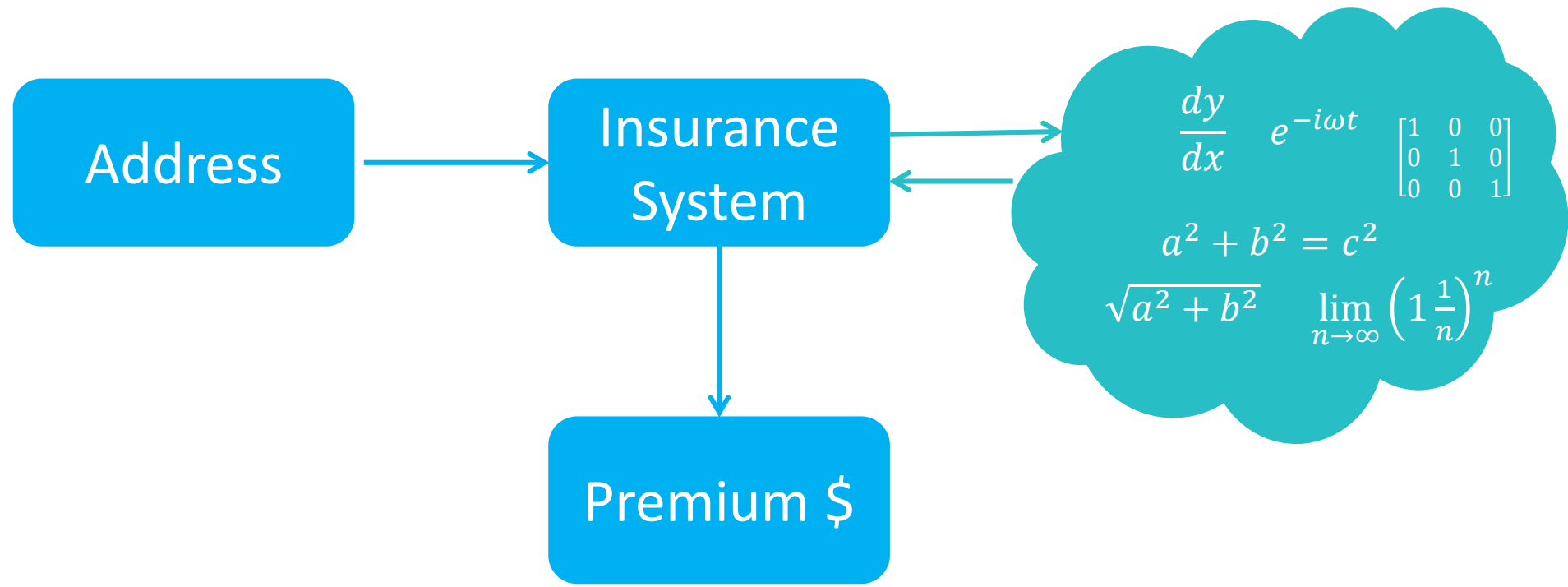
Assess performance of man-made flood defenses



1

Solution #7- Leverage Cloud Computing

Geocoding + Nearest Watercourse + Flood Zoning



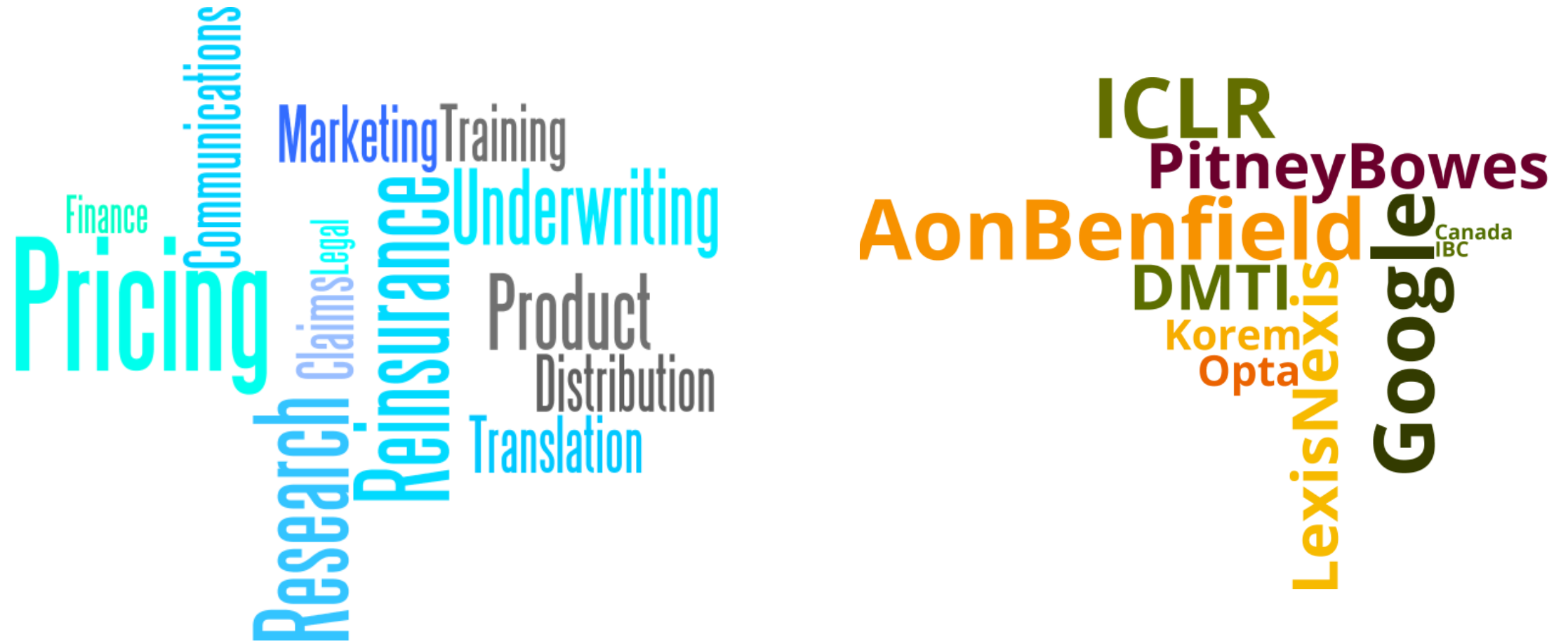
1

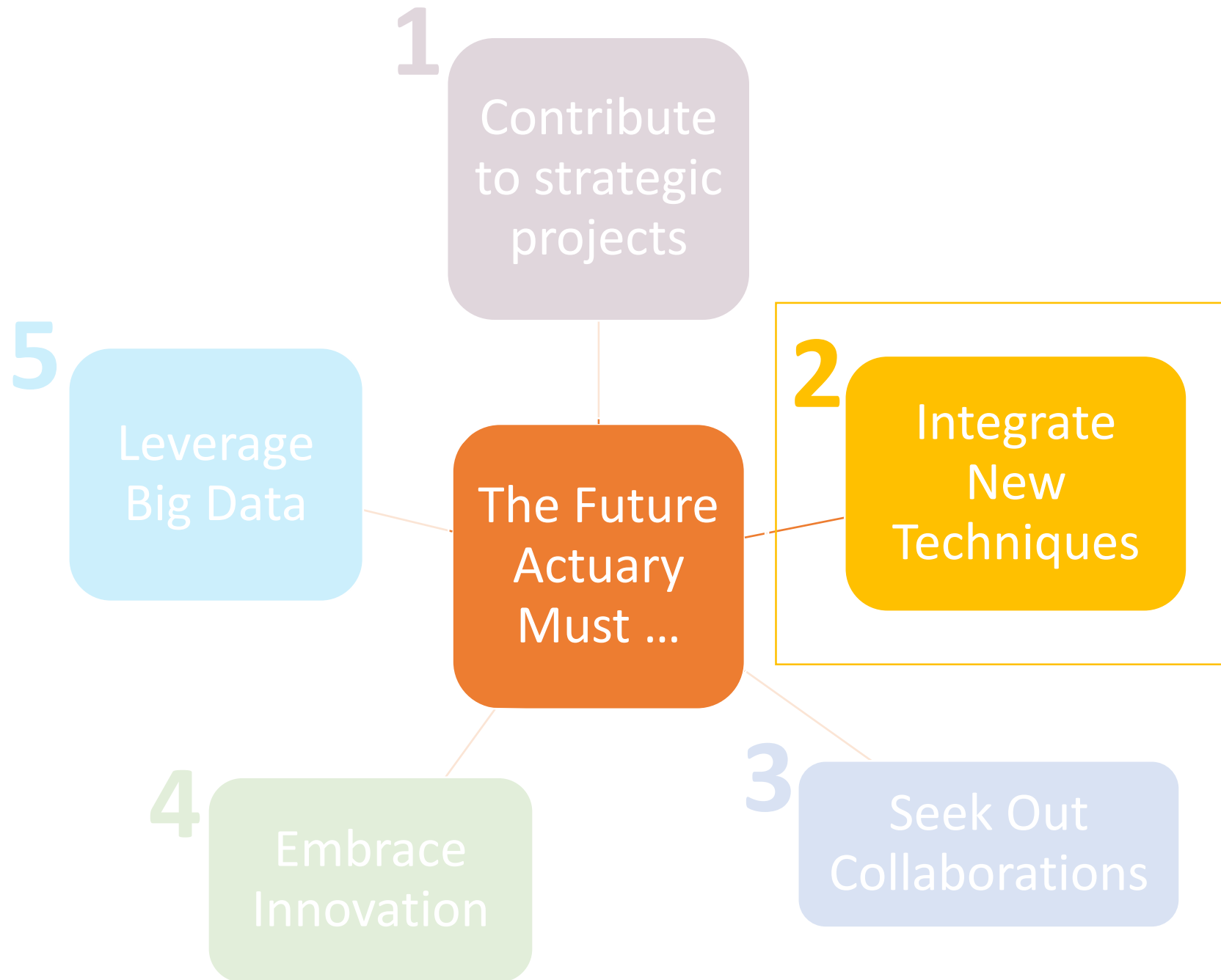
Solution #8- Actuaries & Data Scientists

Convert Risk Indicators into Premiums



Solution #9- Involve All Sectors



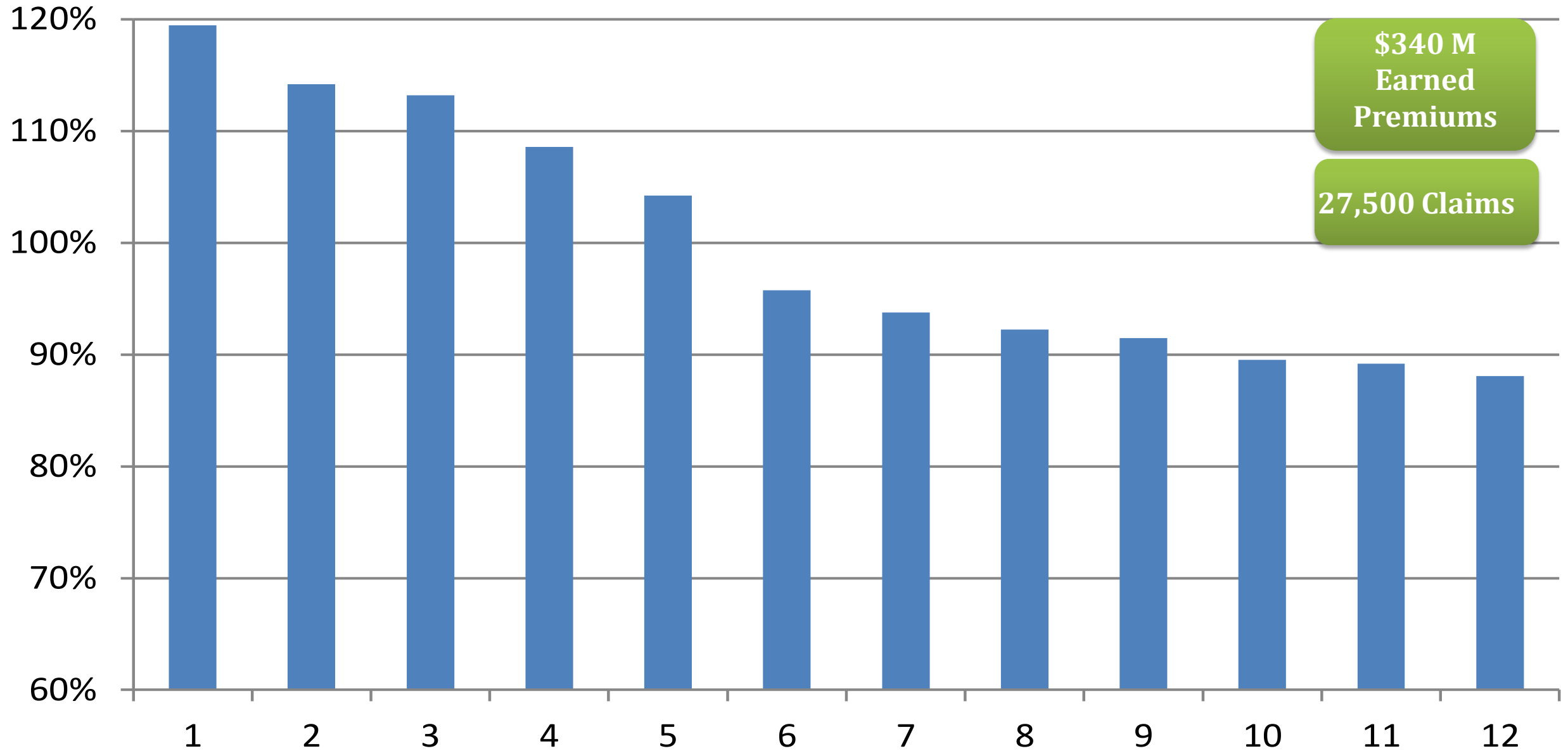




Example 2.1 – Statistical Techniques Credibility vs Overfit testing ?

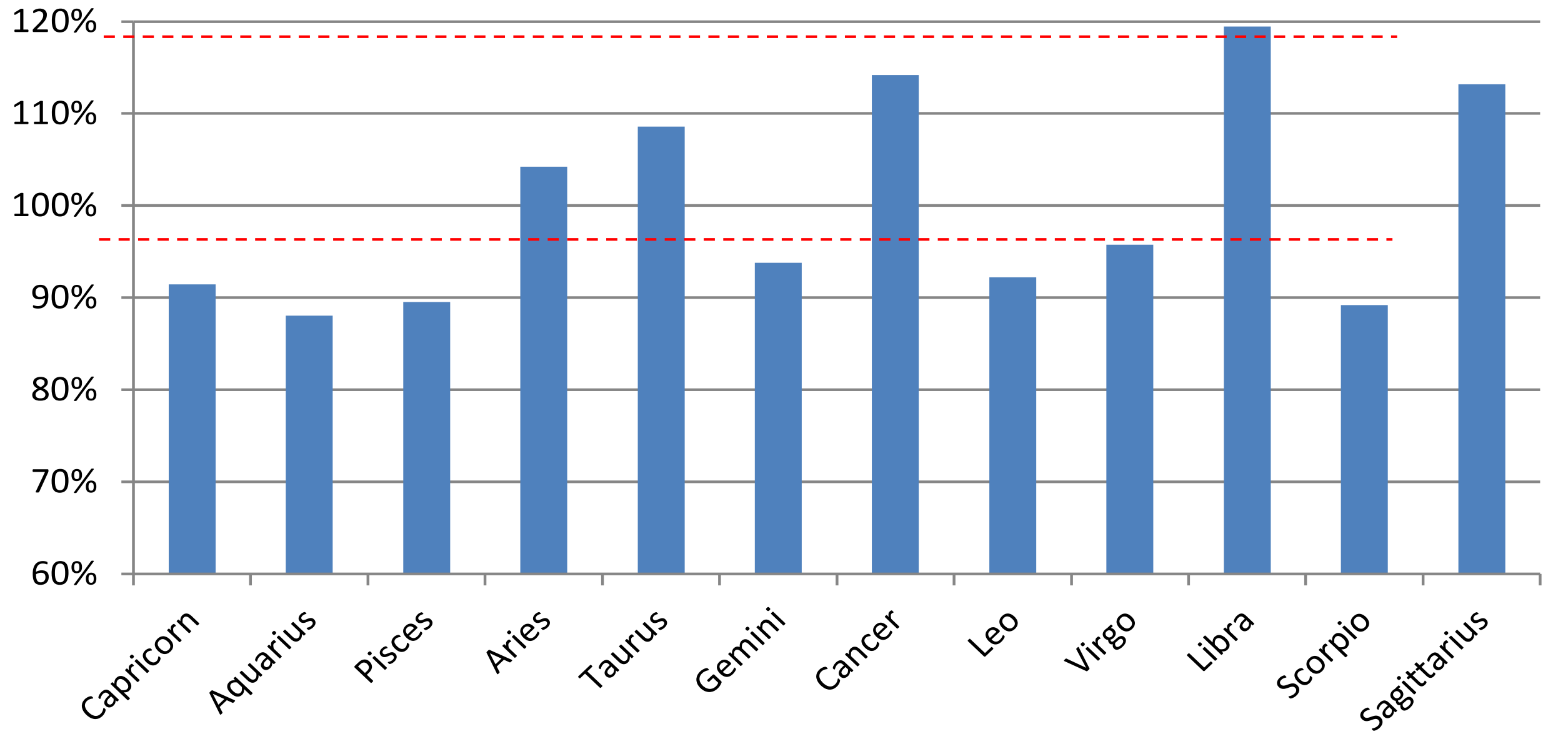


Auto Relative LR by "a new" rating factor





Auto Relative LR by "a new" rating factor





Example 2.2 – Statistical Techniques

List of Methods



Most Frequent

Visual Analytics

- Discovery, « insight », quick results
- 1st step in descriptive analysis

Decision tree

- Easy to understand, very visual.
- Often used as benchmark

Regression

- Forecast as a function of predictors
- Linear, exponential, GLM, logistic (1/0)

Unsupervised analysis

- Create segments
- No target variable

Machine Learning

- Complex models – neural networks, SVM, KNN
- Often used as “benchmark”



Frequent

Sentiment Analysis

- Get tone of a conversation or voice recording.
- E-mails, call centers, Facebook

Decision Forrest

- Numerous trees « vote » for the best decision tree
- Reduce overfitting

Text Mining

- Analyse unstructured information in a text
- « Read » notes, medical reports

Cluster analysis

- Create population segments
- Territories, marketing segmentation

Factor Analysis

- Reduce the number of predictors
- Big data, Unstructured data



Less Frequent

Time Series

- Analyse phenomena having cycles
- Econometrics, river flows,

Multinomial Regression

- Binary regression with many targets
- Analyse consumption baskets

Discriminant Analysis

- Analyse « strength » of predictors
- Analyse surveys, questionnaires

Anova

- Variance analysis
- Validate average differences

Survival analysis

- Truncated data
- Life expectation



Others

Basics Statistics

Distribution fitting

Multidimensional Scaling

Association rules

MARSplines

Non Parametric Statistics

C&R Trees

Classification trees

Canocical Analysis

Chad Analysis

Items Analysis

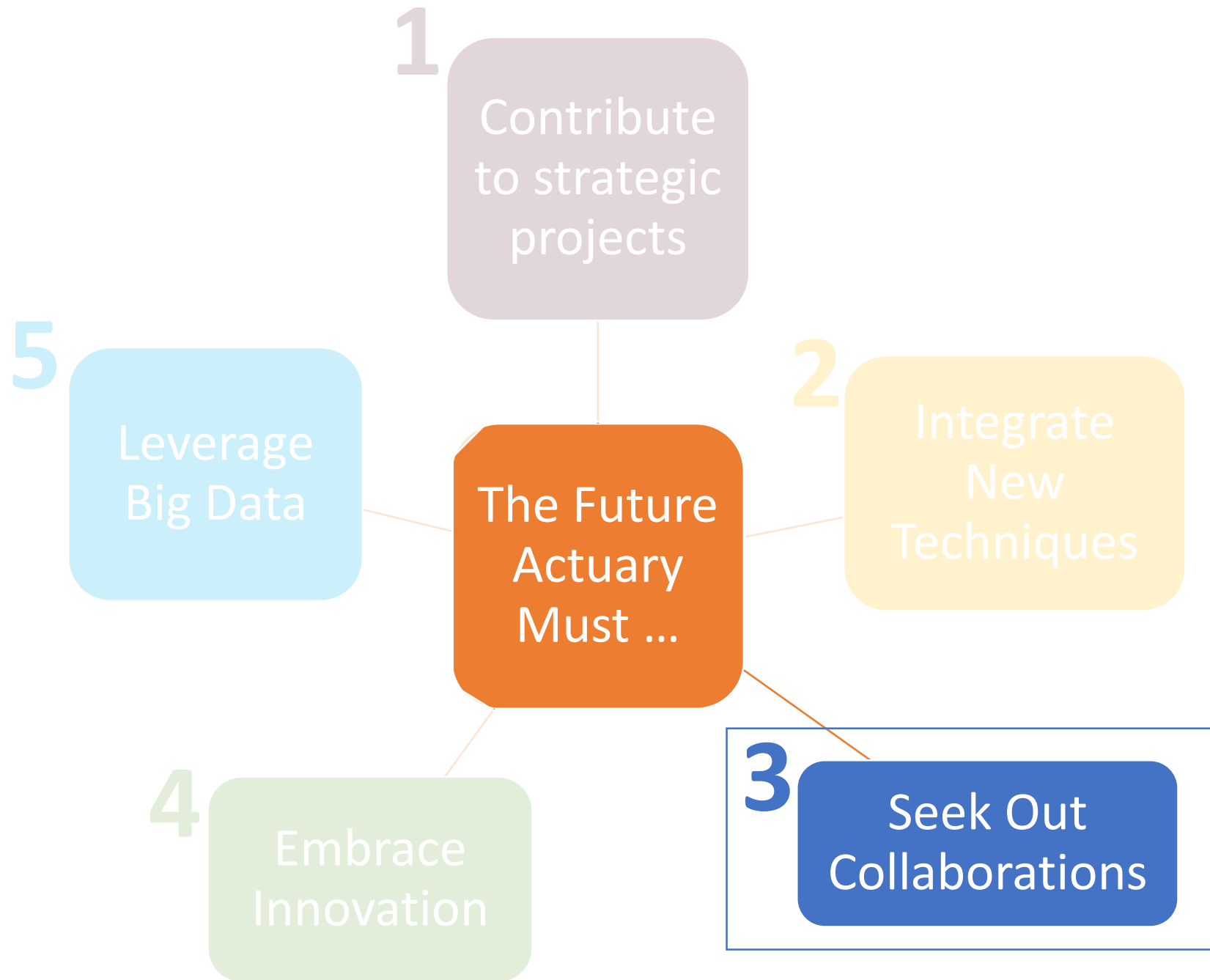
Correspondance Analysis

Experimental Design

Boosting Trees

SEPATH

Power Analysis



Internally



*Internal
Advanced Analytics
Show & Tell*

Analytic Day

Innovation training

*Emerging
Technology
Training*

Externally



CO-OPERATORS C'EST...

- 38 milliards \$ d'actifs sous administration
- Assurance auto, habitation, agricole et commerciale
- Assurance vie et collective & Services Financiers
- Un promoteur actif d'une société durable



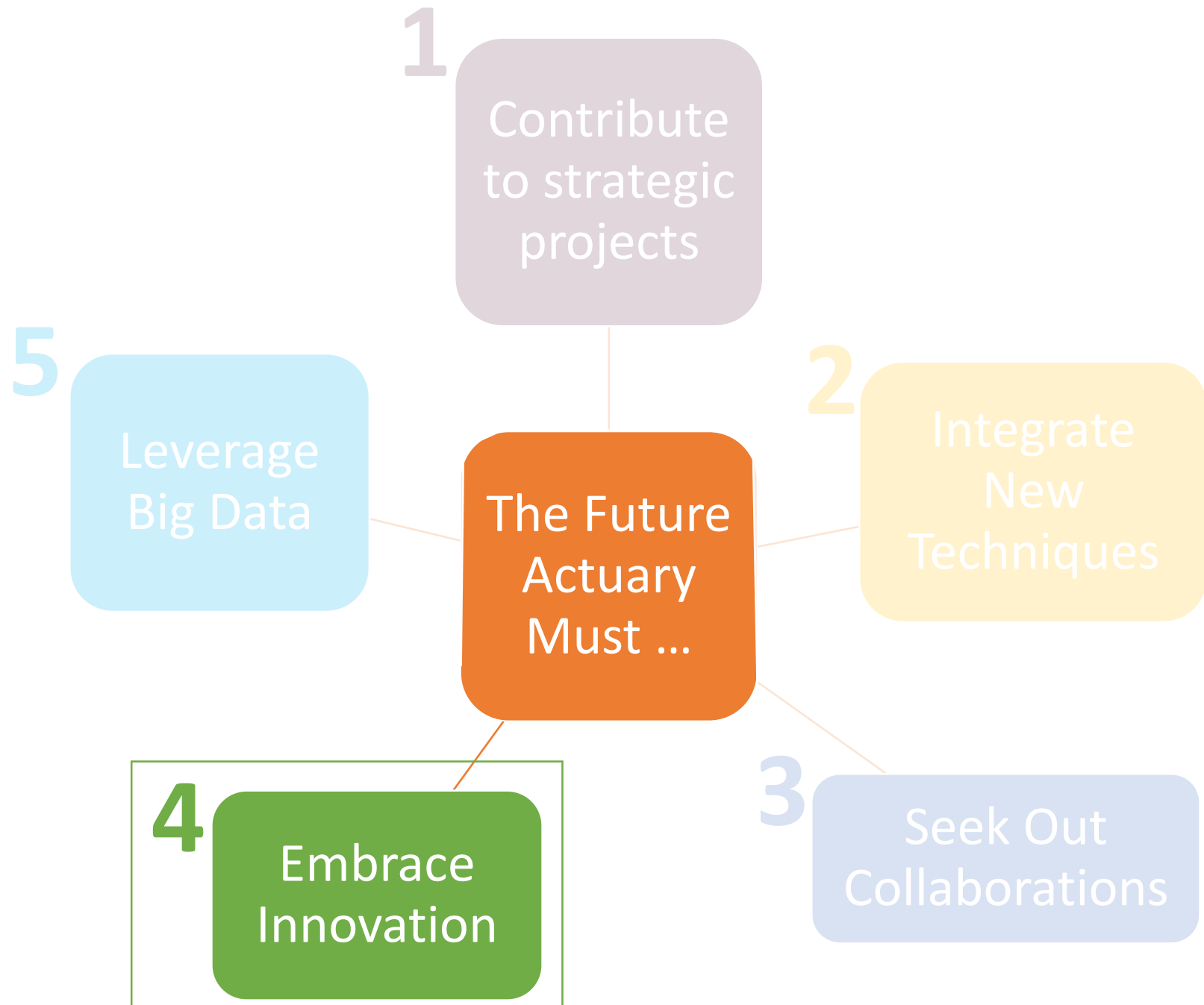
Connexion with Research & Academic Community

Participation at industry events

professional events

conferences

...



Example 4.1 - Emerging Techniques & Tools

Text Mining



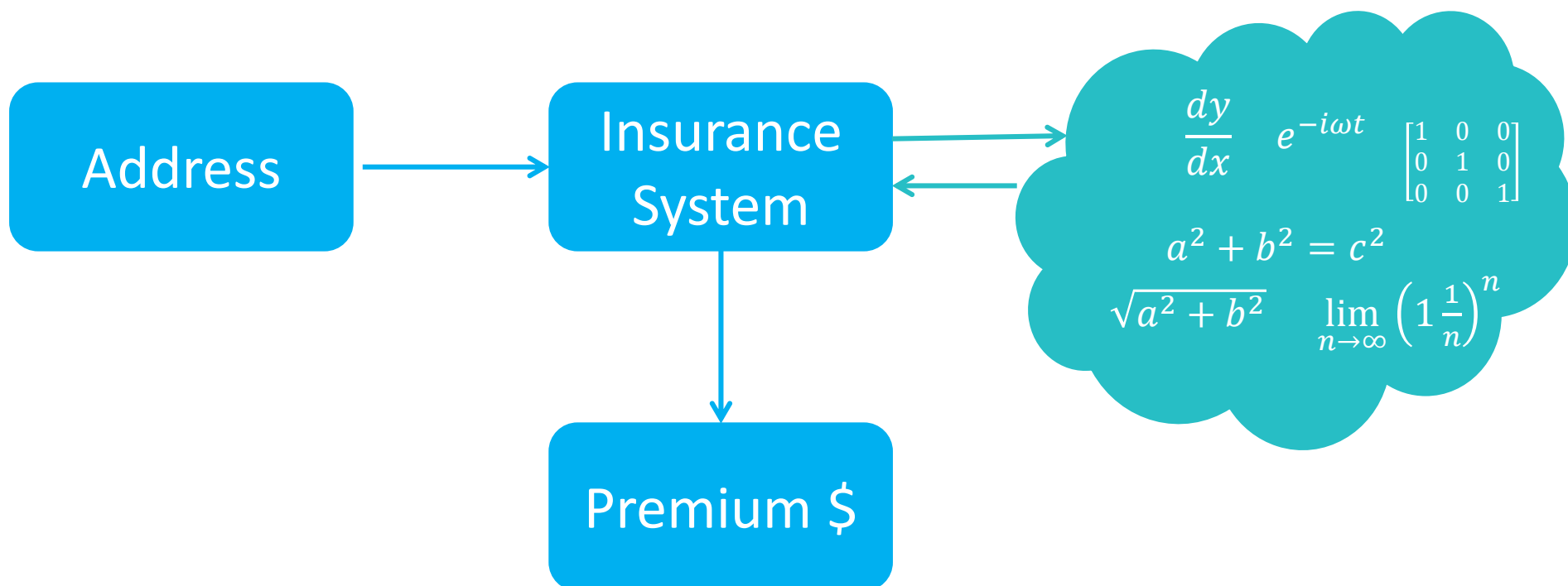
"... severe flooding ...
... 3 feet of water."

"6 inches of water ...
... clear water."

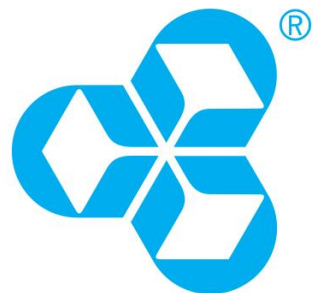
"... 3-4 feet of water ...
... in basement."

Example 4.2 – Emerging Technology

Cloud Geo-Computation



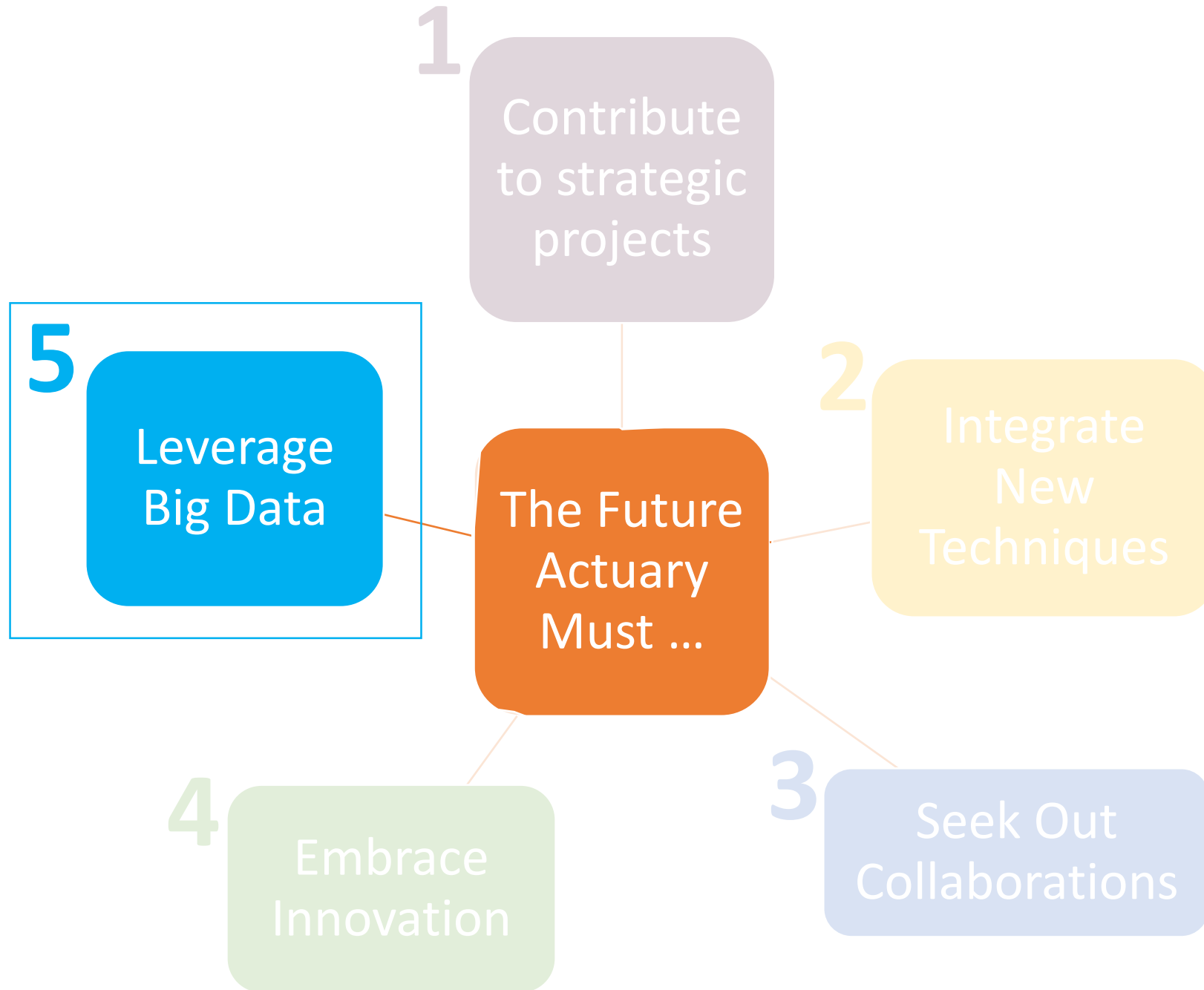
Example 4.3 – Data Scientist Designation



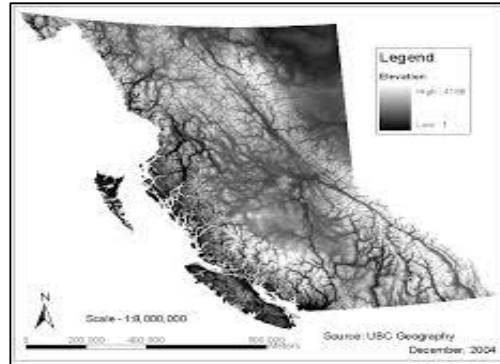
Data Scientist is not a
job description,
it is an expertise

it is an expertise
job description

- **Statistics & Mathematics**
- **Technology & Data**
- **Business sense & Communication**

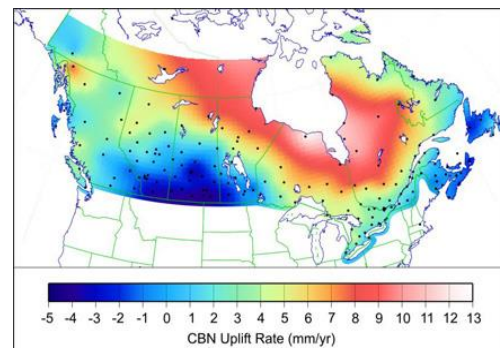
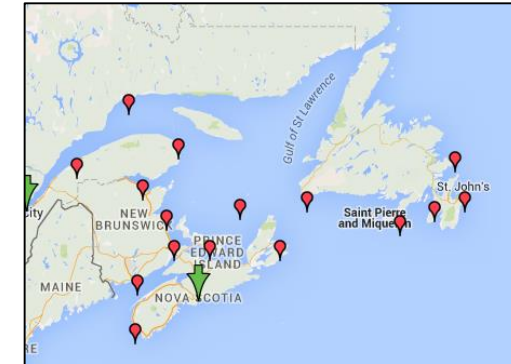


Example 5.1 – In Every Pricing Projects



Canadian Geospatial Databases

Local Tide Prediction Tables



Post-Glacial Rebound Estimates

And many more ...

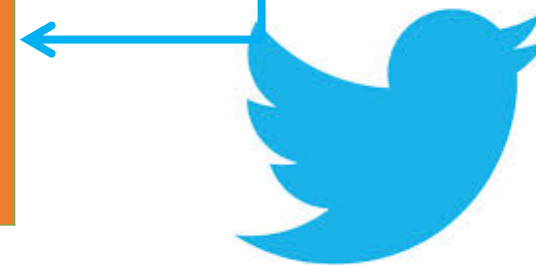
January 2016			February 2016		
Day	Time	Height (m)	Day	Time	Height (m)
1	12:44 AM	1.9	1	1:33 AM	1.9
1	5:59 AM	0.9	1	6:58 AM	1.0
1	1:13 PM	2.0	1	2:04 PM	1.8
1	6:35 PM	0.9	1	7:32 PM	0.9
2	1:35 AM	1.9	2	2:33 AM	1.9
2	6:54 AM	0.9	2	8:35 AM	1.0
2	2:06 PM	1.9	2	3:08 PM	1.7
2	7:39 PM	0.9	2	9:10 PM	0.9
3	2:30 AM	1.9	3	3:38 AM	1.9

Example 5.2 - Social Media

Citizen Williams ([@CitizenWilliams](#))

[12-06-21 08:29](#)

Someone needs to tell the Co-operators on the Bedford Highway that there are two Fs in affect - on both sides of their sign. :S [#Halifax](#)

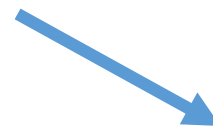


Social Media
analysis



Sign changed in
the afternoon

Example 5.3 - Social Media



Rona Kellogg
Kruger Agropur
Polaroid
Alcan SAQ
Rogers Loblaw
Desjardins
Banque Nationale
Raymond Chabot Grant Thornton

Agropur Nespresso
Métro Olymel
Desjardins

The future is already here,
it's just not very evenly distributed.

- William Gibson