

Are Actuaries New Data or Models Quality Professionals?

Pierre Chatelain, ADDACTIS France and SAF

About the speaker

- **Chatelain Pierre**– *Actuary PhD student, ADDACTIS France and Laboratory of Actuarial and Financial Sciences (SAF).*

Working household insurance :

1. Data quality
2. Climatic risks
3. Geo-localised pricing

SAF

Fifteen lecturers-researchers working on the sciences of management, applied mathematic and economics



addactis
THE RISKTECH FOR INSURANCE

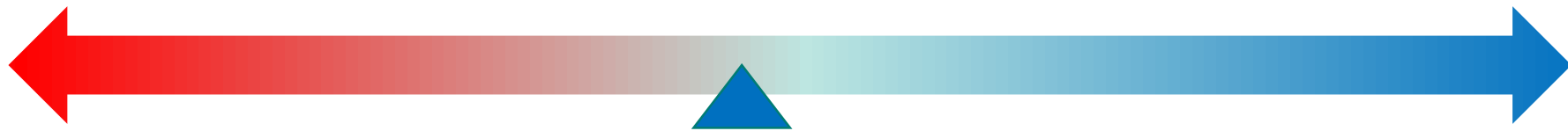
ADDACTIS Worldwide
Delivers our SOLUTIONS and ensures a responsive support to insurers and reinsurers all over the World

ADDACTIS France
Leads the BUSINESS R&D and drives our insurance business knowledge with high level Consulting

ADDACTIS Software
Leads our SOFTWARE R&D and provides first-class Software thanks to technical teams fully dedicated to Insurance Software Development

Data quality
professionnals

Model quality
professionnals



1. Evaluate the data quality

2. Understand the bias of
the data quality

3. Improve the data quality

4. Communicate about it

Are Actuaries
New Data

or

Models
Quality Professionals?

1. Choose data for modeling

2. Evaluate the impact of
the data quality on a model

3. Evaluate the model quality.

4. Communicate about it

For more
details see :

ROGOVA, Galina L. ; BOSSE, Eloi: Information quality in information fusion. In: *2010 13th International Conference on Information Fusion IEEE (Veranst.)*, 2010, S. 1–8

RAMAKRISHNAN, Raghu ; GEHRKE, Johannes: *Database management systems*. McGraw Hill, 2000

Data Base -
Premium, claims,
Individuals
Characteristics,
financial index, ...

Data quality before models quality

(Panko, R. R. 2009)
Figure 2-10

			Accuracy
Mechanical Action	Typing	Pointing	99.5%-99.8%
Computer Program	Formula		95%-98%
Spreadsheet	Program Module	Spreadsheet Module	0%

errors in Excel WorkSheet

IMPRECISION
UNCERTAINTY

More about Data Quality

Standard deviation

Fraud

representative

Censorship
life table
(Efron,1967)

natural
evolution

More about Model Quality

TIMELESS

COMPLETENESS

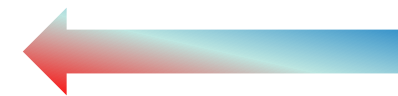
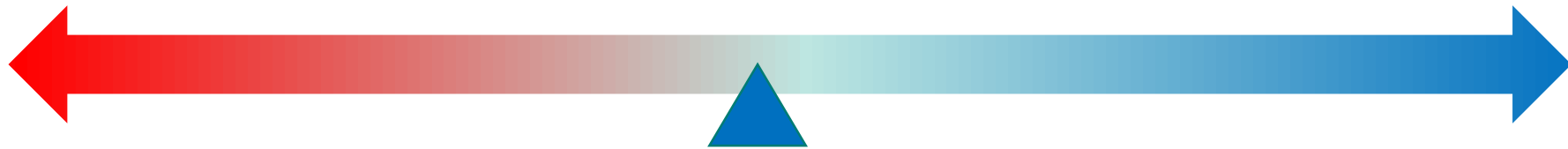
Usually

More about Model Quality

on the (Rubins, 1976) nomenclature

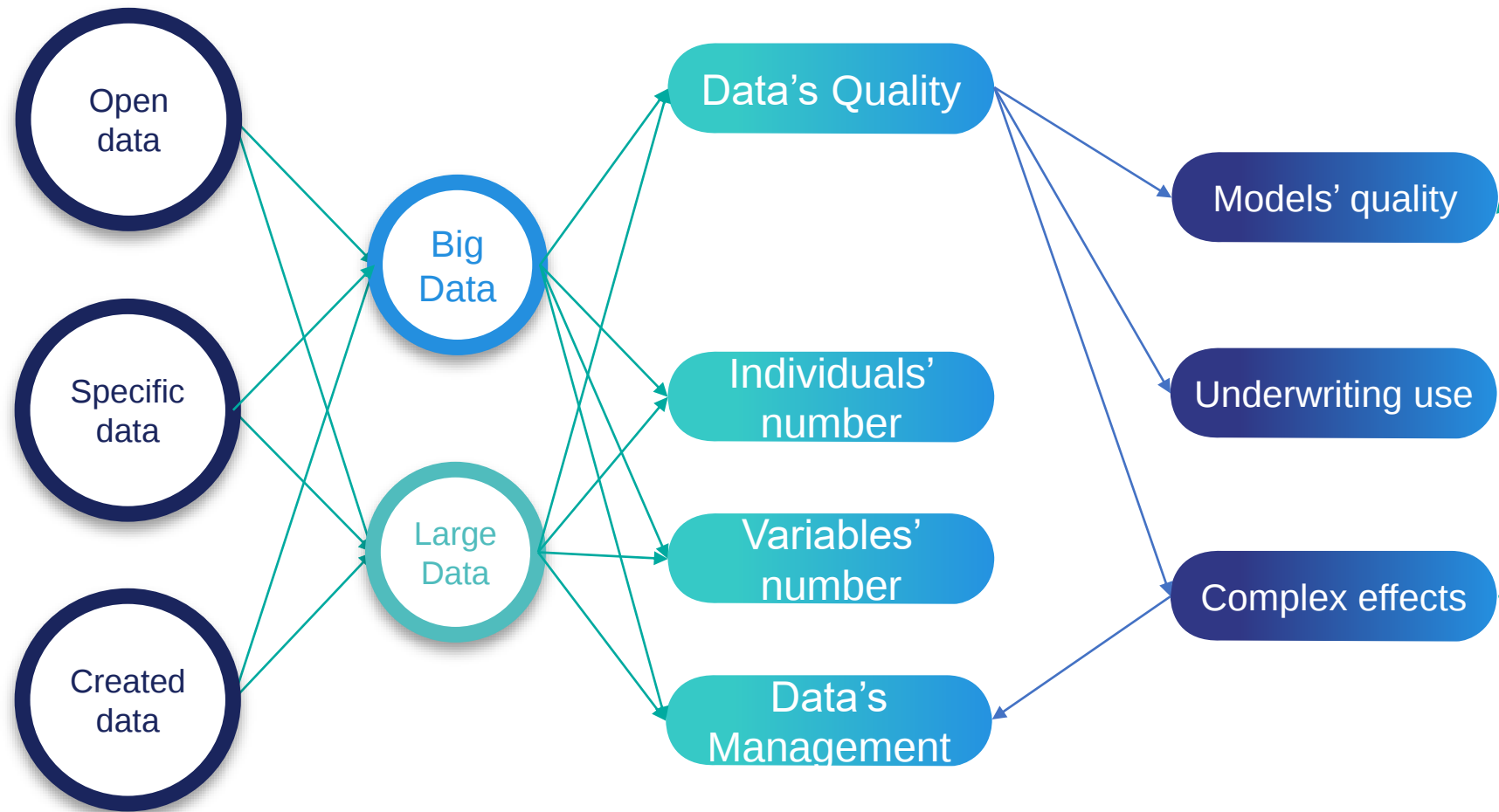
Data quality
professionnals

Model quality
professionnals



More and more data available

Quality and external Data



Often the impacts data's quality are **downplayed or underestimate**

How to integrate quality in models ?

Different impact of the quality on statistical methods

- Impact the model performance
 - Biais the estimate coefficient but not always downward
- Biais other estimates coefficient.
 - Biais the intercept.

Two set of observations of the same variable but of different quality should be considered as **two different variable**

Pricing and external Data

INTEREST

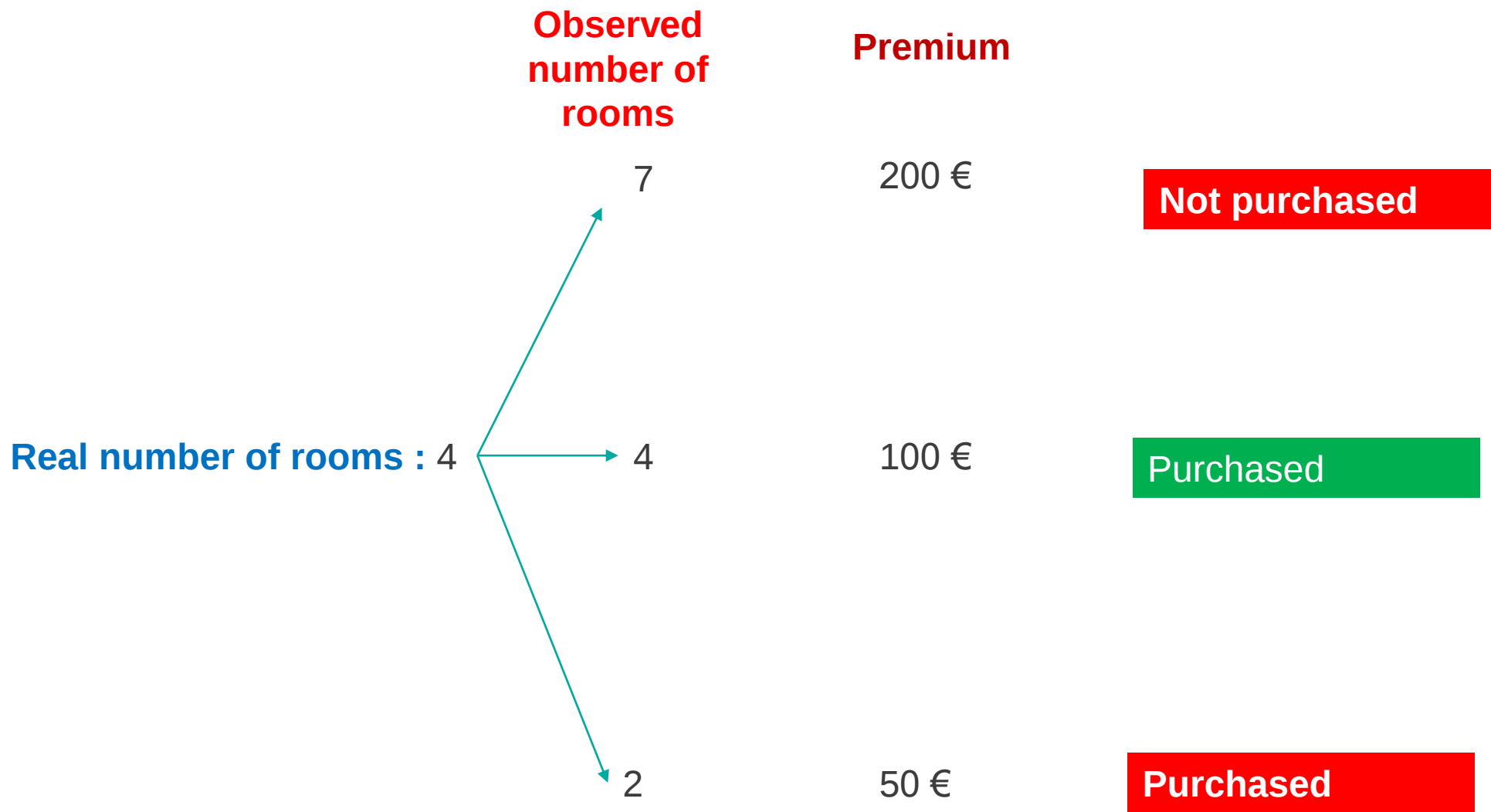
Why ?

1. Increase model performance
2. Decrease the number of questions
3. Create new products or features

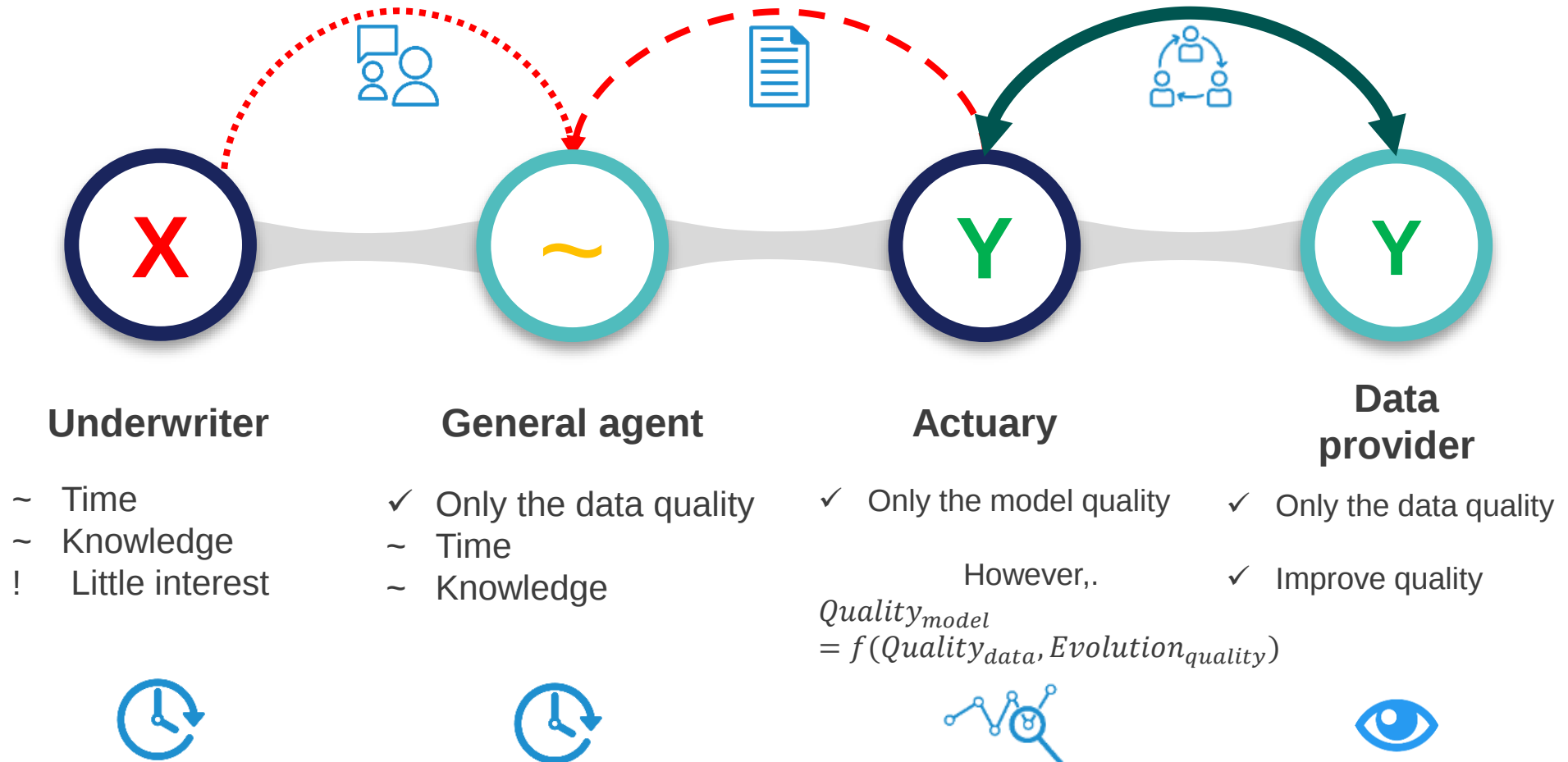
Impact
on a
global
scale



Impact on
a
individual
scale



Who must interact with quality ?



In the futur

Data quality
professionnals

Model quality
professionnals



Knowledgeable on **data quality** but
their main focus will be on models quality.

**Thank you very much for
your attention!**



Pierre CHATELAIN

ADDACTIS France

pierre.chatelain@addactis.fr



References

- Katrien Antonio and Jonas Crevecoeur (2020) [Claims reserving in non-life insurance: old and new adventures](#) One World Actuarial Research Seminar LRisk - KU Leuven and ASE - University of Amsterdam June 17, 2020
- Panko, Raymond R.. (2008) Thinking is bad: Implications of human error research for spreadsheet research and practice. Computer Science
- Breslow, N., & Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *The Annals of statistics*, 437-453.
- Efron, B. (1967, July). The two sample problem with censored data Vol. 4, No. University of California Press, Berkeley, CA, pp. 831-853.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3), 581-592.

