



**SOCIETY OF
ACTUARIES®**

SOA Big Data Seminar

13 Nov. 2018 | Jakarta, Indonesia

Session 1

Big Data Overview – Basic knowledge and perspective from insurance industry

Paul Setio Kartono, FSAI, ASA, MAAA
William Soetrisno, FSA

Big Data Overview

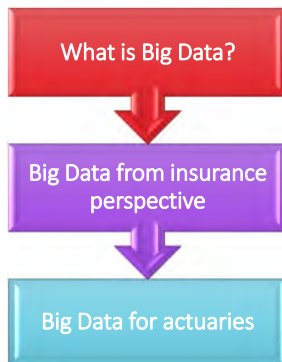
WILLIAM SOETRISNO

FSA, Pricing Officer of Manulife Indonesia

13 November 2018



Agenda



"After careful consideration of all 437 charts, graphs, and metrics, I've decided to throw up my hands, hit the liquor store, and get snookered. Who's with me?!"

What is Big Data?



Fun Facts

If we decided to download all of the data in the internet, put it into CDs (1 giga bytes). Stack those CDs up. How high is the stack?



- A • Jakarta – Singapore (1000 km)
- B • Jakarta – Hongkong (6000 km)
- C • Jakarta – USA (16000 km)

Answer

Internet size is now 14 zettabyte

14 zettabyte = 14,000,000,000,000 gigabyte

Dimension of CD = 1.2 millimeters (thickness)

So, the height of the CD stack is **1.68 million KM**

Definitions

Big Data refers to a process that is used when traditional data mining and handling techniques cannot uncover the insights and meaning of underlying data.

- Technopedia

Big Data represents the information assets characterized by such a high volume, velocity and variety to require specific technology and analytical.

- "A Formal definition of Big Data"

Big Data is a term that describes the large volume of data – both structured and unstructured – that inundates a business on a day-to-day basis.

- SAS

A term used to refer to the study and applications of data sets that are too complex for traditional data-processing application software to adequately deal with.







- Wikipedia

Extremely large data sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions.

- Dictionary



Characteristics of Big Data

	Volume		Veracity
	Variety		Visualization
	Velocity		Value

Big Data Tools

Data Analysis & Platforms HPCC Systems, Storm, Dremel, Hadoop, Apache Drill, IKANOW, BRILLIANT DECISIONS, Hortonworks, Spark, SAMOA	Databases / Data warehousing bigdata, INFOBRIGHT, Cassandra, 4store, H2, GlobalisDB, InfiniDB, riak, Infinispan, HYPERTABLE, Firebird, ORACLE BERKELEY DB, MariaDB, Drizzle, HyperSQL, monetdb, RethinkDB, SQLite	In-Memory Computing GridGain, hazelcast, TERRACOTTA, NMemory, GORAK
ERP BI Solutions talend, Jaspersoft, Palo, spagoobi, pentaho, jedox, BIRT	Business Intelligence openi.org, rapidminer, ANIME, WEKA, togaware, SPDF	Data Mining Solr, elasticsearch, northgate, BASE INTERNATIONAL, QM
Key/Value KERO SPIKE, leveldb, redis, Chordless, Tokyo Cabinet, SCALLEN	Document Store mongoDB, Couchbase, RaptorDB, EIDB, JASDB, SchemaDB, CouchDB	Graph databases Gremlin, HIVE, TriniteGraph, INFO GRID, Neoj, FlockDB, Graph, BrigtetarDB
Object databases db4objects, Zope, mobject, Magma, PERSEVERE, EyeDB, NDatabase, Stirling	Operational VOLTDB, Apache Kafka, ThinkUp, Corona	Multidimensional FIS, SciDB, rasdaman
Multimodel ArangoDB, sedna, LIQUIDBASE	XML Databases existdb, BASE, Qizx	Grid Solutions BIRDFACES, Galaxy

<https://dataflaq.com/big-data-open-source-tools/os-home/>

Big Data from Insurance Perspective



Big Data to Insurance

Big data provides new insights from healthcare data, social networks, telematics sensors, others.

Will impact insurance business from end to end. Product design to selling process, underwriting to claim.

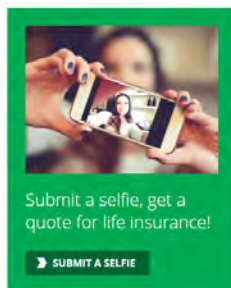
Many insurance companies have made a big commitment and effort in Big Data.

It's only the beginning...

Insurance Underwriting

Example of innovative way to underwrite

Protect your family future in just a snap!



<https://term.lgamerica.com/selfie-quote/#/>

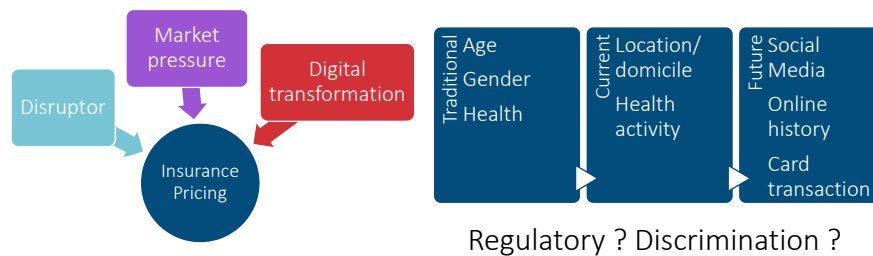
Facial analytics platform to determine:

- Age
- Gender
- BMI

Insurance Underwriting

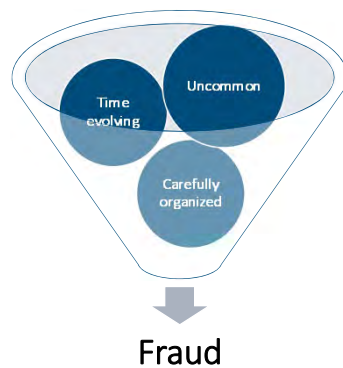
Insurance Pricing

- Predictive modelling is not a new thing
For example, mortality table to generate premium



Claim Management – Fraud detection

Big Data can help to improve claim management by reducing lost due to fraud and improve processing time.



Social Network Analysis

- Cross reference data from multiple sources to develop a pattern
- Technologies used such as text mining, sentiment analysis, content categorization were used to create a fraudulent score

Predictive Analytics model

- Creating a predictive / propensity to fraud model based on information inputted
- This include a long reports written by the claim investigator which will be translated into information feed to the model

Big Data for Actuaries



What Big Data bring to actuary

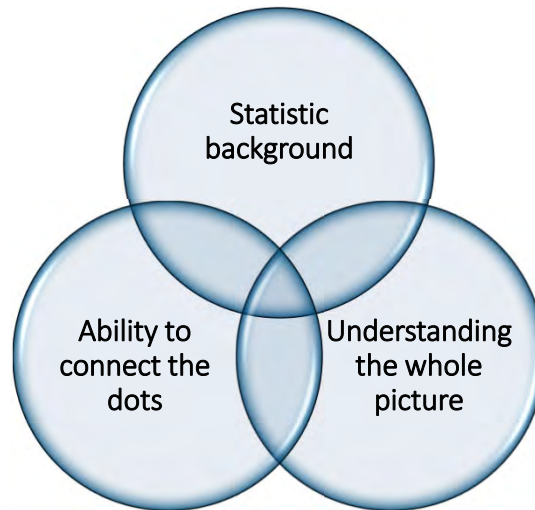
More data for actuarial analysis

Knowledge gives power

Best estimate → stochastic

Explain people behaviors better

What actuary bring to Big Data



Thank you



Big Data in Action

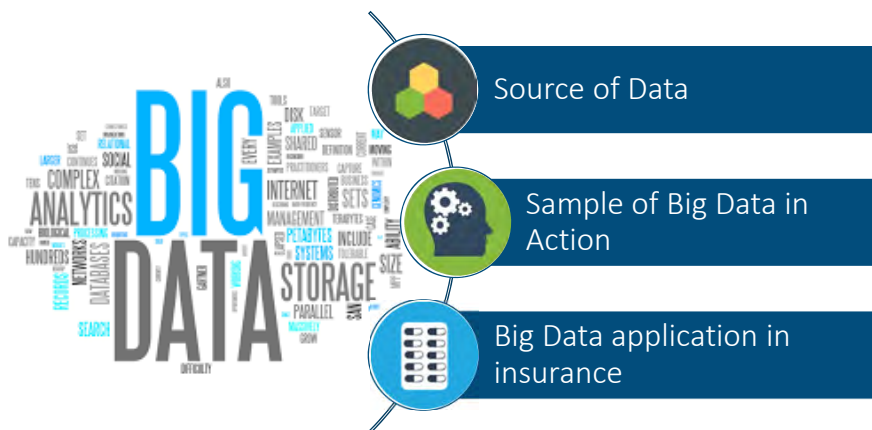
PAUL SETIO KARTONO

FSAI, ASA, MAAA, Director & Chief Strategy Officer FWD Life Indonesia

13 November 2018



Big Data today and around you



Source of Data

Internal

- Customer Demographic profile (age, gender, location, family tree, health condition)
- Payment history
- Contact history
- Campaign response
- Claim (frequency, severity, type)
- Sales Activity (GPS, illustration, contact, training, behavior)
- Policy changes history
- Many more



External

- Customer Demographic profile enrichment with Social Media crawling
- Business Partner data (e.g. financial information)
 - Weather data
 - Financial market data
- Vendor data (hospitals, workshops)
 - Community Activity
 - Civil registration data
 - Many more

Big Data Sample 1



meet Andy

- Age : 35
- Occupation : Marketing manager, JV company
- Marital : married, no kids, wife work
- Area : BSD (home), Sudirman (work)

Policy history

- Group Medical policy from employer
- Has been with company 2 years
- Outpatient claim 10% limit
- Inpatient claim once, 1 year ago due to injury in motorcycle touring
- Read 60% of health article from email

Social Media

- Active in Soc Med (have account in Facebook, Twitter, Instagram, LinkedIn)
- Average 5,000 followers
- 5 posts per day
- Travelling abroad 2 times a year (holiday)
- Follow sports and fitness activity hashtags

Results

Social Status:

- Mid
- Mid Affluent
- Affluent
- HNW

Financial Literacy:

- Illiterate
- Literate
- Savvy

Digital Literacy:

- Illiterate
- Literate
- Savvy

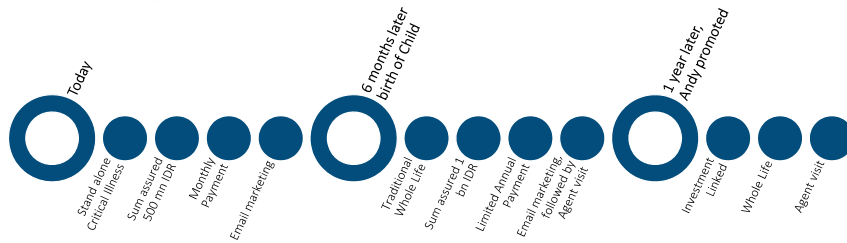
Health Condition:

- Poor
- Good
- Active

Big Data Sample 1



Next Best offer for Andy



Big Data Sample 2



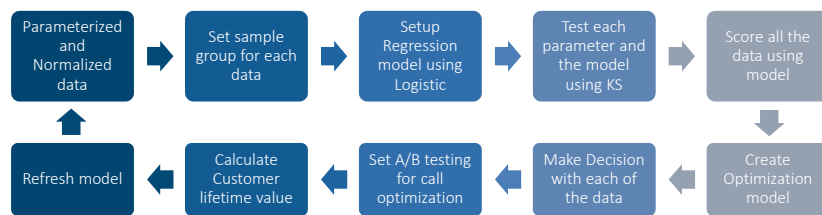
Background

- New partnership with credit card just signed (1 mn database)
- Agreed to offer 3 products
- Credit Card transaction is given
- Telemarketing channel was chosen as distribution method



Objectives and Method

- Maximize premium with given data
- Set Contact management
- Set Product offering campaign
- Use Logistic model for the data with RFM
- Use KS test for data fit
- A/B test for improvement



Big Data Sample 2

Data X

Product	Conversion Probability	Average Ticket Size	Acquisition Cost Loading	VNB Margin
Product A	50%	1,800,000	900,000	40%
Product B	55%	1,500,000	800,000	50%
Product C	30%	2,000,000	1,100,000	45%

Actuary to decide and add value: How to improve the odd? (e.g. change script, alter timing), How to reduce expense? (e.g. use automated call as introduction, use VOIP), case size vs persistency?

Data Y

Product	Conversion Probability	Average Ticket Size	Acquisition Cost Loading	VNB Margin
Product A	3%	1,800,000	900,000	55%
Product B	4%	1,500,000	800,000	40%
Product C	6%	2,000,000	1,100,000	35%



Big Data application in Insurance

Increase Sales

- Propensity modelling
- Next Best Offer
- Automated Underwriting

Reduce Expense

- Fraud Detection
- Automate Processing
- Selective Campaign

Customer Experience

- Personalized Service
- Personalized Offer
- Auto claim payment

Manage Claim

- Preventive health condition
- Health improvement with wearables
- Behavior change

