



A case study:
using models to achieve
100% data quality

VSAE Actuarialcongres 5 maart 2024

Linda de Koter MSc AAG

Agenda

- Welcome
- Combining the toolsets of actuaries and software developers
- Data Issues
- Data Findings
- Future work
- Technical Issues



axini

Combining toolsets

Actuaries and Developers works together



sit amet, consectetur
eiusmod tempor incididunt
magna aliqua. Ut enim ad
nostrud exercitation ullamco

Lorem ipsum dolor

sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum

Lorem ipsum

807 Lorem ipsum dolor
sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut

807-A

1/4

A6

0-100 mts

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum

The dream of combined effort

- Would it be possible to recalculate and check complete pension entitlements?
- What can happen when you bring together
 - Talkers (and big dreamers)
 - Calculators (aka actuaries)
 - Builders (aka software developers)
- And:
 - Start talking about quantifying pension databases
 - Don't take no for an answer



The dream of calculating everything

- Faced with disbelief, this was not possible, or so we heard
- Pension applications are complex, and not easy to replicate
- Would there be a market for this approach?

MODEL PARTS

- aanpakken
 - algemeen
 - dienslijd
 - gronslag
 - prognose
 - tijdsvenredig
 - uitkering
- aanpakkenmodule
- basis
 - filter
 - functies
 - indicators
 - interval
 - personen
 - tijdsbepalling
 - uitsluiting
- koopsom_en_factor
- output
 - includes
 - verwerking_sut
 - verwerking_sut_met_filter
- regeling
 - berekening
- run
 - berekening**
 - includes
 - Initialiseer_tabellen
 - input_specification.dsl
 - output_specificatie_obv_exc...
- vvp
 - vvp_bij_mutatie
 - vvp_factoren
- wijzigingen
- readme
- vb_csv_transformatie
- vragenlijst

```

basis/uitsluiting x | output/verwerking_sut x | output/verwerking_sut_met_filter x | vvp/vvp_bij_mutatie x | bc
output/includes x | run/includes x | run/berekening x | ★ aanspakkenmodule/met_uitsluiting_en_interval x | aar
1 # De bij elke mutatie worden alle aanspakken en onderbouwingsgegevens vastgelegd
2 # een deel van deze informatie wordt vervolgens vastgehouden om te tonen bij de resultaten.
3
4 # De volgorde van vaststellen is: <br />
5 # Start met doorrollen van de vorige mutatie naar de huidige mutatie. Dit proces bestaat uit het updaten var
6 # dienslijden in het verleden en tijdsvenredige aanspakken voor alle pensioenproducten. <br />
7 # Vervolgens wordt met de nieuwe mutatie de dienslijd naar de toekomst en de gronslagen vastgesteld.
8 # met deze dienslijden en gronslag worden vervolgens de pensioenaanspakken prognose vastgesteld.
9 # tenslotte worden de ingegane uitkeringen geupdated.
10
11 # Dit is het resultaat dat op het scherm komt.
12 # per mutatie, wordt de mutatie verwerkt in tijdsvenredige producten, prognose en uitkeringen
13 # de resultaten voor de pensioenberekening wordt weggeschreven en getoond op het scherm als output/
14
15 def verwerk_polis
16   transform_csv_files_model
17
18   mutaties.map.with_index do |mutatie, i|
19     vorige_mutatie = nil
20     vorige_mutatie = mutaties[i-1] if i > 0
21
22     resultaat_mutatie(vorige_mutatie, mutatie)
23   end
24 end
25
26 # In deze method wordt voor 1 mutatie vastgesteld welke aanspakken op het scherm getoond dienen te worden
27 # Indien deze mutatie de 1e is in de test, dan wordt ook een eerste gronslag berekening uitgevoerd. Zodat v
28 # volgende mutatie de pensioengronslag over de voorgaande periode niet is overgenomen uit de data, maar is
29 def resultaat_mutatie(vorige_mutatie, mutatie)
30   if vorige_mutatie
31     #log << "mutatie: #{mutatie_naam_obv_code(mutatie['code'])}"
32     mutatie_verwerken(vorige_mutatie, mutatie)
33     aanspakken = aanspakken_afgerond_zonder_nil(huidige_stand['aanspakken'])
34   else
35     pensioenparameters_inlezen(mutatie)
36     aanspakken = aanspakken_afgerond_zonder_nil(huidige_stand['aanspakken'])
37     gronslagen(mutatie).each {|k, v| @pensioeninfo_berekening['aanspakken'][k] = v}
38   end
39
40   uitsluiting_specifieke_aanspakken(aanspakken)   if @run_met_uitsluitingen
41

```



axini

Data issues

The trouble of corrupt data

- How to write a date?
- Integers turning into booleans
- Percentages written in two different ways

The trouble of incomplete data

- Participants changing status without mutation



The trouble of fuzzy data

- Users of a system find "nice" workarounds
 - Most of the time not "nice" in a data perspective
- Data migrations used fields for different meanings
- Data not following specifications
 - i.e. expected percentage, got a monetary value

Data findings

- 4

	Fondsspecifieke uitruil OP PP	Fondsspecifieke uitruil OP PP
mu	42	42
mu	2019-01-01	2019-01-01
sta	opname met risicodekking dekking	opname met risicodekking dekking
aan	any	...
vvp	any	...

+ 5

+ 6

- 7

	Premievrijmaken - uitdienst	Premievrijmaken - uitdienst
mu	16	16
mu	2019-07-31	2019-07-31
sta	premie vrij na ontslag	premie vrij na ontslag
aanspraken	any	...
vvp	any	...

- 8

	Jaarwerk	Jaarwerk
mutatie	50	50
mutatie_code	2020-01-01	2020-01-01
mutatie_datum	premie vrij na ontslag	premie vrij na ontslag
status_deelnemer		

- aanspraken

pensioengrondslag_voltijd	0	0
pensioengrondslag_bevroren_voltijd	0	0.0
OP_tijdsevenredig	2.33000..4.33000	52.06
OP_prognose	any	0.0
OP_uitkering	0	0
OP_verevend_tijdsevenredig	0	0
OP_verevend_uitkering	0	0
TOP_tijdsevenredig	0	0
TOP_uitkering	0	0
TOP_verevend_tijdsevenredig	0	0
TOP_verevend_uitkering	0	0
PP_tijdsevenredig	1.33000..3.33000	36.45
PP_niet_uitruilbaar_tijdsevenredig	0	0
PP_fictief_tijdsevenredig	any	0.0

Kenmerken:
grondslag: 3588
Deeltijd: 8.0%
Opbouw: 2.0%

Periode: 0.58
Opbouw OP = 3.33
Opbouw PP = 2.33

- 0			
mutatie		Jaarwerk	Jaarwerk
mutatie_code		50	50
mutatie_datum		2018-01-01	2018-01-01
status_deelnemer		actief in regeling	actief in regeling
- aanspraken			
pensioengrondslag_voltijd		18751.26	18751.26
PP_tijdsevenredig		0	0
PP_niet_uitruilbaar_tijdsevenredig		586.84	586.84
PP_fictief_tijdsevenredig		0	0
PP_verzekerd		5078.16	5078.16
PP_uitkering		0	0
- 1			
mutatie	Wijziging voltijd salaris en/of deeltijdpercentage	Wijziging voltijd salaris en/of deeltijdpercentage	Wijziging voltijd salaris en/of deeltijdpercentage
mutatie_code		2	2
mutatie_datum		2018-01-01	2018-01-01
- 2			
mutatie		Jaarwerk	Jaarwerk
mutatie_code		50	50
mutatie_datum		2019-01-01	2019-01-01
status_deelnemer		actief in regeling	actief in regeling
- aanspraken			
pensioengrondslag_voltijd		19413.51	19413.51
PP_tijdsevenredig		0	2048.52
PP_niet_uitruilbaar_tijdsevenredig		586.84	586.84
PP_fictief_tijdsevenredig		163.07	0
PP_verzekerd		0.0	7318.15
PP_uitkering		0.0	0

The trouble of the results

- 3			
mutatie	Overlijden deelnemer -> start PP, WzP als er partner/kind is	Overlijden deelnemer -> start PP, WzP als er partner/kind is	
mutatie_code	19	19	
mutatie_datum	2020-12-25	2020-12-25	
status_deelnemer	ingegaan nabestaandepensioen partner/kind	ingegaan nabestaandepensioen partner/kind	verk
aanspraken	any	...	50
vvp	any	...	1-01
- 4			
mutatie	Jaarwerk	Jaarwerk	oen
mutatie_code	50	50	0
mutatie_datum	2021-01-01	2021-01-01	0.0
status_deelnemer	ingegaan nabestaandepensioen partner/kind	ingegaan nabestaandepensioen partner/kind	0
- aanspraken			
pensioengrondslag_voltijd	0	0	0.0
pensioengrondslag_bevroren_voltijd	0	0.0	18.81
OP_tijdsevenredig	0	0	0
OP_prognose	any	0.0	0
OP_uitkering	0	0	0
OP_verevend_tijdsevenredig	0	0	0
OP_verevend_uitkering	0	0	0
TOP_tijdsevenredig	0	0	0
TOP_uitkering	0	0	19.51
TOP_verevend_tijdsevenredig	0	0	n
TOP_verevend_uitkering	0	0	
PP_tijdsevenredig	0	0	
PP_niet_uitruikbaar_tijdsevenredig	0	0	
PP_fictief_tijdsevenredig	any	0.0	
PP_verzekerd	any	0.0	
PP_uitkering	2188.51000..2190.51000	916.31	
BPP_tijdsevenredig	0	0	

axini



What will the future bring ?

The dream of Data Quality

- It is great to find and fix data issues
- But based on the guidelines of the Pensioenfederatie: stay clean
- It's easy, once done, to repeat this proces at a given interval
- So keep testing and evaluating your data!

The dream of testing the future

- Having the approach ready, it's easy to extrapolate
- Why not test mutations based on the models that are already there
- Expand to also take (subsets) of SPR and FPR calculations
- Generate testcases for new systems

axini

Technical Issues



The trouble of huge data sets

- Normal use of structure data is less suitable
- Reading big data files causes Out-Of-Memory issues
- New input method via csv
 - Field converters made new OOM issues
 - Creating better converters made extra OOM issues
- Second method via csv worked, but slow

The trouble of unknown databases

- No predefined query to get data from the customer
- Complexity of finding all needed data and interpret them
 - What do all the indicator codes mean?
 - Where are the outcomes of mutations?
 - Where are the choices made for specific entitlement distributions?
 - Why does this entitlement "suddenly" appear or disappear?
 - Etc ...

The trouble of running calculations

- With many participants over the years come many calculations
- Available hardware was quite small (the regular laptop had better specs)
- New OOM crashes to deal with for the builders



Questions?



Contact:



Linda de Koter MSc AAG

Linda.de.koter@axini.com

LinkedIn:



axini

