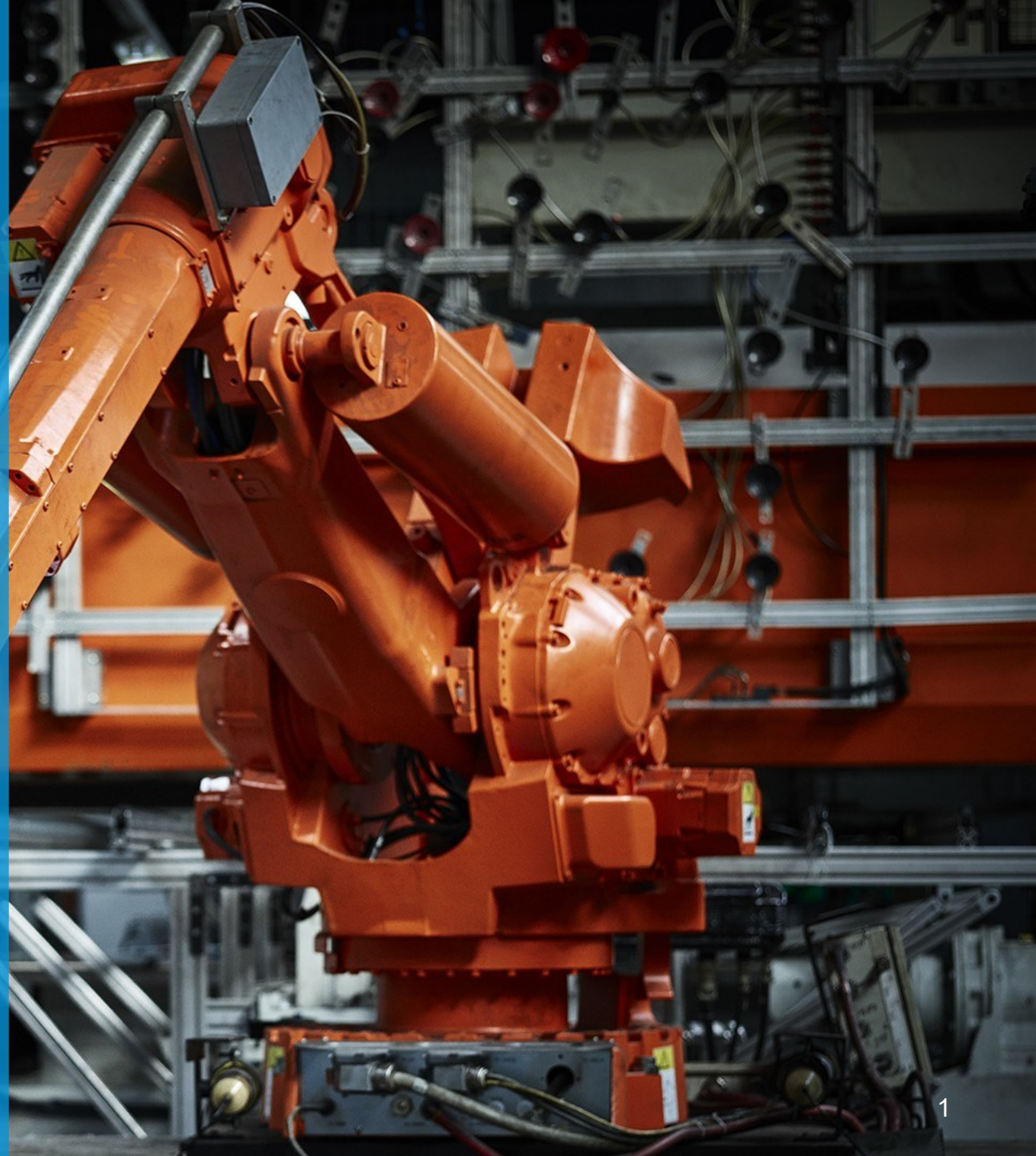
Swiss Re

Automating Reports with R Markdown

Claudio Rebelo, Actuary at Swiss Re
Après Midi October 07, 2020



Disclaimer

The views and opinions expressed in this presentation are solely my own and do not necessarily represent of reflect the views of Swiss Re

What is R Markdown?

R Markdown (.Rmd) is an authoring format that enables easy creation of dynamic documents, presentations, and reports from R.

It combines the core syntax of markdown (an easy to write plain text format) with embedded R code chunks that ar run so their output can be included in the final document

source:
https://rmarkdown.rstudio.com/

Reproducibility and Automation: Shifting Paradigms!

Reproducibility: R Markdown documents are fully reproducible i.e.: automatically regenerated whenever underlying R code or data changes*

Some Key Benefits of Reproducibility:

- Understand what was done months ago;
- Adjust the code or data, even early in the process and re-run all analysis;
- Share with others so they can further extend your research

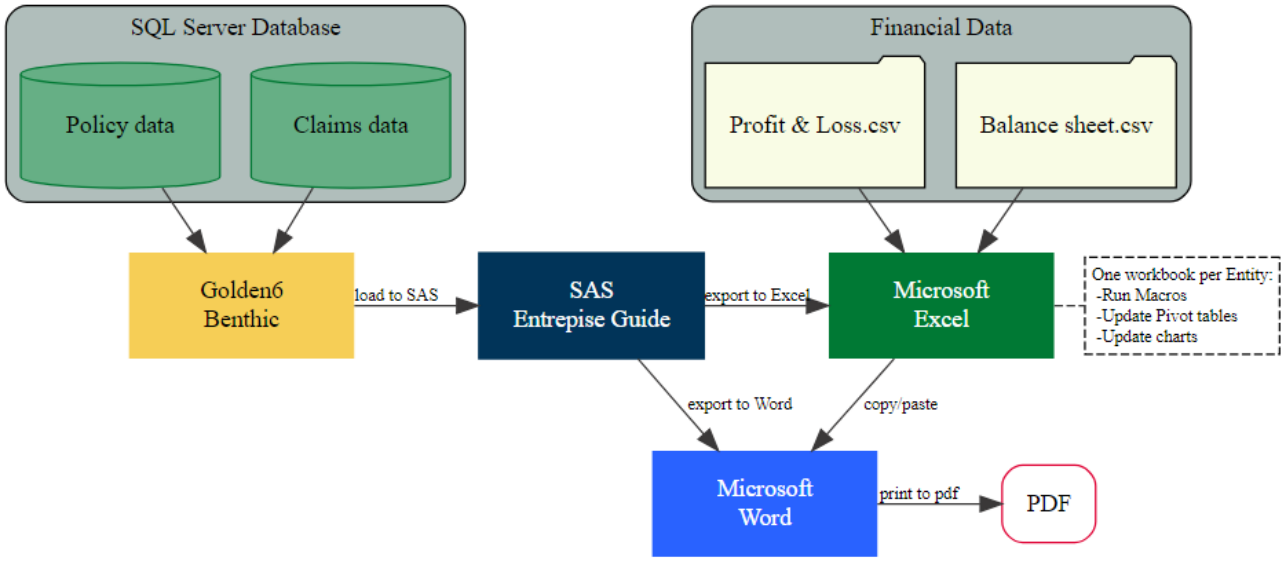
source: R Programing for research

Automating reports

- Reproducibility entails automation but the converse is not always true
- We will focus on automating taks regardless if the reproducibility chain is broken or not;
- A pratical example of a fictional Casualty Insurance Company will be presented as a demo

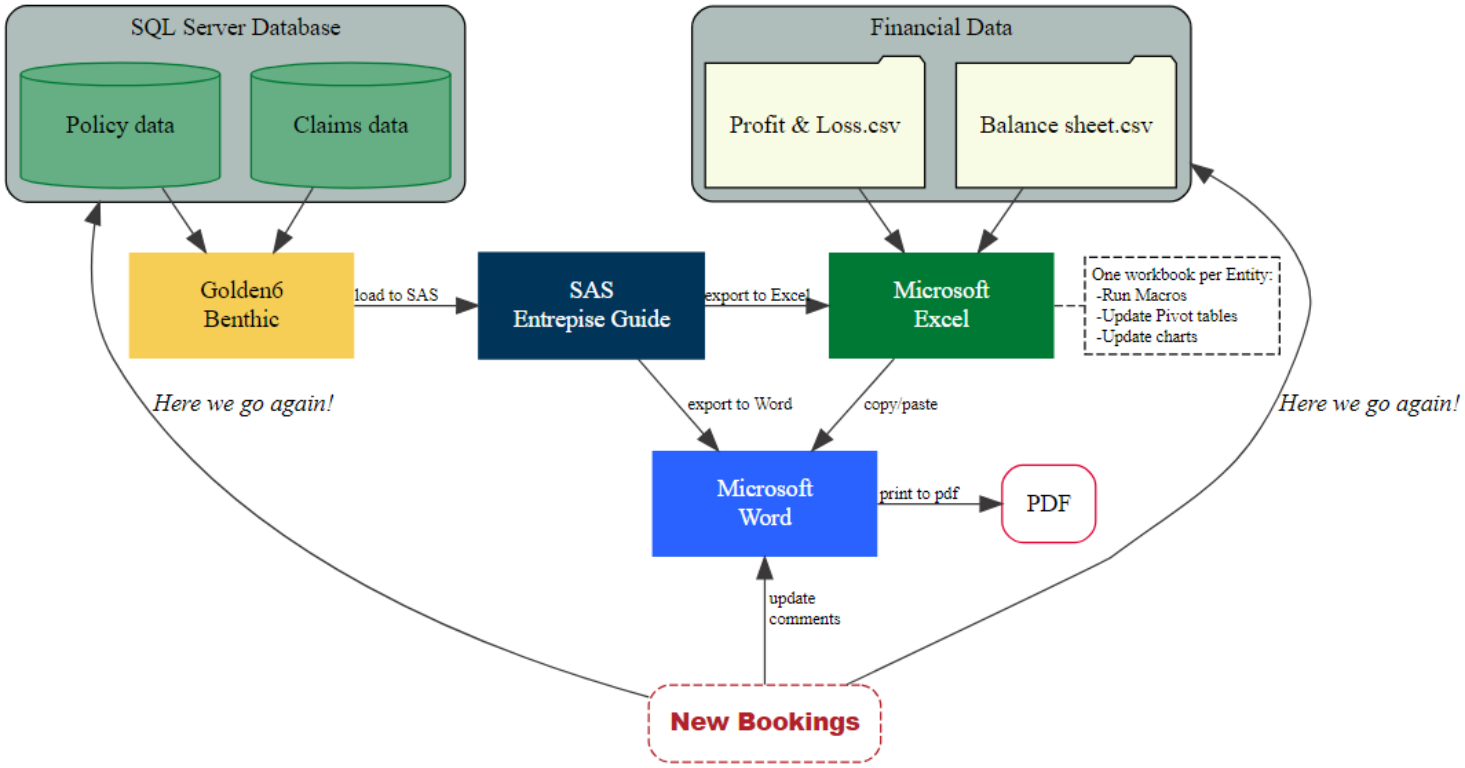
* https://rmarkdown.rstudio.com/

Quarterly Report: Messy Workflow



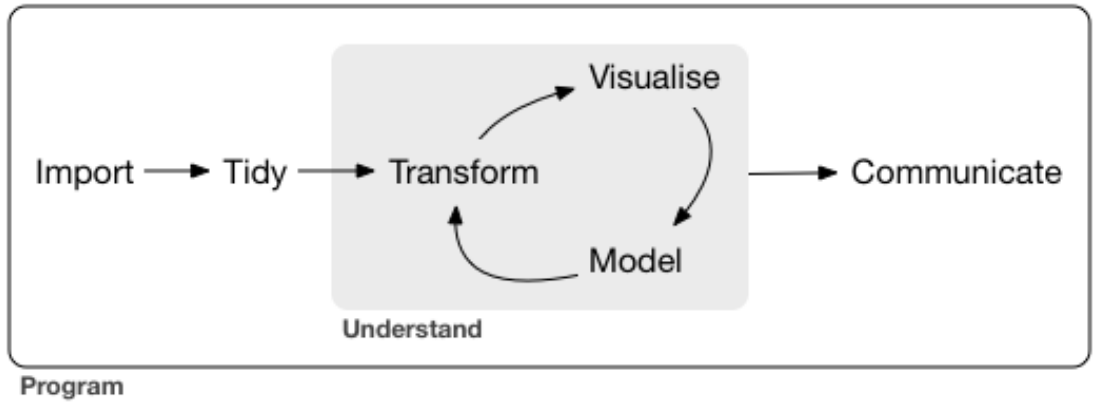
https://rich-iannone.github.io/DiagrammeR/mov/DiagrammeR.mp4

Quarterly Report: Messy Workflow



https://rich-iannone.github.io/DiagrammeR/mov/DiagrammeR.mp4

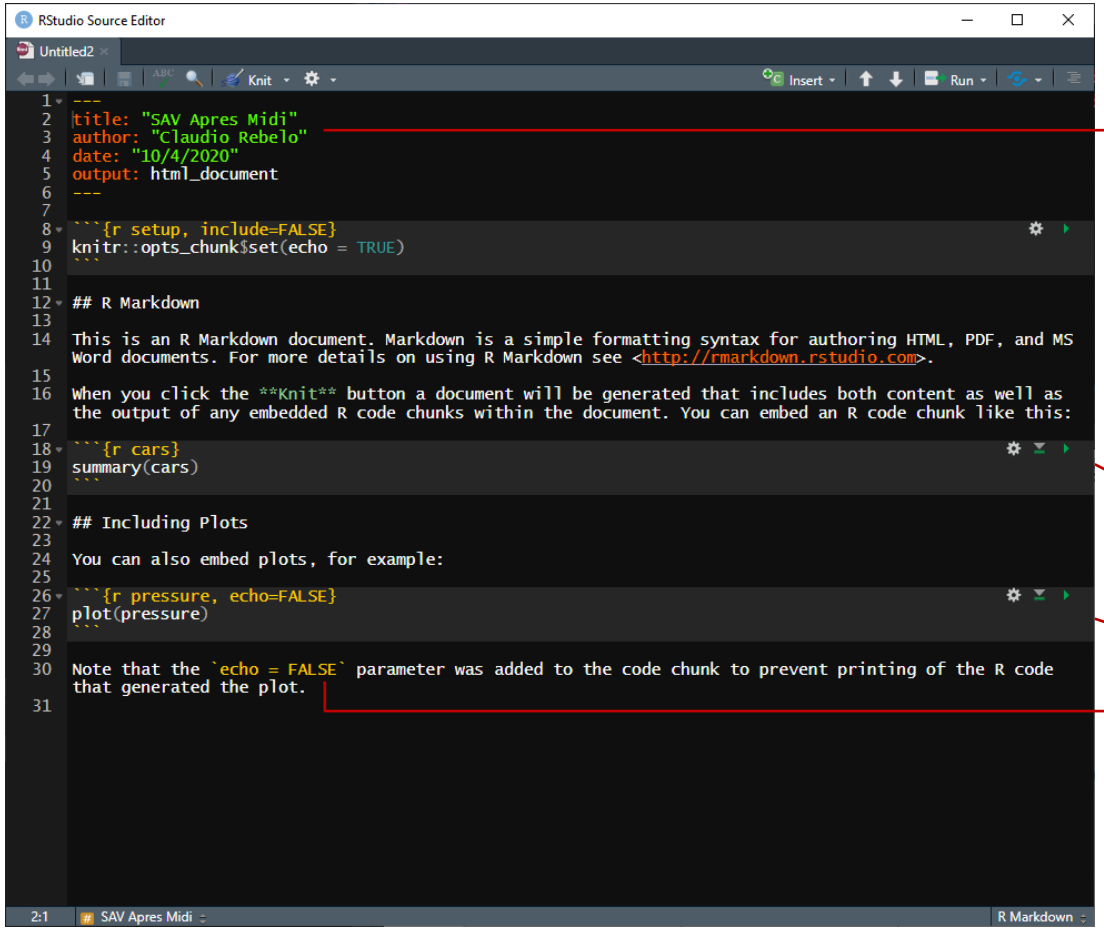
Why not do that all in R?



https://r4ds.had.co.nz/

R Markdown file
Ready to knit?

The three components of an R Markdown file: YAML header; Text & Code chunk



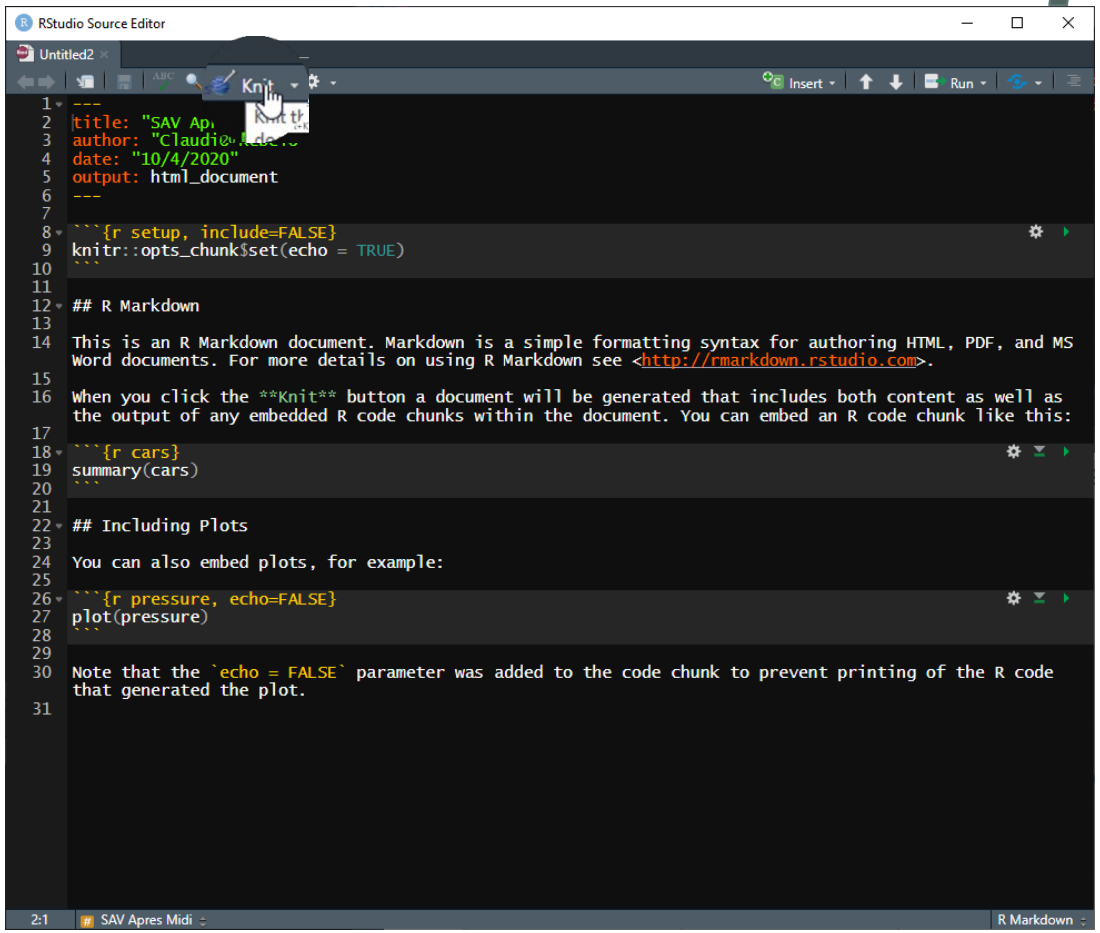
YAML header defines the structure of the file such as defining the output format e.g.: Word, PDF, HTML...

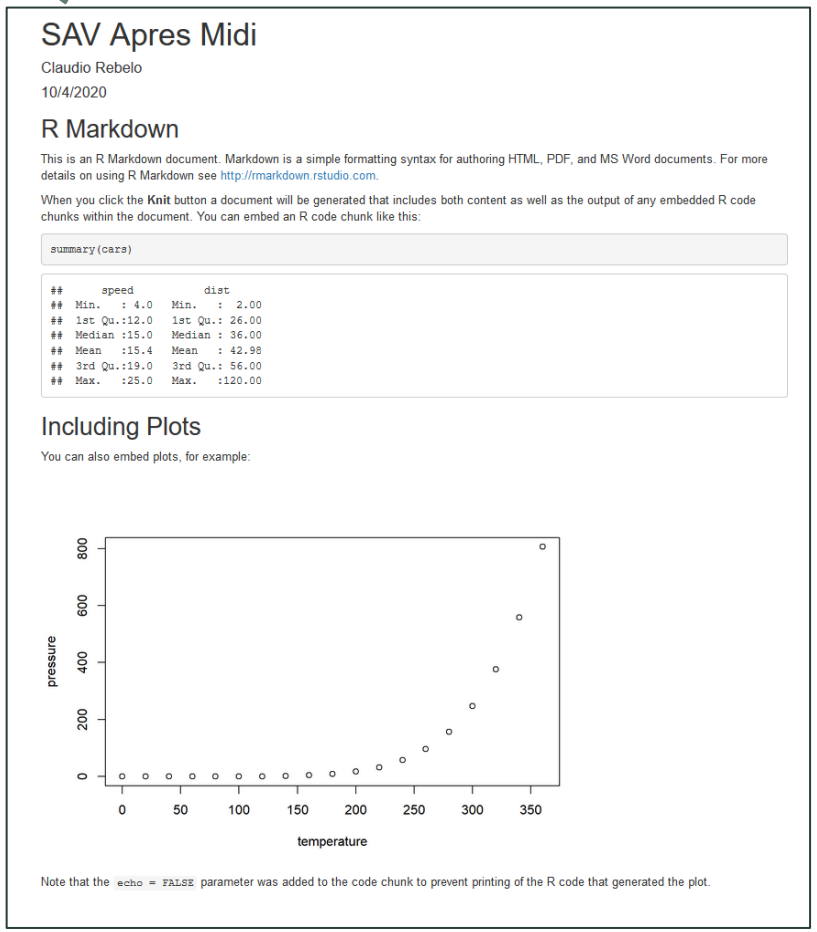
Text written in Markdown. Markdown is designed to be an easy-to-write formatting syntax.

R code either via code chunks surrounded by ```
or via
inline code `r age_of_insured`

Hit the knit! To generate the report

Alternatively use the shortcut Ctrl+Shift+K





What is happening under hood?

When R Markdown renders the .Rmd file it sends it to knitr, which executes the code chunks and creates a new markdown (.md) file

The markdown file is then processed by pandoc which is responsible for creating the finished format.

All this happens with a simple click of a button!

What about my Python/SQL/JavaScript skills?

R Markdown code chunks can execute code in many other languages besides R such as:

- Python
- SQL
- Bash
- Rcpp
- Stan
- JavaScript
- CSS

```
```{python}
x = 'hello, python world!'
print(x.split(' '))
```
```

To get the full list of supported engines; type in the console: `names(knitr::knit_engines$get())`

Aaron Berg:
https://rstudio.com/resources/rstudioconf-2018/beyond-r-using-r-markdown-with-python-sql-bash-and-more/

Automating Reports
Remarkable Re (RR) Study Case

RR's Database: current and last quarter
Policy Data

POL_ID	POL_INCEP	UWY	Insured	LoB	Region	LE	Industry	UP	UAC	PL	CR	IBNR
POL_10	31.03.2013	2013	Ganso	General Liability	Connyland	Remarkable Solutions	Sports	11'965.71	0.00	0.00	0.00	-308.23
POL_10	31.03.2014	2014	Ganso	General Liability	Connyland	Remarkable Solutions	Sports	12'656.65	0.00	0.00	0.00	-604.68
POL_10	20.06.2013	2013	Virgil van Dijk	General Liability	Disneyland	Remarkable Solutions	Mining/Metals	30'640.19	0.00	0.00	0.00	-445.77
POL_10	01.01.2014	2014	Virgil van Dijk	General Liability	Disneyland	Remarkable Solutions	Mining/Metals	57'647.79	0.00	0.00	0.00	-1'297.13
POL_10	01.04.2015	2015	Virgil van Dijk	General Liability	Disneyland	Remarkable Solutions	Mining/Metals	37'758.71	0.00	0.00	0.00	-4'691.57
POL_10	01.04.2016	2016	Virgil van Dijk	General Liability	Disneyland	Remarkable Solutions	Mining/Metals	30'981.87	0.00	0.00	0.00	-6'085.97
POL_10	01.04.2017	2017	Virgil van Dijk	General Liability	Disneyland	Remarkable Solutions	Mining/Metals	30'981.87	0.00	0.00	0.00	-9'852.87
POL_10	01.05.2018	2018	Virgil van Dijk	General Liability	Disneyland	Remarkable Solutions	Mining/Metals	23'642.00	0.00	0.00	0.00	-11'166.90

POL_ID Policy ID a unique policy identifier

POL_INCEP Policy inception date

UWY Underwriting Year

Insured Name of the policy holder/client (football players)

LoB Line of business

Region: region/land of the policy holder

LE Legal entity i.e.: remarkable's entity which underwrote the policy

Industry Industry/Occupancy of the policy holder

UP, UAC, IBNR Ultimate Premium, Ultimate and IBNR respectively as at the valuation date

PL, CR Total (cumulative) Paid Losses and Outstanding Case Reserves as at the evaluation date. Comprises all claims under the policy

Claims data, Initial Expectation (T0) data and Meta data

Claims data

CLAIM_ID	CLAIM_DESC	POL_ID	PL	CR
CLAIM_15	CLNAME1/155005CLNAME2/<NA>	POL_10	-30'012.35	0.00
CLAIM_18	CLNAME1/189601CLNAME2/<NA>	POL_10	-41'705.43	0.00
CLAIM_16	CLNAME1/ALLEGED NON CONFOR	POL_10	0.00	0.00
CLAIM_16	CLNAME1/ALLEGED NEGLIGENCE A	POL_10	0.00	-162'803.05
CLAIM_17	CLNAME1/CLAIMANT FALL OF HIS	POL_10	0.00	0.00

Claim_ID A unique claim identifier

Claim_DESC Brief description of the claim (includes unpleasant redundant characters e.g.: 'CLNAME1 / ')

POL_ID Policy identifier that originated the loss

PL and CR same meaning as in the policy data but for that specific loss as at the evaluation date

T0 Initial Expectation data

POL_ID	UP	UAC	UL
POL_20	1'239.73	-185.96	0.00
POL_21	783.94	-195.98	-331.46
POL_21	783.94	-195.98	-331.46
POL_19	5'613.56	-196.47	0.00

POL_ID, UP and UAC same meaning as in the policy data set

UL stands for Ultimate Loss in this case is the Ultimate Expected Loss since the origin is the T0 data set

Meta data

financial period	date	extraction date	extraction time
cq	31/12/2019	23/01/2020	09:45:03
lq	30/09/2019	23/01/2020	09:55:03

cq current quarter
lq last quarter

What type of reports can be produced with this data set?

An Exampleof a report in Word

https://davidgohel.github.io/officer/ https://davidgohel.github.io/officedown/ https://davidgohel.github.io/flextable/

Remarkable Re's Overview as of Q4 2019

Claudio Rebelo

October 04, 2020

Table of Contents

Remarkable Re

1 Introduction

Remarkable Re (RR) is a wholly fictional company part of the Remarkable Insurance Group (RIG) and it is one of the world's leading providers of Casualty (re)insurance.

Headquartered in Zurich, Switzerland, it was founded in October 2020 for the purpose of the *'Actuarial Data Science Après-Midi'* discussion.

The scope of this study is to provide a monthly helicopter overview of RR's financial situation to its internal key stakeholders.

This report should be read in conjuction with the non-existing detailed document descriping RIG's Swiss Solvency Test (SST) and its internal model.

The current version is only a draft and is not suitable for any other purpose than the set out above.

This **remarkable** report should not be quoted or referred to any third parties other than FINMA's and RR's independent auditors.

All figures are in USD unless stated otherwise.

Data as of December 31, 2019 was extracted on 23/01/2020 at 09:45:03 while data as of September 30, 2019 was extracted on 23/01/2020 at 09:55:03.

Table 1: Quarterly Technical Result in USD millions

USD millions	Underwriting Years 2019	2018 & prior	Total
Premium:			
Written Premium	0.7	0.1	0.8
Acquisition Costs	-0.1	0.0	-0.1
Net Premium	0.6	0.1	0.6
Losses:			
Paid Losses	-0.3	-0.8	-1.0
Case Reserves	-1.0	1.5	0.5
IBNR	1.1	3.4	4.5
Ultimate Losses	-0.1	4.2	4.0
Technical Results	0.4	4.2	4.7

i) Tech. Results = Net Premium + Ultimate Loss

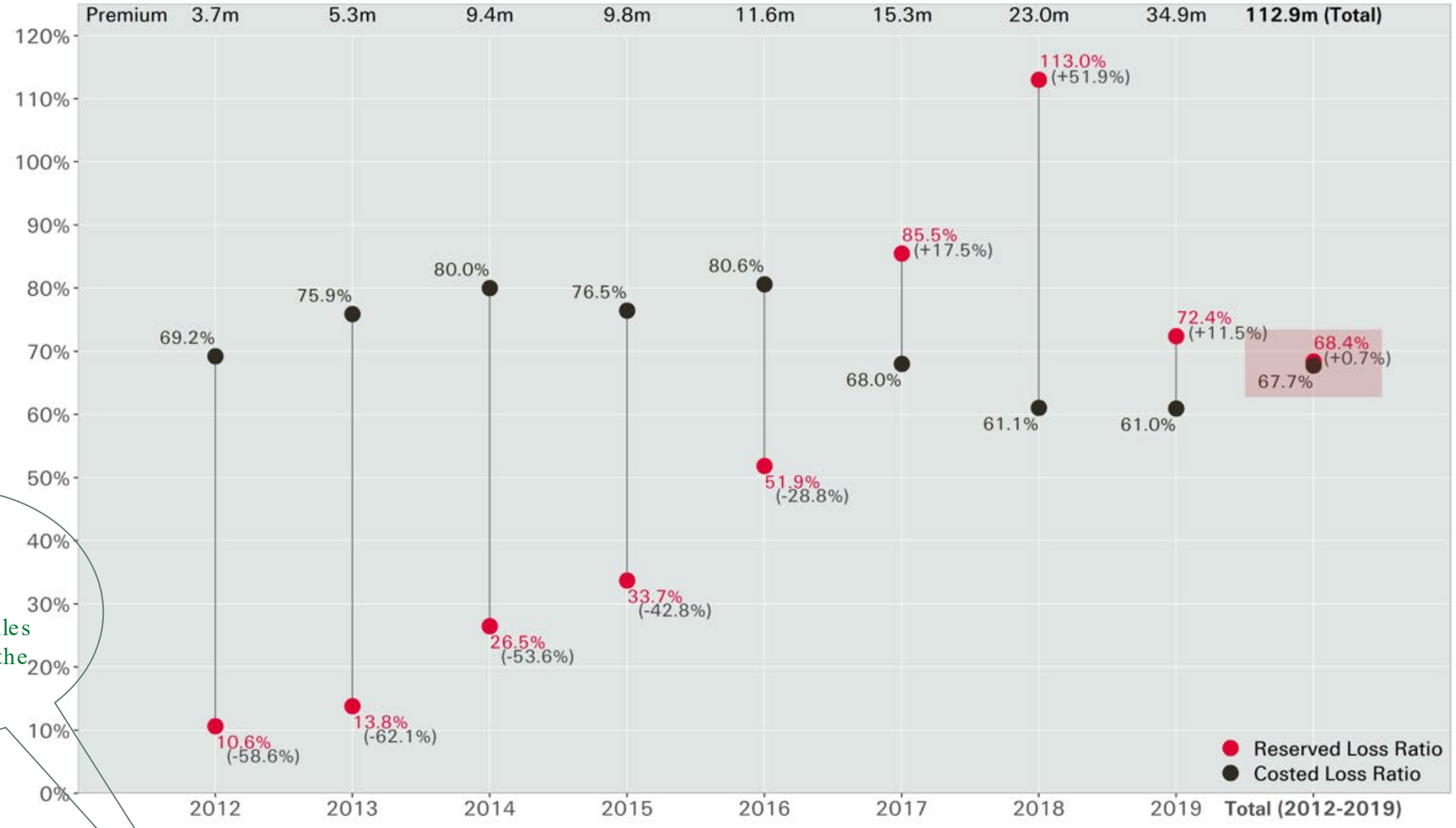
The book experienced a positive technical result of

4.7m

for contract years 2012 to 2019.

Let R do the talking! Add adverbs & adjectives according to predefined rules to make it sound human.

The result is mainly driven by IBNR releases due to low loss emergence whereof the largest 10 loss movements are:



Adverbs and adjectives according to predefined rules to humanize the text!

Figure 1: Costed vs Reserved by Underwriting Year

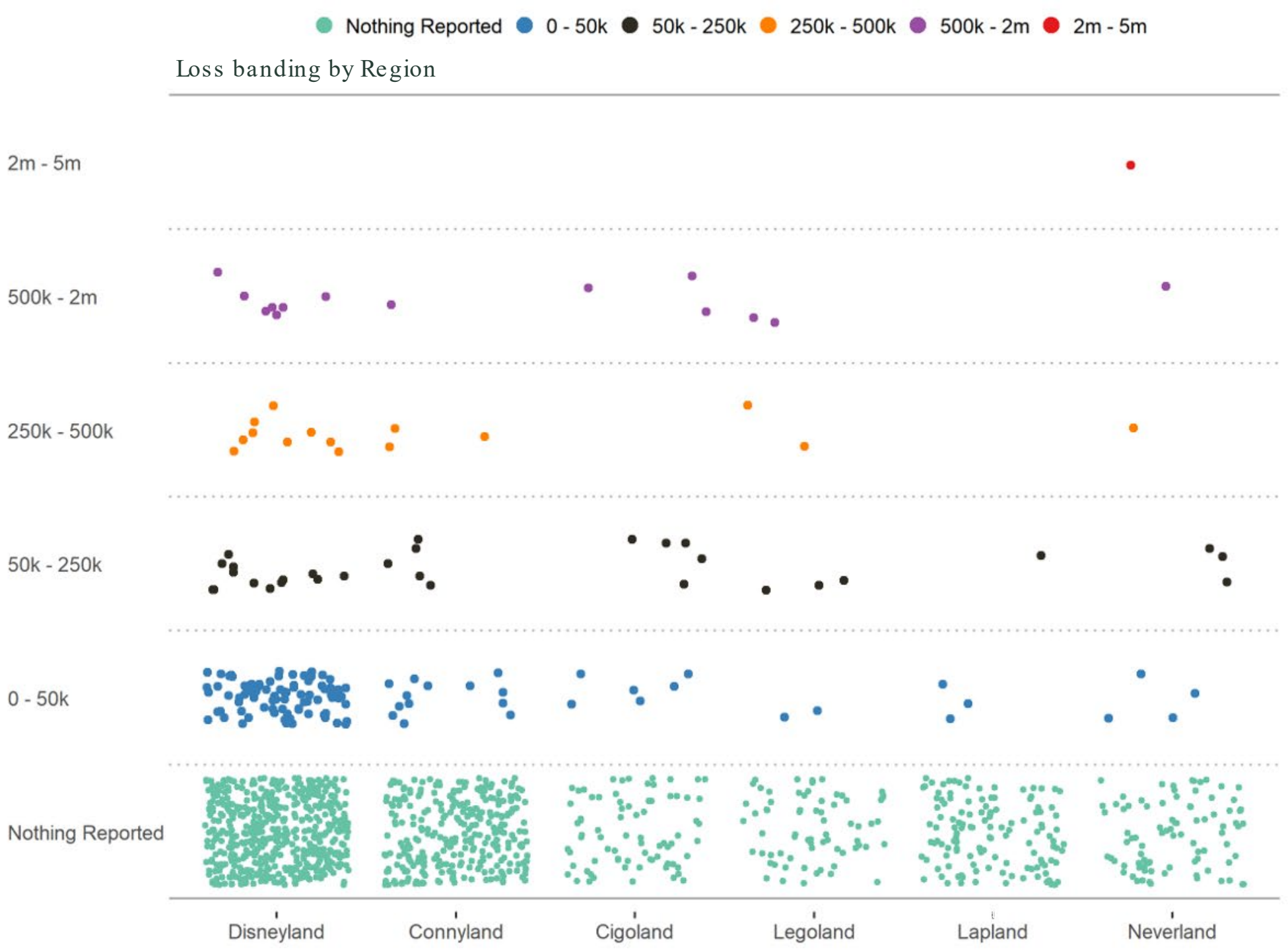
The average costed loss ratio is 67.7% compared to the current average reserved of 68.4%, a difference of only 0.7%, for contract years 2012 to 2019.

By most profitable insured

Insured	Region	Line of Business	Legal Entity	Industry	Underwriting Years	Written Premium	Technical Result
Arturo Vidal	Connyland	General Liability	Remarkable Solutions	Fishery	12, 13, 14, 15, 16, 17, 18, 19	3'204'495	1'465'977
Thibaut Courtois	Connyland	General Liability	Remarkable Solutions	Government Administration	14, 15, 16	1'915'298	1'227'676
Luka Modric	Disneyland	General Liability	RR International	Fishery	12, 13, 14, 15, 16, 17, 18	2'802'174	1'189'076
Sergio Busquets	Legoland	General Liability	RR International	Sports	12, 13, 14, 15, 16, 17, 18, 19	1'677'351	1'146'233
Alexis Sánchez	Connyland	General Liability	Remarkable Solutions	Banking/Mortgage	13, 14, 15, 16, 17, 18, 19	1'885'000	1'001'215

By least profitable insured

Insured	Region	Line of Business	Legal Entity	Industry	Underwriting Years	Claim Ratio	Written Premium	Technical Result
Jérôme Boateng	Neverland	General Liability	Remarkable Solutions	Sports	17, 18, 19	8/6	1'655'850	-5'728'808
Paul Pogba	Cigoland	General Liability	Remarkable Solutions	Automotive	17, 18, 19	2/3	1'117'721	-1'518'722
Claudio Marchisio	Connyland	General Liability	Remarkable Solutions	Entertainment/Movie Production	16, 17, 18, 19	1/4	398'498	-1'176'192
Toni Kroos	Disneyland	General Liability	RR International	Computer Software/Engineering	17	8/1	332'972	-1'163'318
Carlos Idriss Kameni	Disneyland	General Liability	RR International	Computer Software/Engineering	17, 18	1/2	42'259	-896'395


Nothing Reported
0 - 50k
50k - 250k
250k - 500k
500k - 2m
2m - 5m
Loss banding by Region
2m - 5m
500k - 2m
250k - 500k
50k - 250k
0 - 50k
Nothing Reported
Disneyland
Connyland
Cigoland
Legoland
Lapland
Neverland

Focus only on the non-automated sections

While it breaks the chain of reproducibility, the process is by far superior than the messy workflow

Remarkable Re

6 Changes in Actuarial assumptions

Analysist please fill this section.....

6.1 Apriori Loss Ratio adjustments

Analysist please fill this section.....

6.2 Pattern Adjustments

Analysist please fill this section.....

6.3 Data quality issues and other topics

Analysist please fill this section.....

Challenge:
What can be done to (partially) automate this section?

HTML is the richest format for communication that supports a variety features

Example of a report in html

Remarkable Re's Overview as of Q4 2019

Claudio Rebelo

October 05, 2020

1 Introduction

Remarkable Re (RR) is a wholly fictional company part of the Remarkable Insurance Group (RIG) and it is one of the world's leading providers of Casualty (re)insurance.

Headquartered in Zurich, Switzerland, it was founded in October 2020 for the purpose of the *'Actuarial Data Science Après-Midi'* discussion.

The scope of this study is to provide a monthly helicopter overview of RR's financial situation to its internal key stakeholders.

This report should be read in conjuction with the non-existing detailed document descriping RIG's Swiss Solvency Test (SST) and its internal model.

The current version is only a draft and is not suitable for any other purpose than the set out above.

This **remarkable** report should not be quoted or referred to any third parties other than FINMA's and RR's independent auditors.

All figures are in USD unless stated otherwise.

Data as of December 31, 2019 was extracted on 23/01/2020 at 09:45:03 while data as of September 30, 2019 was extracted on 23/01/2020 at 09:55:03.

▼ Description of main quarterly loss movements

Underwriting year 2019:

- USD -0.92m new loss on Dejan Lovren (Region: Disneyland, Industry: Computer Software/Engineering, Entity: Remarkable Solutions)
- USD -0.17m new loss on Cristiano Ronaldo (Region: Connyland, Industry: Package/Freight Delivery, Entity: Remarkable Solutions)
- USD +0.13m loss improvement on Stefan Savic (total to date: 0.00m, Region: Disneyland, Industry: Law Practice/Law Firms, Entity: Remarkable Solutions)

Underwriting year 2018:

- USD +0.60m loss improvement on Claudio Marchisio (total to date: -1.30m, Region: Connyland, Industry: Entertainment/Movie Production, Entity: Remarkable Solutions)
- USD -0.56m loss worsening on Carlos Idriss Kameni (total to date: -0.92m, Region: Disneyland, Industry: Computer Software/Engineering, Entity: Remarkable Solutions)
- USD +0.56m loss improvement on Manuel Neuer (total to date: -0.02m, Region: Cigoland, Industry: Banking/Mortgage, Entity: Remarkable Solutions)
- USD -0.13m loss worsening on Marcelo (total to date: -0.18m, Region: Disneyland, Industry: Sports, Entity: Remarkable Solutions)
- USD -0.11m loss worsening on Bernd Leno (total to date: -0.22m, Region: Cigoland, Industry: Automotive, Entity: Remarkable Solutions)
- USD -0.11m loss worsening on Robert Lewandowski (total to date: -0.40m, Region: Disneyland, Industry: Law Practice/Law Firms, Entity: Remarkable Solutions)

Underwriting year 2015:

- USD +0.64m loss improvement on Thibaut Courtois (total to date: 0.00m, Region: Connyland, Industry: Government Administration, Entity: Remarkable Solutions)

By Underwriting Year

Table 2.2: Quarterly Technical Result in USD millions

UW Year	Premium	Acquisition Costs	Reported Losses	IBNR	Ultimate Loss	Technical Result	Ultimate Loss Ratio	Combined Loss Ratio
2012	3.7	-0.2	0.0	-0.4	-0.4	3.0	11%	17%
2013	5.3	-0.4	0.0	-0.7	-0.7	4.2	14%	20%
2014	9.4	-0.6	-0.3	-2.2	-2.5	6.3	26%	33%
2015	9.8	-1.0	-0.5	-2.8	-3.3	5.5	34%	44%
2016	11.6	-1.4	-1.4	-4.7	-6.0	4.1	52%	64%
2017	15.3	-2.4	-6.6	-6.5	-13.1	-0.1	86%	101%
2018	23.0	-3.7	-15.4	-10.6	-26.0	-6.7	113%	129%
2019	34.9	-5.4	-4.4	-20.9	-25.2	4.2	72%	88%
Total	112.9	-15.2	-28.5	-48.7	-77.3	20.5	68%	82%

= (Ultimate Loss Ratio) + (-Acquisition Costs) / Premium

> The **combined loss ratio** is
>
> # 82%
>
> for contract years 2012 to 2019.

The average loss ratio is 68% for the same period while the worst performing year is 2018 with a combined loss ratio of 129%.

By Reported Loss

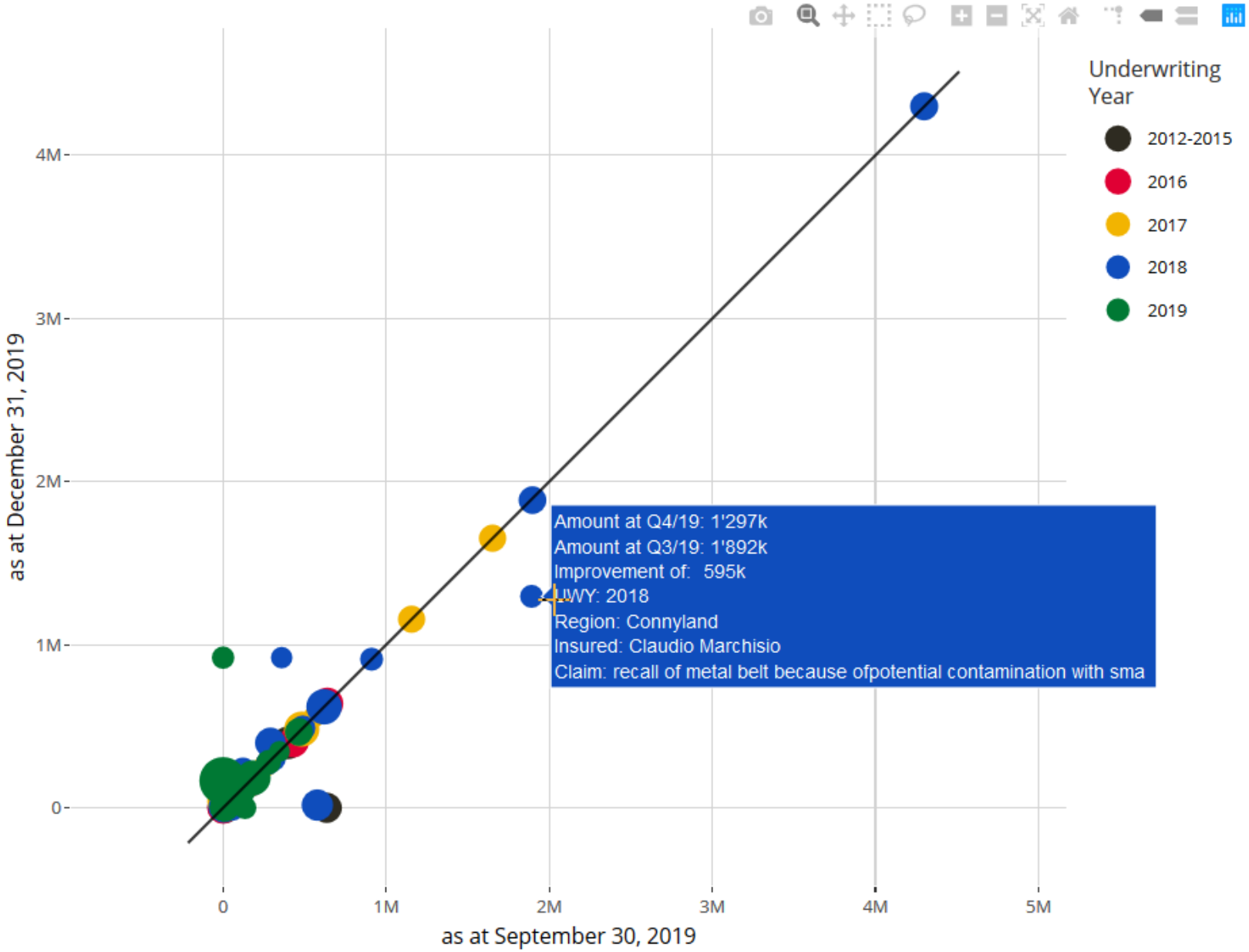


Figure 5.1: Reported Losses as at prior and current quarter

With R you can build web applications with Shiny

Thus, the question to be asked is:

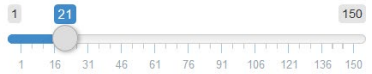
Can we incorporate Shiny in an R Markdown html document?

Mark meets Shiny

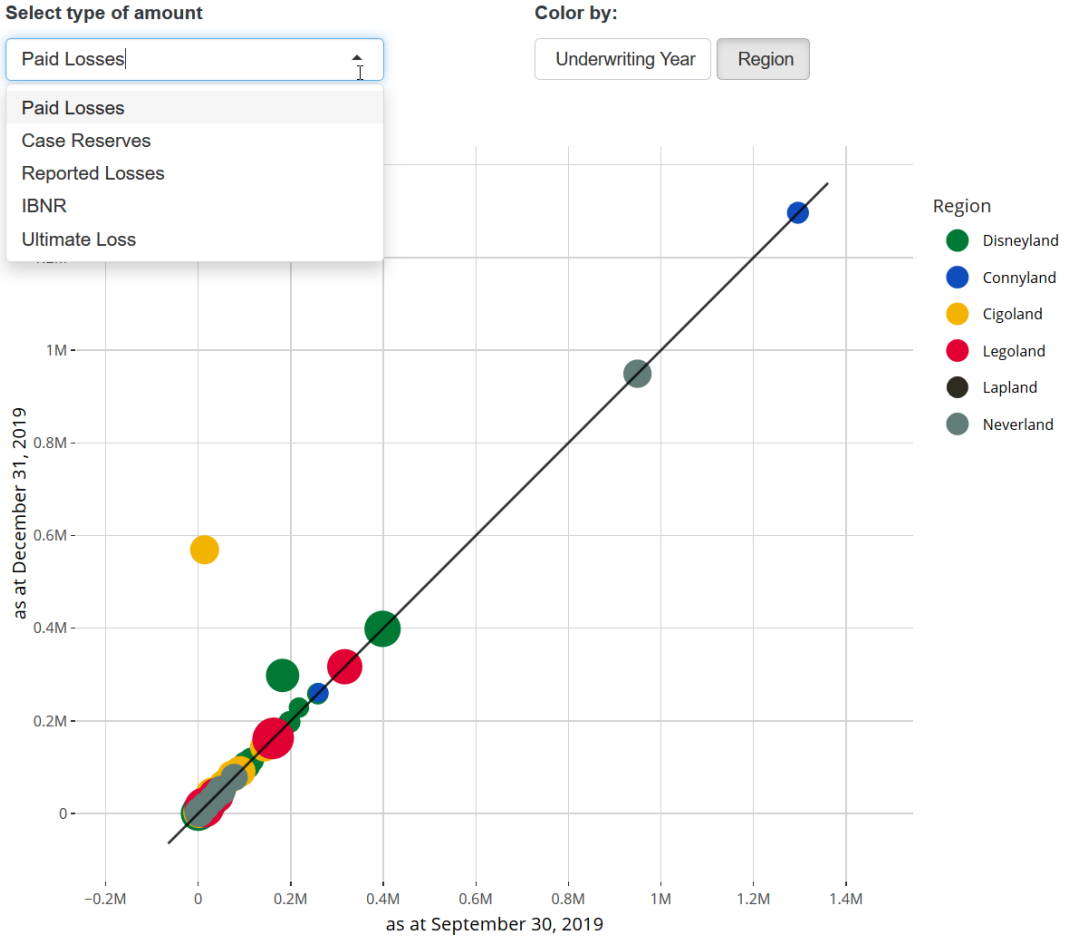
Add in the YAML section runtime: shiny

```yaml
---
author: "Claudio Rebelo"
date: "`r format(Sys.Date(), '%B %d, %Y')`"
output:
   bookdown::html_document2:
     toc: true
     toc_float: true
     toc_depth: 3
     collapsed: false
     number_sections: true
     css: template.css
     fig_caption: true
editor_options:
  chunk_output_type: console
runtime: shiny
---
```

Select number of words to display



3.1 Quarterly movementy by Paid Losses



Note that:

When you add a shiny component to your R Markdown document, the file can no longer be saved locally neither can it be shared as a stand-alone file

Interactive documents require a server side

You can share the file like any other shiny app

My two cents about adding Shiny to an R Markdown report

Keep Shiny components fairly simple

A report is like a story: you – as the writer – should be in control of the narrative

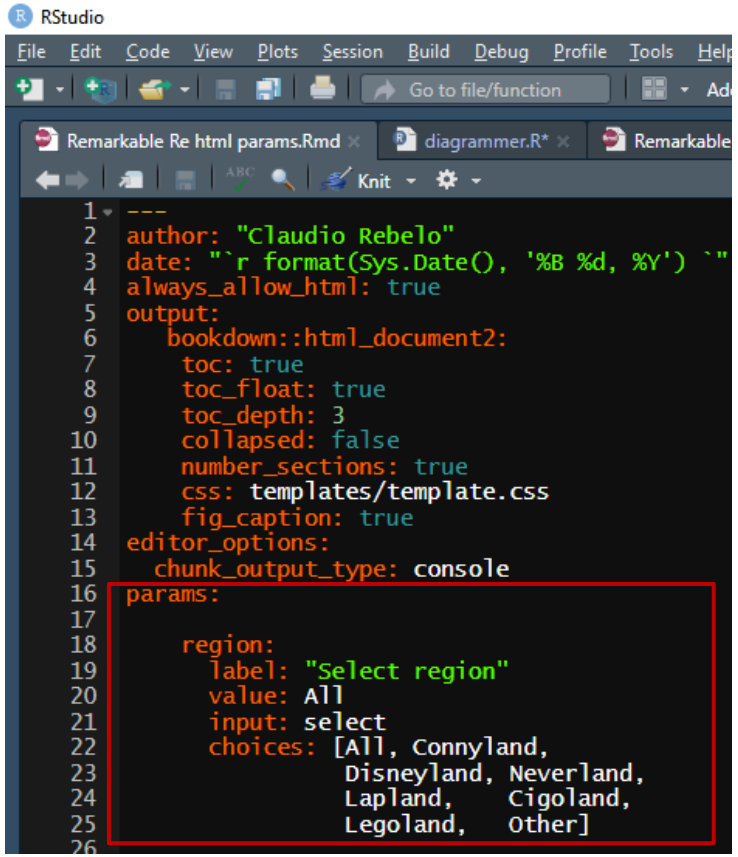
Too many Shiny components is likely to distract the reader from the message you are trying to convey: it should still be a report and not a dashboard

As a suggestion: add a Shiny component (e.g.: dashboard) in the appendix

Parameterized reports
The road to full Automation

Reach the next level of automation with paramaterized reports

Add params to the YAML header (line 16)



Adjust your code

replace the variable with params$region

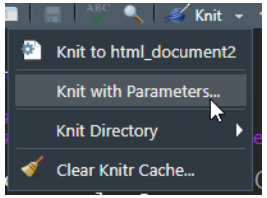
Before:

```
filter(df_RR, Region == "Cigoland")
```

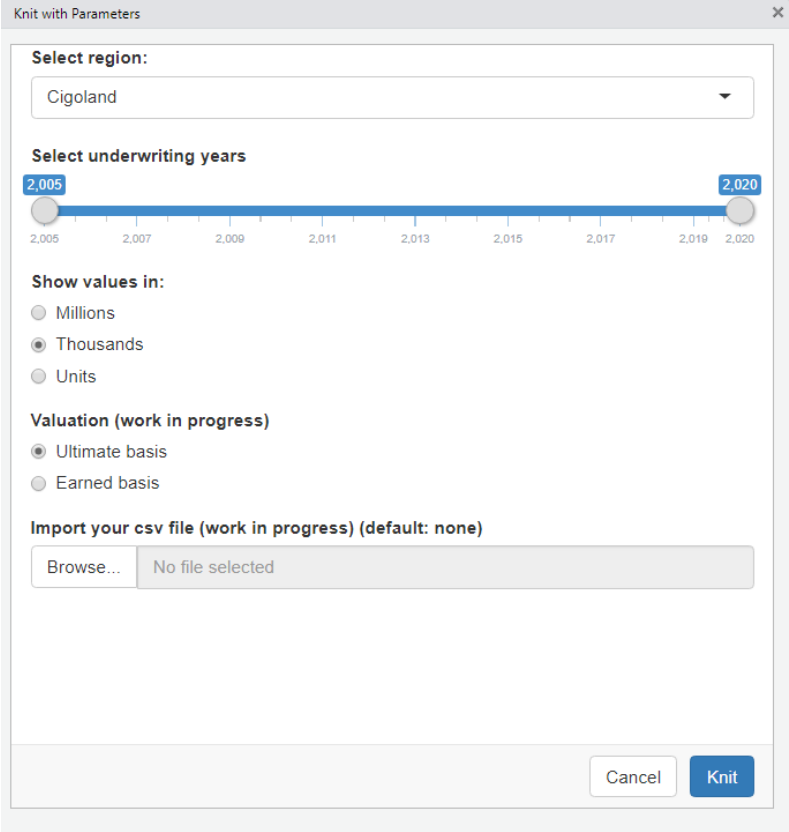
After:

```
filter(df_RR, Region == params$region)
```

Knit with Parameters



Knit with Parameters



The document is now a regional Cigoland report

2 Cigoland's Technical Results

2.1 Quarterly results

The table below displays the quarterly technical result as of December 31, 2019.

Note that all values are acording to their financial impact for example, negative losses represent an increase in loss amount.

Table 2.1: Quarterly Technical Result in USD thousands

	Underwriting Year		
USD thousands	2019	2018 & prior	Total
Premium:			
- Written Premium	11k	0k	11k
- Acquisition Costs	0k	0k	0k
Net Premium	11k	0k	11k
Losses:			
- Paid Losses	-1k	-602k	-604k
- Case Reserves	10k	941k	951k
- IBNR	-10k	343k	333k
Ultimate Losses	-1k	682k	681k
Technical Results	10k	682k	691k

The book of business experienced a positive result of 691k for contract years 2012 to 2019.

The result is mainly driven by improvement in Reported Losses and by IBNR releases .

▼ Description of main quarterly loss movements (click arrow to expand)

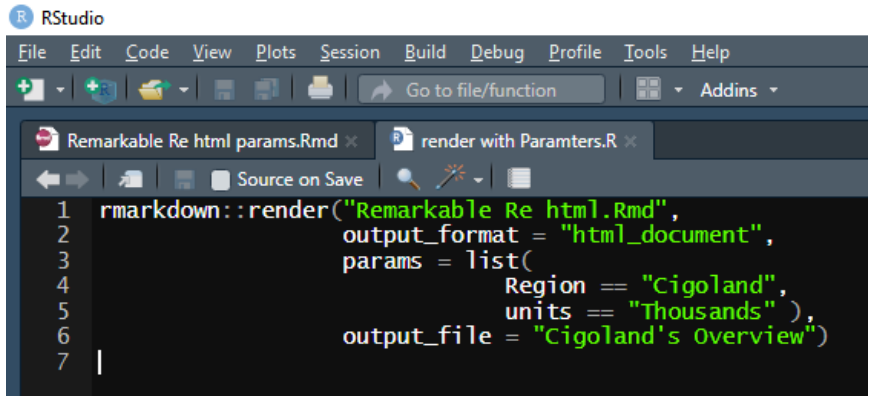
```
Underwriting year 2019:

  • USD +11k loss improvement on Paul Pogba (total to date: 0k, Region: Cigoland,
    Industry: Automotive, Entity: Remarkable Solutions)
```

Can we automate the workflow even further?

Can we automate the workflow even further?

YES!

Use the 'params' argument in the 'render' function to pass the parameters as a list



Next step: create a function inside a 'for loop'

to run as many reports as needed in one go!

The road to Automation

Follow the BBC approach

Develop dedicated corporate packages for:

- Corporate Branding
- Data Extraction
- Visualizations
- ...

To facilitate the implementation of the "R Markdown" workflow

https://bbc.github.io/rcookbook/

With RStudio Connect

RStudio Connect is a paid publishing platform

Each parameterized report can be saved on the platform.

You can add a mailing list to each saved report and instruct RStudio Connect to run and email the reports according to a predefined schedule!

That is full automation!

Faster & Better

From the Messy Workflow to the R Markdown approach we created reports:

- Much faster
- With less resources
- and more importantly, delivered a much better final product!

If cars can drive themselves, so can your reports

Claudio Rebelo,
Actuary, Swiss Re.

Legal notice