



Data Science with R and the PI System

Sjoerd Boersma

Data Scientist

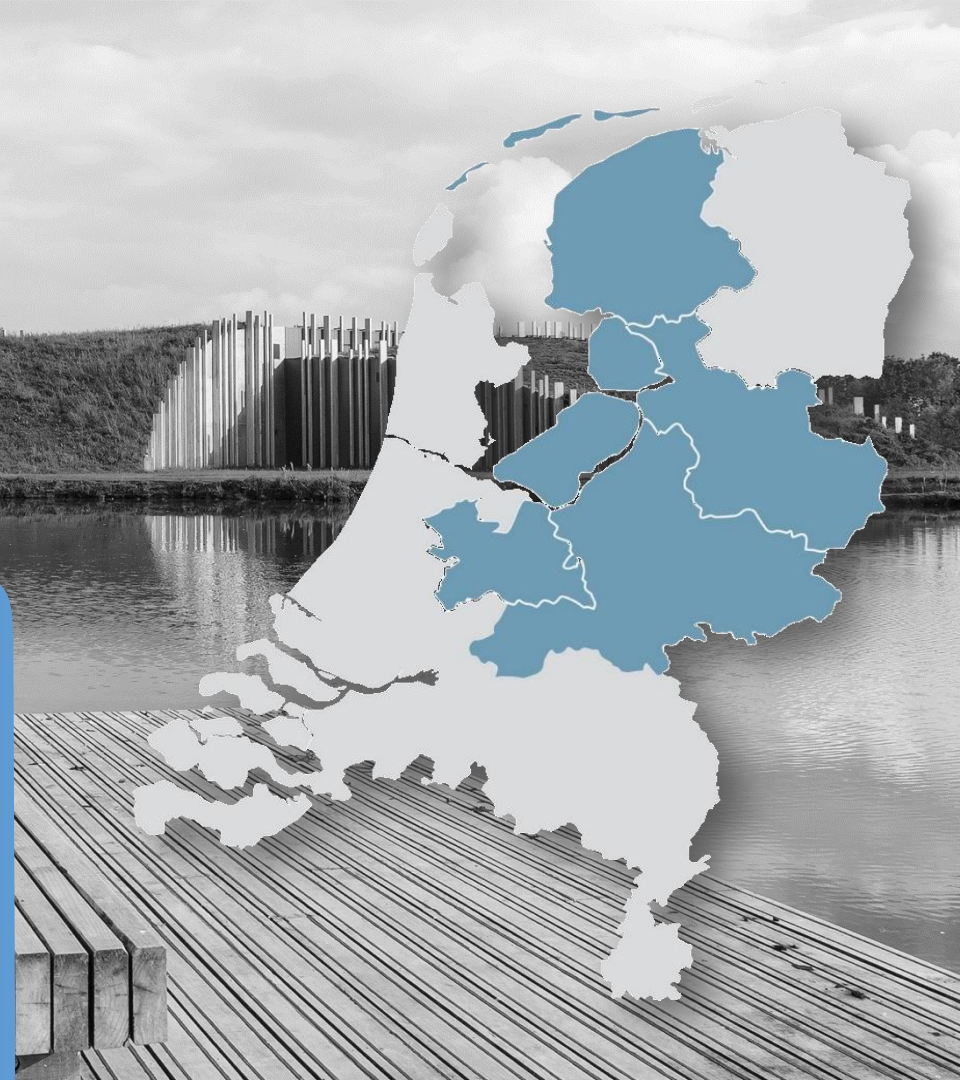




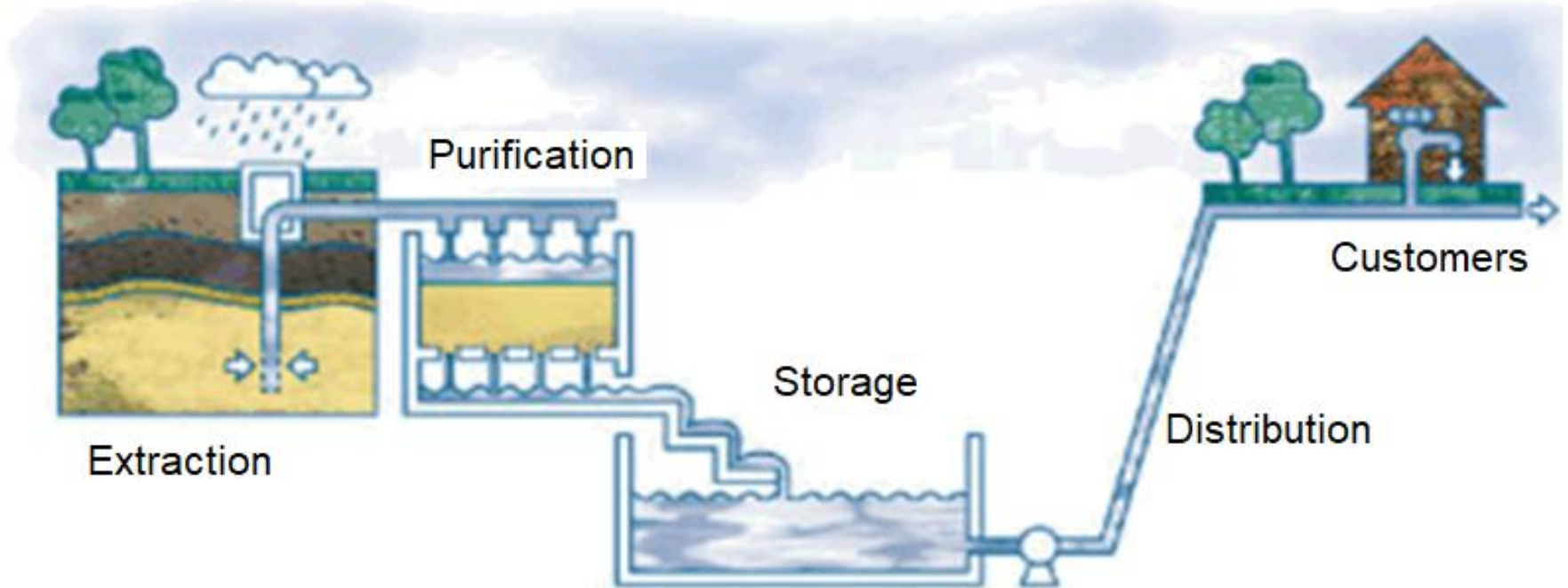
LAAT WATER VOOR JE WERKEN

Vitens in numbers

5	provinces
100	production sites
1,400	employees
50,000	km water mains
5.7 mln.	customers
100 mln.	€ investments
360 mln.	m ³ drinking water per year



Drinking water process



Data Science with R and PI

Data Science

- Collection of methods to use the scientific method, statistics and algorithms to extract knowledge and insights from data

R

- A programming language popular amongst data scientists and statisticians

PI

- ... needs no further introduction

Goal of this presentation

Help you understand:

- Why to use R for Data Science on PI-data
- How to use R for Data Science on PI-data

Show you:

- How easy it is to use R for Data Science on PI-data

Why R?

- Data analysis, data handling, statistics, machine learning
- Is R always the best solution? NO!
- Benefits of R:
 - Open source / community
 - Reproducibility
 - Flexibility

Overview

1. Introductions ✓
2. Getting Data
3. Using Asset Framework
4. Writing data back to PI
5. Doing **Data Science**
 - 2 examples

How to get your data out of the PI System and into R using the PI Web API

GETTING DATA

Getting data

```
library(piwebapi)
```

```
pi = piwebapi$new(pi_server_address, FALSE, username, password, FALSE, TRUE)
```

```
data = pi$data$getInterpolatedValues(path = 'pi:\\\\SR-16635\\FR-MOWE-PT01-meetwaarde',  
                                     startTime = '2018-07-01',  
                                     endTime = '2018-08-01',  
                                     interval = '15m')
```

	timestamp	value	unitsAbbreviation	good	questionable
1	2018-06-30T22:00:00Z	322.4664		1	0
2	2018-06-30T22:15:00Z	331.2576		1	0
3	2018-06-30T22:30:00Z	334.9206		1	0
4	2018-06-30T22:45:00Z	334.7985		1	0
5	2018-06-30T23:00:00Z	337.9731		1	0
6	2018-06-30T23:15:00Z	341.514		1	0

.....

Getting data

```
data = data %>%  
  select(timestamp, value) %>%  
  mutate(timestamp = ymd_hms(timestamp, tz = 'Europe/Amsterdam'))
```

	timestamp	value
1	1-7-2018 00:00	322.4664
2	1-7-2018 00:15	331.2576
3	1-7-2018 00:30	334.9206
4	1-7-2018 00:45	334.7985
5	1-7-2018 01:00	337.9731
6	1-7-2018 01:15	341.514

.....

Getting data

```
get.interpolated = function(tag, startTime = '2018-07-01', endTime = '2018-08-01', interval = '15m') {  
  path = paste0('pi:\\\\SR-16635\\', tag)  
  
  data = pi$data$getInterpolatedValues(path = path,  
                                       startTime = startTime,  
                                       endTime = endTime,  
                                       interval = interval)  
  
  data = data %>%  
    select(timestamp, value) %>%  
    mutate(timestamp = ymd_hms(timestamp, tz = 'Europe/Amsterdam')) %>%  
    rename_(.dots = setNames("value", tag))  
  
  return(data)  
}
```

```
get.interpolated('FR-MOWE-PT01-meetwaarde')
```

Getting data

```
list.of.tags      = list('FR-MHST-PT01-meetwaarde', 'FR-MAMN-PT01-meetwaarde', 'FR-MELE-PT01-meetwaarde')
list.of.timeseries = lapply(list.of.tags, get.interpolated)
dataframe        = list.of.timeseries %>% reduce(left_join, by = "timestamp")
```

	timestamp	FR-MHST-PT01-meetwaarde	FR-MAMN-PT01-meetwaarde	FR-MELE-PT01-meetwaarde
1	1-7-2018 00:00	296.4591	299.1453	369.3529
2	1-7-2018 00:15	295.7265	287.6679	364.8352
3	1-7-2018 00:30	295.4823	296.4591	364.1026
4	1-7-2018 00:45	297.4359	305.9829	364.7131
5	1-7-2018 01:00	301.0989	294.8718	371.3065
6	1-7-2018 01:15	296.5812	294.5055	369.3529

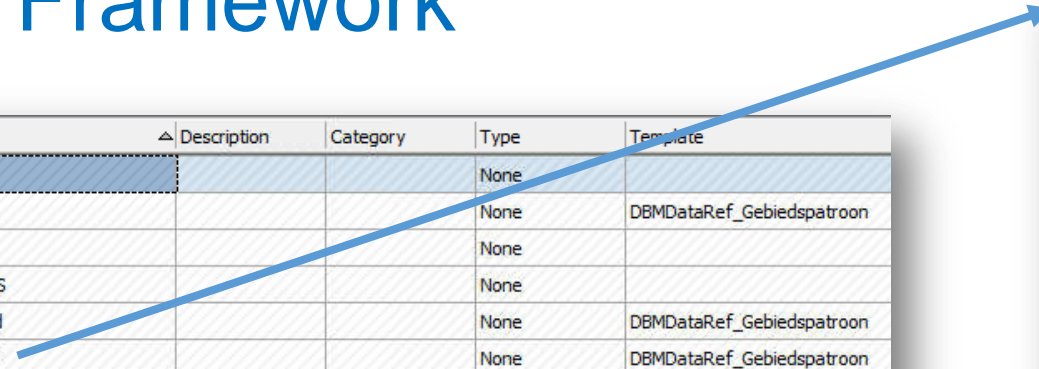
.....

How to load and explore PI Asset Framework-structures in R

ASSET FRAMEWORK

Asset Framework

Name	Description	Category	Type	Template
Configuration			None	
Vitens			None	DBMDataRef_Gebiedspatroom
000 TEST			None	
001 SPECIALS			None	
FLV Flevoland			None	DBMDataRef_Gebiedspatroom
FRL Friesland			None	DBMDataRef_Gebiedspatroom
GEN Gelderland-Noord			None	DBMDataRef_Gebiedspatroom
GEO Gelderland-Oost			None	DBMDataRef_Gebiedspatroom
GEZ Gelderland-Zuid			None	DBMDataRef_Gebiedspatroom
OVN Overijssel-Noord			None	DBMDataRef_Gebiedspatroom
OVZ Overijssel-Zuid			None	DBMDataRef_Gebiedspatroom
URM Utrecht-Randmeren			None	DBMDataRef_Gebiedspatroom
UTR Utrecht-Randmeren			None	DBMDataRef_Gebiedspatroom
UTW Utrecht-West			None	DBMDataRef_Gebiedspatroom
UTZ Utrecht-Zuid			None	DBMDataRef_Gebiedspatroom



FRL Friesland	
FRL-01 Vlieland	
FRL-02 Terschelling	
FRL-03 Ameland	
FRL-04 Schiermonnikoog	
FRL-05 Dokkum-Holwerd	
FRL-06 Buitenpost	
FRL-07 Leeuwarden	
FRL-07-01 Leeuwarden-stad	
FRL-07-01-A Westeinde	
Drukmeting MLDK	
Drukmeting MLJH	
Drukmeting MLKA	
FRL-07-01-B Bilgaard I	
FRL-07-01-C Bilgaard II	
FRL-07-01-D Lekkuum-Zwette	
FRL-07-01-E Camminghab...	
FRL-07-01-F Oranjewijk-Ni...	
FRL-07-01-G Aldlân	
FRL-07-01-H Hemrik	
FRL-07-01-I Goutum	
FRL-07-02 Hurdegaryp	
FRL-08 Grou-Franeker	
FRL-09 Sneek	
FRL-10 Joure-Heerenveen	
FRL-11 Gorredijk	
FRL-12 Oosterwolde-Drachten	

Asset Framework

Explore structure

- Children / parents
- Search full hierarchy

How to write data back into the PI System using the PI Web API

WRITING BACK TO PI

Writing back to PI

```
pi.value = PITimedValue(timestamp = as.character(now()), value = 42)
web.id = pi$data$convertPathToWebId('pi:\\\\SR-16635\\barcelona')
pi$stream$updateValue(webId = web.id, PITimedValue = pi.value)
```

```
pi$data$getRecordedValues(path = 'pi:\\\\SR-16635\\barcelona', startTime = "*-200d", endTime = "**")
```

	timestamp	value	unitsAbbreviation	good	questionable	substituted
1	2018-08-08T09:28:46Z	42		1	0	0

How to analyse, model, predict and plot in R using this PI-data

DATA SCIENCE

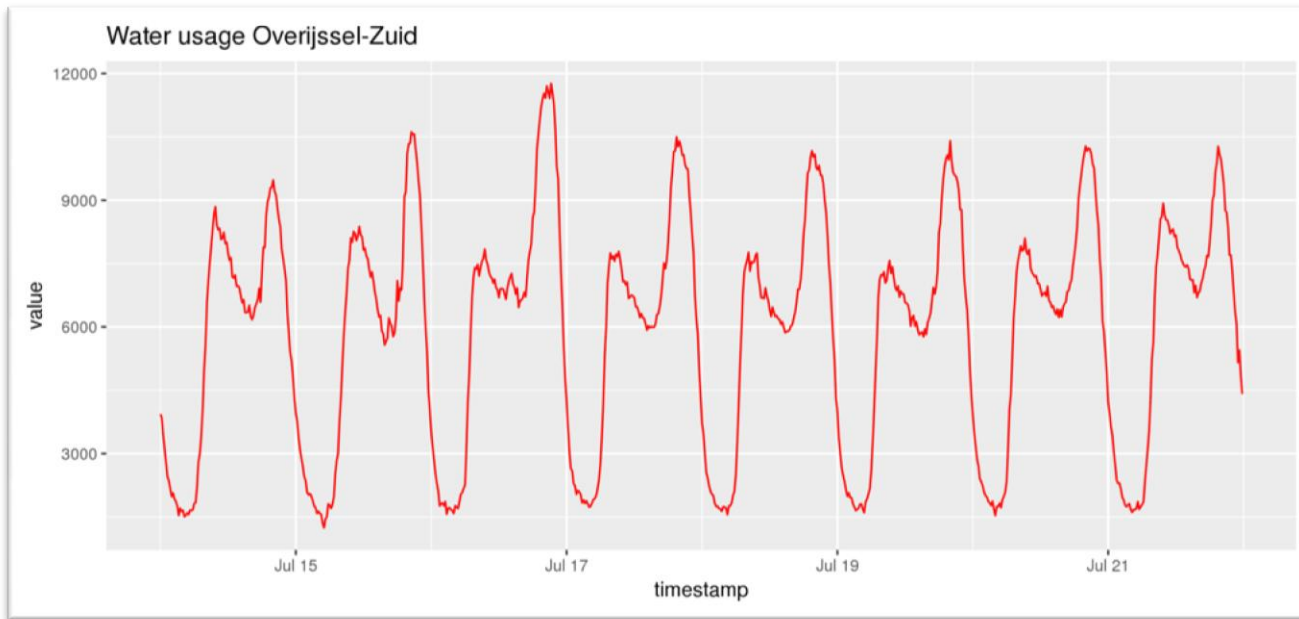
Data Science

Two examples:

- Night's Watch
- Water usage predictions

Night's Watch

Water usage during the night is a good indicator of small leaks



Night's Watch

Email

Rapport Nachtwacht v2

Dagelijkse Afwijkingen

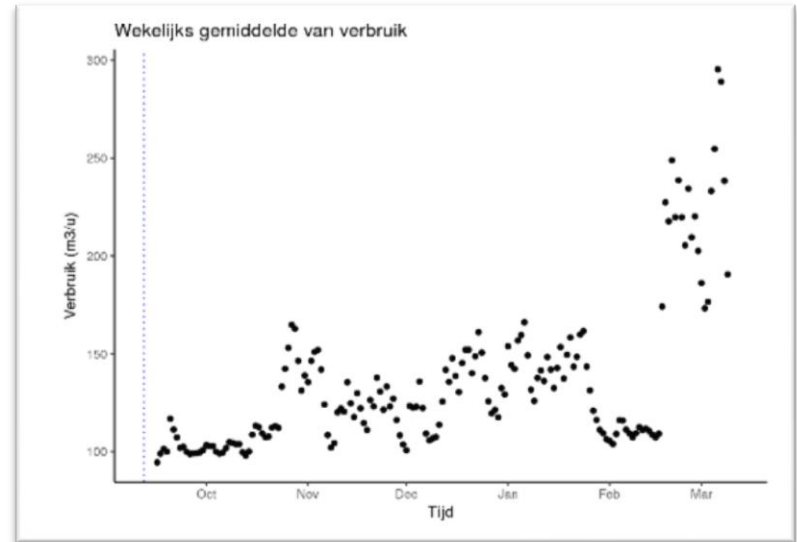
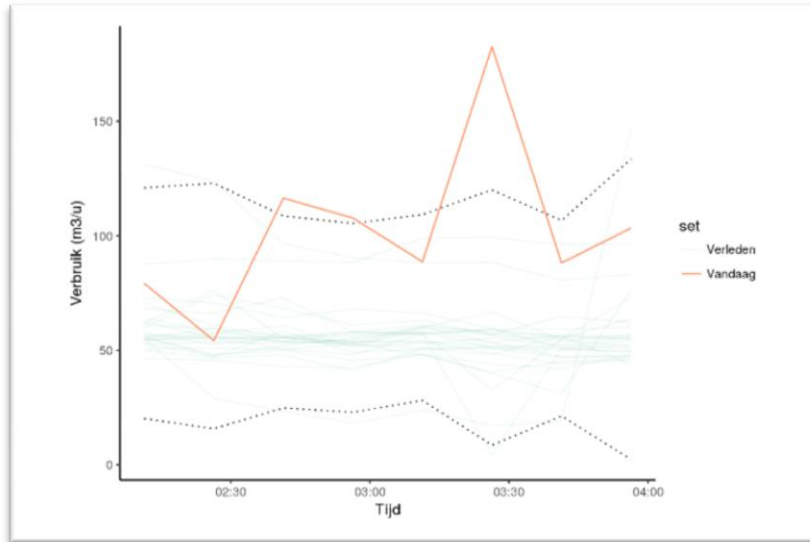
Gebied	Bovengrens Overschreden	Gemeten Gemiddelde Flow (m3/u)	Verwachte Gemiddelde Flow (m3/u)	Afwijking Gemiddelde Flow (m3/u)	P-waarde	Afwijkende punten
OVN-03-04 Vroomshoop-Kloosterhaar	Nee	102.5	66.9	35.6	0.005	3
OVN-03-02 Ommen	Nee	104.1	89.7	14.4	0.092	1
GEN-07-01 Apeldoorn	Nee	408.8	306.0	102.8	0.318	1
FRL-10-05 Munnekeburen	Ja	30.8	26.8	3.9	0.719	0
OVZ-01-01 Hellendoorn-Rijssen	Nee	168.6	123.9	44.8	0.862	1
UTW-04-01 IJsselstein	Nee	30.4	18.1	12.3	0.977	1
GEZ-02-01 Groesbeek-hoog	Nee	5.9	4.3	1.6	0.987	2

Toename in gemiddelde verbruik

Gebied	Vorige Gemiddelde (m3/u)	Nieuwe Gemiddelde (m3/u)	Toename (m3/u)
FRL-09-04 Heeg	-198.8	-35.9	162.9
OVZ-03 Stedenband	437.9	543.5	105.5
OVZ-03-03 Enschede lagezone	188.8	284.6	95.8
UTR-02 Putten	117.8	211.5	93.7
GEN-11-01 Putten	73.4	161.0	87.6
UTR-02-01 Putten	73.9	159.8	85.9
GEN-07 Apeldoorn	271.5	334.4	62.9

Night's Watch

Assessment still done by people, but less work



Water usage predictions

How will water usage change in the next 1-20 years?

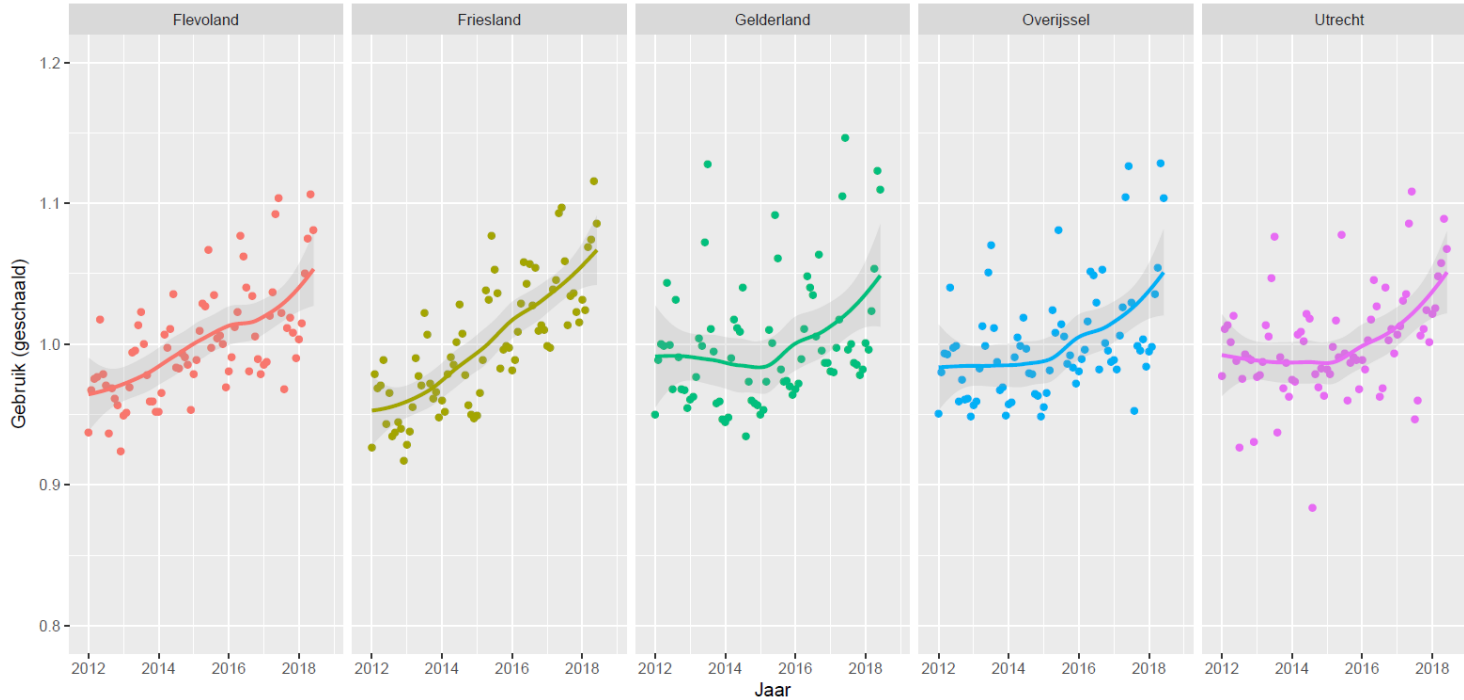
→ Great strategical importance for Vitens

Model water usage based on external data:

- Weather
- Holidays
- Economic growth
- Population growth
- ...

Water usage predictions

Watergebruik per provincie per maand



This presentation is almost over, just give me a few more minutes of your attention

CONCLUSIONS

Conclusions

Using **R** in combination with the **PI** Web API allows us to:

- easily get data from **PI**
- make good use of Asset Framework structures
- write data back into **PI**
- bring your **Data Science** projects to the next level

Conclusions

“Now that you have seen how easy it is,
there is no reason not to give it a try”

Wanna know more?



- Sjoerd Boersma
- Data Scientist
- Vitens N.V.
- Sjoerd.Boersma@vitens.nl

- Go to <https://github.com/osimloeff/PI-Web-API-Client-R>
- Visit <https://PISquare.osisoft.com> for community
- Visit the Data Science booth in the expo



KEEP
CALM
AND
USE
R

Questions?

Please wait for
the **microphone**




State your
name & company

Please rate this session in the mobile app!

**DOWNLOAD
THE MOBILE APP**

- Rate sessions and provide feedback
- Meet and connect with other attendees

Join the conversation and SHARE
what you saw #osisoft #piworld

A smartphone displaying the OSiSoft PIWorld app interface. The screen shows the OSiSoft logo (a stylized atom) and the text "OSiSoft PIWorld".

謝謝 KEA LEBOHA
 TAPADH LEIBH 고맙습니다
 БАЯРЛАЛАА MISAOTRA ANAO
 DZIĘKUJĘ CI NGIYABONGA TEŞEKKÜR EDERIM GRACIES OBRIGADO شكرا SALAMAT
 DANKON TANK TAPADH LEAT
 KÖSZÖNÖM DANKIE TERIMA KASIH GRACIES
 СПАСИБО
 ПАКМЕТ СІЗГЕ
 GO RAIBH MAITH AGAT
 БЛАГОДАРЯ GRACIAS MAHADSANID
 ТИ БЛАГОДАРАМ
 ТАК DANKE
 RAHMAT MERCI
 HATUR NUHUN
 GRACIAS TIBI
 DANK JE EΥΧΑΡΙΣΤΩ GRATIAS TIBI
 АЇЎ SALAMAT MAHALO IĀ 'ŌE TAKK SKALDU HA
 GRAZZI ПAKKA ПЕР
 ПАХМАТ САГА
 FALEMINDERIT
 ありがとうございました
 SIPAS JI WERE TERIMA KASIH
 MATUR NUWUN
 MATUR NUWUN
 UA TSAUG RAU KOJ
 ТИ БЛАГОДАРАМ
 СИПОС
 CẢM ƠN BẠN
 WAZVIITA
 MULŢUMESC
 HVALA FAAFETAI
 ESKERRIK ASKO
 HVALA ХВАЛА ВАМ
 TEŞEKKÜR EDERIM
 GRAZIE
 DI OU MÈSI
 ĐAKUJEM



OSIsoft.

PIWorld

THANK YOU