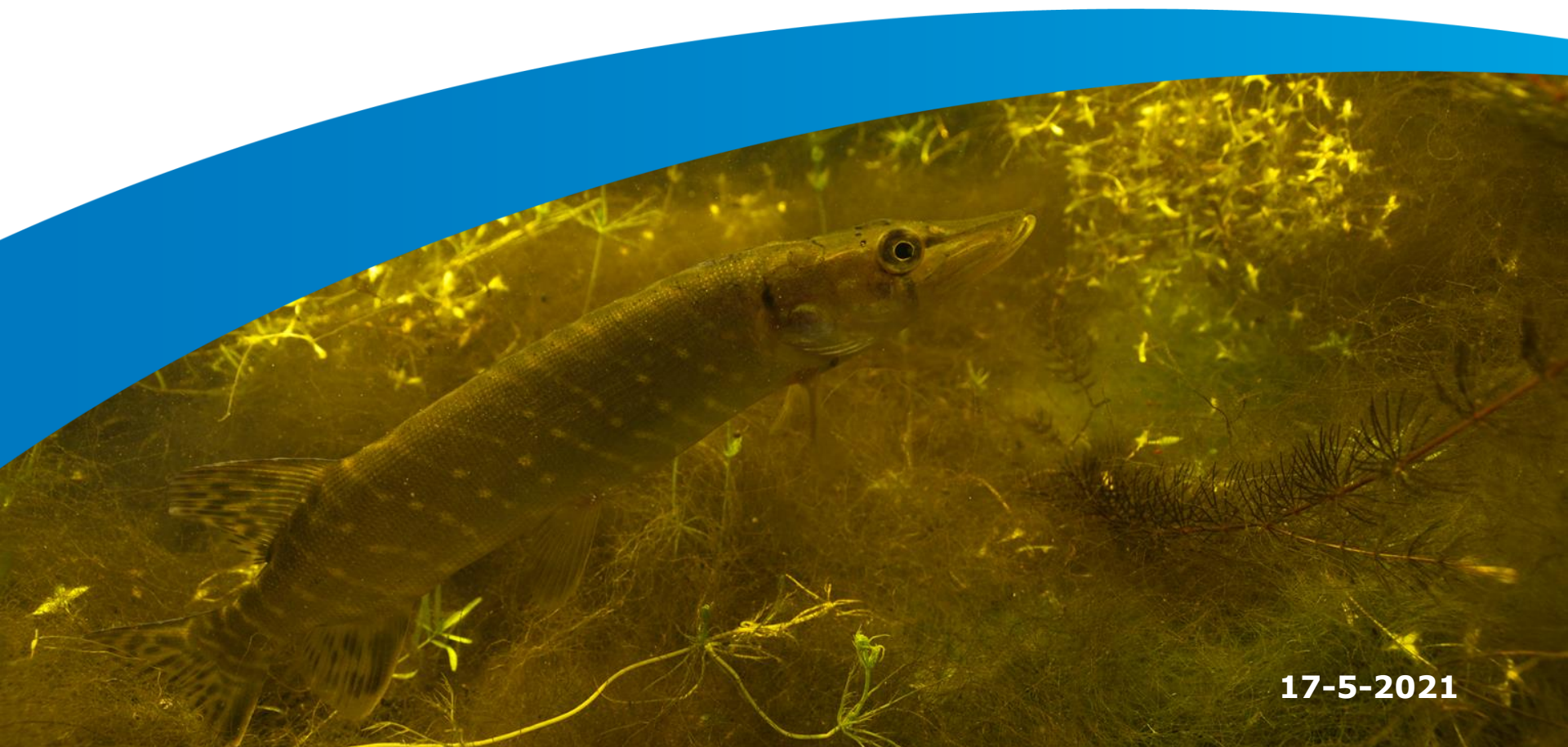


# Analyseren kan je leren

CoP Ecologische Data-analyse



17-5-2021

# Data-analyse

—

## Hoe leer je dat?

# 1. Iets over mijn persoonlijke ontdekkingsstocht

→ Lang, met omwegen en doodlopende wegen

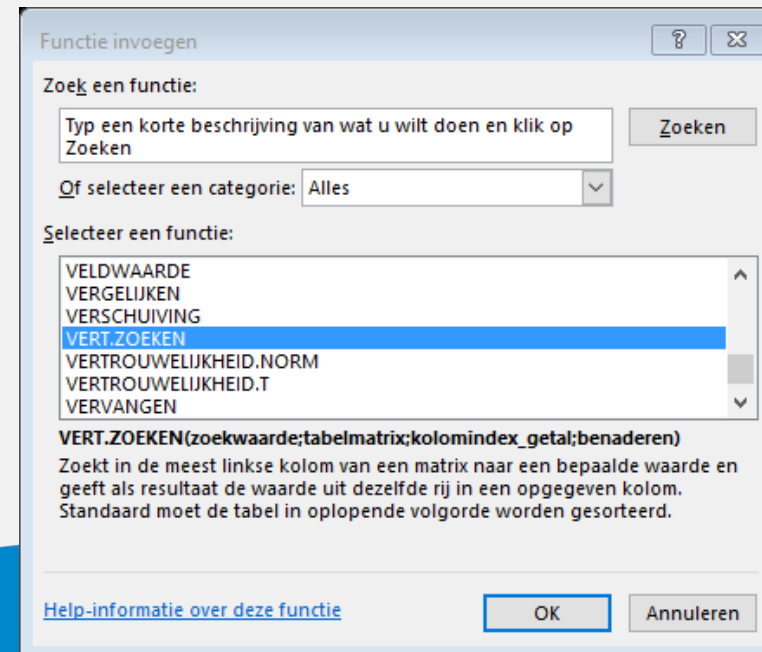
# 2. Wat ik daarvan geleerd heb

## Stage WSHD (2007)

- Bladeren door het maatlattenboek op zoek naar soorten
- Handmatig beoordelen van waterlichamen in Excel

### Nieuwe data-analyse vaardigheden

- Wat ik zelf had ontdekt in Excel
- VERT.ZOEKEN()

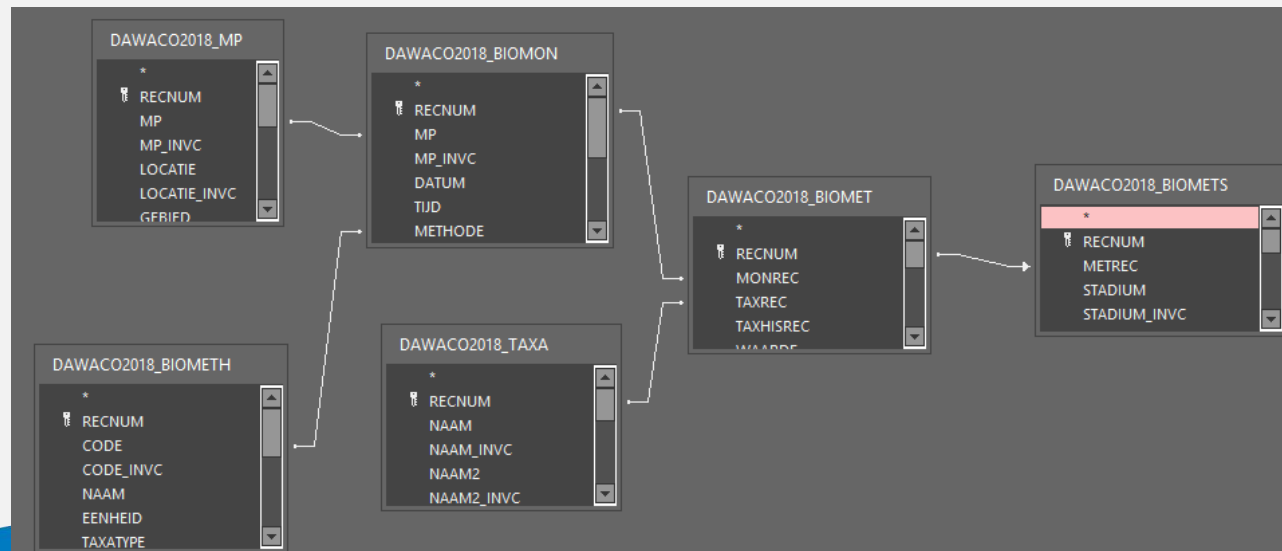


# Afstuderen WUR (2008)

- Analyse data uit de Limnodata Neerlandica
- Gigantische Excel spreadsheets
  - 65536
  - lastig om na een week de spreadsheet nog te begrijpen

## Nieuwe data-analyse vaardigheden

- Koppelen van tabellen\*
- Relationale database (MS Access)



## HHSK (2009 - heden)

- Monitoring, databeheer en analyse

### Nieuwe data-analyse vaardigheden

- Ontdekking van draaitabellen in Excel (2009)\*
- Ontdekking van VBA in Excel (2011)\*
  - macro's
  - zelfgeprogrammeerde sneltoetsen
- SPSS (2011)\*
  - nooit echt een succes
- Canoco (2014)\*
  - erg specifiek

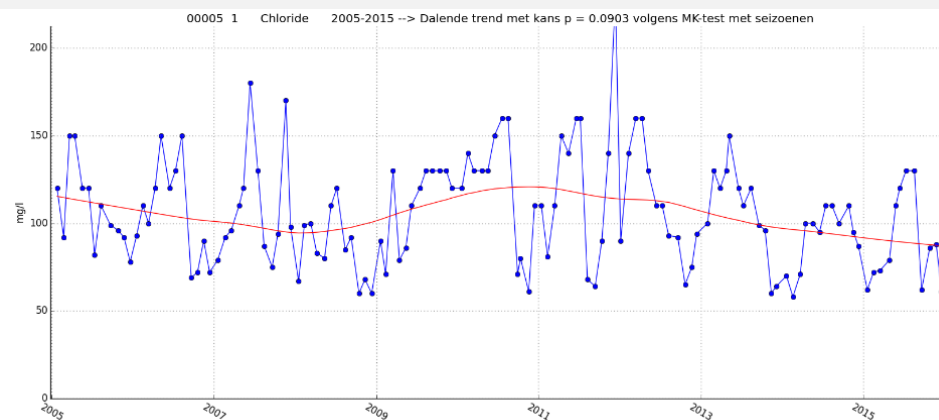
# Ontdekking van programmeren: Python\*

- ik was er niet erg goed in (weet ik nu)
- het voelde alsof ik superkrachten had



## Nieuwe data-analyse vaardigheden

- Grote hoeveelheden data in één keer
  - handwerk was minder beperkend
- 12000 grafieken in 12 uur... 😊
  - maar beter dan met de hand in Excel





## Programmeren in R

- Voelde voor mij veel natuurlijker dan Python
- Cursussen op edX en Coursera hielpen me op weg
  - veel online resources, zie [link](#) of [link](#)

## Nieuwe data-analyse vaardigheden

- Veel data efficiënt en automatisch verwerken op talloze manieren
- R is mijn Zwitsers zakmes
  - Het kan bijna alles
  - GIS, visualisaties, rapportages, interactieve apps etc.



# Mijn lessen voor data-analyse

# 1. Kennis van de mogelijkheden

Al mijn data-analyse kennis is terug te voeren op iemand (of een boek of artikel) die mij informeerde over de mogelijkheden

- Er is veel kennis al beschikbaar  
Maar hoe weet je wat er kan?
- Zoek iemand die je verder kan helpen!  
(bijv. in de CoP EDA)

## 2. Programmeren is een superpower

- Laat de computer het werk voor je doen  
Een beetje luiheid kan geen kwaad
- Je moet eerst investeren maar het verdient zich terug  
Er zijn prima goedkope/gratis online cursussen



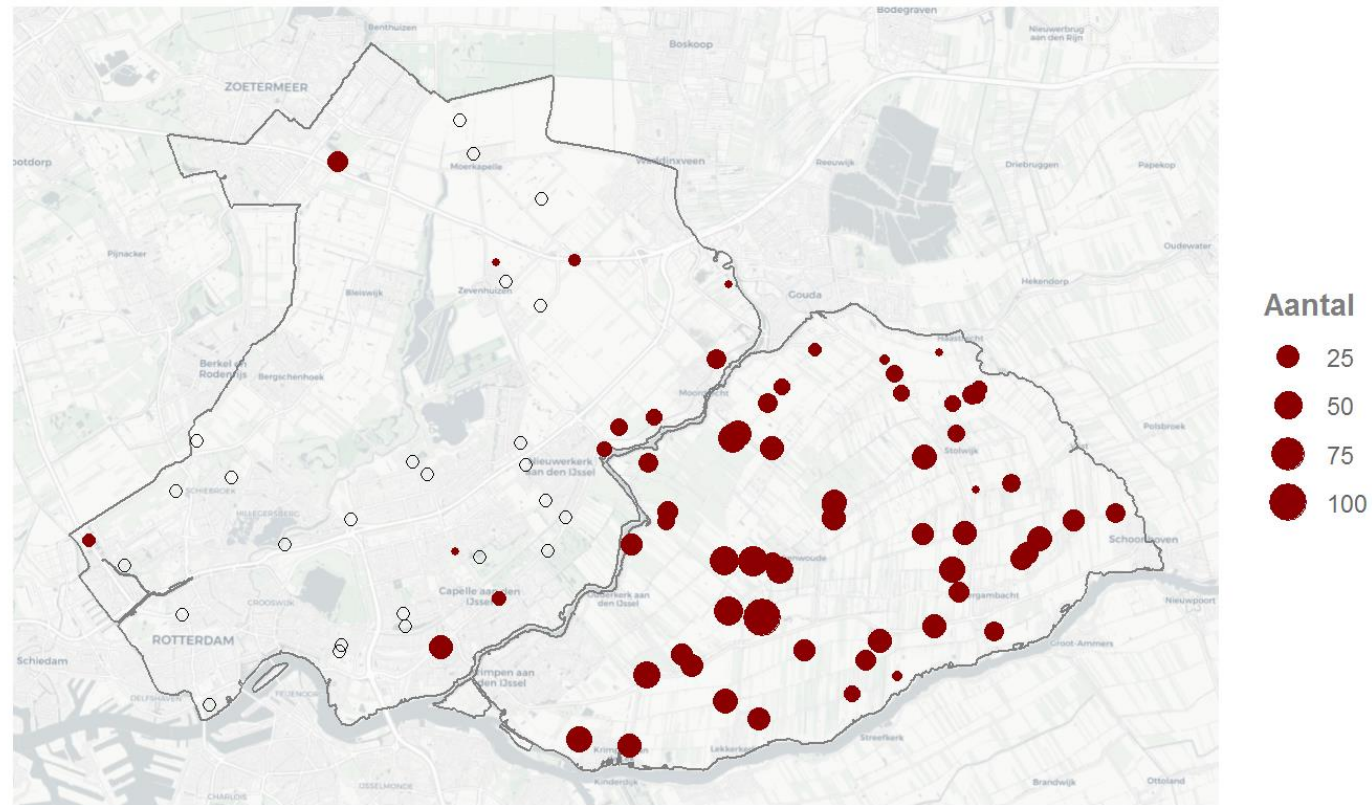
### 3. Snelheid is belangrijk

- Als iets (te) veel tijd kost doe je het niet
- Snelheid in data-analyse opent nieuwe mogelijkheden
- Nog een reden om te leren programmeren!

## 4. Maak het zichtbaar

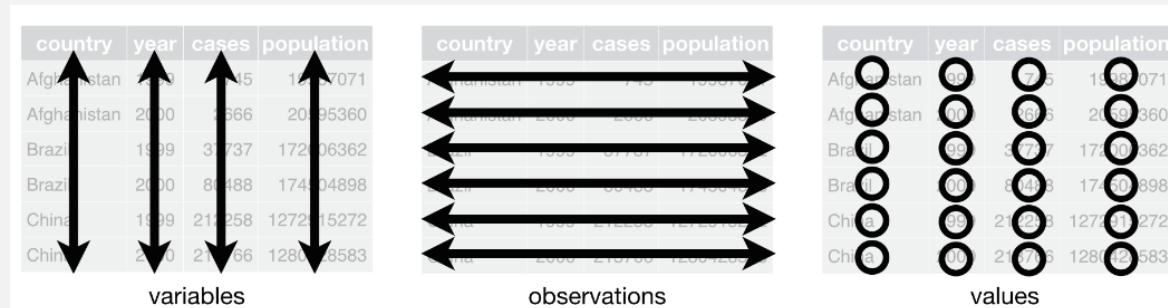
- Een plaatje zegt meer dan 1000 getallen.

Aantal gevangen kreeften



# 5. Zorg voor structuur in je data

- Het netjes maken van data kost vaak het meeste tijd
- Leer hoe een nette dataset is opgebouwd  
zie bijv. tidy data



*"Happy families are all alike; every unhappy family is unhappy in its own way."*  
Leo Tolstoy

*"Tidy datasets are all alike, but every messy dataset is messy in its own way."*  
Hadley Wickham

# Data-analyse leer je ook door te doen

# Data-analyse leer je door te doen

Idealiter met een dataset samen achter een computer

Dat gaat helaas niet

In plaats daarvan...

- een gezamenlijke dataset
- zelf analyseren
- resultaten delen op volgende CoP



# Toelichting dataset

## Kreeftenonderzoek HHSK 2020

- Ze hebben zich behoorlijk uitgebreid
- Brengen schade toe aan het ecosysteem
- Weinig kennis over verspreiding en dichtheden

# Toelichting dataset

## Kreeftenonderzoek HHSK 2020

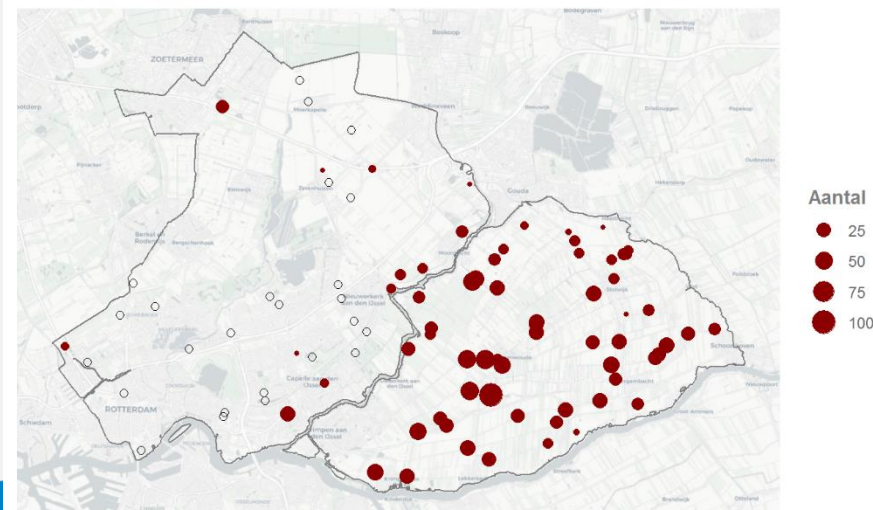
### wijze van onderzoek

- 12 kreeftenkorven met aas
- 1 nacht
- 94 onderzoekslocaties – hele beheergebied
- ook vegetatieonderzoek

### resultaten

- beschreven in korte rapportage  
zie [link](#)

Aantal gevangen kreeften



# Toelichting dataset

Excelbestand met 5 tabbladen  
zie [link](#)

data\_meetpunten

data\_kreeften

data\_kreeften\_detail

data\_vegetatie\_soorten

data\_vegetatie\_kenmerken

# Toelichting dataset

data\_meetpunten

	A	B	C	D	E	F	G
1	mp	x	y	gebied	bodemtype	bodemtype_detail	gemeente
2	QK_001	102283	434749	Krimpenerwaard	Veen	Waardveengronden op bosveen (of eutroof broekveen)	Krimpenerw.
3	QK_013	103699	434573	Krimpenerwaard	Klei	Kalkarme drechtvaaggronden	Krimpenerw.
4	QK_019	107339	435306	Krimpenerwaard	Veen	Weideveengronden op bosveen (of eutroof broekveen)	Krimpenerw.
5	QK_022	113914	437789	Krimpenerwaard	Klei	Kalkloze drechtvaaggronden	Krimpenerw.
6	QK_028	111290	444439	Krimpenerwaard	Veen	Weideveengronden op bosveen (of eutroof broekveen)	Krimpenerw.
7	QK_031	107678	442900	Krimpenerwaard	Veen	Koopveengronden op zeggeveen, rietzeggeveen of (mesotroof) b	Krimpenerw.
8	QK_034	111966	442629	Krimpenerwaard	Veen	Koopveengronden op bosveen (of eutroof broekveen)	Krimpenerw.
9	QK_037	112864	443298	Krimpenerwaard	Veen	Koopveengronden op bosveen (of eutroof broekveen)	Krimpenerw.

# Toelichting dataset

data\_kreeften

	A	B	C	D	E	F
1	mp ▾	n_kreeften_totaal ▾	procambarus_clarkii ▾	orconectes_virilis ▾	procambarus_acutus ▾	orconectes_limosus ▾
2	QK_001	40	40	0	0	0
3	QK_013	32	32	0	0	0
4	QK_019	26	26	0	0	0
5	QK_022	18	15	2	1	0
6	QK_028	10	10	0	0	0

# Toelichting dataset

data\_kreeften\_detail

	A	B	C	D	E	F	G
1	mp ▾	datum ▾	tijd ▾	naam ▾	lengte ▾	waarde ▾	eenheid ▾
2	QK_001	25-5-2020	12:20:00	Procambarus clarkii	8	2	n
3	QK_001	25-5-2020	12:20:00	Procambarus clarkii	9	3	n
4	QK_001	25-5-2020	12:20:00	Procambarus clarkii	10	9	n
5	QK_001	25-5-2020	12:20:00	Procambarus clarkii	11	16	n
6	QK_001	25-5-2020	12:20:00	Procambarus clarkii	12	9	n

# Toelichting dataset

data\_vegetatie\_soorten

	A	B	C	D	E	F
1	mp ▾	datum ▾	taxatype ▾	naam ▾	waarde ▾	eenheid ▾
2	QK_001	14-7-2020	MACFT	Angelica sylvestris	0,5	%
3	QK_001	14-7-2020	MACFT	Epilobium hirsutum	0,5	%
4	QK_001	14-7-2020	MACFT	Glyceria maxima	0,5	%
5	QK_001	14-7-2020	MACFT	Lemna minor	0,5	%
6	QK_001	14-7-2020	MACFT	Mentha aquatica	0,5	%
7	QK_001	14-7-2020	MACFT	Myosotis scorpioides ssp. scorpioides	0,5	%

# Toelichting dataset

data\_vegetatie\_kenmerken

	A	B	C	D	E	F
1	mp	datum	taxatype	kenmerkcode	kenmerknaam	kenmerk_waarde
2	QK_001	14-7-2020	MACFT	BDKDRAADALG	Bedekking draadalgen [%]	0
3	QK_001	14-7-2020	MACFT	BDKDRIJFBLAD	Bedekking drijfbladvegetatie [%]	0
4	QK_001	14-7-2020	MACFT	BDKEMERS	Bedekking emerse laag vegetatie [%]	0,5
5	QK_001	14-7-2020	MACFT	BDKFLAB	Bedekking flab of darmwier [%]	0
6	QK_001	14-7-2020	MACFT	BDKKROOS	Bedekking kroos of kroosvaren [%]	0,5
7	QK_001	14-7-2020	MACFT	BDKSUBMERS	Bedekking submerse laag vegetatie [%]	0
8	QK_001	14-7-2020	MACFT	BREEDTE	Waterbreedte in m	6
9	QK_001	14-7-2020	MACFT	OEVERSOORT	Oeversoort	Natuurlijk
10	QK_001	14-7-2020	MACFT	SCHADUWPERC	Beschaduwning in %	0



# Analyse kreeftendata

## Stappen

- Grasduinen door de data
  - Probeer een beeld te krijgen van de data
- Formuleer een onderzoeksvraag
  - Een goede vraag helpt je op weg
- Probeer je vraag te beantwoorden
- Hoe kan je je resultaat delen?
  - Bijv. tekst of grafieken of een kaart of...

# Analyse kreeftendata

## Suggesties voor analyse

Welke ruimtelijke variatie is er:

- in de verspreiding van kreeften
- in de verspreiding van planten

Is er een relatie tussen:

- kreeftendichtheid en vegetatiedichtheid
- kreeftendichtheid en plantensoorten
- kreeftendichtheid en gebiedskenmerken (bodemtype, landgebruik etc.)

# Analyse kreeftendata

Het gebruik van andere databronnen wordt aangemoedigd!

- **KNMI-data**

Er is een uitschieter in de kreeftenvangst, is die te verklaren door de weersomstandigheden?

- **Andere data van HHSK**

vegetatie in eerdere jaren of chemie  
via [hhsk.nl](http://hhsk.nl) – zie [link](#)

- **Resultaten uit andere kreeftenonderzoeken**

- **Etc.**

# Links

Link naar [dataset kreeften](#)

zie ook: [https://github.com/RedTent/CoP\\_EDA](https://github.com/RedTent/CoP_EDA)

Kreeftenrapportage 2020

[https://www.schielandendekrimpenerwaard.nl/kaart/waterkwaliteit/waterkwaliteitsrapportages/kreeften\\_2020/](https://www.schielandendekrimpenerwaard.nl/kaart/waterkwaliteit/waterkwaliteitsrapportages/kreeften_2020/)

Leren programmeren in R

[https://education.rstudio.com/learn/beginner/Getting\\_Started\\_with\\_R\\_and\\_Rstudio](https://education.rstudio.com/learn/beginner/Getting_Started_with_R_and_Rstudio)

Online boeken over data-analyse en visualisatie

[https://redtent.github.io/my\\_books/](https://redtent.github.io/my_books/)

Vragen? Mail me gerust!

[j.van.tent@hhsk.nl](mailto:j.van.tent@hhsk.nl)