

# FANS Analytics Netværksmøde

25/11-2020 – på Teams

25/11-2020 – på Teams

FANS  
Analytisk  
Netværksmøde

# Dagsorden

- **Velkomst**
  - v. Frans Holm, SAS
- **Analytiske Nyheder i SAS**
  - v. Rune Hjort Nielsen & Rebecca Grüner Hansen, SAS
- **Nyt Nordisk Data Science netværk**
  - v. Frans Holm, SAS
- **Tolkning og Bias i data – Opfølgning fra sidst**
  - v. Rune Hjort Nielsen, SAS
- **Tolkning og Bias i data – Case**
  - v. Ulrik Gerdes, læge, uafhængig dataknuser
- **Pause - Pause**
- **The Analytics Use Case Canvas (in English)**
  - v. Erik Strömngren, SAS
- **Opdag abnormiteter ved hjælp af tidsrækkeanalyse**
  - v. Rebecca Grüner Hansen, SAS
- **Afslutning & lodtrækning**
  - v. Frans Holm, SAS

# Styregruppe

## Styregruppe:

- Louise Boysen (Steno Diabetes Center)
- Rune Hjort Nielsen (SAS)
- Rebecca Grüner Hansen (SAS)
- Frans Holm (SAS)

FANS

Analytisk

Netværksmøde

# Diverse Nyheder

## SAS Forum 2020 Sessions & Præsentationer

<https://www.sas.com/sas/events/nordic/sasforum.html>

<https://communities.sas.com/t5/Nordic-Events-and-Presentations/Presentations-SAS-Forum-Nordics-2020-17-November-Virtual/ba-p/699410>

# Sparrings Sessioner

Se link på Community  
Ingen tilmelding

## Diverse Nyheder

<https://communities.sas.com/t5/Nordic-Events-and-Presentations/FANS-Network-Meeting-Sparrings-sessioner/ba-p/666045>

### 2020

- ~~— Visual Analytics~~
- ~~— 6/11~~
- Programmering
  - 8/12 kl 9-10
- Platform
  - 8/12 kl. 11-12
- **Analytisk**
  - **8/12 kl. 13-14**

# SAS Analytics Explores

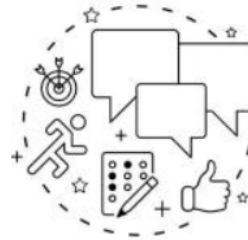
- <https://www.youtube.com/watch?v=IcZvGHCHM10>
- <https://explorers.sas.com/en/home.html>

 SAS®  
ANALYTICS EXPLORERS

Join now and start earning points to claim rewards\*, including SAS goodies, free access to training courses and events!



**GADGETS TO MAKE YOUR  
DAY MORE FUN**



**LEARNING AND  
RECOGNITION**



**NETWORKING AND  
MENTORING**



## GRATIS E-LEARNING

Få adgang til alle vores e-learning kurser gratis i 30 dage

Aktiver SAS Learning Subscription og start din læring i dag

Spørgsmål? Kontakt os på [saskurser@sas.com](mailto:saskurser@sas.com)

# Live Web kurser på dansk i efteråret 2020 med Ole Steen & Barbara Olsen

[SAS Programming 1: Essentials](#)

27. 29. oktober, kl. 09:00-16:00

[SAS Programming 2: Data Manipulation  
Techniques](#)

2. 5. november, kl. 09:00-12:30

[SAS Macro Language 1: Essentials](#)

11. 12. november, kl. 09:00-16:00

[SAS® Advanced DATA Step and Macro Techniques](#)

24. 26. november, kl. 09:00-16:00

[SAS SQL 1: Essentials](#)

30. november – 3. december, kl. 09:00-12:30

25% rabat af liste  
prisen for alle FANS  
medlemmer



# FANS Lodtrækning

Vind en SAS ting!

# FANS Lodtrækning

## Dagens gave!



# FANS lodtrækning

Send mail til:

[frans.holm@sas.com](mailto:frans.holm@sas.com)

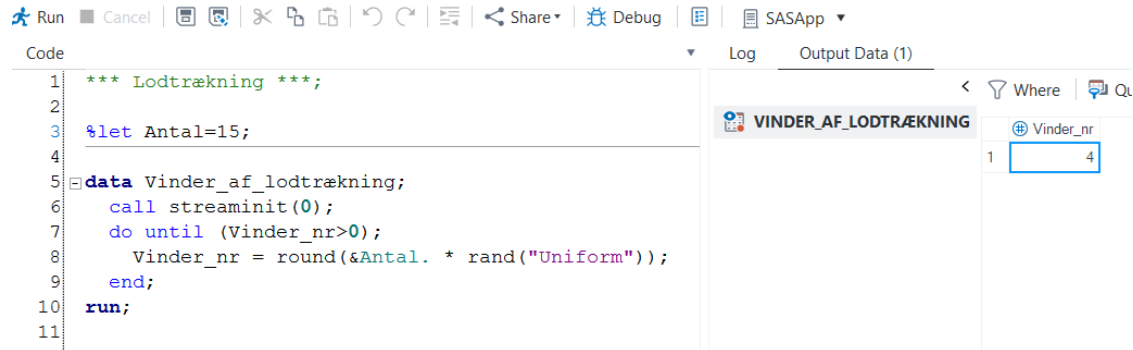
Skriv:

Subject: Lodtrækning

Mailen: Navn + Adresse

T-shirt størrelse: (S,M,L,XL – US-size)

Udtrækning  
Hvordan:



The screenshot shows the SAS Studio interface. The top menu bar includes 'Run', 'Cancel', 'Share', 'Debug', and 'SASApp'. The main area is divided into two panes. The left pane, titled 'Code', contains the following SAS code:

```
1 *** Lodtrækning ***;
2
3 %let Antal=15;
4
5 data Vinder_af_lodtrækning;
6   call streaminit(0);
7   do until (Vinder_nr>0);
8     Vinder_nr = round(&Antal. * rand("Uniform"));
9   end;
10 run;
11
```

The right pane, titled 'Output Data (1)', shows a table with the following data:

VINDER_AF_LODTRÆKNING	
	Vinder_nr
1	4

# Analytiske Nyheder i SAS

v. Rune Hjort Nielsen & Rebecca Grüner Hansen, SAS



# Analytiske nyheder

# Analytiske Nyheder i SAS

SAS 9.4M7

New release:

- SAS Enterprise Guide 8.3
- SAS Studio 3.81
- SAS Enterprise Miner 15.2
- SAS Text Miner 15.2
- SAS Business Rules Manager 3.3
- SAS Decision Manager 3.3
- SAS Forecast Server 15.2
- SAS Model Manager 14.3

# Analytiske Nyheder i SAS

SAS 9.4M7

Retired products:

- SAS Customer Link Analytics
- SAS Contextual Analytics In-Database Scoring for Hadoop
- SAS Master Data Management Advanced and Standard
- SAS Master Data Management Quality Advanced Server and Standard Server
- SAS Risk Management for Banking
- SAS Strategy Management
- SAS Visual Process Orchestration
- SAS Web Studio Server

# SAS and Microsoft Strategic Partnership

SAS and Microsoft have formed an extensive technology and go-to-market strategic partnership:

- Microsoft Azure is the preferred cloud provider for SAS Cloud
- SAS and Microsoft will co-develop integrations across Microsoft's entire cloud portfolio, including Azure, Dynamics 365, Microsoft 365, and Power Platform
- SAS solutions will appear in the Azure Marketplace





## Source-based Engines

### In-Stream



### In-Hadoop



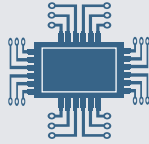
### In-Database



Parallel & Serial, Pub / Sub,  
Web Services, MQs

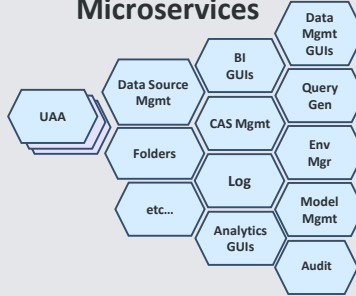
# SAS® Viya™

## In-Memory Runtime Engine



## Cloud Analytic Services (CAS)

## Microservices



## Solutions

### Customer Intelligence



### Analytics



### Risk Management



### Business Visualization



### Fraud and Security Intelligence



### Data Management



## APIs



CLOUDFOUNDRY



docker



kubernetes



ANSIBLE



rpm



Google Cloud Platform



openstack  
CLOUD SOFTWARE

vmware

ORACLE  
CLOUD



# New Nordic Data Science Network

v. Frans Holm, SAS

11/12-2020 – on Teams

FANS Nordic  
Data Science  
Network

# Agenda

9-11 CET

- **Welcome**
  - by Frans Holm, SAS
- **Data Science Use Cases**
  - by Jacob Mardfelt, SAS
- **Ethical AI**
  - by Josefin Rosén, SAS
- **Introduction to Linear Mixed Models**
  - by Ina Conrado, SAS
- **Penalized regressions - where regression analysis meets machine learning**
  - by Rune Hjort Nielsen, SAS
- **Closing & Lottery**
  - by Frans Holm, SAS



*Har du:*

Input?

Forslag?

Ideer?

Dele gode videoer/papers!

Tak  
fordi du  
er taler



Copyrig

# Tolkning og Bias i data

– Opfølgning fra sidst

v. Rune Hjort Nielsen, SAS

## Fra sidste møde

Hvor mange af faldgruberne handler om hvordan vi udvælger datasættet?

## DATA FALLACIES TO AVOID



### CHERRY PICKING

Selecting results that fit your claim and excluding those that don't.



### DATA DREDGING

Repeatedly testing new hypotheses against the same set of data, failing to acknowledge that most correlations will be the result of chance.



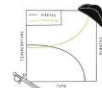
### SURVIVORSHIP BIAS

Drawing conclusions from an incomplete set of data, because that data has 'survived' some selection criteria.



### COBRA EFFECT

Setting an incentive that accidentally produces the opposite result to the one intended. Also known as a Perverse Incentive.



### FALSE CAUSALITY

Falsely assuming when two events appear related that one must have caused the other.



### GERRYMANDERING

The act of manipulating the geographical boundaries used to group data in order to change the result.



### SAMPLING BIAS

Drawing conclusions from a set of data that isn't representative of the population you're trying to understand.



### GAMBLER'S FALLACY

Mistakenly believing that because something has happened more frequently than usual, it's now less likely to happen in future (and vice versa).



### HAWTHORNE EFFECT

The act of monitoring someone can affect their behaviour, leading to spurious findings. Also known as the Observer Effect.



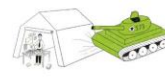
### REGRESSION TOWARDS THE MEAN

When something happens that's unusually good or bad, it will revert back towards the average over time.



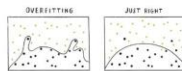
### SIMPSON'S PARADOX

When a trend appears in different subsets of data but disappears or reverses when the groups are combined.



### MCNAMARA FALLACY

Relying solely on metrics in complex situations and losing sight of the bigger picture.



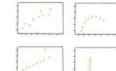
### OVERFITTING

Creating a model that's overly tailored to the data you have and not representative of the general trend.



### PUBLICATION BIAS

Interesting research findings are more likely to be published, distorting our impression of reality.




### DANGER OF SUMMARY METRICS

Only looking at summary metrics and missing big differences in the raw data.


# Fra sidste møde

Data faldgruber i forbindelse med datahåndtering.


## DATA FALLACIES TO AVOID




**CHERRY PICKING**  
Selecting results that fit your claim and excluding those that don't.




**DATA DREDGING**  
Repeatedly testing new hypotheses against the same set of data, failing to acknowledge that most correlations will be the result of chance.




**SURVIVORSHIP BIAS**  
Drawing conclusions from an incomplete set of data.




**COBRA EFFECT**  
Setting an incentive that accidentally produces the opposite result to the intended data point as a positive incentive.




**FALSE CAUSALITY**  
Falsely assuming when two events appear related that one must have caused the other.




**GERRYMANDERING**  
Manipulating the geographical boundaries used to group data in order to change the result.




**SAMPLING BIAS**  
Drawing conclusions from a set of data that isn't representative of the population you're trying to understand.




**GAMBLER'S FALLACY**  
Mistakenly believing that because something has happened more frequently than usual, it's somewhat likely to happen in future (and vice versa).




**HAWTHORNE EFFECT**  
The act of monitoring someone can affect their behaviour, leading to spurious findings. Also known as the Observer Effect.




**REGRESSION TOWARDS THE MEAN**  
When something happens that's certainly good or bad, it will revert back towards the average over time.




**SIMPSON'S PARADOX**  
When a trend appears in different subsets of data but disappears or reverses when the groups are combined.




**MCNAMARA FALLACY**  
Relying solely on metrics in complex situations and losing sight of the bigger picture.




**OVERFITTING**  
Creating a model that's overly tailored to the data you have and not representative of the general trend.



**JUST BURY**



**PUBLICATION BIAS**  
Interesting research findings are more likely to be published, distorting our impression of reality.



**DANGER OF SUMMARY METRICS**  
Only looking at summary metrics and missing big differences in the raw data.

GECKBOARD.COM

Reid Murray, data-liberacy@geckboard.com





# Hvilke faldgruber var der her?

## DATA FALLACIES TO AVOID



### CHERRY PICKING

Selecting results that fit your claim and excluding those that don't.



### DATA DREDGING

Repeatedly testing new hypotheses against the same set of data, failing to acknowledge that most correlations will be the result of chance.



### SURVIVORSHIP BIAS

Drawing conclusions from an incomplete set of data, because that data has 'survived' some selection criteria.



### COBRA EFFECT

Setting an incentive that accidentally produces the opposite result to the one intended. Also known as a Perverse Incentive.



### FALSE CAUSALITY

Falsely assuming when two events appear related that one must have caused the other.



### GERRYMANDERING

Manipulating the geographical boundaries used to group data in order to change the result.



### SAMPLING BIAS

Drawing conclusions from a set of data that isn't representative of the population you're trying to understand.



### GAMBLER'S FALLACY

Mistakenly believing that because something has happened more frequently than usual, it's now less likely to happen in future (and vice versa).



### HAWTHORNE EFFECT

The act of monitoring someone can affect their behaviour, leading to spurious findings. Also known as the Observer Effect.



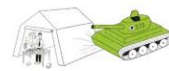
### REGRESSION TOWARDS THE MEAN

When something happens that's unusually good or bad, it will revert back towards the average over time.

	APPLICANT SUCCESS RATE	
	MALES	FEMALES
DEPT A	80%	70%
DEPT B	70%	60%
DEPT C	60%	50%
DEPT D	50%	40%
DEPT E	40%	30%

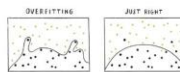
### SIMPSON'S PARADOX

When a trend appears in different subsets of data but disappears or reverses when the groups are combined.



### MCMANAMA FALLACY

Relying solely on metrics in complex situations and losing sight of the bigger picture.



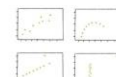
### OVERFITTING

Creating a model that's overly tailored to the data you have and not representative of the general trend.



### PUBLICATION BIAS

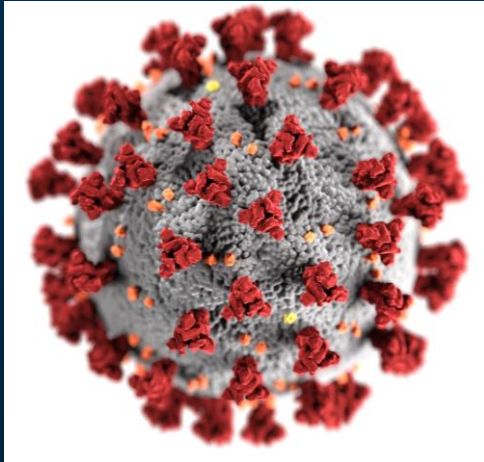
Interesting research findings are more likely to be published, distorting our impression of reality.



### DANGER OF SUMMARY METRICS

Only looking at summary metrics and missing big differences in the raw data.

# Hvilke nævner Ulrik?



## DATA FALLACIES TO AVOID



### CHERRY PICKING

Selecting results that fit your claim and excluding those that don't.



### DATA DREDGING

Repeatedly testing new hypotheses against the same set of data, failing to acknowledge that most correlations will be the result of chance.



### SURVIVORSHIP BIAS

Drawing conclusions from an incomplete set of data, because that data has 'survived' some selection criteria.



### COBRA EFFECT

Setting an incentive that accidentally produces the opposite result to the one intended. Also known as a Perverse Incentive.



### FALSE CAUSALITY

Falsely assuming when two events appear related that one must have caused the other.



### GERRYMANDERING

Manipulating the geographical boundaries used to group data in order to change the result.



### SAMPLING BIAS

Drawing conclusions from a set of data that isn't representative of the population you're trying to understand.



### GAMBLER'S FALLACY

Mistakenly believing that because something has happened more frequently than usual, it's now less likely to happen in future (and vice versa).



### HAWTHORNE EFFECT

The act of monitoring someone can affect their behaviour, leading to spurious findings. Also known as the Observer Effect.



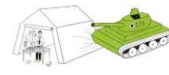
### REGRESSION TOWARDS THE MEAN

When something happens that's unusually good or bad, it will revert back towards the average over time.

	APPLICATION SUCCESS RATE	
	PHYSICS	CHEMISTRY
PHYSICS	80%	70%
CHEMISTRY	70%	80%
PHYSICS & CHEMISTRY	75%	75%
PHYSICS & CHEMISTRY	75%	75%

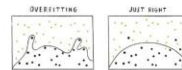
### SIMPSON'S PARADOX

When a trend appears in different subsets of data but disappears or reverses when the groups are combined.



### MCNAMARA FALLACY

Relying solely on metrics in complex situations and losing sight of the bigger picture.



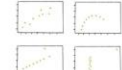
### OVERFITTING

Creating a model that's overly tailored to the data you have and not representative of the general trend.



### PUBLICATION BIAS

Interesting research findings are more likely to be published, distorting our impression of reality.



### DANGER OF SUMMARY METRICS

Only looking at summary metrics and missing big differences in the raw data.

# Tolkning og Bias i data

– Case

v. Ulrik Gerdes, læge, uafhængig datakuser

Se præsentationen her:

[http://www.kliniskbiokemi.net/GerdesGuide/UG SAS FANS 25112020.html](http://www.kliniskbiokemi.net/GerdesGuide/UG_SAS_FANS_25112020.html)

# FANS Lodtrækning

## Dagens gave!

Send mail til:

[frans.holm@sas.com](mailto:frans.holm@sas.com)

Skriv:

Subject: Lodtrækning

Mailen: Navn + Adresse

T-shirt størrelse: (S,M,L,XL – US-size)



# Pause – Pause – Pause - Pause

Vi starter igen kl. 10.45

# The Analytics Use Case Canvas

(in English)

v. Erik Strömngren, SAS

# Opdag abnormiteter ved hjælp af tidsrækkeanalyse

v. Rebecca Grüner Hansen, SAS



# Agenda

- Mange virksomheder bruger ressourcer på at lede efter abnormiteter i data.
  - Det kan være alt fra hvidvask af penge i en bank til anvendelse af medicin på et sygehus.
- Vi sætter her fokus på, hvordan kan man ved hjælp af tidsrækkeanalyse kan forbedre eller komme i gang med den proces.

# Projektbeskrivelse

Kunde:

- En kontrolenhed, der overvåger og kontrollerer flow af varer ind og ud af Danmark.

Formål:

- Finde abnormiteter i vare flowet over tid, samt liste disse efter hvilke, der er mest mistænkelige.
- Assistere analytikerne i deres daglige job.

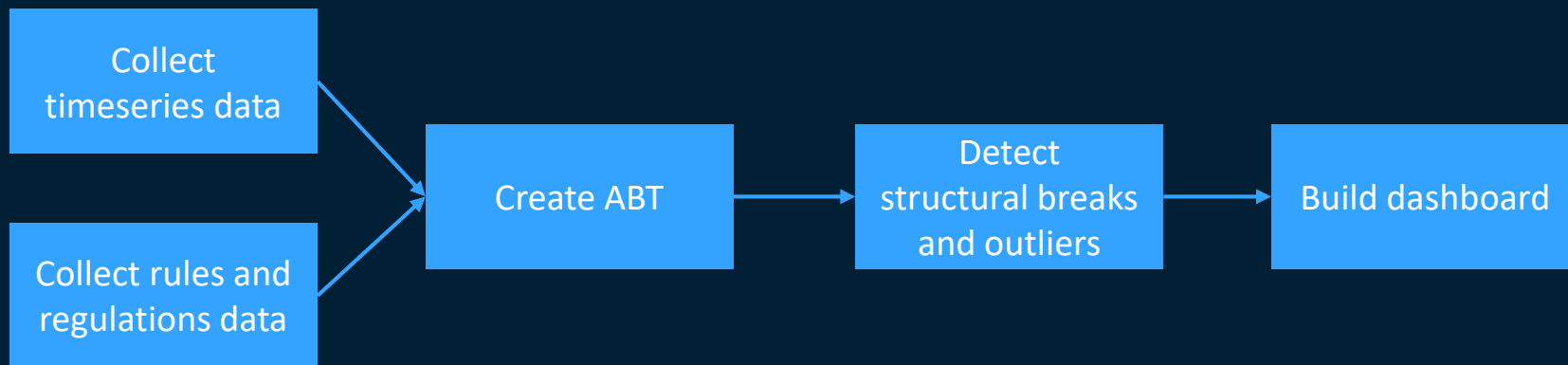
# Hvad og hvorfor?



- Identificer mistænkelige vare flows til yderligere undersøgelse
  - Dvs. stor stigning eller fald i værdien/mængden af flowet af forskellige varer
- Tilføjelse eller ændringer i regler og lovgivninger kan medføre ændringer i vare flows over tid.
- Det kan også medføre ændringer i vare flows for relaterede varer.

# Hvordan?

- Proces overblik:



# Hvordan?



- Betragter vare flows over tid
- Modellerer flowet pba. historik
- Finder strukturelle ændringer i data, herunder niveau skift og outliers, pba. modellen/-erne
- Ranglister de fundene strukturelle ændringer

# DEMO



# Tak!

[Rebecca.Hansen@SAS.com](mailto:Rebecca.Hansen@SAS.com)

[sas.com](https://sas.com)

# Afslutning & lodtrækning



# Sparrings Sessioner

Se link på Community  
Ingen tilmelding

## Diverse Nyheder

<https://communities.sas.com/t5/Nordic-Events-and-Presentations/FANS-Network-Meeting-Sparrings-sessioner/ba-p/666045>

### 2020

- ~~— Visual Analytics~~
- ~~— 6/11~~
- Programmering
  - 8/12 kl 9-10
- Platform
  - 8/12 kl. 11-12
- **Analytisk**
  - **8/12 kl. 13-14**

[sas.com/fans](https://sas.com/fans) -> Events -> All live Events + Webinars

## Netværksmøder

- ~~24/11 Platform Overview – Virtuelt~~
- ~~25/11 Analytisk netværk – Virtuelt~~
- 26/11 Platforms netværk - Virtuelt
- 27/11 Programmerings netværk – Virtuelt
- 1/12 Visual Analytics netværk – Virtuelt
- 11/12 Data Science Network, Nordic – Virtuelt

2021  
Uge 5 & 22

# FANS Lodtrækning

## Dagens gave!



Tak for i dag  
vi ses fans!

*Husk evaluering*

