

ROBERT KOCH INSTITUT



# Dashboards as strategy to integrate multiple data streams for a holistic surveillance

Fabian Eckelmann, Stéphane Ghazzi,  
Alexander Ullrich

Signale Group - Robert Koch Institute

# What is the problem?

## The Situation:

- Complex question
- Heterogenous data sources

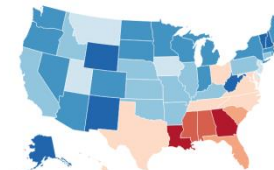
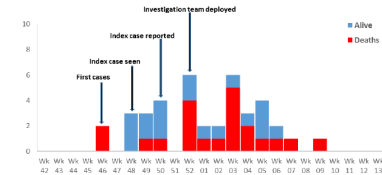


## Goal:

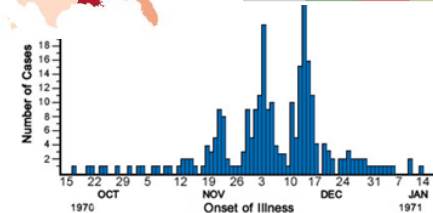
- Informed decision
  - using all relevant sources
- Wholistic understanding of the situation
- Overview for experts

## Solutions:

- Data Visualisations
- Reports
- Dashboards



	2013	2014	2015	2016	2017	2018
Product 1	758	869	955	614	681	643
Product 2	690	761	841	835	892	748
Product 3	935	932	918	935	708	708
Product 4	604	907	850	550	957	600
Product 5	953	545	516	574	533	936
Product 6	941	523	775	720	606	816



# Outline

1. Why Dashboards?
  - Benefits
  - Three sample Dashboards
2. How to develop useful Dashboards?
  - Theory + Experience
3. Lessons Learned

# Why Dashboards?

## Research shows:

- Visual perception is a powerful channel for information acquisition and comprehension
- Dashboards offer a solution to information overload
- Dashboards avoid workflow interruption

## Benefits:

- Allow Integration of multiple sources
- Add Interactivity, Exploration and Real-time

# Example: Severity Dashboard

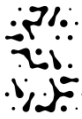
**Goal:** Weekly assessment of the current influenza pandemic regarding transmissibility, seriousness and impact (WHO PISA guidelines)

**Situation:** 4 different sources (viral, 2 x syndromic, lab confirmed)

**Previous Solution:** several Excel tables and Stata scripts

## Dashboard Solution:

- Overview first
  - Two Time Series visualisations (age groups, comparison with historic data)
  - Indicators are explained by different sources (e.g. transmissibility by consultation rate, seriousness by hospitalisation and deaths from syndromic surveillance)
- Detail later
  - E.g. positive rate (determines wave, influences all other analyses / visualisations, threshold can be changed in Dashboard Settings)



Display

indicator type  
Transmissibility

cumul from:  
CW40 epi1

display season  
2017/2018

number of seasons  
3

display date range  
2015-09-28 to 2018-09-30

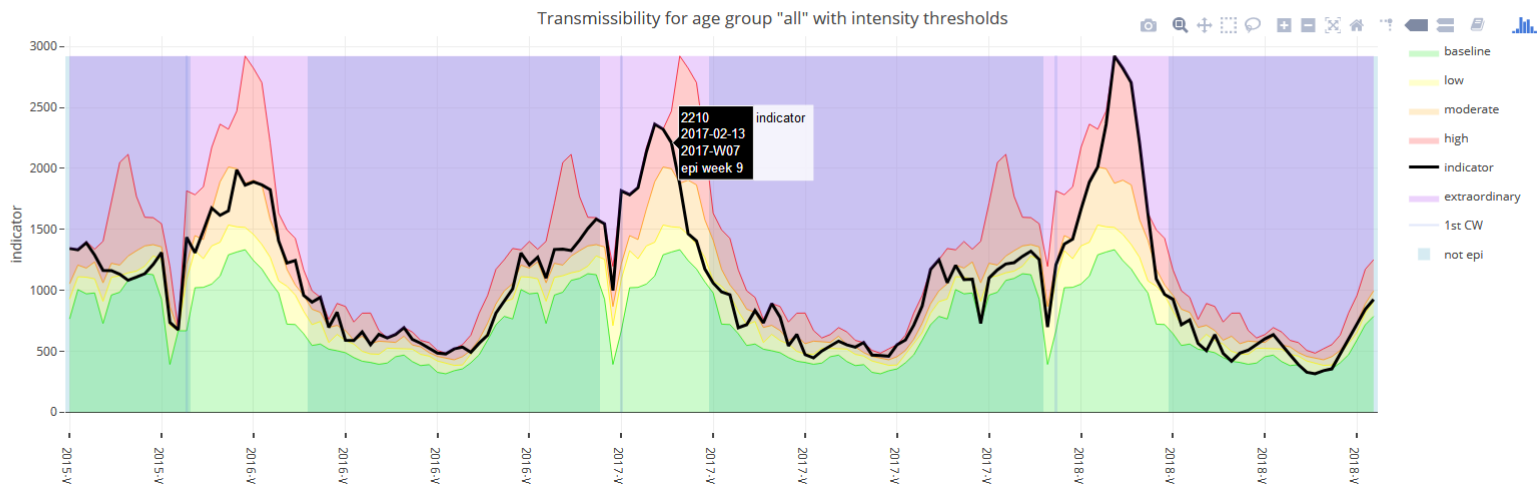
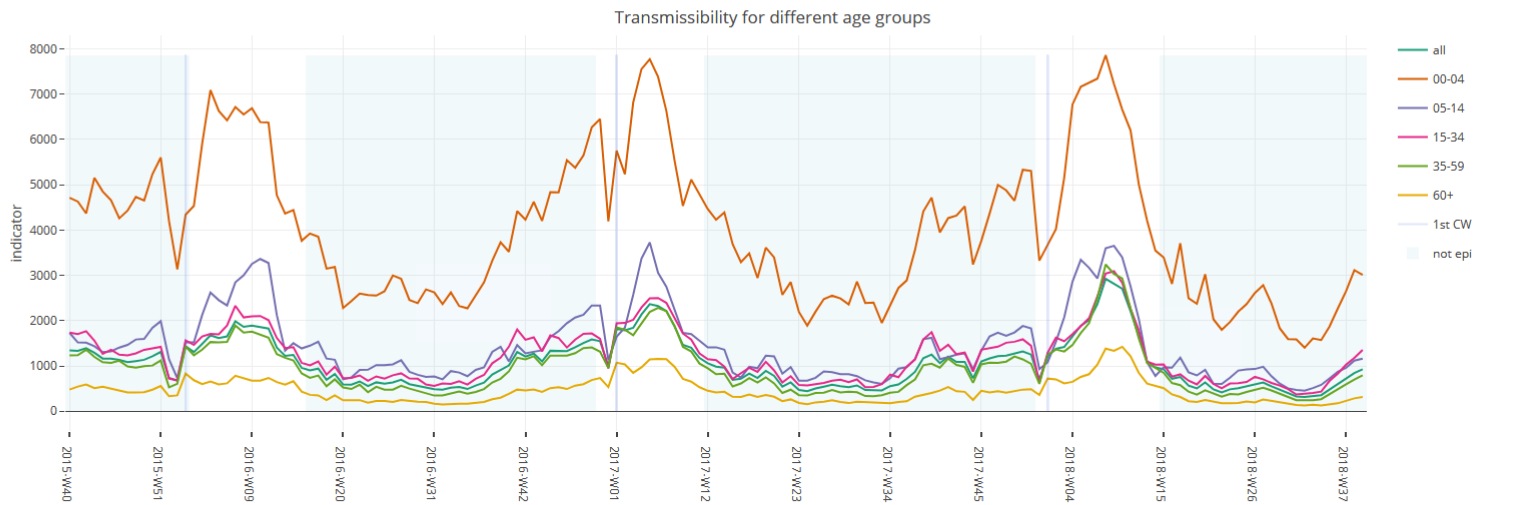
display age group  
all

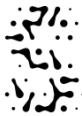
show only epi weeks

Epi options

Intensity thresholds

indicator cumul indicator intensity tables nrz data info





Display

indicator type

Impact 2 BCD

Seriousness 2 TCD

Seriousness 3 BCD

Seriousness 3 SCD

Seriousness 3 TCD

Impact 1

Impact 2 BCD

Impact 2 SCD

Impact 2 TCD

number of seasons

3

display date range

2015-09-28 to 2018-09-30

display age group

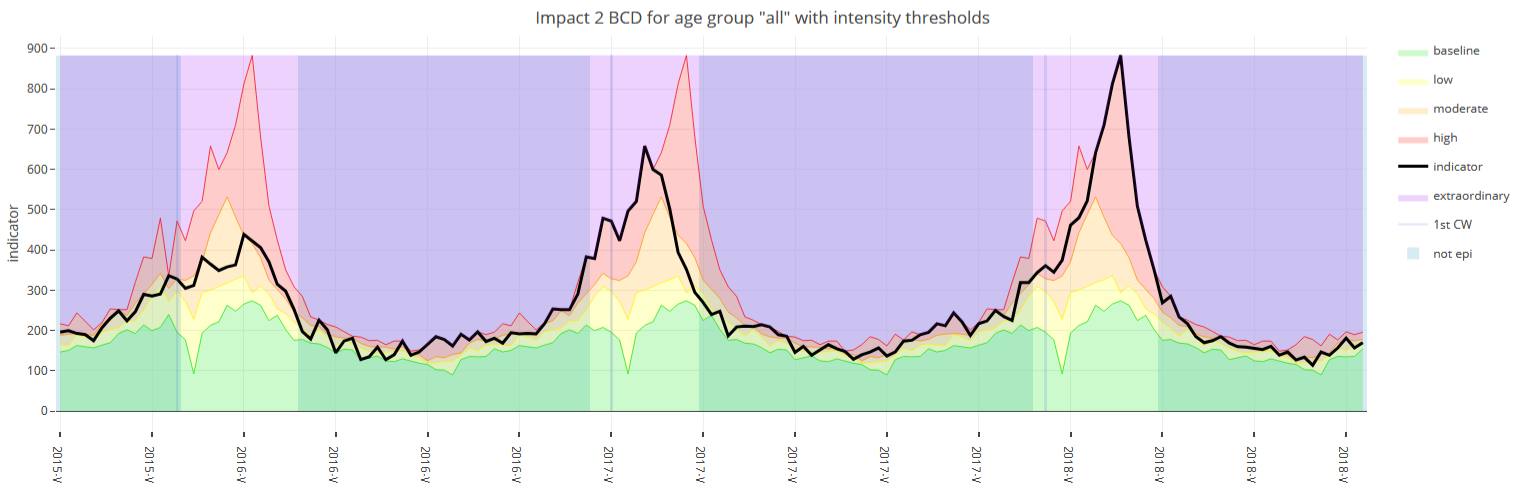
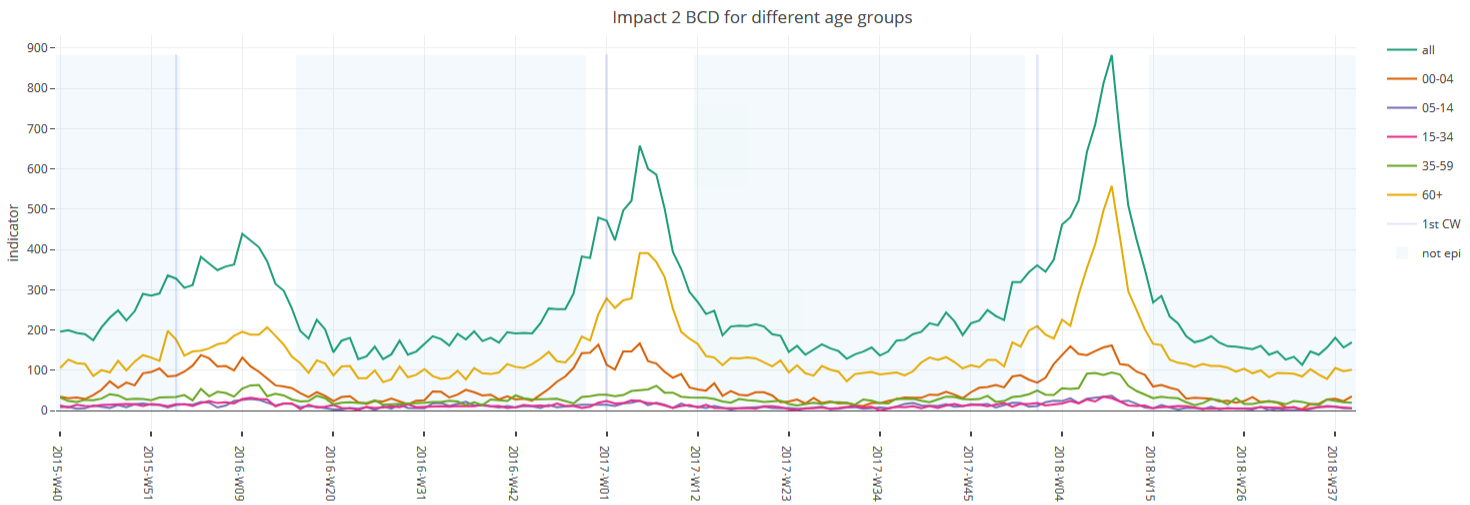
all

show only epi weeks

Epi options

Intensity thresholds

indicator cumul indicator intensity tables nrz data info



# Influenza Dashboard



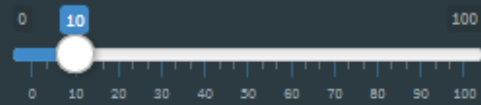
Display <

Epi options >

epi method:

CI, 2 weeks  direct

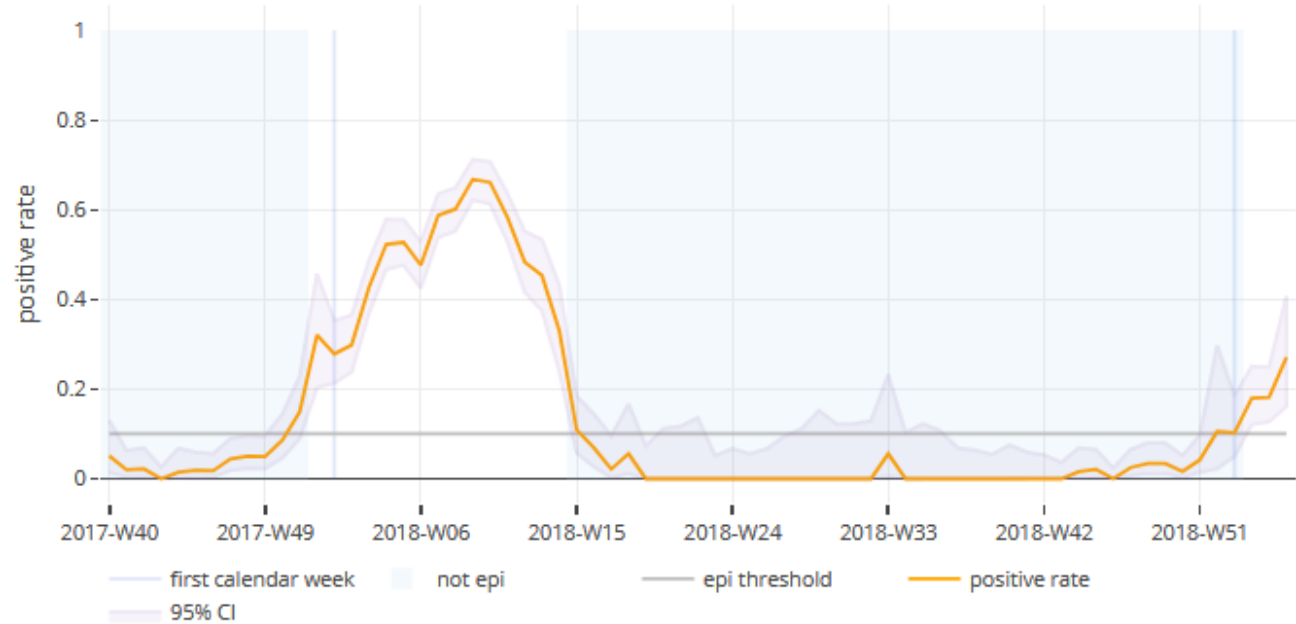
epidemic threshold (%)



Intensity thresholds <

indicator    cumul indicator    intensity tables    **nrz**    data    info

NRZ positive rates and epidemic threshold





# Example: TBE Dashboard

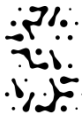
**Goal:** Investigate influence of weather and TBE

**Data sources:**

- 3 vector borne diseases (traditional surveillance)
- ~15 weather variables
- state holidays

**Dashboard Solution:**

- Appropriate visualisations such as scatter plot, cross correlation plots, correlation maps
- Multivariate modeling, regression summaries, predictions



# FSME dashboard

- Overview
- Compare Regions
- Compare with Weather
- Model with Weather
- Controls

**Choose Disease**  
FSV

**min. Incidence**  
0

**Select state**  
Baden-Württemberg  
Bayern

**display season**  
2018-01-01

**# seasons**  
18

**Choose Region**  
State Region County

**Aggregate By**  
Year Quarter Month

**Choose Measure**  
Count Incidence

Combine

## Input Controls

Select weather

Lufttemperatur

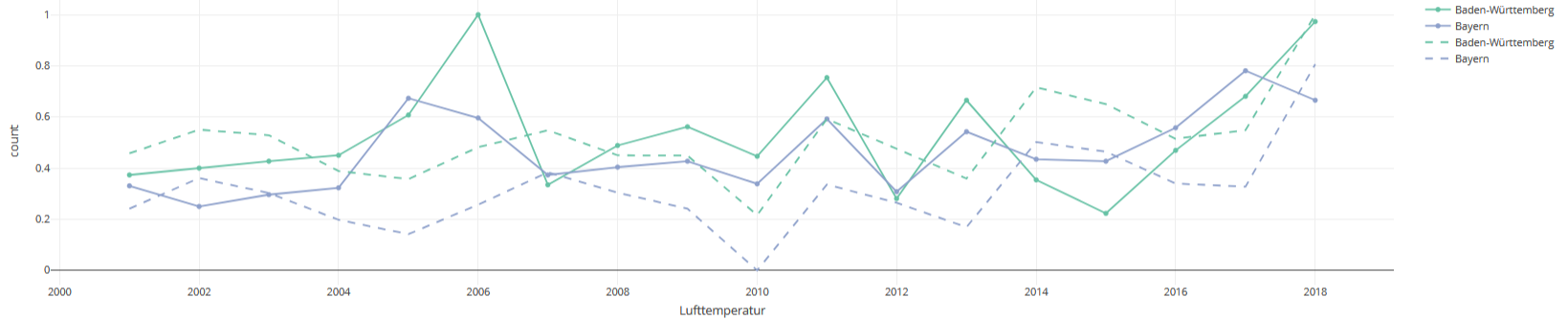
Chose Correlation Period

01-01

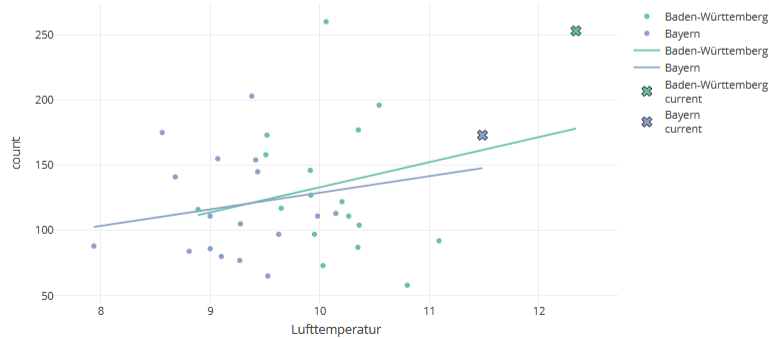
Lag

0

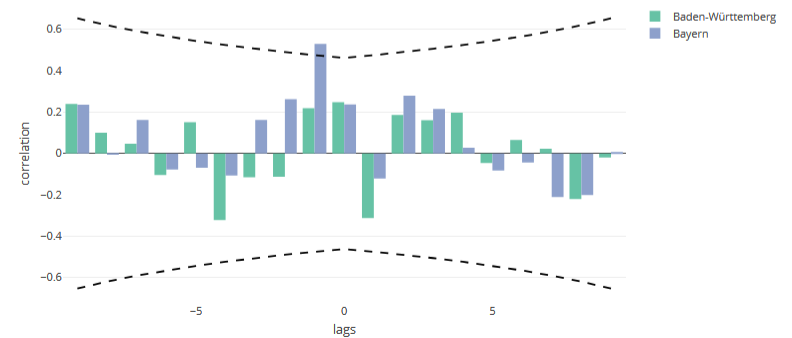
## Time-Series Comparison against weather (normalized [0-1])

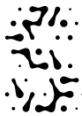


## Scatter Plot with Regression Line



## Cross-Correlation at different lags





- Overview
- Compare Regions
- Compare with Weather
- Model with Weather
- Controls
- Choose Disease: FSV, min. Incidence: 0
- Select state: Baden-Württemberg, Bayern
- display season: 2018-01-01, # seasons: 18
- Choose Region: State, Region, County
- Aggregate By: Year, Quarter, Month
- Choose Measure: Count, Incidence
- Combine

Input Controls

Chose Variables

state x lag\_Lufttemperatur\_01 x lag\_Niederschlagshoehe\_01 x lag\_Schneehoehe\_01 x lag\_Sonnenscheindauer\_01 x lag\_Relative\_Feuchte\_01 x lag\_winterT\_01 x lag\_Tincrease\_01 x

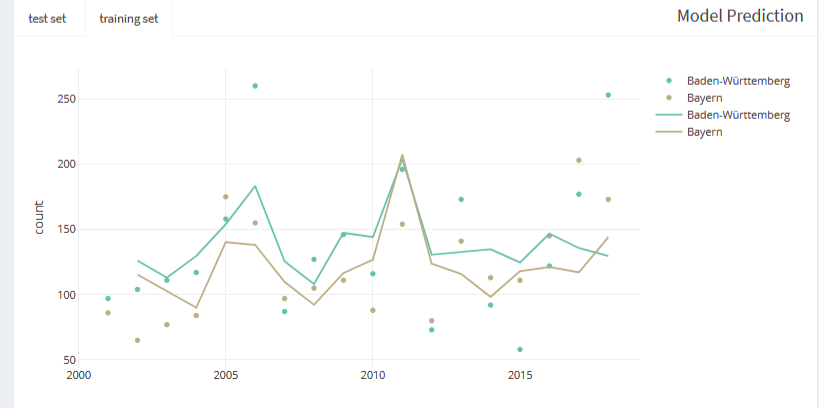
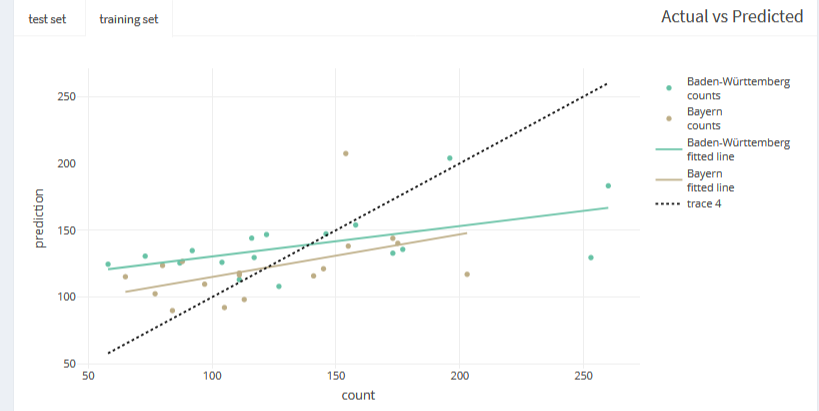
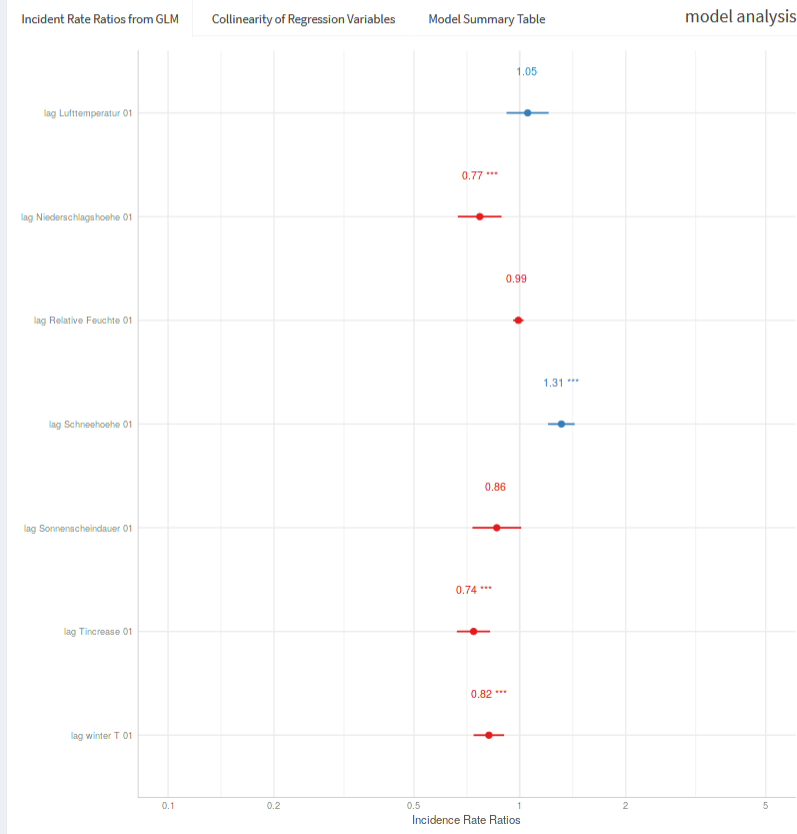
Take all Take none Feature selection (AIC) Feature selection (CV) Feature selection (Prediction error)

Chose Prediction Period

01-01 x

Choose model family

Poisson  Quasi  NegBin



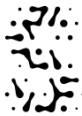
# Example: Signale Dashboard

## **Situation:**

- static weekly reports
- one page per disease
- 5 independent widgets

## **Dashboard Solution:**

- Connected by interactivity
- Exploration through filtering
- Real-Time through Data access via webservice



Filtereinstellungen

Borreliose (BOB)

Nur Fälle mit erfüllter Referenzdefinition

13.12.2018

24.01.2019

Landkreis

- Baden-Württemberg
- Bayern
- Berlin
- Brandenburg
- Bremen
- Hamburg
- Hessen
- Mecklenburg-Vorpommern
- Niedersachsen
- Nordrhein-Westfalen
- Rheinland-Pfalz
- Saarland
- Sachsen
- Sachsen-Anhalt
- Schleswig-Holstein
- Thüringen

Geschlecht

männlich weiblich unbekannt

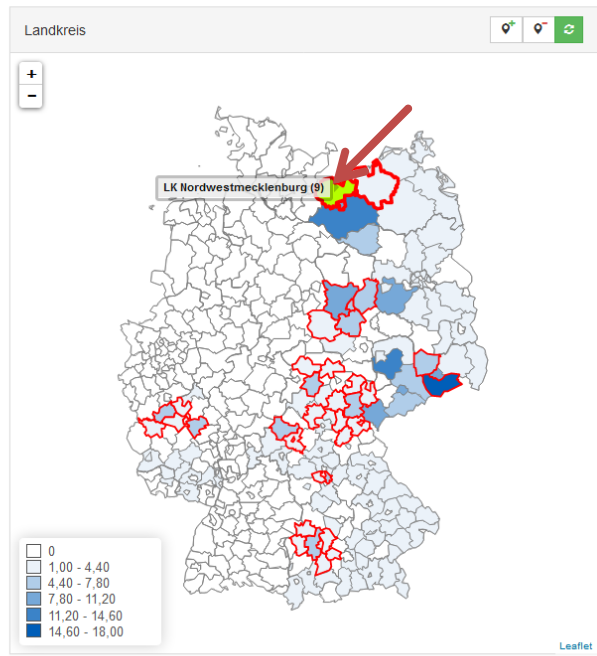
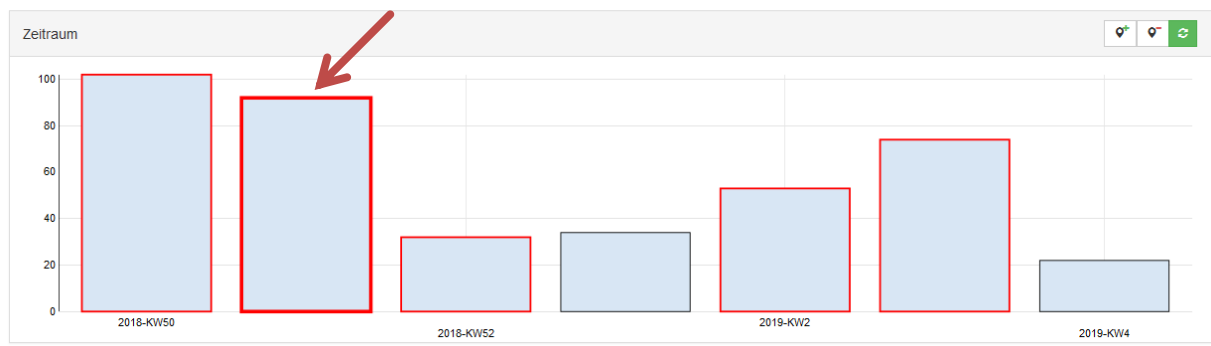
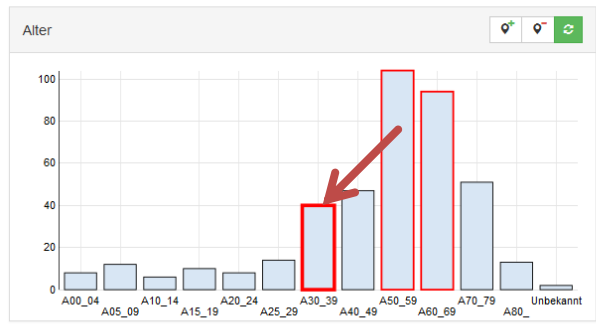
Alter

A00\_04 A05\_09 A10\_14 A15\_19

Subtypen

-andere/sonstige-BOB

Wochen	Label	Score	Aktio...
2019-KW03	1 Signal - LK S...	1	
2019-KW03	1 Signal - LK S...	1	
2019-KW03	1 Signal - LK S...	1	
2019-KW03	1 Signal - LK S...	1	
2019-KW03	1 Signal - LK H...	1	
2019-KW03	1 Signal - LK A...	1	
2019-KW03	1 Signal - LK A...	1	
2019-KW03	1 Signal - LK A...	1	
2019-KW02	1 Signal - LK A...	1	
2018-KW51, 2...	2 Signale - SK ...	1	
2018-KW51, 2...	2 Signale - ...	1	
2019-KW51	1 Signal - LK N...	1	
2018-KW51	1 Signal - LK N...	1	
2018-KW50	1 Signal - LK S...	1	
2019-KW03	1 Signal - SK ...	0.999	
2019-KW03	1 Signal - LK S...	0.999	
2019-KW03	1 Signal - LK S...	0.999	
2019-KW03	1 Signal - LK S...	0.999	
2019-KW03	1 Signal - LK K...	0.999	
2019-KW03	1 Signal - LK H...	0.999	
2019-KW03	1 Signal - LK H...	0.999	
2019-KW03	1 Signal - LK H...	0.999	
2019-KW03	1 Signal - LK H...	0.999	
2019-KW03	1 Signal - LK ...	0.999	
2019-KW03	1 Signal - LK A...	0.999	
2019-KW03	1 Signal - LK A...	0.999	
2019-KW03	1 Signal - LK A...	0.999	
2019-KW02	1 Signal - LK A...	0.999	
2019-KW02	1 Signal - LK A...	0.999	
2018-KW51, 2...	2 Signale - LK ...	0.999	
2018-KW50	1 Signal - LK V...	0.999	





**Filtereinstellungen**

Borreliose (BOB)

Nur Fälle mit erfüllter Referenzdefinition

13.12.2018

01.02.2019

Landkreis

- Baden-Württemberg
- Bayern
- Berlin
- Brandenburg
- Bremen
- Hamburg
- Hessen
- Mecklenburg-Vorpommern
- Niedersachsen
- Nordrhein-Westfalen
- Rheinland-Pfalz
- Saarland
- Sachsen
- Sachsen-Anhalt
- Schleswig-Holstein
- Thüringen

Geschlecht

- männlich
- weiblich
- unbekannt

Alter

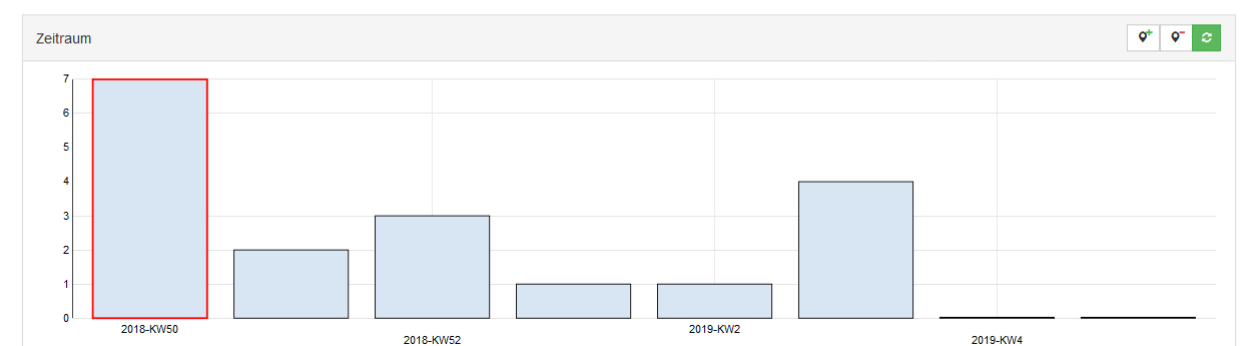
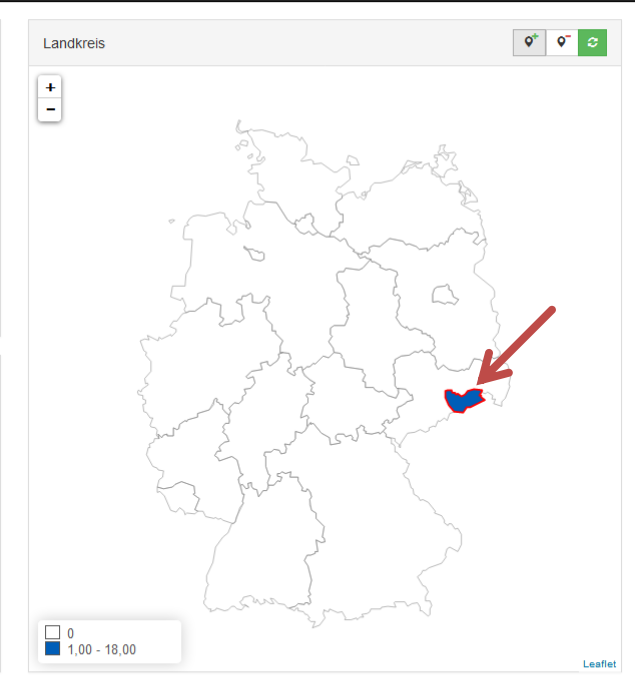
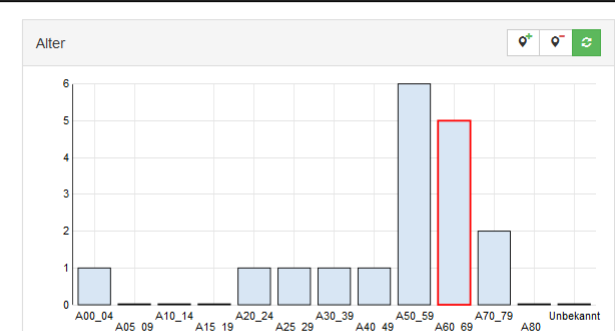
- A00\_04
- A05\_09
- A10\_14
- A15\_19
- A20\_24
- A25\_29
- A30\_39
- A40\_49
- A50\_59
- A60\_69
- A70\_79
- A80\_

Subtypen

- andere/sonstige-BOB

Signale (3)

Wochen	Label	Score	Aktio..
2018-KW50	1 Signal - LK S...	1	
2018-KW50	1 Signal - LK S...	0.996	
2018-KW50	1 Signal - LK ...	0.996	

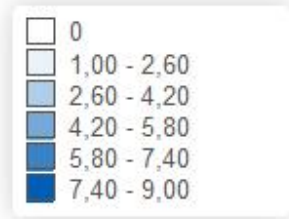
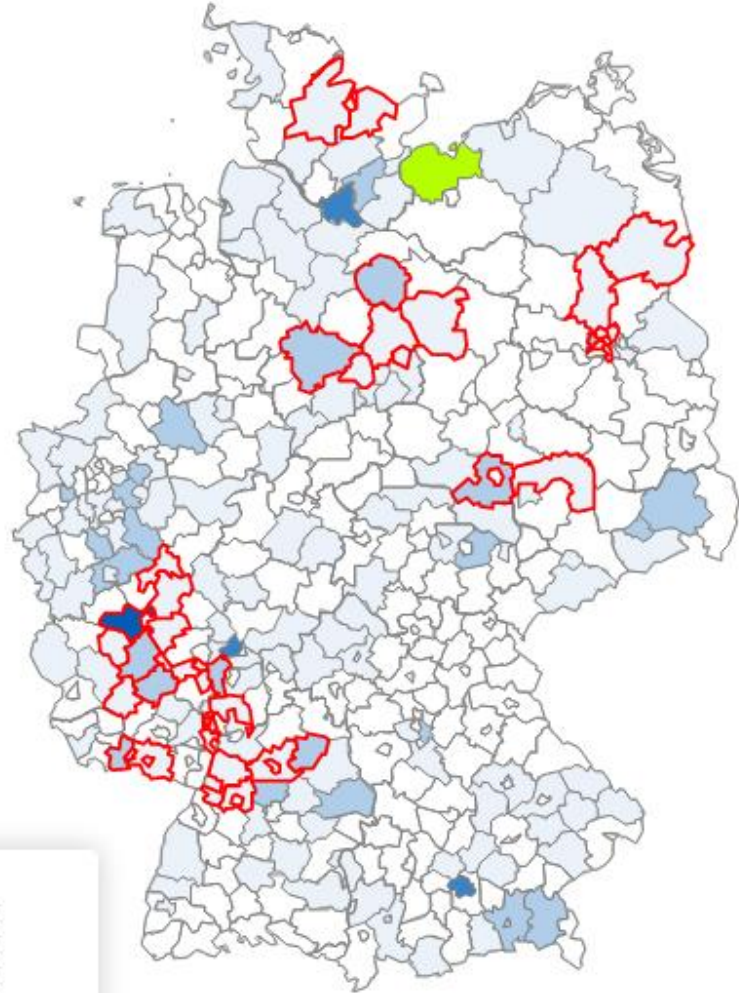




Subtypen



- S.Entebbe
- S.Enteritidis**
- S.Enugu
- S.Epalinges
- S.Frisetos



# How to develop a dashboard?

**Goal:** a dashboard that is useful to the epidemiologist

**Situation:** IT-project with the task of building a dashboard for epidemiologists

- IT group or external provider responsible for development
- Rely on voluntary help from experts/users



# Theory on dashboard design

- Several books and papers on dashboard design

## **My Take home message:**

- Build trust with your users/customers
- Collaborate with the subject expert
- Develop a common language
  - Start with what they expect
  - Stepwise change (e.g. exploratory design, semantic similarity)

# Own Experience

## Team:

- 1-2 Data Scientists, 1-2 epidemiologists
- All participants should have a genuine interest in the success of the product (faster dev, better adaptation, trust in the product)

## Requirements:

- existing analyses and tools
- questionnaires, interviews

## Process:

- Fast initial prototype (R-Shiny)
- Prioritize, add and trim features together
- 5-6 iterations (approx. 2 hour meetings with protocol)

## Side note:

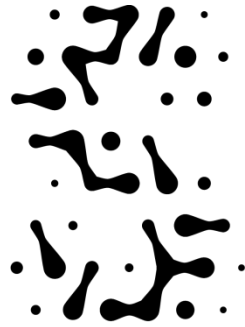
- R-Shiny is a very fast prototyping tool, good for production (internal), easy sharing, ideal for combining analysis and web dev

# Lessons learned

How to develop a dashboard?

**Solution:** continuous collaborative process

1. Work in interdisciplinary teams that have an interest in long-term success
2. Iterative development with early prototypes
3. Collaboration doesn't end with development
4. Start with small simple (one goal) dashboards



SIGNALE

Thanks. Questions?

[signale@rki.de](mailto:signale@rki.de)