

Små tal sett
frå statistikaren
sitt perspektiv

Karl Ove Hufthammer

Korfor samlar me data?

To grunnar til å samla data

Seia noko om **fortida** (sånn **var** det)

Seia noko om **framtida** (sånn **vert** det)

Om fortida



Bokholderi

Om framtida



Kan ta
styring!

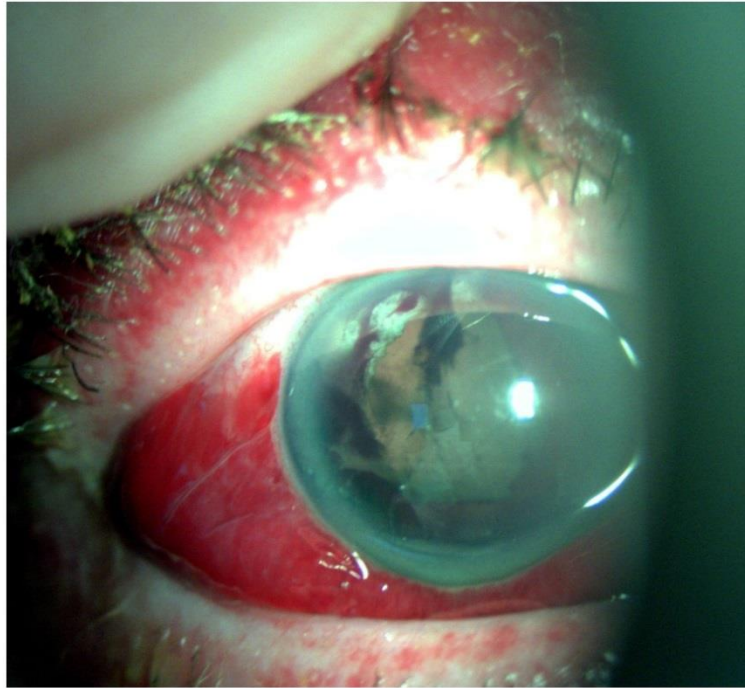
Eksempelstudie

Medisin A: 1-års overleving på 20 %

Medisin B: 1-års overleving på 80 %

Små tal – eksempel

Augeskadar ved fyrverkeri



Denne personen fikk øyeskade av effektbatteri. Foto: Nils Bull

<https://www.vg.no/nyheter/innenriks/i/02xz2/tolv-personer-fikk-oeyeskader-nyttaarsaften-fire-alvorlige>

Eit lite tal

Nyttårsaftanen 2017–2018:

9 augeskadar

Seier dette noko om korleis
det vil gå til neste år?

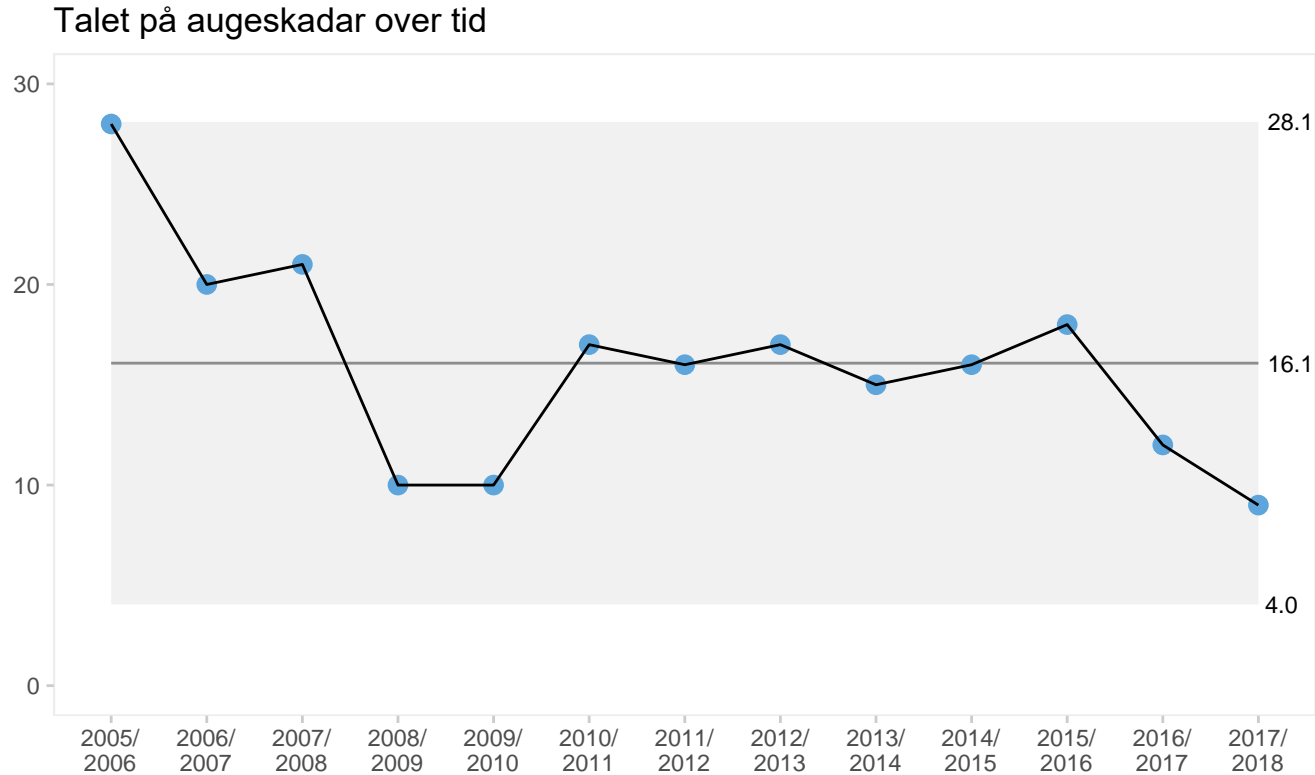
Historiske data

Øyeskader på nyttårsaften

- 2017/18: 9 skader- 3 alvorlige og 6 moderat alvorlige
- 2016/17: 12 skader- 4 alvorlige og 8 moderat alvorlige
- 2015/16: 18 skader –6 alvorlige og 12 moderat alvorlige
- 2014/15: 16 skader – 4 alvorlige og 12 moderat alvorlige
- 2013/14: 15 skader - 7 alvorlige og 8 moderat alvorlige
- 2012/13: 17 skader - 5 alvorlige og 12 moderat alvorlige
- 2011/12: 16 skader - 4 alvorlige og 12 moderat alvorlige
- 2010/11: 17 skader - 6 alvorlige og 11 moderat alvorlige
- 2009/10: 10 skader - 3 alvorlige og 7 moderat alvorlige
- 2008/09: 10 skader - 5 alvorlige og 5 moderat alvorlige
- 2007/08: 21 skader - 6 alvorlige og 15 moderat alvorlige
- 2006/07: 20 skader - 3 alvorlige og 17 moderat alvorlige
- 2005/06: 28 skader – 6 alvorlige og 22 moderat alvorlige

«Dette er andre år på rad med en nedgang i antallet øyeskader.»

Augeskadar over tid



Beste estimat på nivået

- 2017/18: 9 skader- 3
- 2016/17: 12 skader-
- 2015/16: 18 skader
- 2014/15: 16 skader
- 2013/14: 15 skader
- 2012/13: 17 skader
- 2011/12: 16 skader
- 2010/11: 17 skader
- 2009/10: 10 skader
- 2008/09: 10 skader
- 2007/08: 21 skader
- 2006/07: 20 skader
- 2005/06: 28 skader

Totalt 209 augeskader på 13 år.

Nivåestimat: $209/13 \approx 16$

Uvisse i estimata

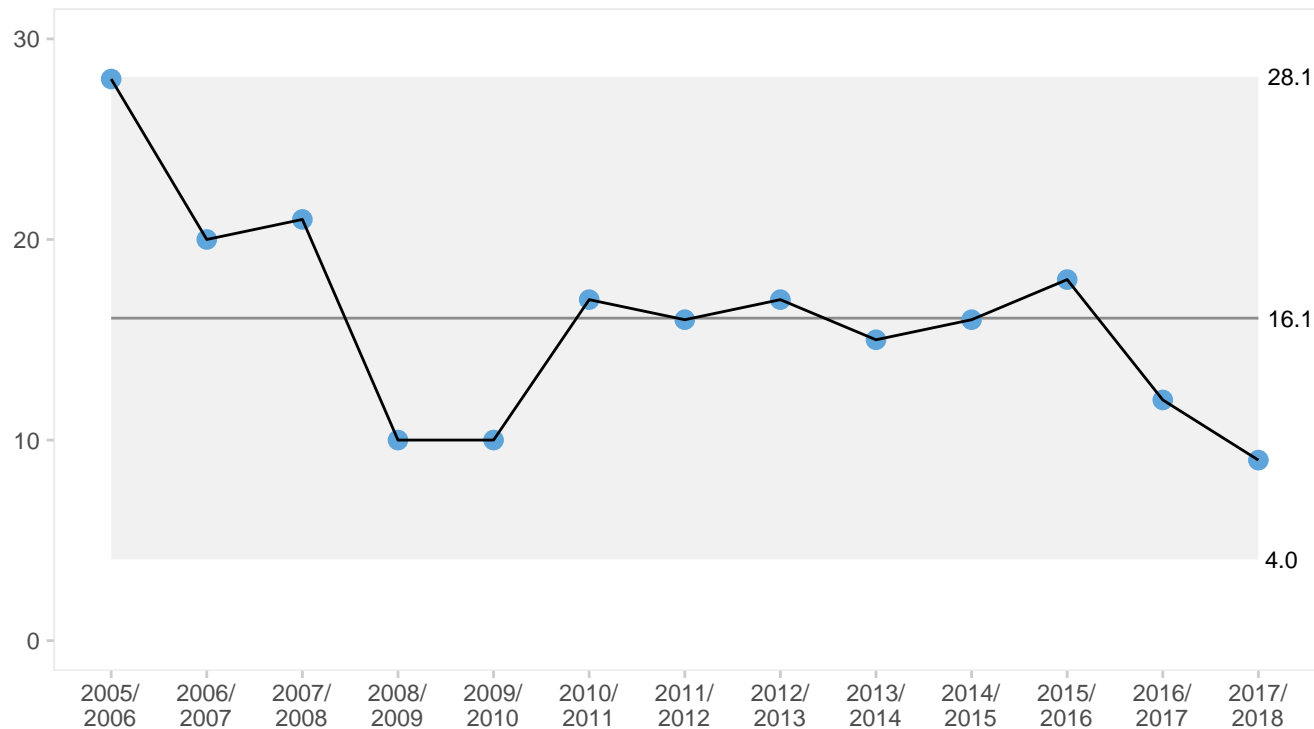
To moglege tilfelle for same estimat (16).

1 år med data, observerer 16 augeskadar:
95 %-konfidensintervall: (9–25)

13 år med data, observerer 16 augeskadar/år:
95 %-konfidensintervall: (14–18)

Estimat på augeskadar neste år

Talet på augeskadar over tid



Fasiten

NYTTÅRSFEIRINGEN:

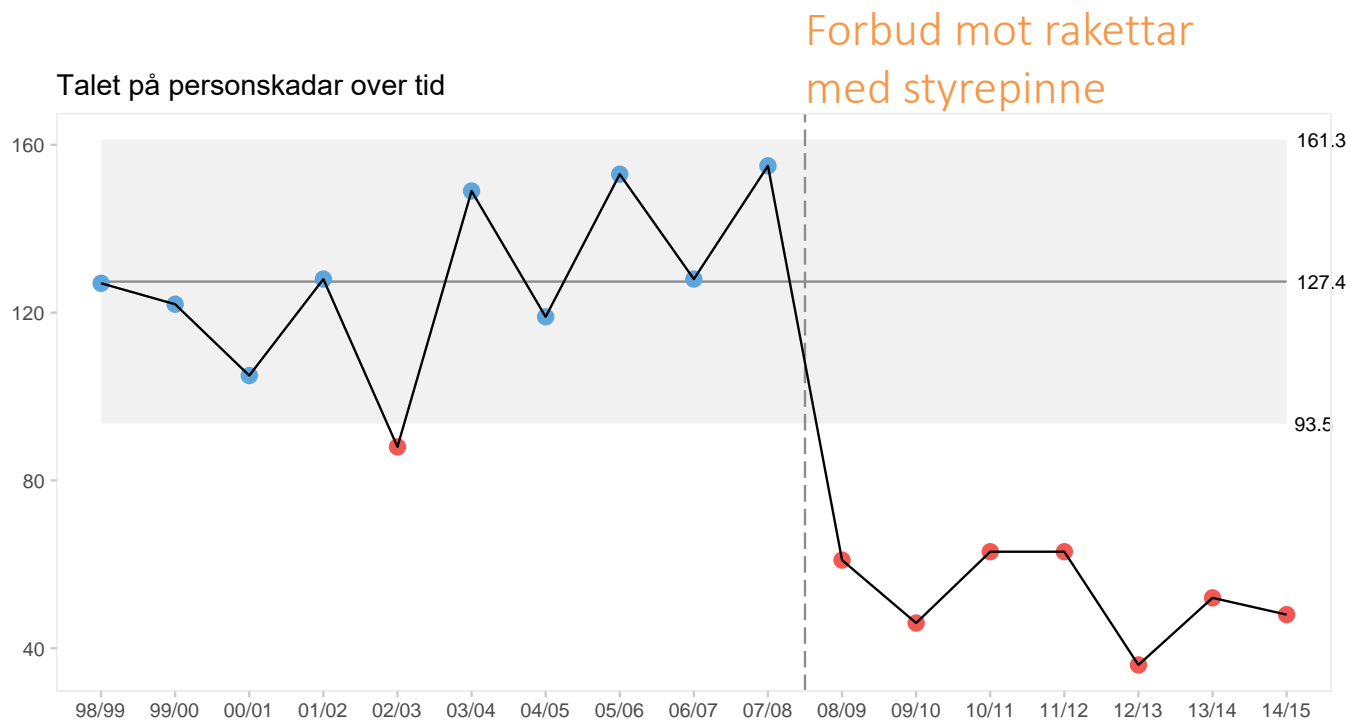
16 personer fikk øyeskader



– Ingen av de skadde brukte vernebriller da ulykken skjedde, sier overlege Nils Bull. Foto: TV 2

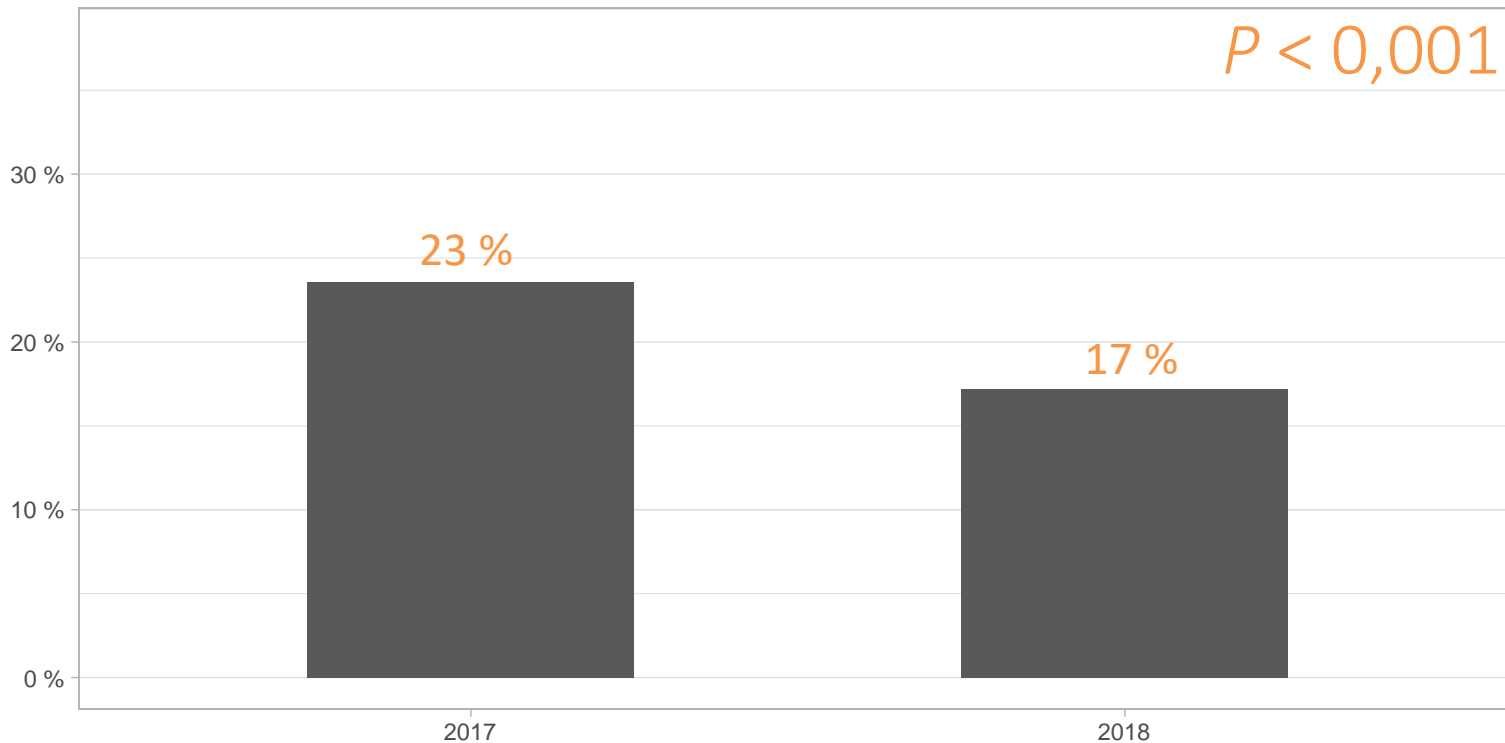
Det var to år med nedgang og jeg trodde vi var inne på en god trend, men nå er vi tilbake på nivået vi lå på i mange år.

Personskadar fyrverkeri

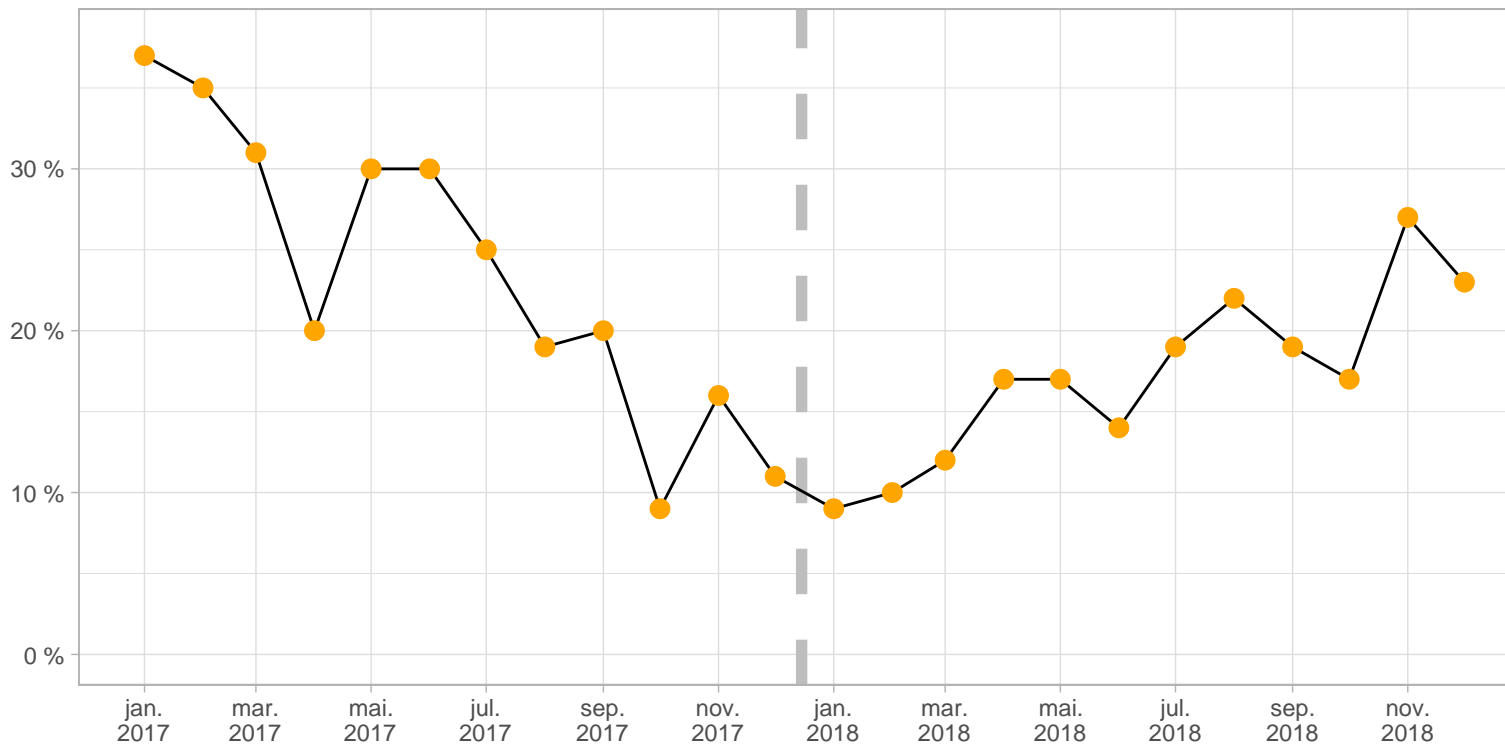


Å måla det som ikkje finst

Nivåforskjell?



Nivåforskjell?



Samhandlingsbarometeret

Mandeloperasjonar

Samhandlingsbarometeret

Kunnskapsb...

DEFINISJON STATISTIKK

Datasett: Mandel operasjonar 0-18 år (I002)

Tal operasjonar pr. 1000 i aldersgruppa NB! Alle rater er rekna om til 1-års rate.

komnrhjem: Blandet År-Tertial: År Helseforetak: HBE-HDS-HFD Måltall: Mandelop

År-Tertial	2009	2010	2011	2012	2013	2014	2015	2016	2017
komnrhjem									
Sogn og Fjordane	7,9	6,7	7,4	7,2	5,6	4,8	5,0	3,3	4,7
Askvoll	4,3	4,4	13,2	14,6	6,0	3,0	12,1	6,1	4,5
Aurland	7,5	7,7	5,3	2,7	-	-	3,0	6,1	3,0
Balestrand	2,9	3,0	-	3,2	6,6	-	-	7,5	3,8
Bremanger	11,8	8,7	10,8	13,2	8,7	5,4	5,6	3,4	5,8
Eid	6,4	3,2	7,8	4,6	5,3	5,3	4,0	2,7	5,9
Fjaler	8,1	-	4,1	11,1	4,1	2,7	6,9	8,4	-
Flora	11,4	6,7	10,9	8,3	8,4	7,5	5,0	2,4	5,0
Førde	12,3	11,3	6,9	9,4	5,8	3,8	5,0	3,9	6,2
Gaular	4,2	15,5	5,7	5,6	7,9	5,3	10,5	1,3	6,6
Gloppen	10,4	7,8	6,4	5,8	8,6	5,1	3,6	2,9	5,1
Gulen	7,2	12,7	5,6	-	1,9	1,9	3,9	7,7	1,9
Hornindal	14,3	-	3,0	3,1	3,1	3,3	3,3	-	3,5

Pussige Balestrand?

	kommune <chr>	tid_tekst <chr>	rate <dbl>	operasjonar <dbl>
1	Balestrand	2009 2.tertial	8.6	3.01
2	Balestrand	2010 1.tertial	9.1	2.99
3	Balestrand	2012 2.tertial	9.6	3.00
4	Balestrand	2013 1.tertial	9.9	3.01
5	Balestrand	2013 2.tertial	9.9	3.01
6	Balestrand	2016 1.tertial	11.2	3.00
7	Balestrand	2016 3.tertial	11.2	3.00
8	Balestrand	2017T3	11.4	3.01

Pussige alle?

3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 72
170 111 63 43 25 23 11 4 4 7 4 2 1 3 2 1

Datasett: Mandel operasjonar 0-18 år (I002)

Tal operasjonar pr. 1000 i aldersgruppa

NB! Alle rater er rekna om til 1-års rate.

komnrhjem: Blandet

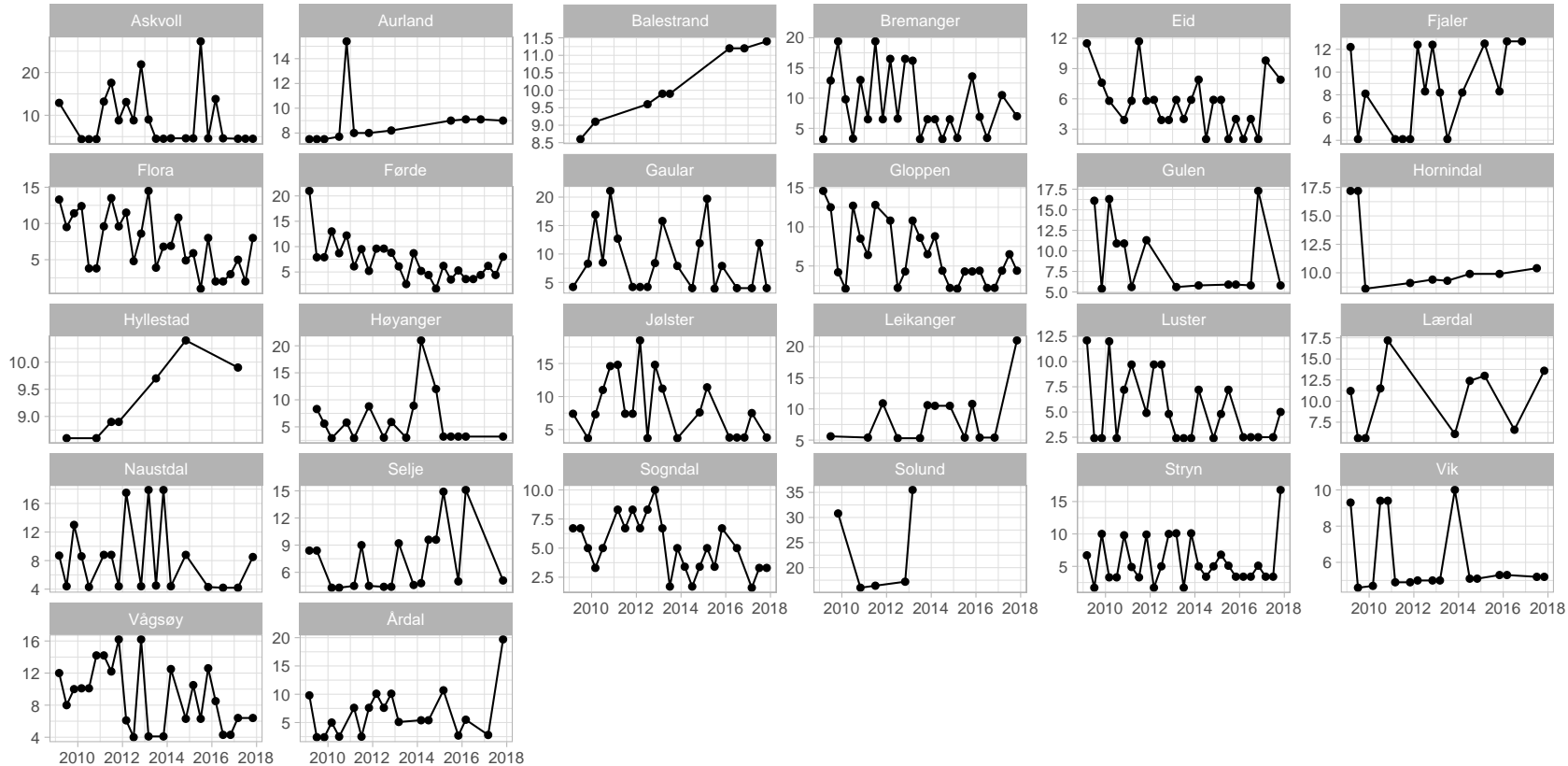
År-Tertial: År

Helseforetak: HBE-HDS-HFD

Måltall: Mandeloperasjonar 0-18 år

År-Tertial	2009	2010	2011	2012	2013	2014	2015	2016	2017
komnrhjem									
Sogn og Fjordane	7,9	6,7	7,4	7,2	5,6	4,8	5,0	3,3	4,7
Askvoll	4,3	4,4	13,2	14,6	6,0	3,0	12,1	6,1	4,5
Aurland	7,5	7,7	5,3	2,7	-	-	3,0	6,1	3,0
Balestrand	2,0	2,0	-	2,2	6,6	-	-	7,5	2,8

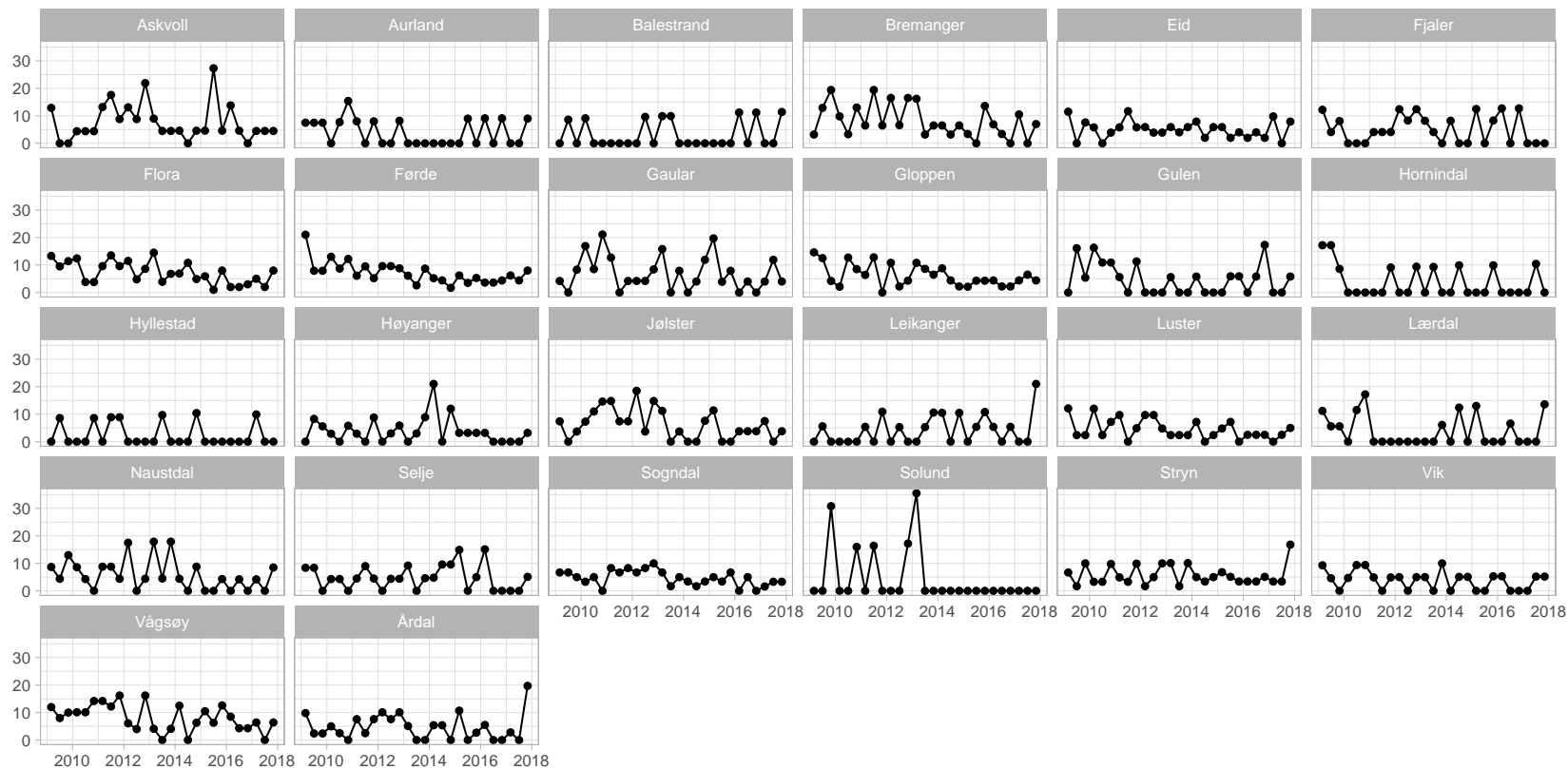
Manglande verdier



Manglande verdier eller 0?

År-Tertial	2009 1.tertial	2009 2.tertial	2009 3.tertial	2010 1.tertial	2010 2.tertial	2010 3.tertial	2011 1.tertial	2011 2.tertial	2011 3.tertial	2012 1.tertial	2012 2.tertial	2012 3.tertial
kommrhjem												
Sogn og Fjordane	10,4	6,3	7,1	7,4	5,3	7,4	7,3	7,6	7,1	7,7	5,3	8,5
Askvoll	12,9	-	-	4,4	4,4	4,4	13,2	17,6	8,8	13,1	8,8	21,9
Aurland	7,5	7,5	7,5	-	7,7	15,4	8,0	-	8,0	-	-	8,2
Balestrand 	-	8,6	-	9,1	-	-	-	-	-	-	9,6	-
Bremanger	3,2	12,9	19,4	9,8	3,3	13,0	6,5	19,4	6,5	16,5	6,6	16,5
Eid	11,5	-	7,6	5,8	-	3,9	5,8	11,7	5,8	5,9	3,9	3,9
Fjaler	12,2	4,1	8,1	-	-	-	4,1	4,1	4,1	12,4	8,3	12,4

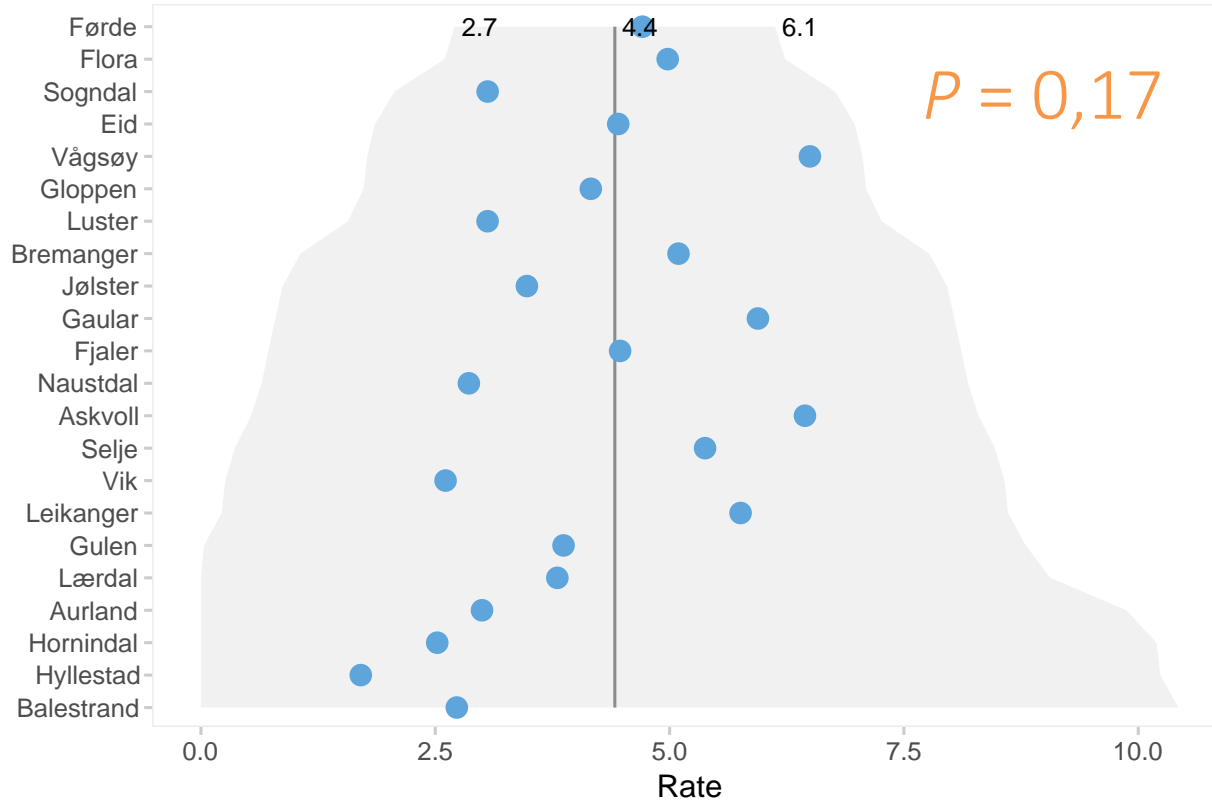
Fiksa datasett (og lik skala)



Stabil (predikerbar) prosess?



Samanliknar dei ein kan



Stikkord

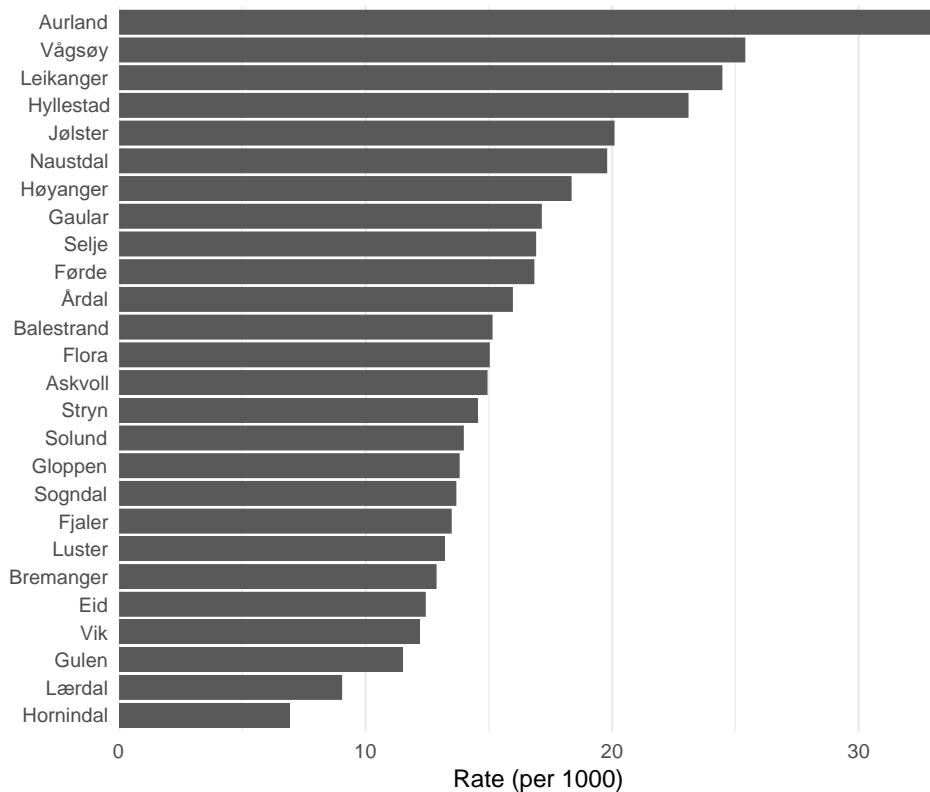
Statistisk prosesstyring (SPC)

Shewhart-diagram

Traktdiagram

Kva ein ikkje bør gjera

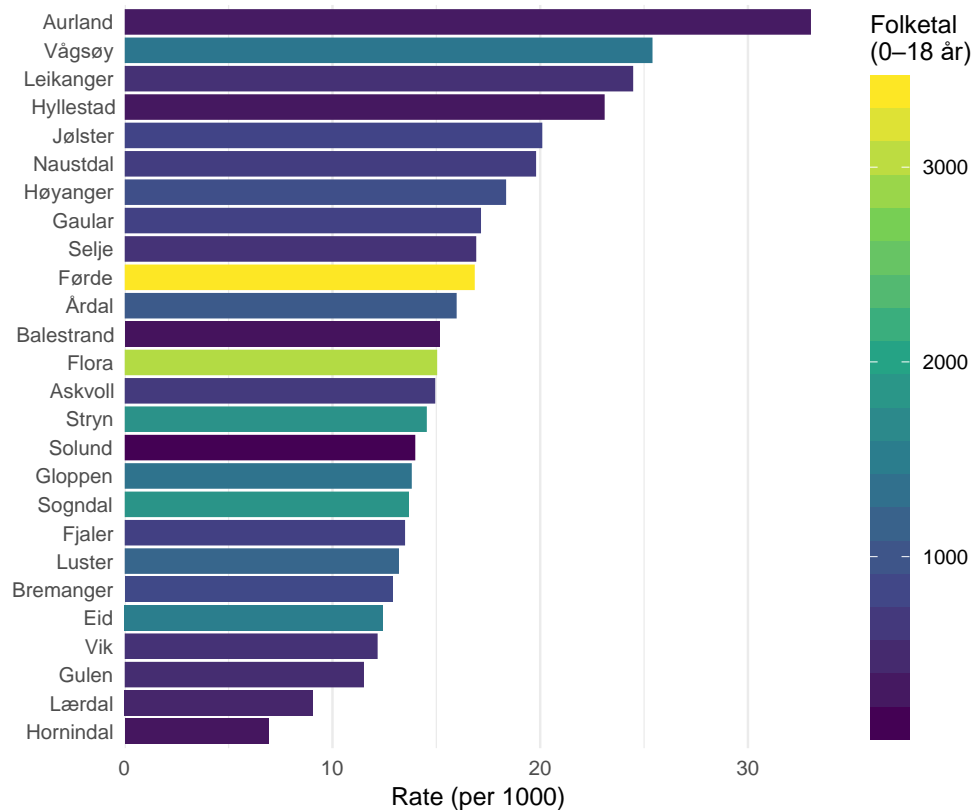
Naiv rangering (andre data)



Korfor på
topp/botn:
signal + støy

Små kommunar:
stor uvisse
(mykje støy)

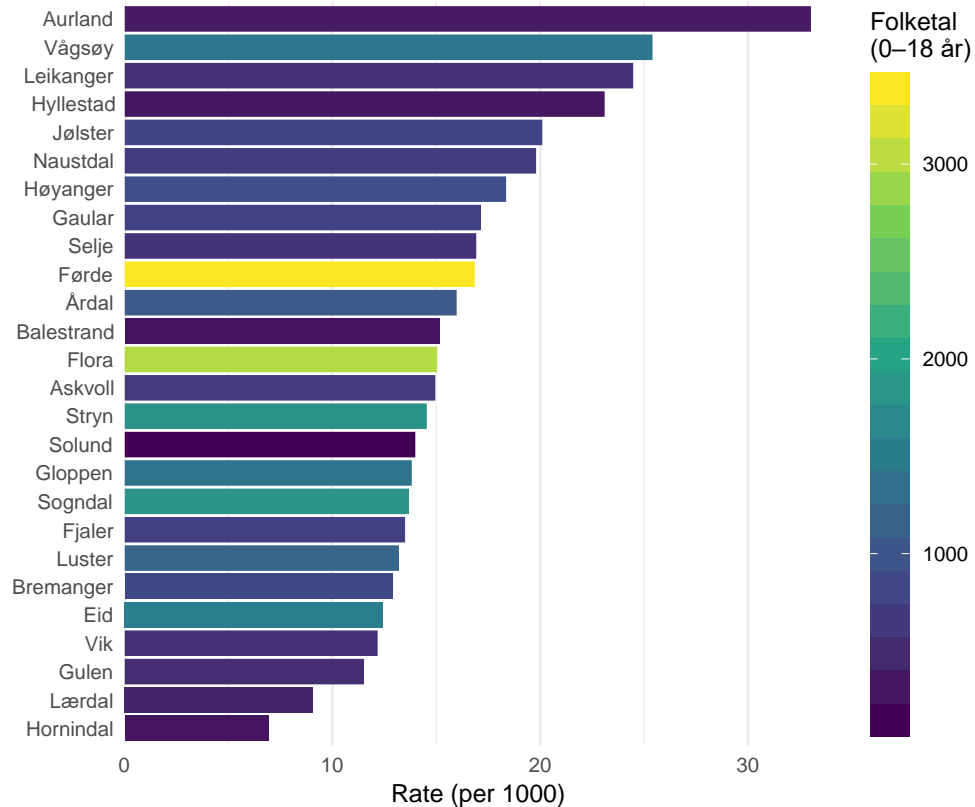
Naiv rangering



Små kommunar
dominerer nesten
alltid øvre/nedre del
av slike rangeringar.

Og er typisk (veldig)
over-/underestimering
av sanne verdier.

Dårleg mål på variasjon

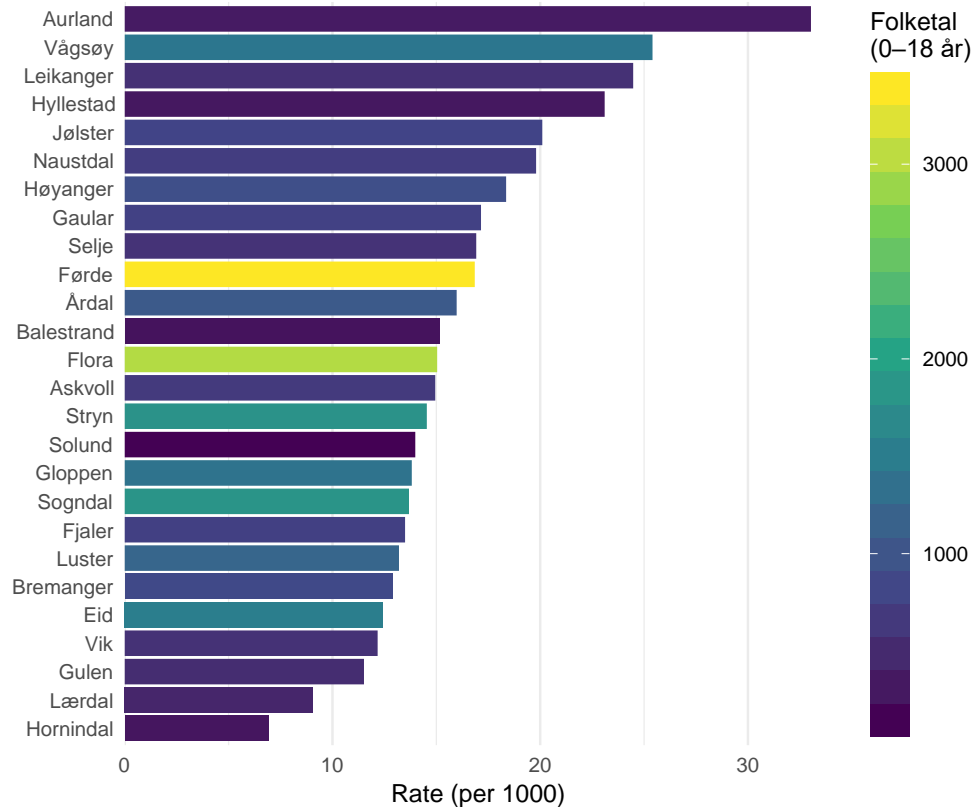


Ser av og til: Høgaste rate delt på lågaste rate.

Basert på tala som

- 1) er aller **mest usikre**
- 2) har **mest over- /underestimering**
- 3) har **dårlege statistiske eigenskapar** i ein slik formel

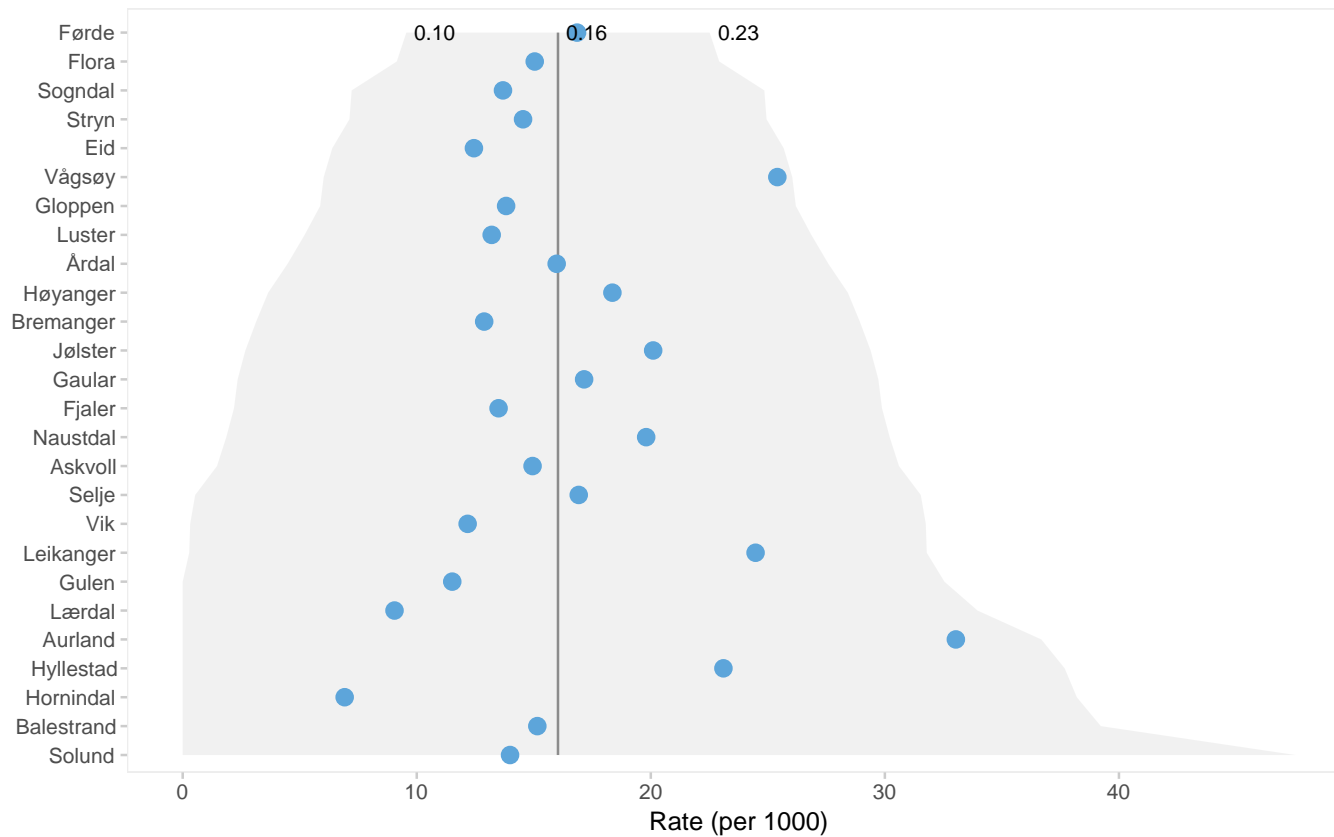
Dårleg mål på variasjon



Ser av og til: Høgaste rate delt på lågaste rate.

Er her Aurland/Hornindal:
 $33/7 = 4,7$

Rettare analyse



Kjapp fiks

Viss ein oppdagar ein
tilsynelatande forskjell:

I år

Aurland: 33 (per 1000)

Hornindal: 7 (per 1000)

$P = 0,02$

I fjor

Aurland: 12 (per 1000)

Hornindal: 27 (per 1000)

$P = 0,16$

Meir avanserte
statistiske metodar

«Betre» estimat

Kan bruke data frå **fleire** kommunar til å estimera variasjonen **mellom** kommunar.

Kan så bruke såkalla statistiske «**krympe-
metodar**» som tar omsyn til denne variasjonen, for å få estimat med **mindre over-/underestimering**.

«Betre» estimat

Byggjer på føresetnaden om at kommunar **liknar** (mykje eller lite) på kvarandre.

Ein «lånar» data frå kvarandre.
Eller «lånar» **styrke/presisjon**.

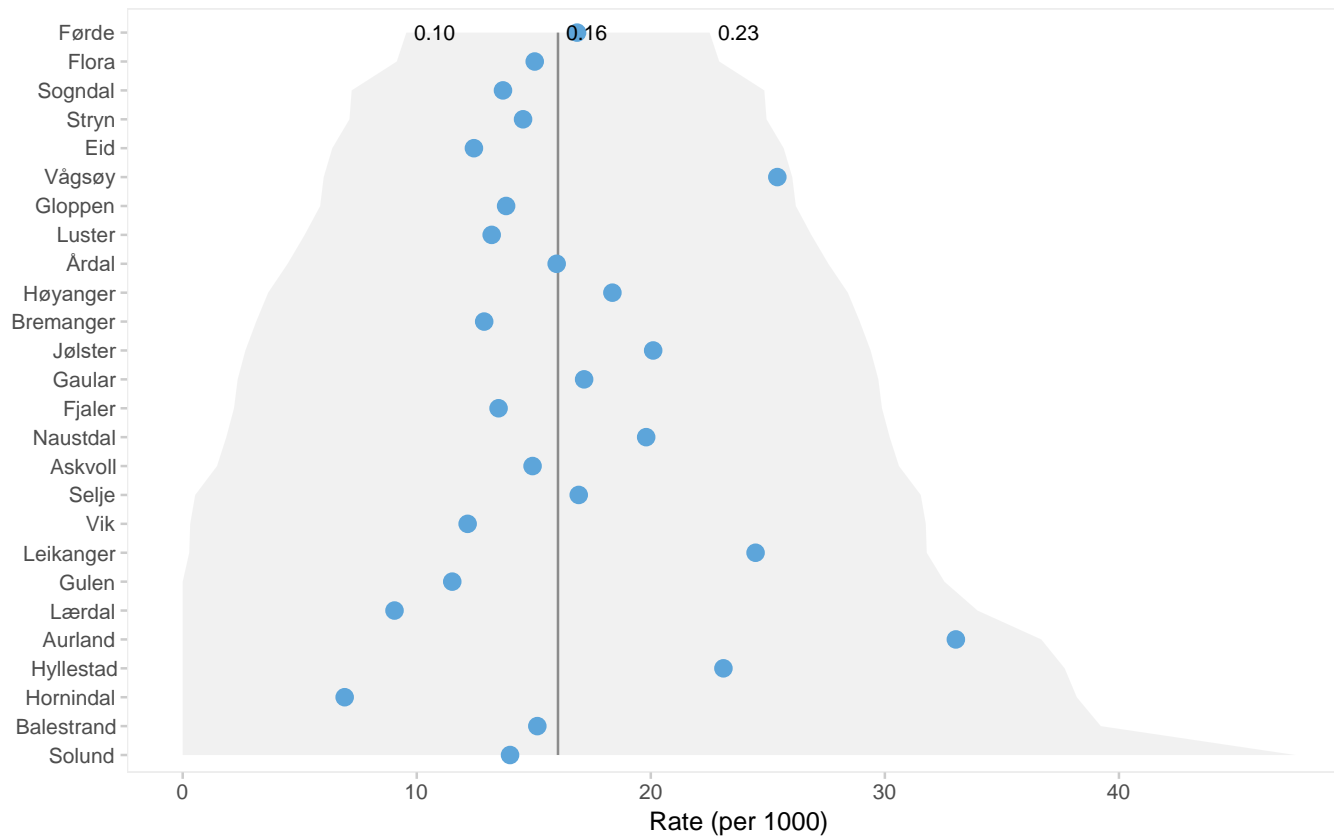
«Betre» estimat

Eksempel på slike krympemetodar er **mixed effects**-modellar og **bayesianske** modellar.

Fungerer **stabiliserande** når ein har lite data.

(Fjernar støy, men kan òg **glatta over signal**.)

Rettare analyse



Oppsummert

Oppsummert

Bruk data til å seia noko om **framtida**
(og **gjer noko** dersom ho ikkje ser lys ut).

Små data kan kanskje **gjerast store** ved å
låna data (i tid og rom).

Ikkje prøv å måla ting som **ikkje finst!**

SLUTT