

SEPTEMBER 2019

ANNA FOLKE LARSEN

KAN VI BRUGE MACHINE LEARNING **TIL AT FINDE HETEROGENE EFFEKTER** **I RANDOMISEREDE FORSØG?**

ROCKWOOL FONDENS INTERVENTIONSENHED – HVORDAN ARBEJDER VI



INDHOLD

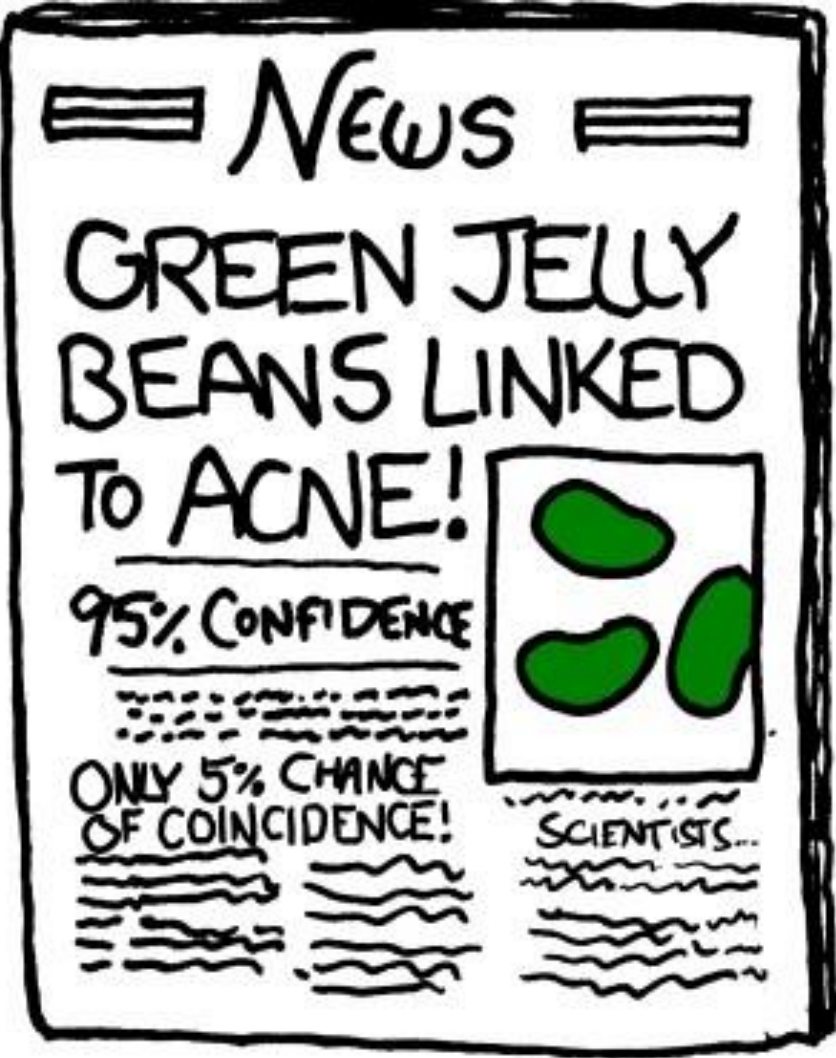
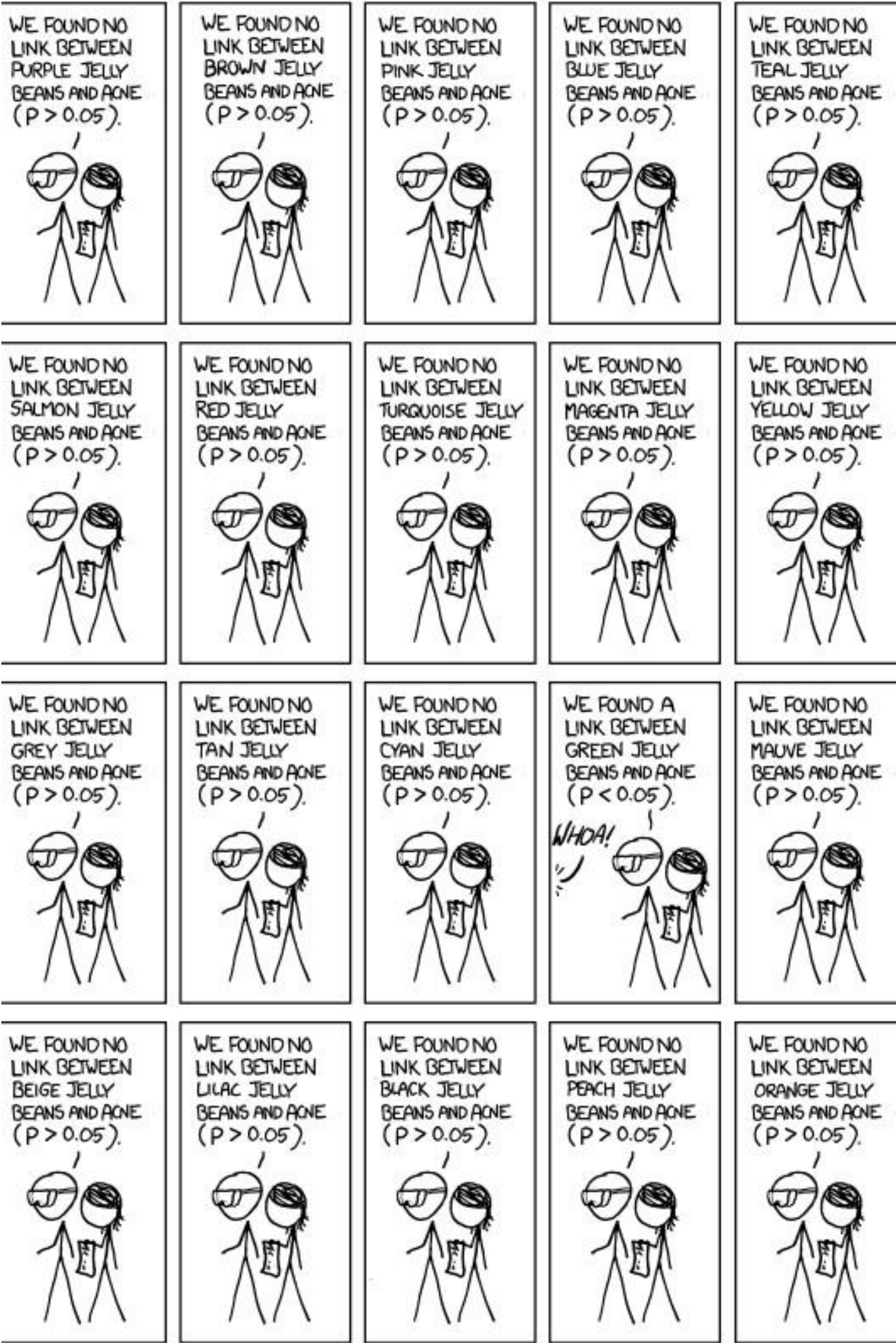
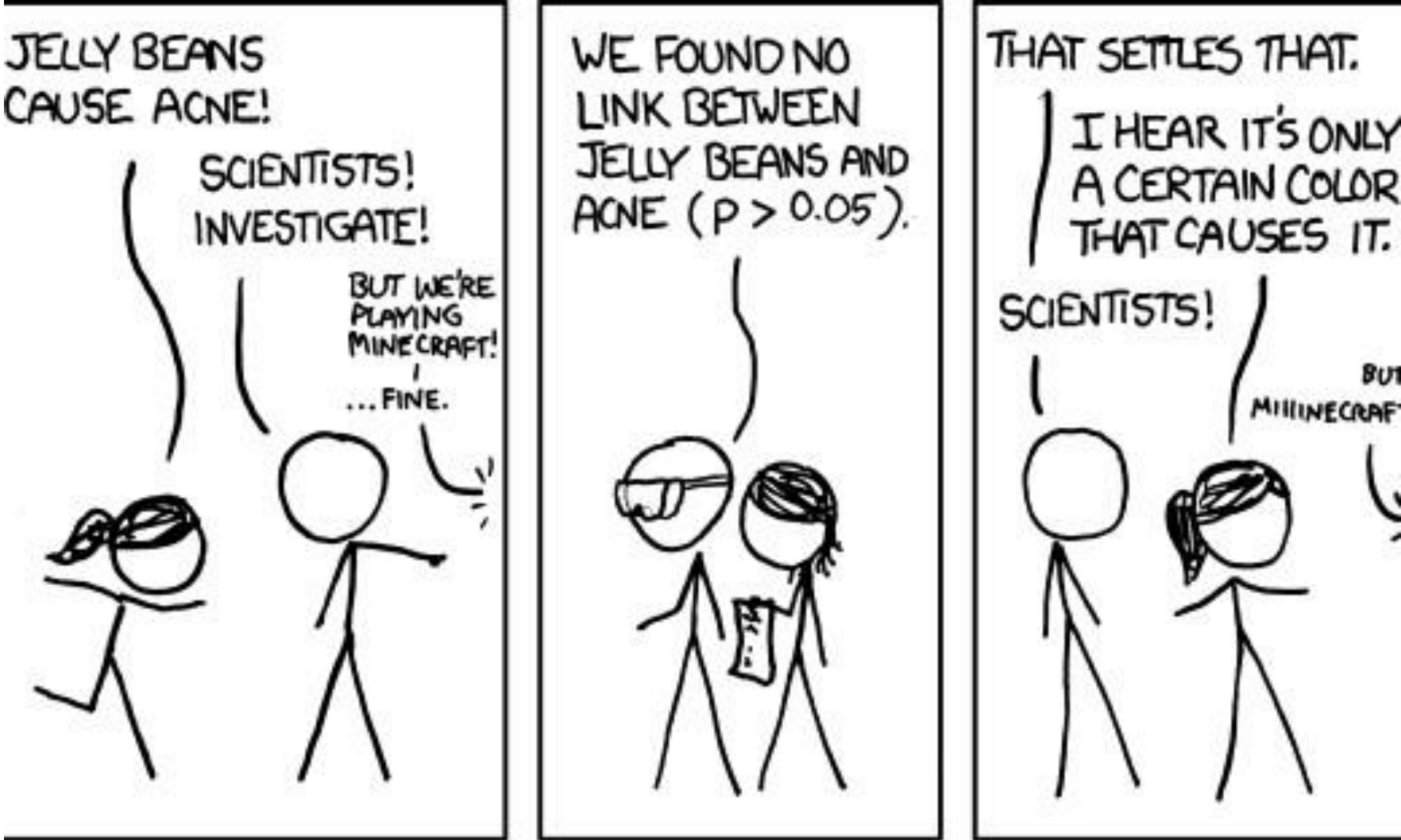
1. Evalueringsdesign og heterogene effekter
2. Machine learning – hvad kan det?
3. Perspekt: Socio-emotionel læringsindsats
4. Diskussion



EVALUERING MED RANDOMISEREDE FORSØG

- Tilfældigt udtrukket indsats- og kontrolgruppe:
 - Sikrer sammenlignelighed
 - Mange elever / klasser / skoler:
 - Øger generaliserbarhed
 - Statistisk styrke til at finde små effekter
 - Med ekstern forsker
 - Sikrer objektivitet og kvalitet
 - Pre-analysis plan
 - Undgår at plukke de bedste resultater
- Gennemsnitseffekt for hele målgruppen + effekter for bestemte undergrupper
 - Hvilke undergrupper er relevante?
 - *Lad os bare prøve en hel masse!*
 - Men mange undergrupper øger sandsynlighed for tilfældigt at finde et falskt positivt resultat
 - (Ellers skal man have et større sample og justere signifikansniveau)

STATISTISK SIGNIFIKANS OG FALSK POSITIV



DILEMMA: HETEROGENE EFFEKTER

Dilemma

- Indsatser har typisk ikke samme effekt for alle
- Svært at vide på forhånd, hvilke undergrupper som har mest gavn af indsatsen
- Vi mister statistiske egenskaber ved hypotesetest hvis vi tester mange undergrupper
- Vigtigt at man i Pre-analysis plan binder sig, så man ikke fisker falske positive resultater bagefter

Hvad kan vi bruge machine learning til?

- I stedet for på forhånd at binde os til at

**Estimere og teste
bestemte heterogene effekter**

- Så binder vi os til en bestemt metode til at

**Estimere og teste
vigtige egenskaber
ved de heterogene effekter**

MACHINE LEARNING METODE

Hvilke egenskaber kan vi estimere og teste:

1. Er der heterogene effekter?
2. Hvor stor er effektforskel mellem de 10% mest påvirkede og de 10% mindst påvirkede?
3. Beskrivende data på de 10% mest påvirkede og de 10% mindst påvirkede

Beskrevet i Chernozhukov et al. (2018)

Metoden kan ikke svare på relative effektstørrelser for forskellige grupper, fx

- Er effekten større for børn af forældre som kun har grunduddannelse?
- Er der signifikant forskel i effekt mellem drenge og piger?

Men at anvende denne ML metode udelukker ikke, at man også laver denne slags test.

PERSPEKT – SOCIO-EMOTIONEL LÆRINGSINDSATS

PERSPEKT 2.0

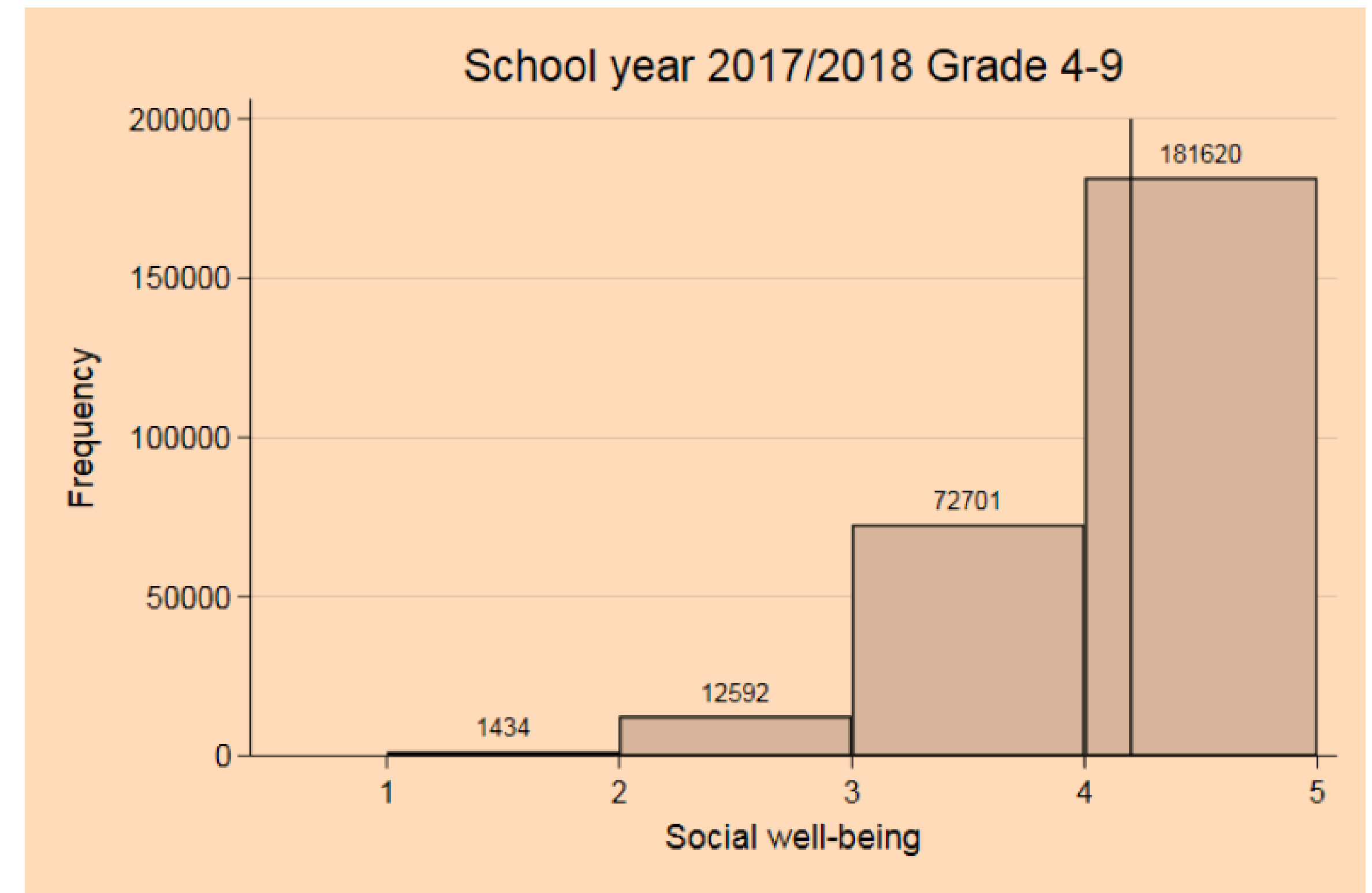
- Udviklet af ROCKWOOL fonden
- Videreudviklet af DCUM i 2018
- Mål: At øge elevernes emotionelle, sociale og personlige færdigheder i grundskolen
- Materialet er 15 kapitler / lektioner
- Fokus bl.a. på børns forskellighed, egne grænser, fællesskab, kommunikation

EVALUERING – et randomiseret forsøg

- 70 skoler deltager med 4. og 5. klassetrin
- 314 klasser, 6500 elever
- Tilfældigt udtrukket om 4. eller 5. klassetrin får indsatsen, det andet klassetrin er kontrol
- Indsatsklasser har brugt PERSPEKT i skoleåret 2018/19
- Ledes af Professor Marianne Simonsen, Århus Universitet
- Komplementeres af en implementerings-evaluering af VIVE

SOCIAL TRIVSEL

- Primært effektmål: Social trivselsindikator fra den nationale trivselsmåling
- Mange elever har høj social trivsel i udgangspunktet
- Derfor forventer vi større effekt blandt elever med lav trivsel året inden indsatsen



HETEROGENE EFFEKTER AF PERSPEKT

Vi har prædefineret følgende undergrupper:

- Baseline social trivsel under medianen
- Baseline social trivsel i 1. kvartil
- Baseline social trivsel i 2. kvartil
- Dreng // Pige
- 4. klasse // 5. klasse
- Forældres oprindelsesland
(begge dansk // én eller begge ikke-dansk)
- Mors uddannelse
(grundskole/faglært/gymnasial // mere udd.)

- Machine learning metode muliggør at vi kan blive klogere!
- Finde undergrupper vi ikke vidste var vigtige
- Fx særlige kombinationer af baggrundsvariable

DISKUSSION

- Man kan ikke undersøge det man ikke ved, er relevant at undersøge
 - Machine learning metode giver mulighed for at opdage noget, som vi ikke vidste var relevant at kigge på – og teste for det
- Statistiske metode vil altid være begrænset til at undersøge
 - det vi kan kvantificere
 - det vi vælger at inkludere
 - Også denne machine learning metode
 - Stadig vigtigt at koble med kvalitative indsigter