

Statistikproduktion 4.0 – betydelse för Sveriges officiella statistik (SCB)

Joakim Malmdin
Gustaf Strandell



Disposition

Hur producerar vi Sveriges officiella statistik idag, sp3.0?

Har förutsättningarna för officiell statistik ändrats? Vad är nytt?

Hur producerar vi Sveriges officiella statistik i framtiden?

Hur utvecklar vi statistikproduktion 4.0?

Vad innebär Statistikproduktion 4.0 för kvaliteten i

-data,

-processer och

-statistik?

Hur producerar vi Sveriges officiella statistik idag, sp3.0?

Designbaserad och modellassisterad urval- och estimationsteori

SCB samlar in data från individer och företag genom digitala blanketter, pappersblanketter och telefonintervjuer

SCB har god tillgång till administrativ data som vi utnyttjar i produktionen av officiell statistik

- Registret över totalbefolkningen
- Företagsdatabasen
- Fastighetsregistret
- Arbetsgivardeklarationer på individnivå, momsdata, företagens skattedeklarationer (sru)



Hur producerar vi Sveriges officiella statistik idag, sp3.0?

Exempel:

Arbetskraftsundersökningarna (AKU):

Bygger på telefonintervjuer med individer. Administrativa uppgifter används i såväl design som estimation (regressionsestimator).

Företagens ekonomi (FEK):

Bygger i första hand på administrativa data (skattedeclarationer). För att specificera uppgifterna skickas blanketter till urval av företag

Omsättningsstatistiken

Momsuppgifter från Skatteverket används som enda källa för alla företag där så är möjligt. Egen insamling begränsas till de största företagen och de där momsuppgifterna inte kommer in snabbt nog till Skatteverket

Har förutsättningarna för officiell statistik ändrats? Vad är nytt?

- Individens vilja att lämna uppgifter sjunker
- Ofantliga mängder digital data
- Plattformer för samarbete om data (utbyte, produktion och användning) – det nationella dataekosystemet
- Allt snabbare teknikutveckling
- Utökad behov av analyser och temporär statistik

Hur producerar vi Sveriges officiella statistik i framtiden

Utökad användning av administrativa data och nya datakällor

- Månatliga arbetsgivardeklarationer på individnivå (AGI)
- Transaktionsdata
- Positionsdata för mobiltelefoner
- Elmätardata
- Jobbannonsdata
- Betalkortsdata

Hur producerar vi Sveriges officiella statistik i framtiden

Minskad börda för uppgiftslämnarna till den officiella statistiken

- Utökad användning av administrativa och nya datakällor ger minskat behov av insamling från uppgiftslämnare
- Insamlingen är designad för att minska uppgiftslämnarbördan
 - > Individer har möjlighet att svara via mobilen
 - > Insamlingen sker direkt från företagens system där så är möjligt
 - > SCB har god kännedom om vilka uppgifter som är svåra att lämna

Hur producerar vi Sveriges officiella statistik i framtiden

Snabb och effektiv statistikproduktion

- Där så är möjligt har manuella moment automatiserats, till exempel med hjälp av ML

Flexibel produktionsprocess

- Kan hantera olika typer av data och nya/förändrade användarbehov

Säker produktionsmiljö

- Spårbarhet, reproducerbarhet, säkrad tillgång till data

Hur producerar vi Sveriges officiella statistik i framtiden

Mer samarbete

- Mellan universitet, producenter av officiell statistik och andra aktörer
- Mellan producenter och användare av officiell statistik

Nya former för redovisning av statistik – mer dynamik

Fortsatt hög medvetenhet om kvalitén i den officiella statistiken

- Fortsatt hög medvetenhet om den statistiska inferensen

Nya metoder och verktyg i verktygslådan

- Samspelet mellan datakällor

Hur utvecklar vi statistikproduktion 4.0?

Ny organisation på SCB från 1 september 2021

- Ny enhet för metod och arkitekturstyrning ansvarar för att utveckla SP4.0

Omvärldsbevakning och samarbete med andra aktörer både inom och utom statistikvärlden - forskare såväl som utförare.

- Det nationella dataekosystemet
- AI Sweden

Prova oss fram utifrån informationsbehov och tillgänglig data – dra generella slutsatser

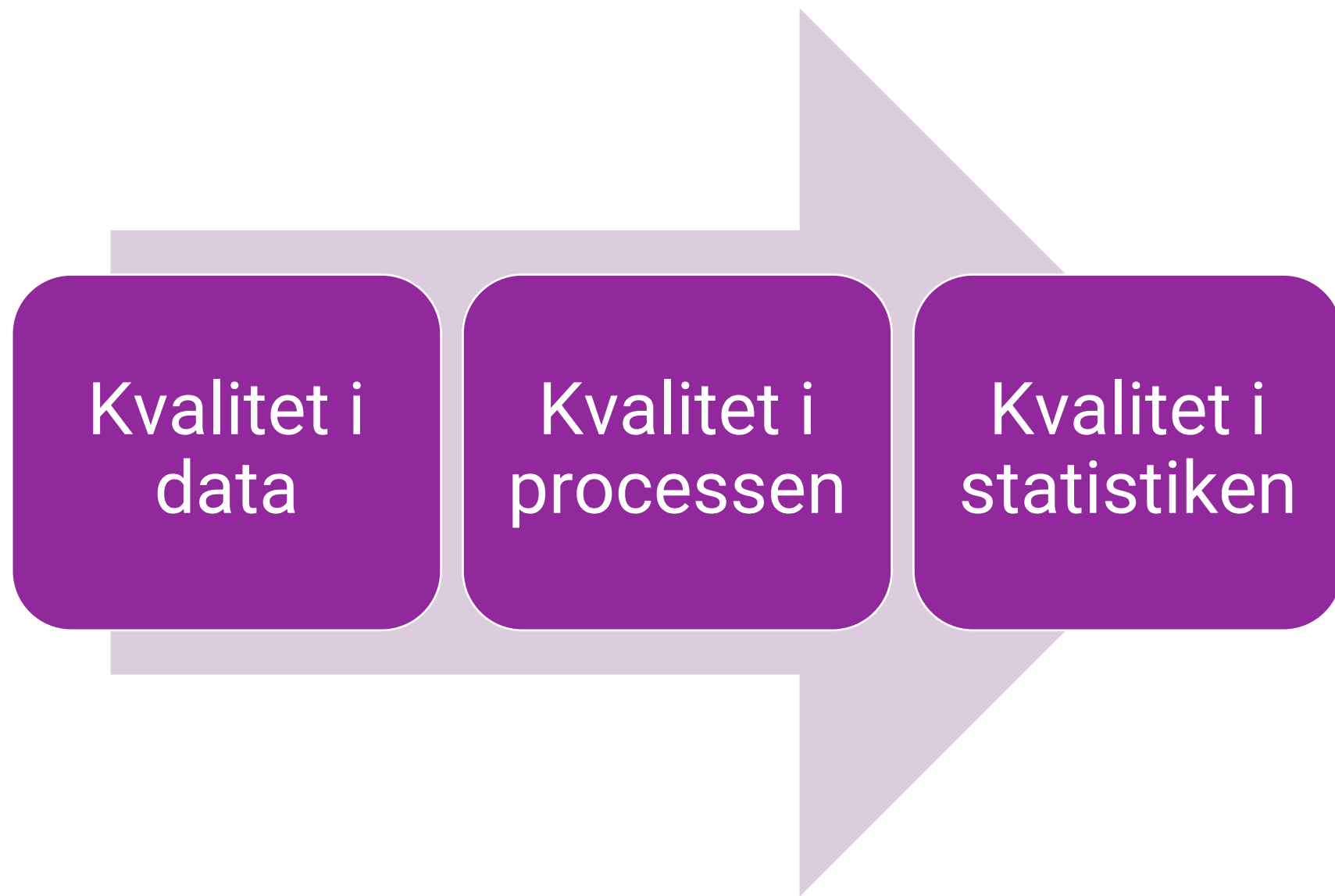
Bygga på det vi kan, men inte vara rädda att omvärdera/testa nytt



***Vad innebär Statistikproduktion 4.0
för kvaliteten i
-data,
-processer och
-statistik?***

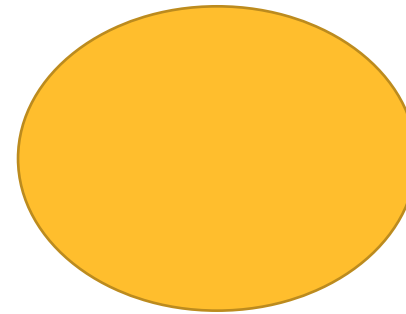
- Förflyttningen mot SP 4.0 handlar, ur ett kvalitetsperspektiv, mestadels om att utveckla och kvalitetssäkra nuvarande statistikproduktionsprocess och anpassa den efter nya förutsättningar.
- Processen har tjänat oss väl genom SP 2.0 och SP 3.0. Nu gäller det att säkerställa den för SP 4.0.
- Vad är då skillnaderna? Och vad betyder det för kvaliteten i statistiken?

- För användarna, som tar del av statistiken, ska det inte innebära någon skillnad alls. Den statistik med ändamålsenlig kvalitet som vi levererar idag ska vi också leverera imorgon. Allt i enlighet med Lagen om officiell statistik.
- Punkten ovan kan ses som något pessimistisk - i och med att kvalitetskomponenten *Relevans* består av bl.a. användarnas informationsbehov, finns goda möjligheter att *höja* den officiella statistikens kvalitet när informationsbehoven i högre grad tillgodoses.
- Hur uppnår vi det sistnämnda?

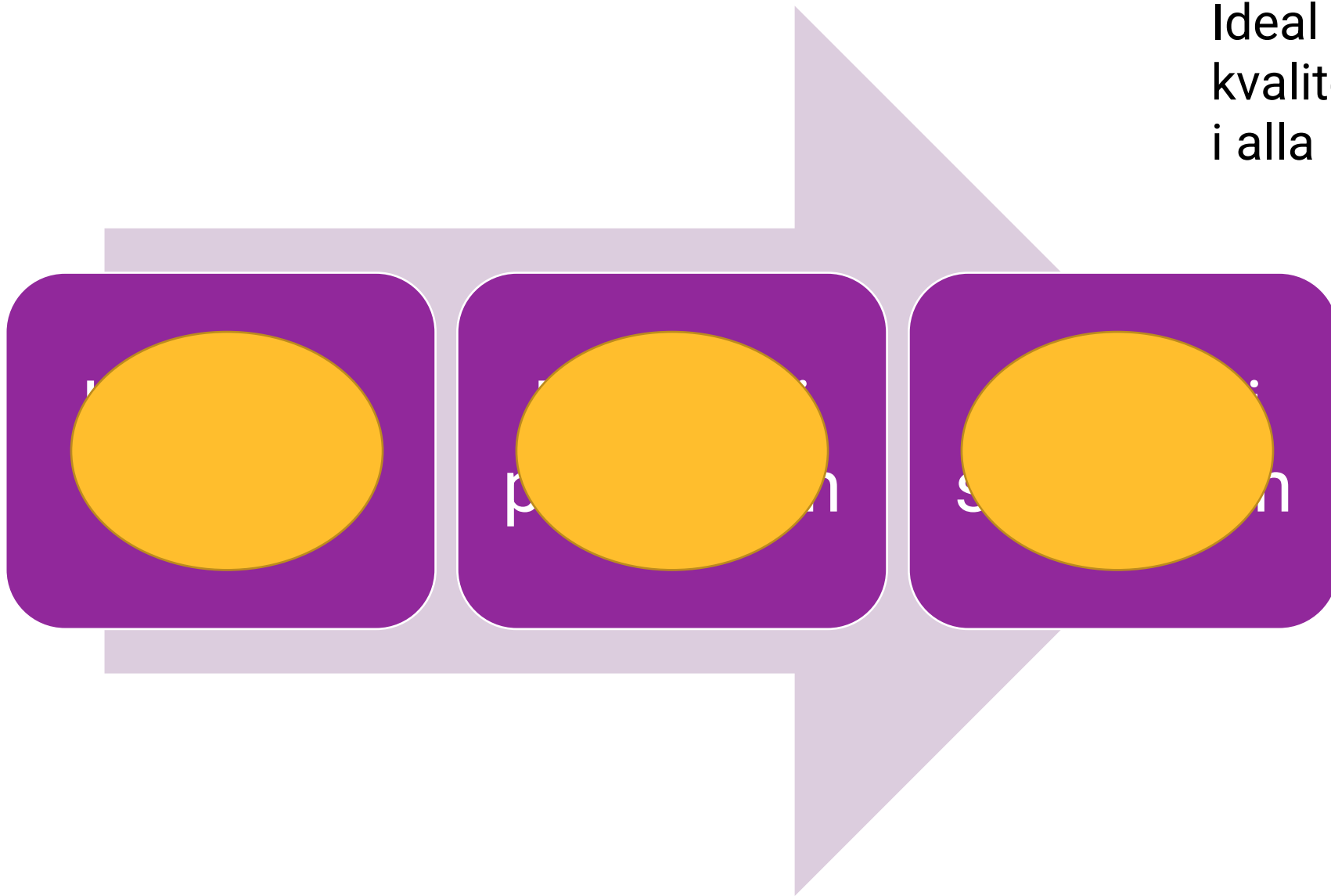


För att kvaliteten i statistiken ska vara anpassad till våra användares behov behöver vi säkerställa att kvaliteten är ändamålsenlig i insamlings- och produktionsprocessen.

Önskad kvalitet och uppnådd kvalitet

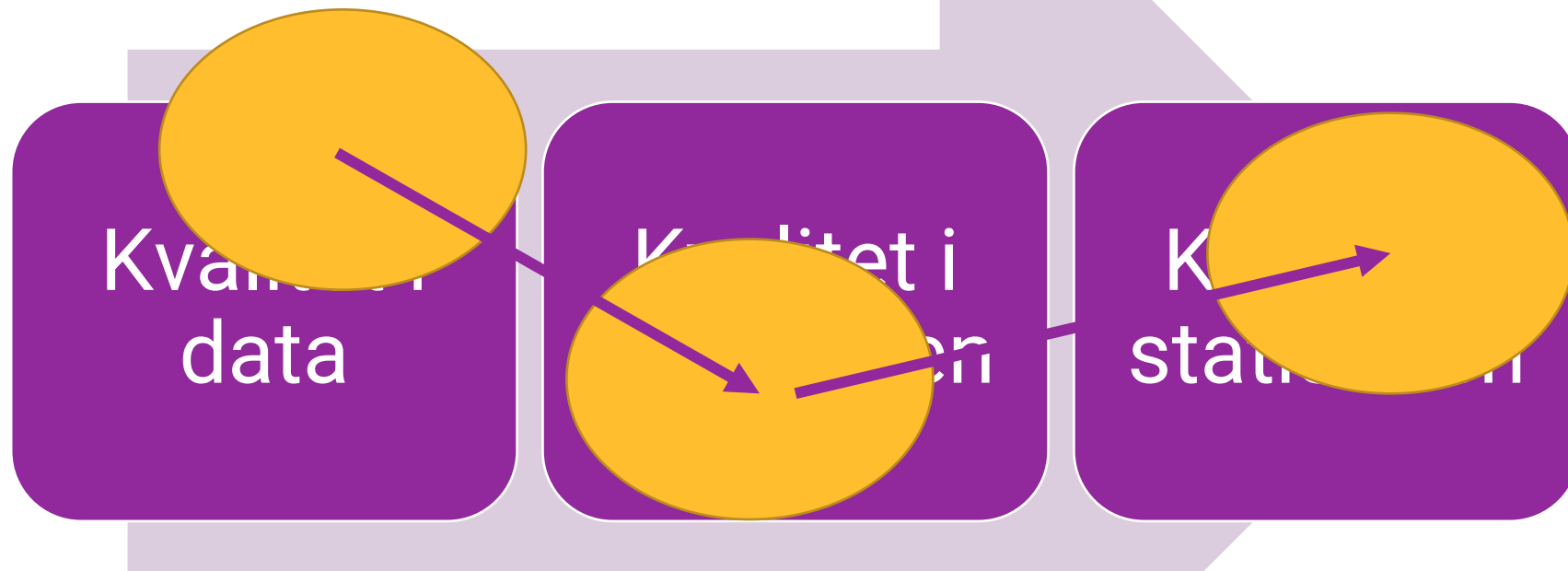


Ideal process –
kvaliteten perfekt
i alla led



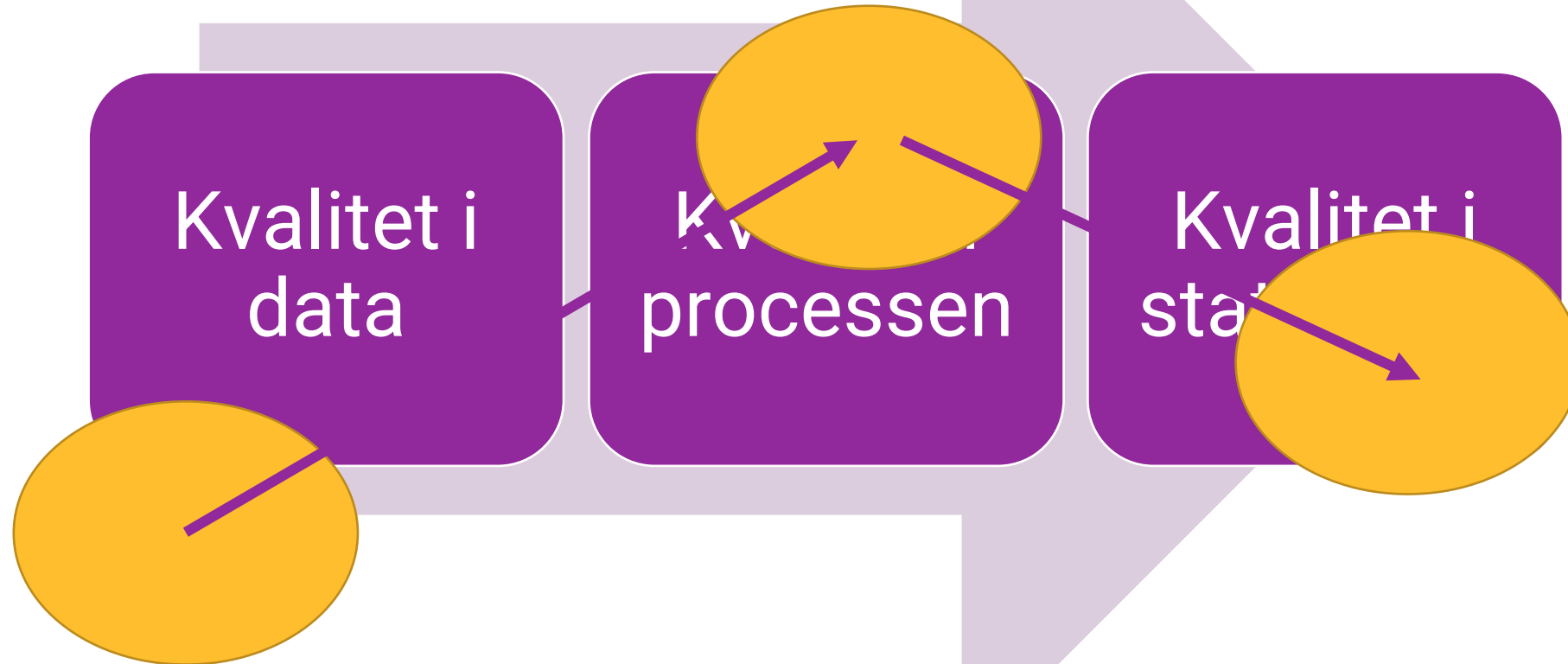
SP 2.0

Statistik baserad på
urvalsundersökningar



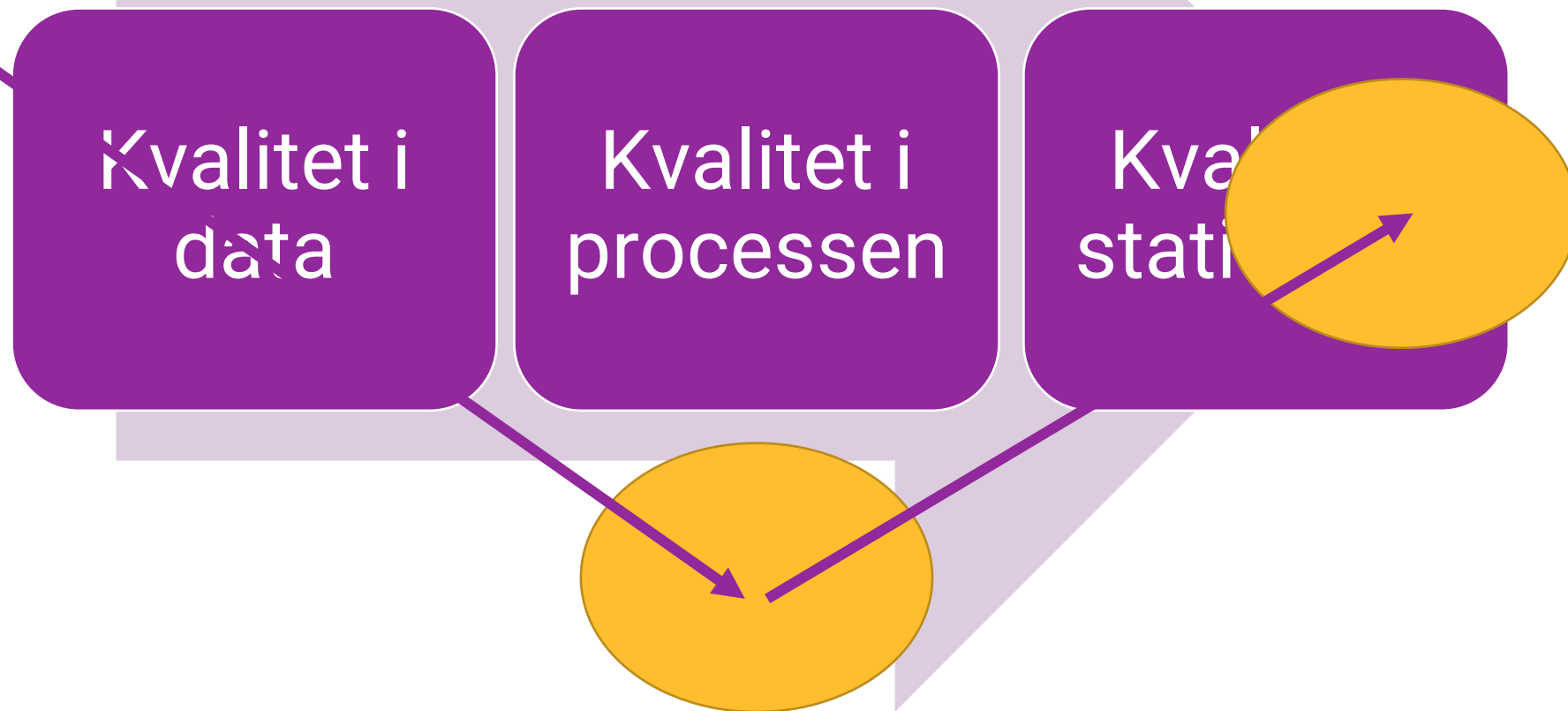
SP 3.0

Statistik baserad på administrativa data och urvalsundersökningar i kombination



SP 4.0

Statistik baserad på administrativa data, urvalsundersökningar och nya datakällor i kombination



- Hur mycket kan den uppnådda kvaliteten i data avvika från den önskade samtidigt som processen kan hantera skillnaden för att i slutändan statistikens kvalitet ska vara anpassad efter användningen?
- En process som ”hanterar skillnaden” i uppnådd och önskad kvalitet i datastadiet är förstuds helt beroende av metodutveckling vad gäller modeller, kombination av källor och annan information som kan vara av betydelse för slutresultatet.

SCB