

Hvordan vi teller

DANIEL INSTEBØ

RÅDGIVER STATISTISK SENTRALBYRÅ



Statistisk sentralbyrå
Statistics Norway

Hvem er vi... egentlig?

- Daniel Instebo, rådgiver
- Maria Østli, rådgiver
- Ingeborg Kool, førstekonsulent
- Inngår i «Seksjon for Kultur- og utdanningsstatistikk»
- Jobber med alt fra datamottak, databearbeiding, analyse, rapportskriving og oppdragshåndtering
- I vår seksjon lager vi statistikk over alt fra barnehage, avslutta grunnskole, nasjonale prøver, videregående opplæring, høyere utdanning [...]



Videregående opplæring og annen videregående utdanning

Videregående opplæring og annen videregående utdanning

- En tverrsnitt-statistikk som viser status inneværende og siste skoleår
- To telletidspunkter:
 - 1. Oktober (inneværende år) → ssbElev, ssbElevFag, ssbKontrakter
 - 1. Juli (siste skoleår) → ssbResultat, ssbResultatFag, ssbAvlagtePrøver
- Ikke det samme som «Gjennomføring i videregående opplæring»
 - Tidsserie – følger et årskull over 5/6 år og gjør opp status over hvor mange som har fullført
- Følger årshjulet lagt frem av Udir
 - Datamottak av søkerfiler på sommeren, elevfiler på høsten og fagopplæring på nyåret
 - Statistikkpublisering i midten av februar



Statistikk om

Videregående opplæring og annen videregående utdanning

Statistikken gir tall for elever, lærlinger, lære kandidater samt deltakere i videregående opplæring og annen videregående utdanning.

Oppdatert: 20. februar 2025

Neste oppdatering: Foreløpig ikke fastsatt



Statistisk sentralbyrå
Statistics Norway

Videregående opplæring			
	2024	Endring i prosent	
		2023 - 2024	2020 - 2024
Elever	202 914	2	2
Læringer/lærekandidater ¹	53 285	2	10
Elever og læringer/lærekandidater	256 199	2	4
Hvorav			
Innvandrere	39 623	3	7
Norskfødte med innvandrerforeldre	17 454	7	27
¹ Elever med fagopplæring i skole er inkludert			

[Standardtegn i tabeller](#)

Kilde

[Statistikkbanken kildetabell 13159](#)



Statistisk sentralbyrå
Statistics Norway

Hva kan vi lese ut fra den tabellen?

- **Konkrete tall:** 202 914 elever og 53 285 lærlinger
- **Totalt:** 256 199 elever og lærlinger til sammen
- **Presisjon:** Ingen slingringsmonn – tallene må stemme nøyaktig
- **Konsistens:** Summen skal være lik uansett om vi bryter ned på fylke, kommune eller skole
- **Krav:** Dette stiller store krav til SSB og andre databehandlere
- **Oppfølging:** Hver tabell vi publiserer er gjenstand for spørsmål



Hvordan sikrer vi god datakvalitet?

Mål: Korrekte, stabile og etterprøvbare data

Administrasjonsperspektiv

- Deltagelse i faggrupper
- Felles variabelliste
- Lik struktur hvert år
- Alle endringer drøftes før de implementeres

Brukerperspektiv

- Grensesnitt må støtte korrekt registrering
- Klare feltbeskrivelser, gode hjelpetekster, logiske valg
- Kodelister må være logiske, stabile og uttømmende
- Unngå overlap, doble koder, tvetydige begreper
- Variabellister må være forståelige og relevante
- Brukere må forstå hva som betyr hva

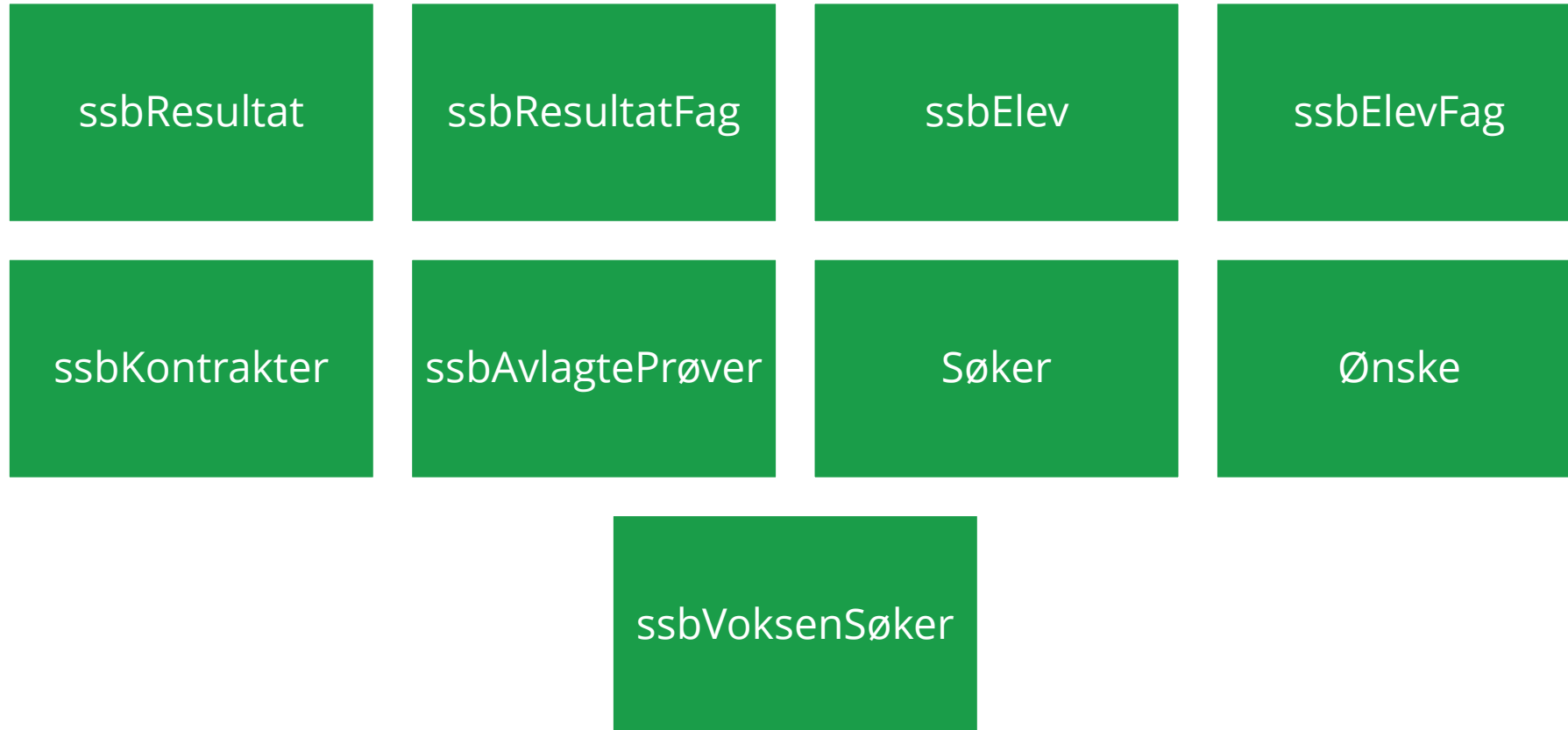
Databehandlerperspektiv

- Kontroller i alle ledd, fra registrering til datamottak til publisering
- Automatisert validering og sammenligning mot tidligere år
- Tett dialog med dataleverandør om feil og avvik
- God dokumentasjon



Fra prinsipper til praksis

Oversikt over de ulike filene



Oversikt over de ulike filene

- Hver fil dekker ulike deler av opplæringsløpet
- Hver fil dekker ulike formål
- Hver fil behandles separat
- Men de må samtidig henge sammen
- Datakvalitet må sikres på tvers – filene skal kunne kobles sammen ved behov



Ulike typer variabler

Persondata

- Fødselsnummer
- Fødselsdato
- Navn
- Skolefylke?

Nøkkelvariabler

- Fødselsnummer
- Programområdekode
- Skolenummer
- Organisasjonsnummer
- Fagkode ?

Kursinformasjon

- Elevstatus
- Kursprosent
- Antall fag
- Karakterer
- Fravær
- Kontraktstype



Persondata

- Formål: Sikre at vi teller reelle personer
- Kontroller:
 - Fødselsnummer og D-nummer valideres automatisk
 - Ugyldige nummer sjekkes mot Skattetaten
 - Eventuelle feil korrigeres automatisk før dataene går videre til bearbeiding og kontroll.
 - Korrekt identifikasjon er avgjørende for å kunne følge elever på tvers av år og utdanningsløp.



Nøkkelvariabler

- Danner grunnlaget for å koble elevdata mot andre datakilder

Variabel	Bruksområde
Fødselsnummer	Kobler på fødselsdato, kjønn, innvandringskategori, første oppholdsdato, foreldres utdanningsnivå [...]
Programområdekode	Knytter til Vigo Kodeverksbase for ytterligere informasjon om programområdet
Skolenummer (evt. orgnr)	Knytter til Vigo Kodeverksbase. Brukes til å identifisere skolens fylke, næringskode og gyldighet.



Kursinformasjon

- Kursinformasjon beskriver elevens opplæringsløp og danner grunnlaget for beregninger av aktivitet, progresjon og gjennomføring.
- Kursaktivitet: Antall fag, kursprosent, årstimer
- Elevstatus: Skiller ordinære elever fra voksne, privatister og elever med AOP
- Resultater: Omfatter karakterer, bestått/ikke bestått, fravær, vurderingsform og fullførtstatus.

En ordinær dataflyt

Kildedata

- Data i sin opprinnelige form
- Arkiveres slik at man alltid kan gå tilbake og lage datagrunnlag på nytt

Inndata

- Anonymisert
- Minimert
- Validering
 - Sjekket opp mot gyldige utfall
 - Sjekket opp mot kodeverk
- Editering (feilretting)

Klargjorte data

- Koblet mot filer fra andre kilder (befolkning, BoF)
- Avsluttende validering
- Slå sammen igang- og avslutta-filer til et selvstendig datagrunnlag

Statistikkdata/Utdata

- Lastes til statistikkbank
- Datagrunnlag lagres/arkiveres for senere bruk
- Opplysninger legges inn i Nasjonal Utdanningsdatabase



Vanlige utfordringer i datagrunnlaget

- Eksempler på feil i personinformasjon:
 - Ugyldige eller feilskrevne navn (testpersoner, dubletter, tekniske feil)
 - Feilregistrert fødselsår (F.eks 1902 i stedet for 2002)
 - Manglende eller ugyldige fødselsnummer/D-nummer



Vanlige utfordringer i datagrunnlaget

- Typiske kontrollfunn:
 - Antall fag: Urealistisk høye verdier (> 40) eller kursprosent på over 100%
 - Programområdekoder: Må samsvare med gyldige kodelister
 - Elevstatus: Feil bruk av statuser kan føre til feil i tabeller og tellinger
- Hva dette betyr i praksis:
 - Selv om det finnes gode kontroller ved datainnsamling oppdages mange feil først når data brukes i statistikken
 - Krever tett dialog med leverandører og kontinuerlig justering av kodeverk og kontroller
 - Feil som ikke rettes blir gjenstand for rapporter og analyse senere
 - Varig høy datakvalitet oppnås kun når kvalitet sikres allerede ved registrering.



Publiseringer fra SSB

- Betydningen av innvandrerbakgrunn for muligheten til å få læreplass (Rapport 2025/26) – Ingeborg Kool
- Utdannings – og yrkeskarrierer for yrkesfagelever (Rapport 2025/16) – Andrine Stengrundet og Daniel Instebø
- Krav til fremmedspråk i studieforberedende utdanningsprogram, fordel eller ulempe for innvandrererelever? (Artikkel) – Kjartan Steffensen og Jiaying Ye
- «Alle» med yrkesfaglig utdanning i Norge er i jobb (Artikkel) – Andrine Stengrundet og Geir Nygård



Datakvalitet

Hvorfor det er viktig

Grunnlag for felles politikkutforming, styring og prioriteringer

Sikrer at tallene faktisk beskriver virkeligheten

Nødvendig for at samfunnet skal kunne stole på og bruke tallene våre

Sikrer dermed tillit til offisiell statistikk som felles kunnskapsgrunnlag

Gjør statistikken etterprøvable og sammenlignbar

Tillit, ansvar og presisjon

Hva det gir oss

Gir et pålitelig grunnlag for å følge elevenes løp over tid

Gjør det mulig å beregne gjennomføring, frafall, progresjon og overgang

Sikrer konsistente data på tvers av filer, år og kilder

Legger grunnlaget for sammenlignbar og robust statistikk

Gir forvaltning, forskning og offentlighet et felles faktagrunnlag

Muligheter og verdiskaping



Udir og SSB – ulike roller, felles grunnlag

- Begge publiserer statistikk og bearbeider de samme filene, men med forskjellig formål og metode
- Separat databearbeiding gir små forskjeller i resultatene
- SSB innhenter data med hjemmel i statistikkloven, som pålegger opplysningsplikt:
«Enhver skal uten hinder av taushetsplikt og etter pålegg fra Statistisk sentralbyrå gi opplysninger som er nødvendige for utvikling, utarbeiding eller formidling av offisiell statistikk.»
- SSB innhenter utdanningsopplysninger fra flere kilder, og har mulighet til å koble disse mot blant annet sysselsettings- og befolkningsdata.



Takk!

