

Henning Holgersen

**Lønnsstatistikk for ansatte i kultur og
underholdning 2010**

Dokumentasjon

© Statistisk sentralbyrå januar 2012 Ved bruk av materiale fra denne publikasjonen skal Statistisk sentralbyrå oppgis som kilde.	Standardtegn i tabeller	Symbol
ISBN 978-82-537-8281-2 Trykt versjon	Tall kan ikke forekomme	.
ISBN 978-82-537-8282-9 Elektronisk versjon	Oppgave mangler	...
ISSN 1891-5906	Oppgave mangler foreløpig	...
Emne: 06.90	Tall kan ikke offentliggjøres	:
Trykk: Statistisk sentralbyrå	Null	-
	Mindre enn 0,5 av den brukte enheten	0
	Mindre enn 0,05 av den brukte enheten	0,0
	Foreløpig tall	*
	Brudd i den loddrette serien	—
	Brudd i den vannrette serien	
	Desimaltegn	,

Forord

Notatet dokumenterer og beskriver prosessen med å klargjøre, kvalitetssikre, analysere og utarbeide lønnsstatistikken for ansatte i kultur og underholdning for år 2010. Statistikken gir en oversikt over lønnsnivået og endringer i lønn per år for ansatte innen kultur og underholdning. Statistikken i sin nåværende form ble første gang produsert i 1997, og er sammenlignbar tilbake til dette året.

Sammendrag

I lønnsstatistikk for kultur og underholdning benyttes et sett kontroll- og revisjonsrutiner som i utgangspunktet er standardiserte for alle lønnsstatistikkene.

Effekten av å kontrollere og revidere rådatamaterialet er som regel betydelig, spesielt hvis delt inn i enkeltgrupper som utdanningsgruppe eller yrkesgruppe. En sammenligning mellom det ferdig reviderte datamaterialet og rådata fra oppgavegiverne viser klare forskjeller både med tanke på gjennomsnittsverdier og standardavvik i materialet.

Av de innkomne observasjonene i 2010 ble om lag 90 prosent godkjent til frigivning. Den tilsvarende prosentandelen i 2009 var om lag 88 prosent. Noe av årsaken til det relativt høye frafallet er at enkelte hele foretak blir holdt utenfor de publiserte tallene. Dette kan skyldes at de ikke leverte tall, eller ble holdt utenfor utvalget året før. Dermed må disse foretakene, for å få sammenlignbarhet i tallene mellom årgangene, også holdes ute i gjeldende årgang.

Datafangsten i 2010 var generelt god. Flere foretak leverte tidligere, noe som trolig er en konsekvens av at det ikke lenger sendes purringer til oppgavegivere som ikke svarer, men direkte varsel om tvangsmulkt. Dette later til å ha unngått at en del foretak spekulerer i innleveringsfrister som blir utsatt utover høsten.

Kultur og underholdning er en forholdsvis divers næring, noe som øker viktigheten av god datainngang og sammenlignbare utvalg fra år til år. Særlig innenfor små undernæringer kan det få stor betydning om et foretak ikke leverer lønnsdata.

Innhold

Forord	3
Sammendrag	4
Innhold	5
1. Innledning	6
2. Populasjon og utvalg	6
3. Datafangsten	6
3.1. Datainnfang	6
3.2. Brutto inkommet	7
4. Behandling av data fra Oracle-basen til endelig produksjonsfil	8
4.1. Tilrettelegging av endelig produksjonsfil	9
Program 1: Nedlesing fra databasen	9
Program 2: Arbeidstidsforhold	9
Program 3: Beregning av månedslønn	10
Program 4: Yrkeskoding	11
Program 5: Godkjenning og heltidsekvivalenter	12
Program 6: Vekting	13
Program 7: Vektjustering	14
4.2. Frafall	14
5. Effekten av kontroll- og revisjonsrutinene	15
5.1. Arbeidstid	15
5.2. Lønnsvariablene	15
Referanser	17
Vedlegg A	18
Figurregister	19
Tabellregister	19

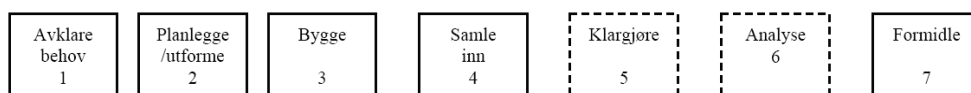
1. Innledning

Dette notatet dokumenterer arbeidsprosessene med å klargjøre, kvalitetssikre og analysere datagrunnlaget til lønnsstatistikken for kultur og underholdning i 2010.

Notatet er todelt, der første del tar for seg datafangsten til lønnsstatistikken mens del to tar for seg kontroll- og revisjonsopplegget og den effekten dette har. Årsaken til denne todelingen er at datafangst og kontroll- og revisjonsrutiner henger tett sammen, og at de til en viss grad er betinget av hverandre. Det vil si at kvaliteten på datafangsten virker inn på hvordan kontroll og revisjon skal legges opp, og motsatt.

I lønnsstatistikken oppfordres oppgavegiverne til å levere data elektronisk via Altinn eller CD. Alternativt kan det benyttes papirskjema. Bearbeidingen av data frem til ferdig reviderte tall kan deles inn i fire trinn: Mottak, registrering, innlesing til Oracle-tabell og etablering av produksjonsfil. Dataene blir kontrollert og revidert på hvert av disse trinnene, men dette notatet konsentrerer seg, i tillegg til beskrivelse av populasjonen, tempo og kvalitet i innkviktering og lagring av data, om den siste delen av kontroll- og revisjonsopplegget. Den skjer etter nedlesing av data fra Oracle-tabellen ved etablering av endelig produksjonsfil (se figur i vedlegg 1).

Statistisk sentralbyrås virksomhetsmodell for standardiserte arbeidsprosesser i statistikkproduksjonen er en modell med 7 statistikkprosesser som bygger på hverandre:



Dette notatet beskriver prosess 5 og 6, klargjøring og analysering av datagrunnlaget.

2. Populasjon og utvalg

Det trekkes et utvalg av foretak fra populasjonen i Bedrifts- og foretaksregisteret (Kristoffersen, Interne notater 2010/28). Foretakene blir stratifisert etter antall sysselsatte. Det er totaltelling blant foretak med sysselsetting over en viss grense, mens trekkandelene er mindre i strata med færre sysselsatte. Foretak med mindre enn fem sysselsatte trekkes ikke. Personutvalget består av alle sysselsatte i bedriftene i de uttrukne foretakene.

3. Datafangsten

Datainngangen per 1. oktober 2010 var relativt god, og endelig status før frigivning av statistikk viste at om lag 99 prosent av de trukne foretakene var kommet inn.

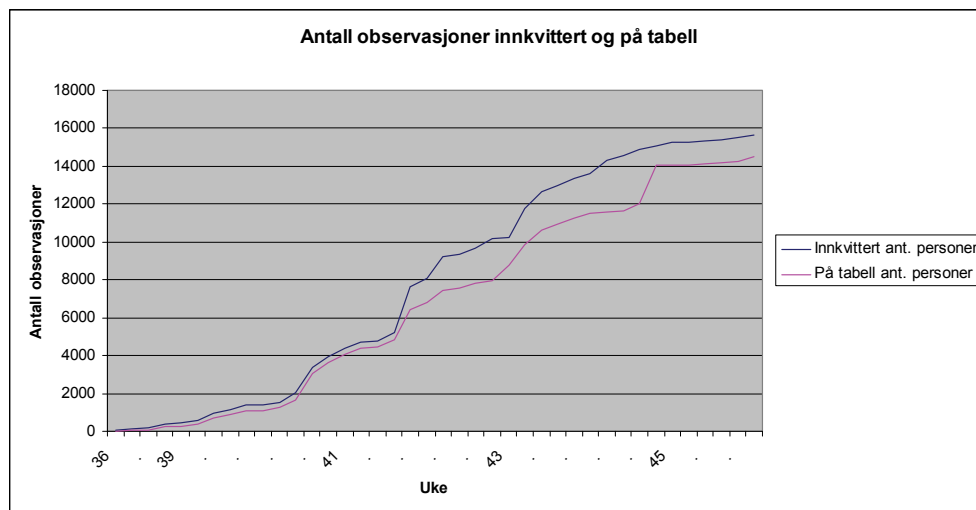
3.1. Datainngang

Datafangsten per 1. oktober 2010

Dato	Hendelse	Antall foretak
17.09.2010	Utsending	Antall foretak i utvalget: 426
12.10.2010	Svarfrist	På tabell pr dato: 4 461 observasjoner
19.10.2010	Varsel om tvangsmulkt	På tabell pr dato: 7 557 observasjoner
29.10.2010	Siste frist for å unngå tvangsmulkt	På tabell pr dato: 11 507 observasjoner
11.03.2010	Frigivning av statistikk	På tabell pr dato: 14 819 observasjoner

Innkvitteringen av foretak innen kultur og underholdning gikk i år greit, noe som trolig delvis skyldes det relativt sene publiseringstidspunktet, og delvis den strengere bruken av tvangsmulkt og varsel om tvangsmulkt i forhold til tidligere år. Det ble gitt unntak for 15 foretak.

Figur 1. Antall observasjoner innkvittert og lest inn i Oracle-tabellen, 2010



To dager etter svarfristens utløp i slutten av uke 41 var 6 642 observasjoner, eller om lag 43 prosent av utvalget lagt på tabell.

3.2. Brutto inkommet

I tabell 1 finner vi endelig antall foretak, bedrifter og sysselsatte som er innkvittert i kultur i 2009 og 2010. De absolutte tallene ligger forholdsvis stabilt, men med en liten oppgang i både antall foretak, bedrifter og sysselsatte. At over 100 prosent er på tabell i 2009 er et tegn på at Bedrifts- og foretaksregisteret mest sannsynlig ikke var helt oppdatert på trekketidspunktet. En annen medvirkende årsak kan være at foretak sender inn lønnsdata uten å være med i utvalget. De foretakene dette gjelder behandles og kontrolleres etter samme kriterier som resten av materialet, og tas med dersom hensiktsmessig. Dette kan være på grunn av endringer i populasjonen, eller for å kompensere for noe av frafallet ellers.

Tabell 1. Endelig antall foretak, bedrifter og observasjoner på tabell i kultur 2009 og 2010

Kultur	2010		2009	
	Antall på tabell	Prosent på tabell	Antall på tabell	Prosent på tabell
Antall foretak	394	96	368	103
Antall bedrifter ...	591	74	547	76
Antall obs.	15 026	97	14 735	104

Tabell 2 er en oversikt over andelen som benytter seg av de ulike innrapporteringsformatene i 2009 og 2010. Altinn er dominerende som leveringsmåte, noe som har vært et uttalt mål å få til.

Tabell 2. Leveringsmåter etter foretak, 2010 og 2009

Innkomet per foretak	2010		2009	
	Antall	Prosent	Antall	Prosent
Totalt	399	100,0	368	100,0
Altinn skjema	329	82,5	285	77,4
Batch	-	-	15	4,1
CD	3	0,8	9	2,4
Diskett	-	-	7	1,9
E-post	36	9,0	29	7,9
Papirskjema	26	6,5	23	6,3
Skjerm	5	1,3	-	-

Altinn er både den dominerende og foretrukne leveringsmåten for lønnsdata, og bruken av Altinn har økt i forhold til tidligere år. Tabell 2 ser på innkvitterte foretak, mens tabell 1 tar utgangspunkt i antall foretak på produksjonstabell. Denne forskjellen fører til at totalt antall foretak er forskjellig i de to tabellene. Fem foretak har blitt innkvittert, men ikke blitt tatt med i produksjonstabellen. Den vanligste årsaken til dette er store feil i tallgrunnlaget.

4. Behandling av data fra Oracle-basen til endelig produksjonsfil

Å innhente lønnsinformasjon på individnivå fra næringslivet har vist seg å medføre at alle opplysninger som kommer inn ikke nødvendigvis er korrekte. Det kan med andre ord være ulike typer feil i materialet, fra relativt ubetydelige unøyaktigheter til mer krevende tilfeller som medfører at observasjonen må forkastes. Det er derfor utarbeidet et system med kontroll- og feilrettingsrutiner for lønnsstatistikken, der hovedmålsettingen er å avdekke eventuell systematikk i typen feil og mangler i det som blir innrapportert. I tillegg er det en målsetting å minimere muligheten for at enkeltobservasjoner med åpenbare feil og mangler skal få uønsket innvirkning i statistikken. Systemet er bygget opp med enkle logiske mikrokontroller først (tabell 3 til tabell 9), før det kontrolleres mer på makronivå etterpå. Allerede ved innlesing av rådata, enten det måtte være fra Altinn-portalene fra filer basert på kravspesifikasjonen for elektronisk innrapportering, eller fra optisk leste skjema blir det kontrollert for logiske feil og mangler. Dette skjer litt forenklet sagt ved at observasjonen merkes i Oracle-basen, gradert etter betydningen av feilen/mangelen. Ved betydelige feil og mangler blir oppgavegiveren kontaktet med det samme, og oppfordret om å levere nye data.

Denne dokumentasjonen tar for seg kvalitetssikringen av datamaterialet etter at den innledende mikro-kvalitetssikringen er gjennomført i databasen. Det er med andre ord allerede ”ryddet opp” i en del av de mest åpenbare feil og mangler.

Tilrettelegging av data fra Oracle-basen frem til endelig produksjonsfil gjøres ved hjelp av fastlagte SAS-rutiner. Rutinene er lagt opp slik at de foregår i fem trinn eller operasjoner, som i stor grad er standardisert og tilnærmet like fra næringshovedområde til næringshovedområde. Opplegget består av:

- Rutiner for *merking* av observasjoner med mulige feil
 - Rutiner for *justering/korrigerings* av observasjoner med opplagte feil
 - Rutiner for å *holde utenfor* observasjoner med opplagte feil
- I korte trekk er rutinene bygget opp som følger:

Trinn 1: Nedlasting fra database og tilrettelegging av SAS-fil. Observasjoner uten verdier på noen av de relevante variable tas ut.

Trinn 2: Grunnleggende kontroll og justering av bakgrunnsvariablene arbeidstid og avlønningsperiode.

Trinn 3: Etablering av lønnsbegrepene (lønn per måned).

Trinn 4: Godkjenning av yrkeskoder.

Trinn 5: Endelig godkjenning av observasjoner. Tar ut observasjoner med ekstremverdier i en eller flere lønnsvariable, og beregner deltidsansatte som heltidsekvivalenter.

Deretter følger trinn 6, der materialet vektet og sjekkes opp mot populasjonen (Bedrifts- og foretaksregisteret).

Se vedlegg 1 for skjematisk oversikt over disse rutinene.

4.1. Tilrettelegging av endelig produksjonsfil

Program 1: Nedlesing fra databasen

Tabell 3 viser gyldige og ugyldige fødselsnummer i kultur- og underholdning i 2009 og 2010. D-nummer er fødselsnummer til utenlandske arbeidere, og av totalen utgjør utenlandske arbeidere 2 prosent i 2010. Alle observasjoner tas med videre i statistikkgrunnlaget, uavhengig om de har gyldig fødselsnummer eller ikke.

Tabell 3. Gyldige og ugyldige fødselsnummer, 2009 og 2010

Fødselsnummer og d-nummer	Antall	
	2010	2009
Totalt	14 457	14 344
Ugyldige fødselsnummer og ikke d-nummer	17	21
Gyldige fødselsnummer og ikke d-nummer	14 150	13 972
Mangelfullt eller ugyldig fødselsnummer og gyldig d-nummer	290	351

Program 2: Arbeidstidsforhold

Program 2 kontrollerer og justerer lønnsperioder og arbeidstid. Kontrollene viser at enkelte observasjoner plasseres i feil lønnsperiode og må omplasseres slik at lønnsopplysningene får riktig håndtering videre i kontrollene.

Alle ansatte har en avlønningsperiode:

- Lønnsperiode 1: Månedslønn 12 ganger i året.
- Lønnsperiode 2: 14-dagerslønn.
- Lønnsperiode 3: Lønn hver uke.
- Lønnsperiode 4: Månedslønn 12,5 ganger i året.
- Lønnsperiode 5: Månedslønn 13 ganger i året.
- Lønnsperiode 6: Timelønn.
- Lønnsperiode 7: Provisjons- eller prosentlønn.

Tabell 4. Status etter kontroller og endringer

Status observasjoner	Frekvens	Prosent
Totalt	14 457	100
Disse må sjekkes. Fastlønte med timelønn i tillegg	51	0,4
Disse må sjekkes. Høy timefortjeneste og lav arbtid	20	0,1
Disse må sjekkes. Timelønte med fastlønn i tillegg	2	0,0
Månedslønte som mangler arbtid	11	0,1
Månedslønte uten fastlønn som rettes til lønnp 6	56	0,4
Obs med lønnsperiode 6 som endres til 1	4	0,0
Observasjoner med annen lønnsperiode som gis lønnsperiode 6	4	0,0
Observasjoner som er uendret	13 896	96,1
Sjekkes. Heltid og 0<fastlønn<10000 (kommafeil?)	132	0,9
Sjekkes. Timelønte med arbtid>45	16	0,1
Timelønte uten arbtid som gis arbtid=timeant/4.33	265	1,8

Tabell 4 viser at 96 prosent av materialet slipper igjennom den første kontrollen som gjøres. Største feilpost er timelønte uten arbeidstid.

Tabell 5. Antall observasjoner fordelt på lønnsperiode før og etter korreksjon av arbeidstid

Lønnsperiode	Antall	
	Før	Etter
Totalt	14 457	14 457
1 - Månedslønn 12 ganger i året	8 178	8 127
2 - 14-dagerslønn	5	5
3 - Lønn hver uke	7	7
4 - Månedslønn 12,5 ganger i året	26	25
5 - Månedslønn 13 ganger i året	6	5
6 - Timelønn	6 220	6 276
7 - Provisjons- eller prosentlønn	15	12

Tabell 5 viser observasjoner fordelt etter lønnsperiode, før og etter korreksjon. Kun observasjoner hvor det er åpenbart hva lønnsperioden skulle ha vært blir korrigert.

Program 3: Beregning av månedslønn

Dette programmet foretar først og fremst beregning av månedsverdier for fastlønn, bonus og uregelmessige tillegg, på grunnlag av de innrapporterte størrelsene. Det foretas også korreksjon og justering av overtidstimer og overtidsgodtgjørelse.

Månedslønnsbegrepet består først og fremst av en summering av avtalt lønn, uregelmessige tillegg og bonus. For å beregne månedslønn må det derfor først gjøres en beregning av de øvrige lønnsselementer som månedslønnen bygger på:

Avtalt lønn

Avtalt lønn er beregnet lønn siste måned for alle, både fastlønte, timelønte og provisjonslønte. Dette for å få sammenlignbare størrelser uavhengig av lønnsperiode. Innrapportert lønn siste lønnsperiode knyttet til variabel for avlønning benyttes i utregningen av avtalt lønn:

Lønnsperiode 1	=	Oppgitt fastlønn siste periode (månedslønn)
Lønnsperiode 2	=	Oppgitt fastlønn per 14-dager multiplisert med 2,165 (antall 14-dagers perioder pr måned).
Lønnsperiode 3	=	Oppgitt fastlønn per uke multiplisert med 4,33 (antall uker i en måned).
Lønnsperiode 4	=	Oppgitt fastlønn per uke multiplisert med (12,5/12).
Lønnsperiode 5	=	Oppgitt fastlønn per uke multiplisert med (13/12).
Lønnsperiode 6	=	Hovedregel er timesats (enten beregnet eller direkte oppgitt) multiplisert med antall timer i en måned (avtalt arbeidstid per uke multiplisert med antall uker i en måned).
Lønnsperiode 7	=	Hovedregel: (bonus hittil i år + bonus 4. kvartal i fjor)/12 + eventuelt oppgitt fastlønn siste periode. Bonusfeltene settes til 0 etter beregning av avtalt lønn.

Uregelmessige tillegg

Innhentes som aggregert størrelse fra 1. januar til tellingstidspunktet. Gjennomsnittsberegning per måned for denne perioden.

Bonus

Innhentes både fra 4. kvartal i fjor, samt aggregert størrelse fra 1. januar til tellingstidspunktet. Gjennomsnittsberegning per måned for hele denne 12-månedperioden.

Månedslønn

Månedslønn fremkommer ved en summering av de innrapporterte verdiene for avtalt lønn, uregelmessige tillegg og bonus.

Overtidsgodtgjørelse

Inngår ikke i månedslønnsbegrepet. Innhentes som aggregert størrelse fra 1. januar til tellingstidspunktet. Gjennomsnittsberegning per måned for denne perioden.

For overtid gjøres en spesiell kontroll av forholdet mellom innrapporterte overtidstimer i forhold til overtidsgodtgjørelsen.

Tabell 6. Oversikt over overtidstimer og -betaling

	Antall	Prosent
Totalt	14 457	100,0
Observasjoner uten overtidstimer eller betaling	11 558	80,0
Observasjoner med flere overtidstimer enn betaling	13	0,1
Observasjoner med verdier på betaling	2 150	14,9
og av disse har mangler det timer på	736	5,1

På basis av denne kontrollen blir overtidsgodtgjørelsen justert på observasjoner som åpenbart bare har fått innrapportert *overtidstillegg*, og ikke hele godtgjørelsen, for utført overtidarbeid.

Dette gjøres ved at det beregnes en lønn per time innen normal arbeidstid for den aktuelle observasjonen, som så *legges til* overtidsgodtgjørelsen per time.

Program 4: Yrkeskoding

I program 4 gjennomføres en logisk kontroll av gyldighet av innrapporterte yrkeskoder. Oppgavegiverne skal rapportere yrkeskoder i henhold til Standard for yrkesklassifisering, C521 (STYRK). I kvalitetssikringsarbeidet før frigivning brukes relativt mye tid på å kvalitetssikre yrkeskodingen. Blant annet sjekkes forholdet mellom yrkeskode og yrkesbetegnelse, og ved vesentlige feil og mangler kontaktes oppgavegiverne. Første siffer i yrkeskoden betegnes som yrkesfelt:

1. Lederyrker
2. Akademiske yrker
3. Høyskoleyrker
4. Kontor- og kundeserviceyrker
5. Salgs- og serviceyrker
6. Yrker innen jordbruk, skogbruk og fiske (ikke i bruk)
7. Håndverkere og lignende
8. Operatører og sjåførere
9. Yrker uten krav til utdanning

Tabell 7. Fordeling av observasjoner på 1-siffer STYRK, 2009 og 2010

	2010		2009	
	Frekvens	Prosent	Frekvens	Prosent
Totalt	14 457	100,0	14 344	100,0
1 - Lederyrker	1 141	7,9	1 152	8,0
2 - Akademiske yrker	1 791	12,4	1 871	13,0
3 - Høyskoleyrker	5 979	41,4	5 764	40,2
4 - Kontor- og kundeserviceyrker	1 790	12,4	1 753	12,2
5 - Salgs- og serviceyrker	1 998	13,8	2 093	14,6
6 - Jordbruk, skogbruk og fiske (ikke i bruk)	91	0,6	82	0,6
7 - Håndverkere og lignende	438	3,0	564	3,9
8 - Operatører og sjåførere	111	0,8	164	1,1
9 - Yrker uten krav til utdanning	1 072	7,4	849	5,9
X - Ukjent	46	0,3	52	0,4

Verdien X er samle verdien for ugyldige innrapporterte yrkeskoder.

Det kobles på 7-sifret yrkeskode fra Arbeidstaker-/arbeidsgiverregisteret (AA-registeret) med oppdatering nærmest mulig tellingstidspunktet. Informasjonen fra AA-registeret brukes til kontroll av innrapportert yrkeskode til lønnsstatistikken. I noen tilfeller henter vi yrke fra AA-registeret ved mangelfull innrapportering til lønnsstatistikken. Det viser seg å være noe avvik mellom hva som ligger i AA-registeret og hva som rapporteres til lønnsstatistikken (se tabell 8). Noe av

forskjellen kan forklares i ulike tidspunkt for rapporteringen. Samtidig har lønnsstatistikken tilgang til viktig informasjon om lønn som hjelp i yrkeskodningen. Og foreløpig ser det ut til at det er en relativt større andel av uoppgitte i AA-registeret enn i lønnsstatistikken. Uoppgitte yrker i AA-registeret gis 1. siffer styrkkode lik 9, mens uoppgitte i lønnsstatistikken gis x som yrkeskode. Dette gjør samlet sett at vi i utgangspunktet stoler mest på yrkeskodene som innrapporteres til lønnsstatistikken.

Tabell 8. 1-siffer STYRK rapportert til hhv. SSB og NAV sitt AA-register

Innrapportert	AA-registeret										Totalt
	0	1	2	3	4	5	6	7	8	9	
1	4	869	56	101	27	16	:	:	:	61	1 141
2	:	43	1 421	142	24	12	:	:	:	137	1 791
3	152	164	325	3 558	218	419	:	49	25	1 061	5 979
4	11	25	50	158	1 096	130	:	:	:	303	1 790
5	16	26	18	107	71	1 120	:	17	:	610	1 998
6	:	:	:	:	:	:	54	:	:	25	91
7	:	:	:	11	:	16	:	345	:	48	438
8	:	:	:	14	:	:	:	:	69	15	111
9	:	:	:	25	17	69	:	20	:	913	1 072
X	:	:	:	:	:	:	:	:	:	43	46
Totalt	204	1 137	1 886	4 126	1 453	1 789	75	450	121	3 216	14 457

9 945 av totalt 14 457 observasjoner har samme 1-siffer yrkeskode i AA-registeret og lønnsstatistikken. Dersom oppgavegiver har rapportert stillingsbetegnelse kan det være mulig å gi stillingskode til observasjoner som mangler dette, forutsatt at stillingsbetegnelsen er beskrivende nok til å identifisere riktig stillingskode.

Program 5: Godkjenning og heltidsekvivalenter

Program 5a: Godkjenning av observasjoner

I program 5 lages en oversikt over godkjente observasjoner og observasjoner som forkastes i statistikken. Lønnsopplysninger har maksimums- og minimumsgrenser. Disse grensene er justert noe fra 2009 til 2010. Hovedtankegangen bak dette kontrollopplegget er som følger:

- Det skal settes romslige grenser for å sikre at observasjoner med antatt korrekte verdier beholdes.
- Grensene skal likevel være satt slik at observasjoner med åpenbart ekstreme verdier, som med stor grad av sikkerhet er feilaktig, holdes utenfor datagrunnlaget.
- Grenseverdiene settes på grunnlag av en rekke forhold, som blant annet erfaring fra tidligere år, kommunikasjon med forhandlingsverdenen (for eksempel minstelønn innen de ulike tariffområdene) og oppgavegivere, samt statistiske mål for spredning.
- Underveis i kvalitetssikringsarbeidet sjekkes observasjoner som faller ut med feilkoder i så stor grad som mulig.

Tabell 9. Oversikt over OK-koder

OK-kode	Antall	Prosent
Totalt	14 457	100,0
ok= 0 - Godkjente observasjoner	13 043	90,2
Ok= 1 - Lave verdier på lønnsvariablene	1	0,0
Ok= 3 - Timelønnte med utilstrekkelig info om timer	7	0,1
Ok= 4 - Alder=missing	1	0,0
Ok=10 - Heltidsansatte med ialt < 11 500 kroner	218	1,5
Ok=12 - Heltidsansatte, timelønnte med avtalt lønn over 200 000	16	0,1
Ok=14 - Heltidsansatte, st1=1 og ialt > 200 000 kroner	5	0,0
Ok=15 - Heltidsansatte, st1=1 og ialt < 21 000 kroner	5	0,0
Ok=16 - Heltidsansatte, st1 = 2,3 og ialt > 115 000 kr	55	0,4
Ok=17 - Heltidsansatte, st1 = 4,5 og ialt over 75 000 kroner	5	0,0
Ok=18 - Heltidsansatte, st1 > 5 og ialt over 65 000 kroner	5	0,0
Ok=19 - Heltidsansatte, Overtid (ot2) over 20 000 kroner	3	0,0
Ok=20 - Heltidsansatte, tillegg (tl) over 65 000 kroner	5	0,0
Ok=21 - Heltidsans, bonus (bon)> 100 000 kr	7	0,1
Ok=30 - Deltidsansatte med arbtid<=1	278	1,9
Ok=31 - Deltidsansatte, st1=1 og timefortjeneste>600	2	0,0
Ok=33 - Deltidsansatte, st1=1 og ialt/(arbtid/37.5) > 200 000	3	0,0
Ok=34 - Deltidsansatte, st1=1 og ialt/(arbtid/37.5) < 21 000	4	0,0
Ok=35 - Deltidsansatte, st1=2,3 og timefortjeneste > 500	37	0,3
Ok=36 - Deltidsansatte, st1=2,3 og timefortjeneste < 80	227	1,6
Ok=37 - Deltidsansatte, st1=4,5 og timefortjeneste > 250	142	1,0
Ok=38 - Deltidsansatte, st1=4,5 og timefortjeneste < 60	104	0,7
Ok=39 - Deltidsansatte, st1=7,8 og timefortjeneste > 250	6	0,0
Ok=40 - Deltidsansatte, st1=7,8 og timefortjeneste < 60	3	0,0
Ok=41 - Deltidsansatte, timefortjeneste > 250	32	0,2
Ok=42 - Deltidsansatte, timefortjeneste < 50	5	0,0
Ok=90 - Heltidsdubletter	95	0,7
Ok=91 - Deltidsdubletter	13	0,1
Ok=92 - Heltid-deltidsdubletter	4	0,0
Ok=100 - Manuelt holdt utenfor	126	0,9

Den vanligste "feiltypen" er deltidsansatte uten arbeidstid. Ofte gjelder dette tidligere ansatte, ansatte som er i permisjon eller som ikke vært satt opp til å jobbe den aktuelle perioden. Hvis vedkommende får etterbetalt lønn, bonus eller annen ytelse vil disse trolig komme med i datagrunnlaget. En annen årsak til denne typen feil er oppgavegivere som oppgir arbeidstid som stillingsbrøk, som da blir mindre enn 1. Hvis dette gjelder mange i samme foretak blir oppgavegiver kontaktet for å avklare.

Heltidsansatte med lav lønn er en annen forholdsvis stor gruppe. En av årsakene til dette er personer som slutter i jobben eller tar fri hele/deler av rapporteringsperioden, og som da blir rapportert med et timeantall som tilsvarer en heltidsstilling, men lavere lønn enn de ellers ville hatt.

Program 5b: Etablering av heltidsekvivalenter

Heltidsekvivalenter betegner lønnen for deltidsansatte justert for arbeidstid slik at den tilsvarer en fulltidsstilling. Dette gir et mål for deltidsansattes lønn som er sammenlignbart over ulike næringer og år.

Program 6: Vekting

Programmet stratifiserer populasjonen etter antall ansatte i foretakene og beregner vekter. Vektemetoden er basert på justert etterstratifisert invers trekkesannsynlighet, beskrevet i notat av Grini (2003/74).

Tabell 10. Oversikt over stratum og vekter

Næring	Stratum	Antall 2010		
		Foretak i utvalg	Foretak i pop	Antatte i pop
90	1	18	20	3 050
	2	10	24	627
	3	51	160	1 433
91	1	56	56	3 000
	2	10	14	161
	3	11	28	184
93	1	102	112	7 210
	2	70	134	2 244
	3	66	278	2 329

Program 7: Vektjustering

Prinsippet for justering av vektene er at sum oppblåst antall i lønnsstatistikken skal reflektere sum antall ansatte fra Bedrifts- og foretaksregisteret (BoF) for hver næring (Grini, notater 2003/74). Dette betyr at gjennomsnittlig lønn for et næringshovedområde skal reflektere et veid gjennomsnitt av hver av næringene som området består av. For å oppnå dette må hver av næringenes relative andel av den estimerte sysselsettingen i undersøkelsen gjenspeile populasjonsandelene fra Bedrifts- og foretaksregisteret. I tillegg må sammenhengen mellom de to siste årganger av lønnsstatistikken kontrolleres. Denne justeringen er spesielt viktig i områder med store lønnsforskjeller mellom næringer, der avvik vil få større betydning for det endelige lønnsnivået.

Tabell 11. Vektjustering 2010, heltids- og deltidsansatte

Næringsområde	Antall i BoF	Vektet antall	BoF Andel	Vektet andel	Samvekt
Total	23 250	17 884	1	1	1
Kunst og underholdning	4 296	2 905	0.184774	0.162484	1.137187
Biblioteker, museer etc.	3 442	2 544	0.148043	0.142298	1.040377
Lotteri- og totalisatorspill ...	1 885	1 500	0.081075	0.083927	0.966023
Sport og fritid etc.	13 627	10 932	0.586108	0.611292	0.958801

For å sikre et riktig størrelsesforhold mellom undernæringene gjøres det en samvekting, som tar utgangspunkt i Bedrifts- og Foretaksregisteret (BoF), og størrelsesforholdet mellom undernæringene og næringen totalt. Undernæringer som er underrepresentert i utvalget blir tillagt større vekt, mens undernæringer som er overrepresentert tillegges lavere vekt. Denne samvekten er undernæringens andel i BoF, delt på undernæringens vektete andel på tabell. Hvis disse andelene er like vil samvekten bli én, og ingen justering gjøres.

Tabell 12. Effekt av vekting 2010. Gjennomsnittlig månedslønn for heltidsansatte

	Antall	Månedslønn		
		Uvektet	Ujustert vekt	Justert vekt
Totalt	8 125	35 523	35 008	34 994
90.0 - Kunst og underholdning	1 747	34 745	34 763	34 763
91.0 - Biblioteker, museer etc.	1 623	34 891	34 521	34 521
92.0 - Lotteri og totalisatorspill	679	37 748	35 995	35 995
93.1 - Sportsaktiviteter	3 602	37 565	36 461	36 461
93.2 - Fritidsaktiviteter	474	21 854	25 286	25 286

4.2. Frafall

Vi har to hovedtyper frafall, enhetsfracfall og partielt frafall. Enhetsfracfall defineres her som frafall av foretak, mens partielt frafall er frafall av personer. Enhetsfracfall skyldes at foretak ikke har innrapportert til oss, eller at det er feil i datamaterialet vi har fått innrapportert. Under enhetsfracfall av foretak kommer også de foretakene vi har valgt å holde ute manuelt. I 2010 ble ett foretak holdt utenfor manuelt.

Partielt frafall kan blant annet skyldes mangelfull innrapportering, feil i data eller at personer ikke er blitt innrapportert. Dette frafallet blir beskrevet i tabell 9.

5. Effekten av kontroll- og revisjonsrutinene

Måling av effekten av kontrollopplegget gjøres her ved å sammenligne datagrunnlaget før og etter at alle programrutinene er kjørt. Dette kapitlet tar for seg hvilken virkning rutinene har på de mest sentrale variablene i lønnsstatistikken.

5.1. Arbeidstid

Arbeidstiden er grunnleggende bestemmende for to forhold: a) om en person er hel- eller deltidsansatt, og b) arbeidstiden benyttes for å beregne månedslønn for en representativ måned. Det legges derfor relativt mye ressurser inn i arbeidet med denne variabelen, og det er et absolutt krav i innrapporteringen at feltet må være utfyllt. Mye av kontrollene går av den grunn ut på å sjekke sammenheng mellom arbeidstid og lønn.

Tabell 13: Effekt av kontroll og revisjon av avtalt arbeidstid pr uke

	Urevidert		Revidert	
	Antall obs	Timeantall	Antall obs	Timeantall
Totalt	14 457	25,61	13 043	26,35
Heltid	8 125	37,33	7 578	37,31
Deltid	6 332	10,58	5 465	11,15

Som det går frem av tabellen er det flere observasjoner på den ureviderte filen (rådata) enn på den reviderte filen (publiseringsfilen) på grunn av at alle ikke-godkjente observasjoner her er holdt utenfor etter kriteriene beskrevet i kapittel 4.1. I tillegg til dette er også observasjoner helt uten innrapportert arbeidstid holdt utenfor den ferdig reviderte filen. Disse observasjonene registreres som deltidsansatte inntil de sjaltes ut, derav nedgangen i antall deltidsansatte fra urevidert til revidert fil. Nedgangen i antall observasjoner fra før til etter revisjon er i stor grad forklart av at en del foretak holdes manuelt utenfor det ferdige datasettet, som nevnt i punkt 4.1.

5.2. Lønnsvariablene

Lønnsvariablene som frigis i den offentlige statistikken er de vi ønsker å måle revisjonseffekten på. Dette gjelder først og fremst månedslønn, som består av avtalt lønn, uregelmessige tillegg og bonus. I tillegg kommer tall for overtidsgodtgjøring, men dette inngår ikke i månedslønnen. Revisjonseffekten på disse variablene defineres her som forskjellen mellom revidert og urevidert verdi, dvs. forskjellen mellom verdi på endelig publiseringsfil og verdi på rådatafil. Problemet i dette tilfellet er at responsvariablene ikke finnes direkte fra oppgavegiverne. De er avledede variable som beregnes i løpet av kontrollrutinene på grunnlag av innrapporterte variable.

For å forenkle vurderingen av revisjonsmålingen tar vi bare for oss heltidsansatte i det følgende. Dette gjøres i hovedsak for at fremstillingen skal bli mer oversiktelig, og subsidiært fordi resultatene blir noe vanskeligere å tolke når lønnen for de deltidsansatte skal inkluderes. Først ser vi næringen som helhet, og tar med alle lønnsvariable, med verdier for gjennomsnitt, median, nedre og øvre kvartil, samt minimum og maksimum.

Tabell 14. Effekt av kontroll og revisjon av lønnsvariable. Heltidsansatte, kroner.

	Urevidert	Revidert	Avvik
Antall (N)	8 125	7 578	-547
Gjennomsnitt (Kr)	Urevidert	Revidert	Avvik
Månedslønn i alt	35 523	35 055	-468
Fast månedslønn	34 389	34 250	-139
Bonus pr. mnd	804	545	-258
Tillegg pr. mnd	330	260	-70
Overtidsbetaling pr. mnd	448	462	14
Median (Kr)	Urevidert	Revidert	Avvik
Månedslønn i alt	32 550	33 000	450
Fast månedslønn	31 991	32 355	364
Bonus pr. mnd	0	0	0
Tillegg pr. mnd	0	0	0
Overtidsbetaling pr. mnd	0	0	0
1. Kvartil (Kr)	Urevidert	Revidert	Avvik
Månedslønn i alt	27 279	28 200	921
Fast månedslønn	26 981	27 800	819
Bonus pr. mnd	0	0	0
Tillegg pr. mnd	0	0	0
Overtidsbetaling pr. mnd	0	0	0
3. Kvartil (Kr)	Urevidert	Revidert	Avvik
Månedslønn i alt	39 158	39 242	84
Fast månedslønn	38 446	38 601	155
Bonus pr. mnd	0	0	0
Tillegg pr. mnd	0	0	0
Overtidsbetaling pr. mnd	0	0	0
Minimum (Kr)	Urevidert	Revidert	Avvik
Månedslønn i alt	0	11 575	11 575
Fast månedslønn	0	6 667	6 667
Bonus pr. mnd	0	0	0
Tillegg pr. mnd	0	0	0
Overtidsbetaling pr. mnd	0	0	0
Maximum (Kr)	Urevidert	Revidert	Avvik
Månedslønn i alt	781 250	174 167	-607 083
Fast månedslønn	781 250	174 167	-607 083
Bonus pr. mnd	192 560	57 917	-134 643
Tillegg pr. mnd	111 111	32 667	-78 444
Overtidsbetaling pr. mnd	23 610	19 802	-3 808

Den største kronemessige effekten av kontroll ligger i at en del observasjoner med opplagt for lav lønn blir fjernet. I seg selv påvirker dette både 1. kvartil, median og 3. kvartil. Også maksimum blir redusert kraftig av kontrollene.

Tabell 15. Effekt av revisjon på gjennomsnittlønn, etter yrkesgrupper

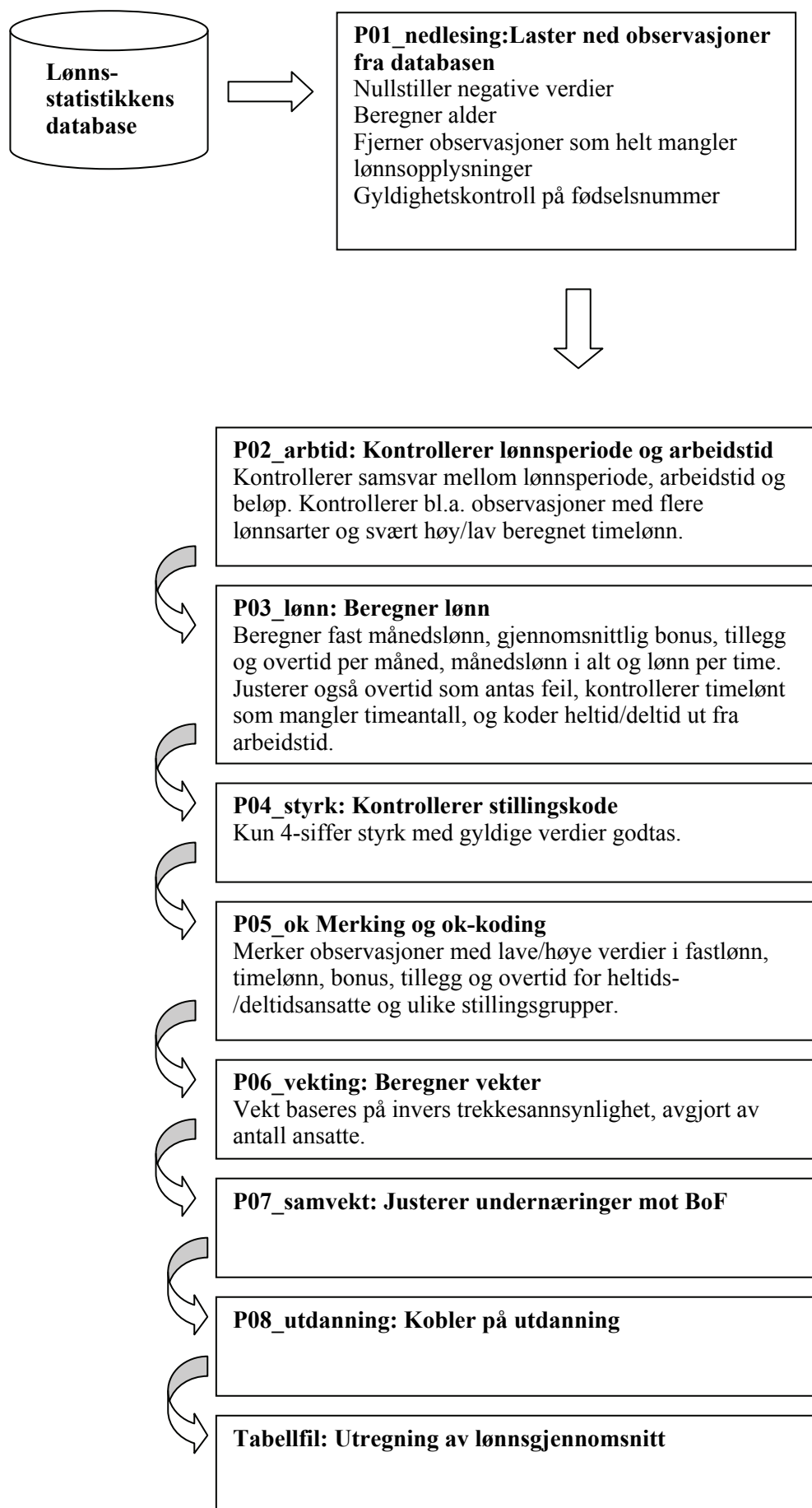
	Urevidert		Revidert	
	N	Ialt (Kr)	N	Ialt (Kr)
Totalt	8 125	35 523	7 578	35 055
1 Lederyrker	1 007	49 357	987	47 426
2 Akademiske yrker	1 569	37 653	1 549	37 260
3 Høgsk.yrker	2 929	38 118	2 722	35 705
4 Kontoryrker	839	28 062	806	28 288
5 Salg og serviceyrker	744	26 377	711	26 875
6 Andre yrker	69	26 349	68	26 030
7 Håndverkere	360	28 319	330	29 854
8 Operatører og sjåf.	90	30 002	87	29 862
9 Uten krav til utd.	486	19 561	296	24 311
Ukjent	32	25 286	22	27 708

Referanser

Kristoffersen (2010): Lønnsstatistikk 2010. Etablering av populasjon og utvalg. Statistisk sentralbyrå, interne notater 2010/28

Grini, Knut Håkon (2003): Lønnsstatistikk privat sektor 1997 – 2001. Dokumentasjon av utvalg og beregning av vekter. Statistisk sentralbyrå, notater 2003/74

Vedlegg A



Figurregister

Figur 1. Antall observasjoner innkvittert i delregisteret og lest inn i Oracle-tabellen, 2010	7
---	---

Tabellregister

Tabell 1. Endelig antall foretak, bedrifter og observasjoner på tabell i kultur 2009 og 2010	7
Tabell 2. Leveringsmåter etter foretak, 2010 og 2009	8
Tabell 3. Gyldige og ugyldige fødselsnummer, 2009 og 2010	9
Tabell 4. Status etter kontroller og endringer	9
Tabell 5. Antall observasjoner fordelt på lønnsperiode før og etter korreksjon av arbeidstid	10
Tabell 6. Oversikt over overtidstimer og -betaling	11
Tabell 7. Fordeling av observasjoner på 1-siffer STYRK, 2009 og 2010	11
Tabell 8. 1-siffer STYRK rapportert til hhv. SSB og NAV sitt AA-register	12
Tabell 9. Oversikt over OK-koder	13
Tabell 10. Oversikt over stratum og vekter	14
Tabell 11. Vektjustering 2010, heltids- og deltidsansatte	14
Tabell 12. Effekt av vekting 2010. Gjennomsnittlig månedslønn for heltidsansatte	14
Tabell 13. Effekt av kontroll og revisjon av avtalt arbeidstid pr uke	15
Tabell 14. Effekt av kontroll og revisjon av lønnsvariable. Heltidsansatte, kroner.	16
Tabell 15. Effekt av revisjon på gjennomsnittlønn, etter yrkesgrupper	16