

Liv Tove Hytjan

**Lønnsstatistikk for ansatte i bygge- og
anleggsvirksomhet 2010**

Dokumentasjon

© Statistisk sentralbyrå januar 2012 Ved bruk av materiale fra denne publikasjonen skal Statistisk sentralbyrå oppgis som kilde.	Standardtegn i tabeller	Symbol
ISBN 978-82-537-8277-5 Trykt versjon	Tall kan ikke forekomme	.
ISBN 978-82-537-8278-2 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 bygge- og anleggsvirksomhet for 2010. Statistikken gir en oversikt over lønnsnivået og endringer i lønn per år for alle ansatte i næringen. Statistikken, i sin nåværende form ble produsert første gang i 1997, og er sammenlignbar tilbake til dette året.

Sammendrag

I lønnsstatistikk for bygge- og anleggsvirksomhet benyttes et opplegg kontroll- og revisjonsrutiner som i utgangspunktet er standardiserte for alle lønnsstatistikkene. Effekten av å kontrollere/revidere rådatamaterialet er i de fleste tilfeller betydelig, spesielt innen enkelte yrkesgrupper. En sammenligning mellom det ferdig reviderte datamaterialet og rådata fra oppgavegiverne viste klare forskjeller både med tanke på gjennomsnittsverdier og standardavvik i materialet. Maskinelt gjennomgås både mikro- og makrobaserte kontroller, og det er makrorutinene som det brukes mest tid og ressurser på i kvalitetssikringen.

Av de innkomne observasjonene i 2010 var det om lag 97,7 prosent som ble godkjent til bruk i frigivning.

Populasjonen utgjør 5 830 foretak med 130 114 sysselsatte per 3. kvartal 2010. Blant foretakene med 50 eller flere ansatte trekkes alle foretakene (fulltelling), blant foretakene med mellom 12 og 49 ansatte trekkes halvparten og for de foretakene med mellom 6 og 11 ansatte trekkes 20 prosent som kommer med i utvalget.

Innhold

Forord	3
Sammendrag	4
Innhold	5
1. Innledning	6
2. Populasjon og utvalg	6
3. Datafangsten	6
3.1. Datainnngang	6
3.2. Brutto inkommet	7
4. Behandling av data fra Oracle-basen til endelig produksjonsfil	8
4.1. Tilrettelegging av endelig produksjonsfil	9
4.1.1 Program 1 nedlesing fra Oracle-basen	9
4.1.2 Program 2 arbeidstidsforhold	9
4.1.3 Program 3 beregning av månedslønn	9
4.1.4 Program 4 yrkeskoding	10
4.1.5 Program 5 godkjenning av observasjoner	11
4.1.5 Program 5b, etablering av heltidsekvivalenter	12
4.1.6 Program 6 vekting	13
4.1.7 Program 7 justering av vektorer	13
4.1.8 Frafall	14
5. Effekten av kontroll- og revisjonsrutinene	14
5.1. Arbeidstid	14
5.2. Lønnsvariablene	14
Referanser	18
Vedlegg A: Prosesdiagram lønnsstatistikk	19
Figurregister	20
Tabellregister	20

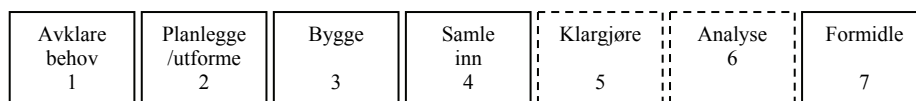
1. Innledning

Dette notatet dokumenterer arbeidsprosessene med å klargjøre, kvalitetssikre og analysere datagrunnlaget til lønnsstatistikken for bygge- og anleggsvirksomhet i 2010.

Notatet er i utgangspunktet todelt, der første del tar for seg datafangsten til lønnsstatistikken, mens del 2 tar for seg kontroll- og revisjonsopplegget og den effekt 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. Dvs. kvaliteten på datafangsten virker inn på hvordan kontroll- og revisjonen skal legges opp, og motsatt.

I lønnsstatistikken er det anledning til å svare enten elektronisk (Altinn eller på diskett/CD) eller på papirskjema. Bearbeidingen av data frem til ferdig reviderte tall kan deles inn i fire trinn: Mottak, registrering, innlesning 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 datafangsten, om den siste delen kontroll- og revisjonsopplegget. Den skjer etter nedlesing av data fra Oracle-tabellen ved etablering av endelig produksjonsfil (se figur 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 (BoF), (Sigrun Kristoffersen, Interne notater 2010/28). Foretakene blir stratifisert etter antall sysselsatte. Det er totaltelling blant foretak med sysselsetting over en viss grense, mens trekkeandelene er mindre i strata med færre sysselsatte. Foretak med mindre enn seks sysselsatte trekkes ikke. Personutvalget består av alle sysselsatte i bedriftene i de uttrukne foretakene.

3. Datafangsten

Datainngangen per 1. oktober 2010 var god, og endelig status før frigivning av statistikk viste at om lag 93,6 prosent av de trukne foretakene var kommet inn.

3.1. Datainngang

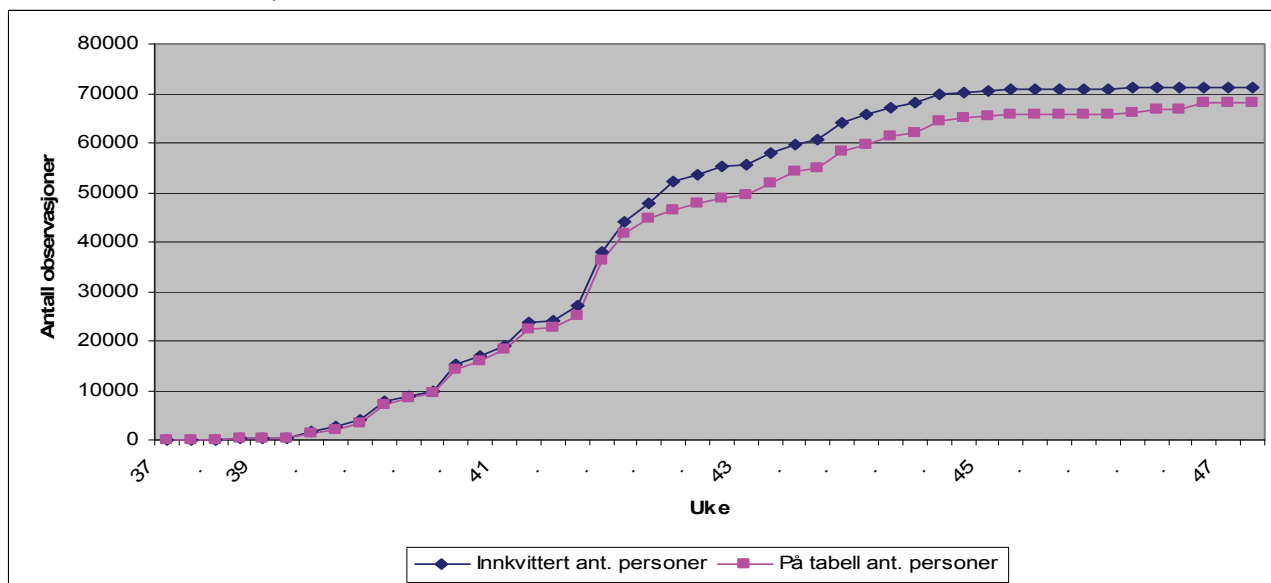
Datafangsten per 1. oktober 2010

Dato	Hendelse	Antall foretak (trekkeenheter)
17.09.10 (uke 37)	Utsending	1 696
08.10.10 (uke 40)	Svarfrist	327 (19,3 prosent)
21.10.10 (uke 42)	Vedtak om tvangsmulkt	719
29.10.10 (uke 48)	Siste frist for å unngå tvangsmulkt	90
13.01.11 (uke 2)	Frigivning av statistikk	1 588 (93.6 prosent)
Antall fritak gitt	Ingen	

Ved svarfrist i uke 40 var om lag 19,3 prosent av foretakene, 19,8 prosent av bedriftene og 23,2 prosent av de ansatte innkvitterte.

Figur 1 viser datainngangen fra uke 37 til uke 47.

Figur 1. Datainngang i bygge- og anleggsvirksomhet. Antall innkvitterte og lest inn i Oracle-tabellen, 2010.



Som vist på figuren er det for 2010 et avvik mellom innkvitterte og innleste observasjoner fra uke 37 til 41, mens det fra uke 41 er avviket minimalt.

3.2. Brutto inkommet

I tabell 1 finner vi endelig antall foretak, bedrifter og sysselsatte som er innkvittert i bygge- og anleggsvirksomhet i 2009 og 2010. Andelen innkvitterte foretak, bedrifter og sysselsatte i 2010 var om lag på samme nivå som i 2009.

Tabell 1. Endelig antall foretak, bedrifter og sysselsatte innkvittert i bygge- og anleggsvirksomhet i 2010 og prosentvis andel innkvittert i 2009 og 2010

Enhet	Datainngang		
	2010		2009
	Antall Innkvittert	Prosent Innkvittert	Prosent Innkvittert
Foretak	1 588	93,6	92,5
Bedrifter	2 190	95,2	94,5
Sysselsatte	71 128	97,7	97,6

Tabell 2 er en oversikt over andelen som benytter seg av de ulike innrapporteringsformatene i perioden 2009 til 2010. Innrapportering via Altinn er stadig økende, mens de andre innrapporteringsformene avtar. Det er fremdeles en stor andel av de som innrapporterer via kravspesifikasjon.

Tabell 2. Andel observasjoner etter innrapporteringsformat i perioden 2009 til 2010

KODE	Prosent	
	2 010	2 009
I alt	100,0%	100,0%
Excelfiler	3,4%	4,1%
Altinn	52,6%	45,5%
Kravspesifikasjon	43,8%	49,5%
Manuelt registrert	0,3%	0,0%
Optisk lest	1,8%	1,9%

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 være ulike typer feil i materialet, fra relativt ubetydelige unøyaktigheter, til mer graverende 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 3-7), før det kontrolleres mer på makronivå etterpå (bl.a. tabulering). Det er arbeidet på makronivå som det klart brukes mest tid på, og det er der vi oppdager systematiske feil/skjevheter i datamaterialet. 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 observasjoner merkes i Oracle-basen, gradert etter betydningen av feilen/mangelen. Ved betydelige feil og mangler blir oppgavegiveren kontaktet med det samme, og oppfordret til 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 fram til endelig produksjonsfil gjøres ved hjelp av fastlagte SAS-rutiner. Rutinene er lagt opp slik at de foregår i 5 trinn eller operasjoner, som i stor grad er standardisert og tilnærmet like fra næringsområde til næringsområde. 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: Nedlesing fra Oracle og tilrettelegging av sasfil. Observasjoner uten verdier på 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

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

Se vedlegg 1 for skjematisk oversikt over disse rutinene.

4.1. Tilrettelegging av endelig produksjonsfil

4.1.1 Program 1 nedlesing fra Oracle-basen

Tabell 3 viser gyldige og ugyldige fødselsnummer for ansatte i bygge- og anleggsvirksomhet i 2009 og 2010. D-nummer er fødselsnummer til utenlandske arbeidere. Alle observasjoner tas med videre i statistikkgrunnlaget, uavhengig av om de har gyldig fødselsnummer eller ikke.

Tabell 3. Gyldige og ugyldige fødselsnummer. 2009 og 2010

Fødselsnummer	D- nummer	Antall	
		2010	2009
I alt		68 141	68 194
Ugyldige fødselsnummer og	ikke d-nummer	27	36
Gyldige fødselsnummer og	ikke d-nummer	65 824	65 193
Mangelfullt eller ugyldig fødselsnummer og	gyldig d-nummer	2 674	2 948
Mangelfullt eller ugyldig fødselsnummer og	ikke d-nummer	12	17

4.1.2 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.

4.1.3 Program 3 beregning av månedslønn

Dette programmet foretar først og fremst en beregning månedslønn på grunnlag av innrapporterte verdier av fastlønn, bonus og uregelmessige tillegg. Det foretas også korreksjon og justering av overtidstimer og overtidsgodtgjørelse.

Månedslønnsbegrepet består av en summering av avtalt lønn, uregelmessige tillegg og bonus. For å beregne månedslønn må det derfor først gjøres en beregning av de øvrige lønselementer som månedslønna 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 avlønningsperiode knyttet til variabel for månedslønn benyttes i utregningen av avtalt lønn:

Avlønningsperiode 1	= Oppgitt månedslønn siste periode
Avlønningsperiode 2	= Oppgitt månedslønn per 14-dager multiplisert med 2.165 (antall 14-dagers-perioder i en måned).
Avlønningsperiode 3	= Oppgitt månedslønn per uke multiplisert med 4.33 (antall uker en måned).
Avlønningsperiode 4	= Oppgitt månedslønn per uke multiplisert med (12,5/12) .
Avlønningsperiode 5	= Oppgitt månedslønn per uke multiplisert med (13/12) .
Avlønningsperiode 6	= Hovedregel er timesats (enten beregnet eller direkte oppgitt) multipliser med antall timer i en måned (avtalt arbeidstid per uke multiplisert med antall uker i en måned).
Avlønningsperiode 7	= Hovedregel: (bonus hittil i år + bonus 4. kvartal i fjor)/12 + evt. oppgitt fastlønn siste periode. Bonusfeltene tømmes 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 for 4. kvartal i fjor, samt som aggregert størrelse fra 1. januar til tellingstidspunktet. Gjennomsnittsberegning per måned for hele denne 12-månedperioden.

Månedslønn

Månedslønna framkommer 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 4. Oversikt over overtidstimer og overtidbetaling

Status observasjoner	Frekvens	Prosent
I alt	68 141	100,0
Observasjoner uten overtidstimer eller betaling	31 539	46,3
Observasjoner med flere overtidstimer enn betaling	216	0,3
Observasjoner med verdier på betaling	35 206	51,7
og av disse mangler det timer på	1 180	1,7

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. Alt dette gjøres maskinelt uten tidsbruk for saksbehandleren.

4.1.4 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 (styrk). I kvalitetssikringsarbeidet før frigivning brukes relativt mye tid på å kvalitetssikre yrkeskodingen. Blant annet sjekkes forholdet mellom yrkeskode og yrkesbetegnelse, og ved feil og mangler kontaktes oppgavegiverne i så stor grad som mulig.

- 1. 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 o.l.
 8. Operatører og sjåfører
 9. Yrker uten krav til utdanning

Innrapporteringa av yrkeskoder så slik ut i 2009 og 2010:

Tabell 5. Fordeling av observasjoner på 1-siffer styrk, 2009 og 2010

St1	2009		2010	
	Frekvens	Prosent	Frekvens	Prosent
I alt	68 194	100,0	68 141	100,0
1	4 484	6,6	4 592	6,7
2	3 346	4,9	3 425	5,0
3	5 181	9,1	5 117	9,2
4	2 834	4,2	2 822	4,1
5	458	0,7	429	0,6
6	5	0,0	14	0,0
7	41 546	60,9	41 031	60,2
8	5 776	8,4	6 115	9,0
9	3 477	5,1	3 451	5,1
X	87	0,1	34	0,1

Det kobles på 7-sifret yrkeskode fra AA-registeret (Arbeidstaker-/arbeidsgiverregisteret) 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 6). Noe av forskjellen kan forklares i ulike tidspunkt for rapporteringen. Samtidig har lønnsstatistikken tilgang til viktig informasjon om lønn som hjelp i yrkeskodingen. Og foreløpig ser det ut til at det er en relativt større andel av uoppgitte yrker i AA-registeret enn i lønnsstatistikken. Dette gjør samlet sett at vi i utgangspunktet stoler mest på yrkeskodene som innrapporteres til lønnsstatistikken.

Tabell 6. Fordeling av antall med 1-siffer styrk i lønnsstatistikken sammenlignet med AA-registeret

Lønns- statistikken	I alt	AA- registeret									
		1	2	3	4	5	6	7	8	9	X
I alt	68 141	3 891	3 772	5 312	2 579	740	77	39 631	5 881	8 832	126
Av dette											
1	4 592	3 244	393	208	65	18	:	502	50	107	5
2	3 425	120	2 601	295	41	:	:	215	5	136	3
3	6 228	229	506	4 149	105	51	:	881	35	260	12
4	2 822	106	98	158	2 214	37	:	57	24	122	5
5	429	25	:	26	19	311	:	12	:	33	:
6	14	:	:	:	:	:	10	:	:	:	:
7	41 031	145	160	409	93	235	48	34 171	700	4 987	83
8	6 115	16	3	12	17	21	13	475	4 976	575	7
9	3 451	6	7	53	25	59	3	614	91	2 585	8
X	34	:	:	2	:	:	:	:	:	26	3

Det er 126 ansatte som ligger uten yrkeskode i AA-registeret, mens det er 34 ansatte som ligger uten kode i lønnsstatistikken.

En direkte sammenligning av den 4-sifrede yrkeskoden fra den ferdige reviderte fila fra lønnsstatistikken og fra AA-registeret, viser at 54 261 observasjoner hadde identisk yrkeskode i disse to kildene.

4.1.5 Program 5 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 ikke blitt endret i 2010 i forhold til 2009. 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 forhandlingsverden (for eksempel minstelønn innen de ulike tariffområdene) og oppgavegivere, samt statistiske mål på spredning. Grenseverdiene gjennomgås og justeres år for år.

- Underveis i kvalitetssikringsarbeidet sjekkes observasjoner som faller ut med feilkoder, i så stor grad som mulig.

Tabell 7. Antall observasjoner fordelt etter koder for grenseverdier (ok-koder) i 2010

Ok-kode	Frekvens	Prosent
I alt	68 141	100
Ok=0 - Godkjente observasjoner	64 852	95,2
Ok=1 - Lave verdier på lønnsvariablene	:	:
Ok=2 - Avtalt lønn 0 kroner	24	0,0
Ok=3 - Timelønte med utilstrekkelig info om timer	34	0,1
Ok=4 - Alder missing	:	:
Ok=10 - Heltidsansatte med i alt < 10 500	2 124	3,1
Ok=11 - Heltidsansatte, fmlonn over 200 000 kroner	:	:
Ok=12 - Heltidsansatte, timelønte med fmlonn over 70 000 kroner ..	61	0,1
Ok=14 - Heltidsansatte, st1=1 and i alt for ledere > 250 000 kroner ..	23	0,0
Ok=15 - Heltidsansatte, st1=1 og i alt for ledere < 20 000	43	0,1
Ok=16 - Heltidsansatte, st1 = 2,3 og i alt over 135 000	12	0,0
Ok=17 - Heltidsansatte, st1 = 4,5 og i alt over 75 000	14	0,0
Ok=18 - Heltidsansatte, st1 > 5 og i alt over 80 000	112	0,2
Ok=19 - Heltidsansatte, overtid (ot2) over 40 000	34	0,1
Ok=20 - Heltidsansatte, tillegg (tl) over 65 000	:	:
Ok=30 - Deltidsansatte med arbtid <= 1	11	0,0
Ok=31 - Deltidsansatte, st1=1 og timefortjeneste > 600	:	:
Ok=32 - Deltidsansatte, st1=1 og timefortjeneste < 120	:	:
Ok=33 - Deltidsansatte, st1=1 i alt / (arbtid/37,5) > 250 000	:	:
Ok=34 - Deltidsansatte, st1=1 i alt / (arbtid/37,5) < 20 000	10	0,0
Ok=35 - Deltidsansatte, st1=2,3 og timefortjeneste > 500	:	:
Ok=36 - Deltidsansatte, st1=2,3 og timefortjeneste < 80	19	0,0
Ok=37 - Deltidsansatte, st1=4,5 og timefortjeneste > 500	102	0,2
Ok=38 - Deltidsansatte, st1=4,5 og timefortjeneste < 60	19	0,0
Ok=39 - Deltidsansatte, st1=7,8 og timefortjeneste > 250	88	0,1
Ok=40 - Deltidsansatte, st1=7,8 og timefortjeneste < 60	115	0,2
Ok=41 - Deltidsansatte og timefortjeneste > 250	18	0,0
Ok=42 - Deltidsansatte og timefortjeneste < 50	35	0,1
Ok=43 - Deltidsansatte og ot2 > 40 000	:	:
Ok=90 - Heltidsdubletter	344	0,5
Ok=91 - Deltidsdubletter	:	:
Ok=42 - Heltids - deltidsdubletter	:	:

4.1.5 Program 5b, etablering av heltidsekvivalenter

Vi inkluderer heltidsekvivalenter i de næringsspesifikke frigivingene. Program 5b regner ut ekvivalentvekter og ekvivalentlønsarter for disse.

Opgavegivere som leverer data elektronisk ved å sende inn fil ifølge kravspesifikasjonen for innrapportering av lønnsstatistikk eller via Altinn rapporterer en kode for heltid eller deltid. Vi ser denne koden i sammenheng med rapportert avtalt arbeidstid og lønnsperiode, se pkt. 4.1.2. Vi kontrollerer også alle lønsarter for både heltids- og deltidsansatte i forhold til maksimums- og minimumsgrenser, se pkt. 4.1.5.

Grenseverdiene sees også i sammenheng med hvilket yrke den ansatte har. Observasjoner som faller ut blir kontrollert for om mulig å kunne rettes opp eller bli bekreftet riktig slik at de kan godkjennes og benyttes i statistikken.

Følgende oversikt viser resultatet etter kontrollene vi gjør, altså hvor mange heltids- og deltidsansatte som ikke blir godkjent.

	Heltidsansatte	Deltidsansatte
Totale obs i næringen	64 424	3 717
Godkjente obs	61 374	3 244
Ikke godkjente obs	3 050	473
% ikke godkjente obs	4,7	12,7

Vi får en unik forståelse av heltidsekvivalenten for næringen i og med at vi beregner gjennomsnittlig antall avtalte timer per uke for heltidsansatte. Dette gjennomsnittet blir utgangspunktet for videre behandling og omregning av deltidsansattes lønn.

Det er dermed kun de deltidsansattes månedslønn vi regner om. Utgangspunktet for omregning er å multiplisere lønna til den enkelte deltidsansatte til hva den ville vært dersom han eller hun jobbet heltid.

4.1.6 Program 6 vekting

Programmet stratifiserer populasjonen etter antall ansatte i foretakene og beregner vekter. Vektemetoden er basert på justert etterstratifisert invers trekkesannsynlighet (Grini, notater 2003/74). Vi mottar også data fra bedrifter/foretak som ikke er med i utvalget. Disse behandles etter samme kriterier og kontroll som resten av materialet. Og vurderes om de skal tas med i datagrunnlaget. Endringer i populasjonen kan være en grunn til at slike foretak tas med.

Tabell 8. Antall foretak i populasjon og utvalg og invers trekkesannsynlighet (Vekt)

Stratum	Antall			Vekt 2010
	Foretak i utvalget	Foretak i pop.	Ansatte i pop.	
4101	26	26	9 787	1,0000
4102	105	145	9 852	1,3810
4103	370	2 077	27 920	5,6135
4201	19	14	6 665	0,7368
4202	34	51	2 876	1,5000
4203	30	165	2 217	5,5000
4311	148	174	8 332	1,1757
4312	70	181	2 730	2,5857
4313	51	285	2 466	5,5882
4321	27	25	14 437	0,9259
4322	136	188	11 829	1,3824
4323	516	2 650	36 344	5,1357

4.1.7 Program 7 justering av vekter

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 sysselsettinga i undersøkelsen, gjenspeile populasjonsandelene fra BoF. I tillegg må sammenhengen mellom de to siste årganger av lønnsstatistikken kontrolleres. Denne justeringen er spesielt viktig i områder med store lønnsforskjeller mellom undernæringer, der avvik vil få større betydning for det endelige lønnsnivået.

Tabell 9. Vektjustering 2010.

Næring	Vektet antall		Vektet andel		BoF andel	Justeringsfakt or mot BoF
	2009	2010	2009	2010		
Totalt	123 276	123 710	1	1	1	
Oppføring av bygninger	43 406	43 410	0,3442	0,3509	0,3477	0,9908
Anleggsvirksomhet	10 163	10 215	0,0731	0,0826	0,0748	0,9056
Riving og grunnarbeid	11 273	12 258	0,0998	0,0991	0,1049	1,0591
Spesialisert bygge- og anleggsvirksomhet	58 434	57 827	0,4829	0,4674	0,4726	1,0111

Vektjusteringen er blitt gjort direkte mot populasjonen i BoF.

Effekten av vekteprosedyrene på det beregnede lønnsnivået vises i tabell 10. Vi ser at vektjusteringen som gjøres, medfører at det samlede lønnsnivået justeres ned med 1 907 kroner.

Tabell 10. Effekten av vekting 2010. Gjennomsnittlig månedslønn for heltidsansatte. Kroner

	Månedslønn 2010. Heltidsansatte. Kroner		
	Uvektet	Ujustert vekt	Justert vekt
Totalt	36 879	35 023	34 972
Oppføring av bygninger	38 564	35 467	35 467
Anleggsvirksomhet	40 889	39 179	39 179
Riving og grunnarbeid	33 736	33 140	33 140
Spesialisert bygge- og anleggsvirksomhet	35 182	34 342	34 342

4.1.8 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 foretak vi har valgt å holde ute manuelt. Totalt sett var enhetsfracfallet på 7 foretak i 2010.

Partielt frafall kan blant annet skyldes mangelfull innrapportering, feil i data eller at personer ikke er blitt innrapportert. Denne type frafall er omtalt i kapittel 4.1.8 hvor det er blitt satt minimums- og maksimumsgrenser for de ulike lønnsvariablene. I bygge- og anleggsvirksomhet er de fleste observasjonene som ikke ble godkjent, deltidsansatte med arbeidstid mindre eller lik 1. Det partielle frafallet var totalt sett på mindre enn 3 prosent av observasjonene i 2010.

5. Effekten av kontroll- og revisjonsrutinene

Måling av effekten av kontrollopplegget gjøres her ved å sammenligne datagrunnlaget før og etter at alle sas-rutinene 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 utfylt. Mye av kontrollene går av den grunn ut på sjekke sammenheng mellom arbeidstid og lønn.

Tabell 11. Effekt av kontroll og revisjon av avtalt arbeidstid per uke

	Antall observasjoner		Gjennomsnittlig avtalt arbeidstid	
	Før revisjon	Etter revisjon	Før revisjon	Etter revisjon
I alt	68 141	64 618	36,5	36,5
Heltidsansatte	64 424	61 374	37,6	37,6
Deltidsansatte	3 717	3 244	18,3	18,5

Som det går fram av tabellen er det flere observasjoner på den ureviderte fila (rådata) enn på den reviderte fila (publiseringsfila). Grunnen er at alle ikke-godkjente observasjoner her er holdt utenfor etter kriteriene beskrevet i kapittel 4.1.5. I tillegg til dette er også observasjoner helt uten innrapportert arbeidstid holdt utenfor den ferdig reviderte fila. Disse observasjonene registreres som deltidsansatte inntil de sjaltes ut, derav nedgangen i antall deltidsansatte fra urevidert til revidert fil.

5.2. Lønnsvariablene

Lønnsvariablene som frigis i den offisielle statistikken er de vi ønsker å måle revisjonseffekten av. 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 månedslønna.

Revisjonseffekten på disse variablene defineres her som forskjellen mellom revidert og urevidert verdi, dvs. forskjellen på verdi på endelig publiseringsfil og verdi på rådatafil. Problemet i dette tilfellet er at publiseringsvariablene ikke finnes direkte på rådatafila fordi de ikke innrapporteres 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. Det er relativt sett få deltidsansatte i denne næringa, og i tillegg blir resultatene noe vanskeligere å tolke når lønna for de deltidsansatte skal inkluderes. Først ser vi næringa som helhet, og tar med alle lønnsvariable, med verdier for gjennomsnitt, median, nedre og øvre kvartil, samt minimum og maksimum.

Tabell 12. Effekt av kontroll og revisjon av lønnsvariable. Heltidsansatte. Kroner

	Revidert	Urevidert	Avvik			
Antall	61 374	66 175	-4 801			
	Gjennomsnitt			Median		
	Revidert	Urevidert	Avvik	Revidert	Urevidert	Avvik
Månedslønn	34 972	32 608	2 364	32 800	32 042	758
Avtalt lønn	33 437	31 343	2 094	31 582	30 851	731
Uregelmessige tillegg	894	933	-39	0	0	0
Bonus	641	644	-3	0	0	0
Overtid	113	937	-824	0	0	0
	1. kvartil			3. kvartil		
	Revidert	Urevidert	Avvik	Revidert	Urevidert	Avvik
Månedslønn	28 959	27 170	1 789	38 000	37 413	587
Avtalt lønn	28 002	26 630	1 372	36 210	35 723	487
Uregelmessige tillegg	0	0	0	1 056	1 060	-4
Bonus	0	0	0	0	0	0
Overtid	0	0	0	952	681	271
	Min			Max		
	Revidert	Urevidert	Avvik	Revidert	Urevidert	Avvik
Månedslønn	10 519	0	10 519	247 664	670 000	-422 336
Avtalt lønn	1 174	0	1 174	196 125	670 000	-473 875
Uregelmessige tillegg	0	0	0	44 831	78 314	-33 483
Bonus	0	0	0	145 833	167 227	-21 394
Overtid	0	0	0	40 000	71 917	-31 917

Denne tabellen viser kun tall på aggregert nivå, dvs. næringshovedområdet som helhet. Kontroll- og revisjonsopplegget er imidlertid bygget opp og tilpasset til at den endelige lønnstatistikken skal utarbeides på et mer disaggregert nivå. Det er derfor tatt høyde for at verdi på lønnsvariablene vil variere etter gitte kjennemerker, f. eks yrke.

Ved å se på forskjellen mellom revidert og urevidert verdi for lønnsvariablene fordelt på yrke, vil man få et bedre bilde av effekten av kontroll- og revisjonsopplegget.

Tabell 13. Effekt av kontroll og revisjon av månedslønn etter næring og yrke. Heltidsansatte. Antall personer og kroner.

	Antall observasjoner			Gjennomsnittlig månedslønn		
	Revidert	Urevidert	Avvik	Revidert	Urevidert	Avvik
Bygge- og anleggsvirksomhet						
totalt						
Totalt	61 374	66 175	-4 801	34 972	32 608	2 364
Lederyrker	4 153	4 438	-285	53 568	55 793	-2 225
Akademiske yrker	3 238	3 295	-57	50 808	50 722	86
Høgskoleyrker	5 818	6 039	-221	43 111	41 641	1 470
Kontoryrker	2 017	2 114	-97	33 370	32 746	624
Salgs-/serviceyrker	310	342	-32	32 744	31 101	1 643
Håndverkere	37 202	40 745	-3 543	32 218	29 036	3 182
Operatører og sjåførere	5 821	6 000	-179	31 844	30 729	1 115
Yrker uten krav til utdanning	2 775	3 158	-383	27 989	24 638	3 351
Andre yrker	40	44	-4	29 085	25 205	3 880
Oppføring av bygninger						
Totalt	18 753	20 245	-1 492	35 467	33 499	1 968
Lederyrker	1 373	1 538	-165	56 386	59 984	-3 598
Akademiske yrker	1 385	1 408	-23	54 863	53 812	1 051
Høgskoleyrker	1 913	1 954	-41	45 778	45 021	757
Kontoryrker	780	804	-24	33 168	33 368	-200
Salgs-/serviceyrker	75	81	-6	33 558	32 009	1 549
Håndverkere	11 718	12 783	-1 065	31 525	28 782	2 743
Operatører og sjåførere	453	480	-27	33 011	30 558	2 453
Yrker uten krav til utdanning	1 046	1 186	-140	27 558	23 619	3 939
Andre yrker	10	11	-1	36 144	22 359	13 785
Anleggsvirksomhet						
Totalt	9 033	9 574	-541	38 317	37 674	1 505
Lederyrker	550	612	-62	55 254	58 564	395
Akademiske yrker	616	634	-18	46 849	54 645	223
Høgskoleyrker	784	795	-11	40 013	46 157	354
Kontoryrker	222	230	-8	29 743	33 362	1 566
Salgs-/serviceyrker	41	55	-14	31 834	25 298	5 532
Håndverkere	5 007	5 300	-293	29 876	35 743	1 973
Operatører og sjåførere	1 354	1 426	-72	28 540	31 955	1 177
Yrker uten krav til utdanning	451	512	-61	26 047	27 279	2 924
Andre yrker	8	10	-2	30 158	28 956	-1 471
Spesialisert bygge- og anleggsvirksomhet						
Totalt	33 588	36 356	-2 768	34 120	31 383	2 737
Lederyrker	2 230	2 288	-58	51 158	52 632	-1 474
Akademiske yrker	1 237	1 253	-16	45 150	46 274	-1 124
Høgskoleyrker	3 121	3 290	-169	40 563	38 434	2 129
Kontoryrker	1 015	1 080	-65	33 341	32 191	1 150
Salgs-/serviceyrker	194	206	-12	32 518	31 203	1 315
Håndverkere	20 477	22 662	-2 185	32 099	28 473	3 626
Operatører og sjåførere	4 014	4 094	-80	31 396	30 438	958
Yrker uten krav til utdanning	1 278	1 460	-182	27 824	24 831	2 993
Andre yrker	22	23	-1	26 171	25 496	675

50 091 heltidsobservasjoner får ikke endret månedslønn i revisjonsprosessen. Det vil si at 18,4 % av heltidsobservasjonene på produksjonsfila er justert.

Justeringene fordeler seg per næring på følgende måte:

	Antall ikke endrede obs	Antall obs på produksjonsfila	Prosent endrede obs
Totalt	50 091	61 374	18,4
Oppføring av bygninger	15 101	18 753	6,0
Anleggsvirksomhet	7 548	9 033	2,4
Spesialisert bygge- og anleggsvirksomhet	27 442	33 588	10,0

Justeringene fordeler seg per yrke på følgende måte:

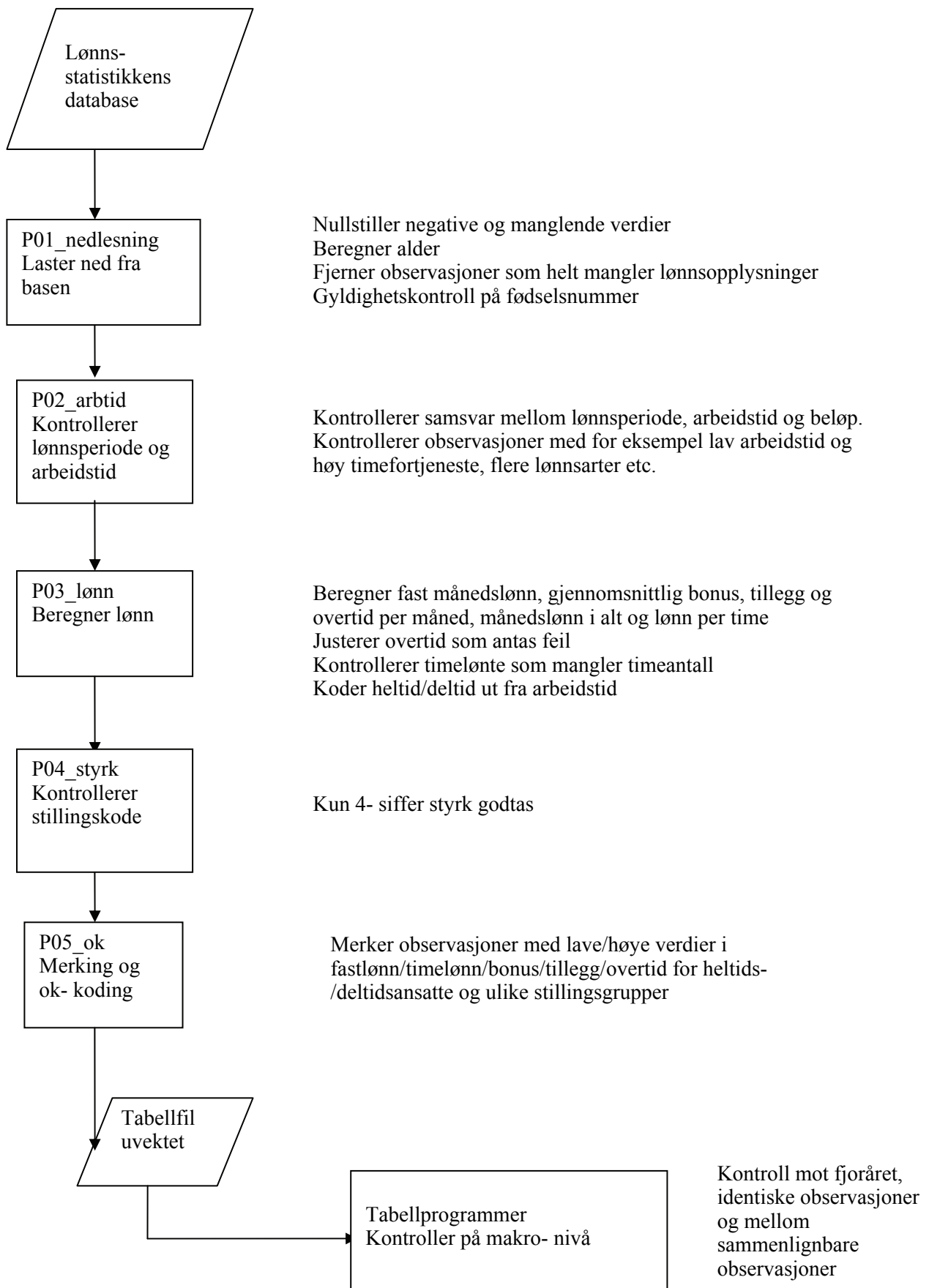
	Antall ikke endrede obs	Antall obs på produksjonsfila	Prosent endrede obs
Totalt	50 091	61 374	18,4
Lederyrker	3 790	4 153	0,6
Akademiske yrker	2 868	3 238	0,6
Høgskoleyrker	5 101	5 818	1,2
Kontoryrker	1 654	2 017	0,6
Salgs-/serviceyrker	239	310	0,1
Håndverkere	29 037	37 202	13,3
Operatører og sjåførere	5 315	5 821	0,8
Yrker uten krav til utdanning	2 060	2 775	1,2
Andre yrker	27	40	0,0

Referanser

Sigrun 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 utval og berekning av vekter. Statistisk sentralbyrå, notater 2003/74.

Vedlegg A: Prosesdiagram lønnsstatistikk



Figurregister

Figur 1.	Datainnngang i bygge- og anleggsvirksomhet. Antall innkvitterte og lest inn i Oracle-tabellen, 2010.	7
----------	---	---

Tabellregister

Tabell 1.	Endelig antall foretak, bedrifter og sysselsatte innkvittert i bygge- og anleggsvirksomhet i 2010 og prosentvis andel innkvittert i 2009 og 2010	7
Tabell 2.	Andel observasjoner etter innrapporteringsformat i perioden 2009 til 2010	8
Tabell 3.	Gyldige og ugyldige fødselsnummer. 2009 og 2010	9
Tabell 4.	Oversikt over overtidstimer og overtidsbetaling	10
Tabell 5.	Fordeling av observasjoner på 1-siffer styrk, 2009 og 2010	11
Tabell 6.	Fordeling av antall med 1-siffer styrk i lønnsstatistikken sammenlignet med AA- registeret	11
Tabell 7.	Antall observasjoner fordelt etter koder for grenseverdier (ok-koder) i 2010	12
Tabell 8.	Antall foretak i populasjon og utvalg og invers trekkesannsynlighet (Vekt)	13
Tabell 9.	Vektjustering 2010.....	13
Tabell 10.	Effekten av vekting 2010. Gjennomsnittlig månedslønn for heltidsansatte. Kroner.....	14
Tabell 11.	Effekt av kontroll og revisjon av avtalt arbeidstid per uke	14
Tabell 12.	Effekt av kontroll og revisjon av lønnsvariable. Heltidsansatte. Kroner.....	15
Tabell 13.	Effekt av kontroll og revisjon av månedslønn etter næring og yrke. Heltidsansatte. Antall personer og kroner.	16