



*Håvard Hungnes Lien*

## **Lønnsstatistikk for ansatte i industri 2008**

Klargjøring og analyse



# Innhold

<i>Sammendrag</i> .....	2
<b>1. Innledning</b> .....	<b>2</b>
<b>2. Populasjon og utvalg</b> .....	<b>3</b>
<b>3. Datafangsten</b> .....	<b>3</b>
3.1. Datainnngang .....	3
3.2. Brutto inkommet .....	4
<b>4. Behandling av data fra Oracle- basen til endelig produksjonsfil</b> .....	<b>5</b>
4.1. Tilrettelegging av endelig produksjonsfil .....	6
4.1.1. Program 1 nedlesing fra Oracle-basen.....	6
4.1.2. Program 2 arbeidstidsforhold .....	6
4.1.3. Program 3 beregning av månedslønn .....	7
4.1.4. Program 4 yrkeskoding.....	9
4.1.5.1 Program 5 godkjenning av observasjoner.....	10
4.1.5.2 Program 5b heltidsekvivalenter.....	12
4.1.6. Program 6 vektning.....	12
4.1.7. Program 7 Vektjustering.....	13
4.1.8. Frafall.....	14
<b>5. Effekten av kontroll- og revisjonsrutinene</b> .....	<b>15</b>
5.1. Arbeidstid.....	15
5.2. Lønnsvariablene .....	15
<b>6. Vedlegg</b> .....	<b>23</b>
Vedlegg 1. Prosessdiagram lønnsstatistikk .....	23
Vedlegg 2. Spredningsmål i lønnsstatistikken .....	25
Litteraturliste .....	27

Figurer:

Figur 1 Datainnngang i industri. Antall innkvitterte og lest inn i Oracle-tabellen, 2007 og 2008.....	4
---	---

## **Sammendrag**

I lønnsstatistikk for industri benyttes et sett kontroll- og revisjonsrutiner som i utgangspunktet er standardiserte for alle lønnsstatistikene.

Effekten av å kontrollere og revidere rådatamaterialet er i de fleste tilfeller betydelig, spesielt innen enkelte yrkesgrupper. 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 2008 var det om lag 91 prosent som ble godkjent til frigivning. Den tilsvarende prosentandelen i 2007 var om lag 94 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. Denne problematikken vil bli løst ved neste publisering av lønnsstatistikk, da reviderte 2008-tall blir presentert sammen med 2009-årgangen etter ny standard for næringsgruppering (NACE), SN2007. Da vil de reviderte 2008-tallene også omfatte foretak som tidligere har blitt ekskludert i datagrunnlaget.

Datafangsten i 2008 var god med tanke på hastighet og samlet datainnngang, men det er en del relativt store foretak som går igjen flere år på rad med dårlige eller ingen leverte data. I datafangsten høsten 2009 vil det ikke bli sendt purringer til oppgavegivere som ikke svarer. SSB går nå direkte til varsel om tvangsmulkt, for å unngå at oppgavegivere spekulerer i innleveringsfrister som blir utsatt flere ganger utover høsten.

Et forhold som bør følges opp særskilt ved utarbeiding av lønnsstatistikk for industri er datarevisjon spesifikt for næringsområder. Dette er fordi en del foretak som for industri (altså på næringshovedområdenivå) tatt under ett er små, likevel vil kunne ha stor innvirkning i de minste næringsområdene. Eksempelvis vil et foretak med 200 ansatte ha liten betydning for hovedtallene for industri, men det kan være helt avgjørende for lønnsutviklingen innen næringsområdet 'Produksjon av tekstil- og bekleddingsvarer'.

### **1. Innledning**

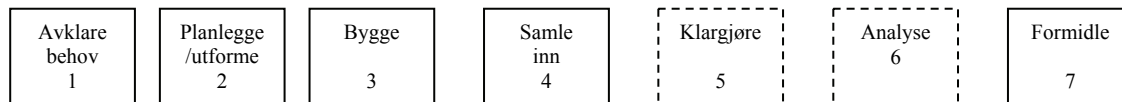
Dette notatet dokumenterer arbeidsprosessene med å klargjøre, kvalitetssikre og analysere datagrunnlaget til lønnsstatistikken for industri i 2008.

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 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 revisjonen skal legges opp, og motsatt.

I lønnsstatistikken oppfordres oppgavegiverne til å levere data elektronisk via Altinn eller cd/diskett, og det finnes også en backup-løsning med levering 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 beskrivelse av populasjonen,

tempo og kvalitet i innkvittring og lagring av data, 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 analysing av datagrunnlaget.

## 2. Populasjon og utvalg

Det trekkes et utvalg av foretak fra populasjonen i Bedrifts- og foretaksregisteret (Beyrer, Kristoffersen og Tyholt, Interne notater 2008/27). 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 fem sysselsatte trekkes ikke. Personutvalget består av alle sysselsatte i bedriftene i de uttrukne foretakene.

## 3. Datafangsten

Datainngangen per 1. oktober 2008 var relativt god, og endelig status før frigivning av statistikk viste at om lag 94 prosent av de trukne foretakene var kommet inn.

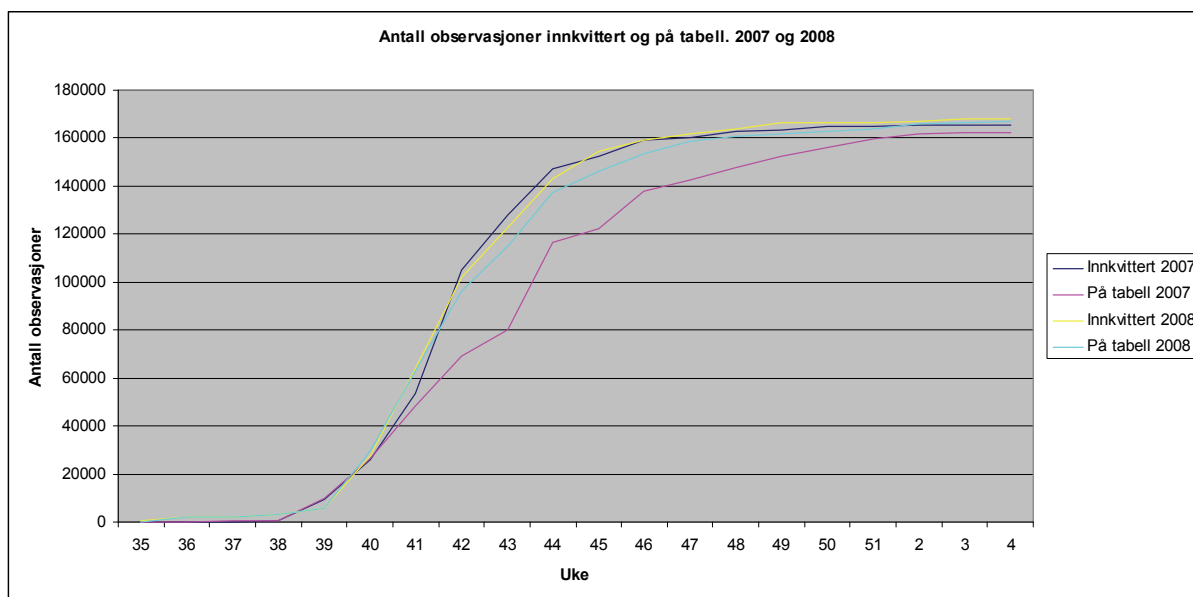
### 3.1 Datainnngang

#### Datafangsten per 1. oktober 2008

Dato	Hendelse	Antall foretak
19.09.08	Utsending	Antall foretak i utvalget: 1 699
10.10.08	Svarfrist	Innkomet data pr dato : 487 (28,7%)
16.10.08	Purring	Innkomet data pr dato : 852 (50,1%)
24.10.08	Svarfrist purring	Innkomet data pr dato : 1 078 (63,4%)
30.10.08	Varsel om tvangsmulkt	Sendt varsel til: 349 (20,5%)
07.11.08	Siste frist for å unngå tvangsmulkt	Innkomet data pr dato : 1 465 (86,2%)
09.02.09	Frigivning av statistikk	Innkomet data pr dato : 1 598 (94,1%)

Innkvittringen av foretak innen industri gikk i hovedsak bra til slutt, men det er svært betenkelig at det skal være nødvendig å purre over halvparten av foretakene og varsle om tvangsmulkt til en femtedel av dem. Et strengere regime for bruk av tvangsmulkt blir innført fra høsten 2009, med forhåpentligvis en god effekt på tempoet i datainngangen. Innlesing til Oracle har holdt et høyt og jevnt tempo gjennom hele prosessen.

**Figur 1** Datainnngang i industri. Antall innkvitterte og lest inn i Oracle-tabellen, 2007 og 2008.



Ved svarfristens utløp i uke 41 var om lag 64 000 observasjoner, eller om lag 38 % av utvalget innkvittert, og 63 000 av disse var innlest på tabell.

### 3.2 Brutto inkommet

I tabell 1 finner vi endelig antall foretak, bedrifter og sysselsatte som er innkvittert i industri i 2007 og 2008. Andelen foretak og personer innkvittert i 2008 var på omtrent samme nivå som i 2007, mens andelen innkvitterte bedrifter er noe lavere.

**Tabell 1 Endelig antall foretak, bedrifter og sysselsatte innkvittert i industri i 2008 og prosentvis andel innkvittert i 2007 og 2008**

Enhet	Datainnngang		
	2008		2007
	Antall Innkvittert	Prosent Innkvittert	Prosent Innkvittert
Foretak	1 592	94	95
Bedrifter	2 563	87	91
Sysselsatte	164 083	93	93

Tabell 2 er en oversikt over andelen som benytter seg av de ulike innrapporteringsformatene i perioden 2006 til 2008. Vi ser at bruken av Altinn er økende, men det er fortsatt et stort forbedringspotensiale i å fronte bruken av denne leveringsmåten.

**Tabell 2 Andel observasjoner etter innrapporteringsformat i perioden 2006 til 2008**

KODE	Prosent		
	2008	2007	2006
I alt	100	100	100
Excellfiler	5	6	7
Altinn	21	17	3
Kravspesifikasjon	72	74	81
Manuelt registrert	0	1	1
Optisk lest	1	1	8

#### 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 til tabell 9), før det kontrolleres mer på makronivå etterpå. Allerede ved innlesing av rådata, enten det måtte være fra Altinn-portalen, 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 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 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 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: Nedlesing fra Oracle og tilrettelegging av SAS-fil. 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, 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

### 4.1.1 Program 1 nedlesing fra Oracle-basen

Tabell 3 viser gyldige og ugyldige fødselsnummer i industri i 2007 og 2008. D-nummer er fødselsnummer til utenlandske arbeidere, og av totalen utgjør utenlandske arbeidere 2,9 prosent i 2008. Alle observasjoner tas med videre i statistikkgrunnlaget, uavhengig av om de har gyldig fødselsnummer eller ikke.

**Tabell 3 Gyldige og ugyldige fødselsnummer. 2007 og 2008**

Fødselsnummer og d-nummer	Antall	
	2007	2008
Gyldig fødselsnummer og ikke gyldig d-nummer	159 621	162 071
Gyldig d-nummer og ikke gyldig fødselsnummer	2 765	4 684
Ugyldig fødselsnummer og ugyldig d-nummer	345	310

### 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.

**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**

<b>Status observasjoner</b>	<b>Frekvens</b>	<b>Prosent</b>
<b>I alt</b>	<b>167 065</b>	<b>100</b>
Disse mangler fortsatt lønnsperiode	3	0
Disse må sjekkes. Fastlønte med timelønn i tillegg	1 769	1,06
Disse må sjekkes. Høy timefortjeneste og lav arbtid	18	0,01
Disse må sjekkes. Timelønte med fastlønn i tillegg	65	0,04
Månedslønte som mangler arbtid	112	0,07
Obs med lønnsperiode 6 som endres til 1	36	0,02
Observasjoner med annen lønnsperiode som gis lønnsperiode 6	10 354	6,2
Observasjoner som er uendret	152 793	91,46
Observasjoner uten lønnsperiode som gis lønnsperiode 1	43	0,03
Sjekkes. Heltid og fastlønn<10000 (kommafeil?)	944	0,57
Sjekkes. Timelønte med arbtid>45.	111	0,07
Timelønte uten arbtid som gis arbtid=timeant/4.33	817	0,49

Tabell 4 viser at 91 prosent av materialet slipper igjennom den første kontrollen som gjøres. Største feilpost er observasjoner med innrapportert feil lønnsperiode.

**Tabell 5 Antall observasjoner fordelt på lønnsperiode før og etter kontroll og korreksjoner av arbeidstid**

<b>Lønnsperiode</b>	<b>Antall</b>	
	<b>Før</b>	<b>Etter</b>
<b>Totalt</b>	<b>167 018</b>	<b>167 062</b>
0	103	13
1	123 790	111 596
2	2 979	1 135
3	4	2
4	230	25
5	282	270
6	39 560	53 965
7	68	55
9	2	1

Lønnsperiodene 0 og 9 er feilrapporteringer, da disse ikke finnes som verdier i SSBs variabel. Feilrapporterte lønnsperioder som har verdier på lønnsvariablene blir kodet om til enten lønnsperiode 1 (hvis fastlønn er større enn timelønn), eller til lønnsperiode 6.

#### **4.1.3 Program 3 beregning av månedslønn**

Dette programmet foretar først og fremst en beregning 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 lønnsselementer 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 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 i en 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 + 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 beregnede månedsverdiene 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 overtidsbetaling**

Status observasjoner	Frekvens	Prosent
<b>I alt</b>		<b>100,0</b>
Observasjoner uten overtidstimer eller betaling		37,4
Observasjoner med flere overtidstimer enn betaling		0,1
Observasjoner med verdier på betaling og av disse mangler det timer på		61,1
		1,4

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

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

#### 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, C521 (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. 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 o.l.
8. Operatører og sjåførere
9. Yrker uten krav til utdanning

**Tabell 7 Fordeling av observasjoner på 1-siffer STYRK, 2007 og 2008**

St1	2007		2008	
	Frekvens	Prosent	Frekvens	Prosent
I alt	162 731	100	167 065	100
1	12 381	7,6	12 492	7,5
2	9 771	6,0	10 981	6,6
3	28 159	17,3	30 430	18,2
4	12 793	7,9	12 344	7,4
5	5 187	3,2	4 737	2,8
6	241	0,2	140	0,1
7	32 337	19,9	35 536	21,3
8	53 957	33,2	50 991	30,5
9	6 929	4,3	6 381	3,8
X	976	0,6	3 033	1,8

Verdien X er samleverdien for ugyldige innrapporterte yrkeskoder.

Som en kontroll av innrapporterte yrkeskoder til lønnsstatistikken kobles det på 7-sifret yrkeskode fra AA-registeret (Arbeidstaker-/arbeidsgiverregisteret) med oppdatering nærmest mulig tellingstidspunktet. Det viser seg å være relativt store avvik mellom hva som ligger i AA-registeret og hva som rapporteres til lønnsstatistikken for noen yrkesgrupper (se tabell 8).

**Tabell 8 Fordeling av antall med 1-siffer STYRK i lønnsstatistikken sammenlignet med AA-registeret**

Lønns- statistikken	AA- registeret										
	I alt	1	2	3	4	5	6	7	8	9	X
<b>I alt</b>	<b>167 065</b>	<b>12 211</b>	<b>10 816</b>	<b>27 745</b>	<b>11 260</b>	<b>5 039</b>	<b>205</b>	<b>29 988</b>	<b>47 631</b>	<b>20 986</b>	<b>1 184</b>
1	<b>12 492</b>	<b>9 517</b>	476	904	201	143	1	260	425	440	125
2	<b>10 981</b>	466	<b>7 690</b>	1 520	157	39	1	163	179	642	124
3	<b>30 430</b>	977	2 022	<b>22 409</b>	619	556	12	869	617	2 093	256
4	<b>12 344</b>	350	262	674	<b>8 751</b>	374	6	212	615	961	139
5	<b>4 737</b>	174	48	680	227	<b>2 840</b>	4	79	98	557	30
6	<b>140</b>	5		3	2		<b>34</b>	16	34	45	1
7	<b>35 536</b>	193	82	632	274	265	26	<b>25 472</b>	3 167	5 232	193
8	<b>50 991</b>	267	82	513	691	552	97	2 209	<b>41 071</b>	5 286	223
9	<b>6 381</b>	16	14	34	138	162	15	336	753	<b>4 866</b>	47
X	<b>3 033</b>	246	140	376	200	108	9	372	672	864	<b>46</b>

Overensstemmelsen mellom AA-registeret og innrapporterte yrkeskoder i lønnsstatistikken er i all hovedsak akseptabel. Det er imidlertid verdt å legge merke til at 14,7 prosent av de som i lønnsstatistikken er oppgitt til å være håndverkere er i AA-registeret oppgitt til å arbeide i yrker uten krav til utdanning.

En direkte sammenligning av den 4-sifrede yrkeskoden fra den ferdige reviderte fila fra lønnsstatistikken og fra AA-registeret viser at 100 718 av 152 323 godkjente observasjoner (66%) på produksjonsfila har identiske verdier på variablene STYRK og AA-STYRK. I 2007 var det samme forholdet 98 396 av 153 384, eller 64%.

#### 4.1.5.1 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 justert opp 6% fra 2007 til 2008. 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 9 Antall observasjoner fordelt etter koder for grenseverdier (ok-koder) i 2008**

Ok-kode	Antall observasjoner	Prosentvis andel
<b>Totalt</b>	<b>176 410</b>	<b>100</b>
Ok=0 - Godkjente observasjoner	159 935	90,7
Ok=1 - Lave verdier på lønnsvariablene	3	0
Ok=2 - Fastlønn 0 kroner	26	0
Ok=3 - Heltidsansatte med ialt<10 500 kroner	3 565	2
Ok=5 - Timelønnte med fast månedslønn (fmlønn) over 68 900 kroner	99	0,1
Ok=6 - Provisjonslønte med bonus over 68 900 kroner	1	0
Ok=7 - Ledere (heltid) som ikke har 21 200<ialt<212 000	336	0,2
Ok=8 - St1 = 2, 3 og i alt over 116 600 kroner	83	0
Ok=9 - St1 = 4, 5 og i alt over 79 500 kroner	35	0
Ok=10 - St1 > 5 og ialt over 68 900 kroner	228	0,1
Ok=11 - Overtid (ot2) over 31 800 kroner	270	0,2
Ok=12 - Tillegg (tl) over 68 900 kroner	8	0
Ok=13 - Bonus (bon) over 106 000 kroner	27	0
Ok=14 - Deltidsansatte med arbtid<=1	189	0,1
Ok=15 - Deltidsansatte st1=1 og timefortjeneste >636	14	0
Ok=16 - Deltidsansatte st1=1 og timefortjeneste<120	14	0
Ok=17 - Deltidsansatte st1=2, 3 og timefortjeneste>530	22	0
Ok=18 - Deltidsansatte st1=2, 3 og timefortjeneste<80	84	0
Ok=19 - Deltidsansatte st1=4, 5 og timefortjeneste>265	158	0,1
Ok=20 - Deltidsansatte st1=4, 5 og timefortjeneste<60	74	0
Ok=21 - Deltidsansatte st1=7, 8 og timefortjeneste>265	244	0,1
Ok=22 - Deltidsansatte st1=7, 8 og timefortjeneste<60	144	0,1
Ok=23 - Deltidsansatte og timefortjeneste>265	64	0
Ok=24 - Deltidsansatte og timefortjeneste<50	35	0
Ok=25 - Timelønnte med utilstrekkelig info om timer	167	0,1
Ok=26 - Alder=missing	9	0
Ok=27 - 0 i fastlønn og 0 i timelønn	30	0
Ok=97 - Heltidsdubletter	1 109	0,6
Ok=98 - Deltidsdubletter	76	0
Ok=99 - Heltid-deltidsdubletter	16	0
Ok=100 - Manuelt holdt utenfor	9 345	5,3

Den vanligste ”feiltypen” er observasjoner som er valgt å holdes manuelt utenfor, enten på grunn av dårlig innrapportert data eller fordi de har ekstremverdier for lønn for sitt næringsområde og samtidig ikke har vært inkludert i datasettene foregående år. De fleste av disse er foretak som leverte data for sent i 2007, og som derfor har blitt holdt utenfor analysen i 2008 også for å skape mest mulig sammenlignbare datasett fra år til år. Et stort foretak med mange høytlønte som er ute av datamaterialet forrige år kan ikke bare uten videre tas inn igjen året etter, fordi effekten av dette vil bli en kunstig høy lønnsøkning for næringsområdet. Revidering av fjorårets tall er en av flere måter å håndtere dette problemet på, men så lenge det stort sett er de samme foretakene som må holdes ute er ikke dette en betydelig utfordring. Overgangen til ny standard for næringsgruppering, NACE rev 2 (SN 2007), i publiseringen høsten 2009 vil også bidra til å få ned antall observasjoner som må holdes ute, fordi det er en anledning til å ta med alle data igjen, i og med at det likevel vil bli et tidsseriebrudd.

#### 4.1.5.2 Program 5b, etablering av heltidsekvivalenter

Inkludering av heltidsekvivalenter i de næringsspesifikke frigivingene er nytt av året, og program 5b regner ut ekvivalentvekter og ekvivalent-lønnsarter for disse.

#### 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, beskrevet i notat av Grini (2003/74).

**Tabell 10 Antall foretak og personer i populasjonen og antall foretak i utvalget, samt justert etterstratifisert invers trekkesannsynlighet (Vekt). Pr strata.**

Stratum	Antall 2008			Vekt 2008	Vekt 2007
	Foretak i utvalget	Foretak i pop.	Ansatte i pop.		
<b>Totalt</b>	<b>1 446</b>	<b>5 751</b>	<b>256 732</b>		
DA	208	904	46 991		
DA01	19	19	20 931	1,0375	1,0531
DA02	19	38	7 394	1,4726	1,223
DA03	170	847	18 666	4,3685	3,5927
DB	61	195	4 328		
DB01	16	16	1 500	0,9588	0,9904
DB02	14	27	1 157	1,3839	1,3055
DB03	31	152	1 671	3,4226	4,256
DC	10	10	186		
DC01	10	10	186	0,9684	1,1319
DD	129	524	14 620		
DD01	15	15	4 165	0,823	0,8713
DD02	19	37	3 126	1,0369	0,9275
DD03	95	472	7 329	3,5672	4,0718
DE	182	786	27 428		
DE01	14	14	8 145	1,188	0,9793
DE02	21	41	6 391	1,0176	1,1215
DE03	147	731	12 892	3,7505	3,4487
DF	4	4	1 120		
DF01	4	4	1 120	1,6064	1,2406
DG	41	110	13 150		
DG01	14	14	7 578	1,3782	1,3492
DG02	12	23	3 858	1,4976	1,5811
DG03	15	73	1 714	6,343	6,9494
DH	63	161	5 250		
DH01	13	13	2 044	1,2048	1,2406
DH02	33	65	2 344	2,1165	2,1893

DH03	17	83	862	6,2499	5,7304
DI	78	262	10 631		
DI01	17	17	5 116	1,2048	1,2406
DI02	19	38	2 764	1,6956	1,5746
DI03	42	207	2 751	5,5421	5,1361
DJ	200	897	33 815		
DJ01	14	14	9 363	1,4169	1,2392
DJ02	15	30	5 557	1,4544	1,4836
DJ03	171	853	18 895	4,2042	4,0223
DK	142	597	25 442		
DK01	17	17	10 572	0,9835	0,8259
DK02	15	30	3 751	1,4463	1,2389
DK03	110	550	11 119	4,5893	4,1593
DL	110	438	19 897		
DL01	13	13	7 989	1,1281	0,9085
DL02	20	40	4 981	1,0519	0,9959
DL03	77	385	6 927	3,3079	3,018
DM	126	482	43 601		
DM01	27	27	22 965	1,3531	1,1817
DM02	13	25	6 487	1,708	1,4058
DM03	86	430	14 149	5,0324	4,8338
DN	92	381	10 273		
DN01	8	8	2 516	0,8021	0,8504
DN02	15	30	2 512	1,426	1,0353
DN03	69	343	5 245	4,1063	4,1103

#### 4.1.7 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 sysselsettinga 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 2008. Heltids- og deltidsansatte**

Næring	Vektet antall		Vektet andel		BoF andel	Justeringsfaktor mot BoF
	2007	2008	2007	2008		
<b>Totalt</b>	<b>256 509</b>	<b>267 866</b>	<b>1</b>	<b>1</b>	<b>1</b>	
Nærings- og nytelsesm.	54 667	50 195	0,2131	0,1874	0,1944	1,0375
Tekstil- og bekl.	4 530	4 378	0,0177	0,0163	0,0142	0,8716
Trelast- og trev.	13 291	15 373	0,0518	0,0574	0,0441	0,7681
Treforedling	7 677	7 077	0,0299	0,0264	0,0309	1,171
Forlag og grafisk ind.	19 377	24 554	0,0755	0,0917	0,0739	0,8067
Kjemisk og min. ind. mv.	26 844	22 678	0,1047	0,0847	0,102	1,2048
Kjemiske råvarer	10 897	7 073	0,0425	0,0264	0,0396	1,4988
Metallindustri	16 749	14 123	0,0653	0,0527	0,0653	1,2388
Verkstedindustri	57 250	78 527	0,2232	0,2932	0,2544	0,8678
Bygging av skip og oljep.	36 427	34 197	0,142	0,1277	0,1521	1,1911
Møbelind. og annen ind.	8 799	9 691	0,0343	0,0362	0,029	0,8021

I 2008 ble samvektingen gjort direkte mot Bedrifts- og foretaksregisteret, fordi det er merkbare strukturendringer innen industri i forhold til i 2007. Det har konsekvent blitt flere ansatte i høytlønnsnæringene og færre i lavtlønnsnæringene. Dette betyr at lønnsveksten for industri som et hele er høyere i 2008 enn hva som ville vært tilfelle hvis antallet ansatte i undernæringene hadde vært konstant fra 2007 til 2008, selv om lønnsutviklingen for de enkelte undernæringene er relativt stabil fra år til år. Strukturendringen er også målt av Nasjonalregnskapet (tabell 6173, Lønn og sysselsetting, i Statistikkbanken), noe som gir støtte til vår konklusjon.

**Tabell 12 Effekt av vekting 2008. Gjennomsnittlig månedslønn for heltidsansatte. Kroner**

	Månedslønn 2008. Heltidsansatte. Kroner		
	Uvektet	Ujustert vekt	Justert vekt
<b>Totalt</b>	<b>34 852</b>	<b>33 952</b>	<b>33 977</b>
Nærings- og nytelsesm.	30 634	30 001	29 852
Tekstil- og bekl.	28 849	29 039	29 039
Trelast- og trev.	29 552	28 694	28 694
Treforedling	32 439	32 295	32 295
Forlag og grafisk ind.	39 287	37 006	37 006
Kjemisk og min. ind. mv.	35 199	34 555	34 555
Kjemiske råvarer	39 784	39 346	39 346
Metallindustri	35 325	35 148	34 984
Verkstedindustri	36 467	35 077	35 077
Bygging av skip og oljep.	36 985	36 500	36 029
Møbelind. og annen ind.	30 556	30 178	30 178

#### 4.1.8 Frafall

Vi har to hovedtyper frafall, enhetsfracfall og partielt frafall. Enhetsfracfall defineres her som frafall av bedrifter, mens partielt frafall er frafall av personer.



Enhetsfracfall skyldes at bedrifter ikke har innrapportert til oss, eller at det er feil i datamaterialet vi har fått innrapportert. Under enhetsfracfall av bedrifter kommer også de bedriftene vi har valgt å holde ute manuelt. Totalt sett var enhetsfracfallet på 101 foretak i 2008.

Partielt fracfall kan blant annet skyldes mangelfull innrapportering, feil i data eller at personer ikke er blitt innrapportert. Denne type fracfall er omtalt i kapittel 4.1.5.1 hvor det er blitt satt minimums- og maksimumsgrenser for de ulike lønnsvariablene. Hvis vi ikke inkluderer de observasjonene som blir holdt utenfor manuelt (kode 100 i tabell 9), er de fleste observasjonene som ikke ble godkjente heltidsansatte med mindre enn 10 500 kroner i månedslønn. Det partielle fracfallet var totalt sett på 7 130 observasjoner.

## 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 utfyllt. Mye av kontrollene går av den grunn ut på sjekke sammenheng mellom arbeidstid og lønn.

**Tabell 13 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	166 880	152 139	35,54	35,72
Heltidsansatte	152 165	138 899	36,83	37,06
Deltidsansatte	14 715	13 240	22,23	22,92

Som det går fram av tabellen er det flere observasjoner på den ureviderte fila (rådata) enn på den reviderte fila (publiseringsfila) grunnet at alle ikke-godkjente observasjoner her er holdt utenfor etter kriteriene beskrevet i kapittel 4.1.5.1. 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. 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 4.1.5.1. Revideringen viser seg å ha flyttet 23 observasjoner fra deltid til heltid, og 2 observasjoner fra heltid til deltid.

### 5.2 Lønnsvariablene

Responsvariablene i lønnsstatistikken er lønnsvariablene som frigis i den offisielle statistikken, og er derfor også de vi først og fremst 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 responsvariablene 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 14 Effekt av kontroll og revisjon av lønnsvariable. Heltidsansatte. Kroner**

	Revidert	Urevidert	Avvik			
Antall	138 899	152 526	-13 627			
	Gjennomsnitt			Median		
	Revidert	Urevidert	Avvik	Revidert	Urevidert	Avvik
Månedslønn	33 977	32 498	1 479	30 625	30 394	231
Avtalt lønn	31 985	30 377	1 608	28 635	28 403	232
Uregelmessige tillegg	1 252	1 377	-125	0	0	0
Bonus	741	803	-62	0	0	0
Overtid	1 862	1 918	-56	415	407	8
	1. kvartil			3. kvartil		
	Revidert	Urevidert	Avvik	Revidert	Urevidert	Avvik
Månedslønn	26 291	25 412	879	37 556	37 785	-230
Avtalt lønn	24 863	24 196	667	35 200	35 347	-147
Uregelmessige tillegg	0	0	0	1 289	1 348	-59
Bonus	0	0	0	625	625	0
Overtid	0	0	0	2 320	2 328	-7
	Min			Max		
	Revidert	Urevidert	Avvik	Revidert	Urevidert	Avvik
Månedslønn	10 502	0	10 502	211 166	717 647	-506 480
Avtalt lønn	567	0	567	208 333	630 700	-422 367
Uregelmessige tillegg	0	0	0	57 778	446 797	-389 019
Bonus	0	0	0	103 027	354 610	-251 583
Overtid	0	0	0	31 768	126 452	-94 684

Som det går fram av tabellen gjelder det for alle variablene at de største *kronemessige* endringene skjer i den øvre enden av skalaen, men i og med at revidert månedslønn er høyere enn urevidert månedslønn betyr det at det er et større *antall* observasjoner som er endret i nedre enn øvre del av skalaen. Et problem med denne tabellen er imidlertid at den tar for seg

tall på et aggregert nivå, det vil si 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, for eksempel næring og yrke.

Ved å se på forskjellen mellom revidert og urevidert verdi for lønnsvariablene fordelt på næring og yrke, vil man få et bedre bilde av effekten av kontroll- og revisjonsopplegget. For å forenkle tar vi i tabell 15 bare for oss månedslønn.

**Tabell 15 Effekt av kontroll og revisjon av månedslønn etter næring og yrke. Heltidsansatte. Antall personer og kroner.**

	Antall uvektede observasjoner			Gjennomsnittlig månedslønn		
	Urevidert	Revidert	Avvik	Urevidert	Revidert	Avvik
Industri totalt						
<b>Totalt</b>	<b>152 526</b>	<b>138 899</b>	<b>-13 627</b>	<b>32 498</b>	<b>33 977</b>	<b>1 479</b>
Lederyrker	12 153	10 894	-1 259	57 794	56 166	-1 628
1210 Adm. dir.	1 400	1 181	-219	80 378	74 115	-6 263
1300 Ledere i små bedrifter	1 012	949	-63	48 788	44 629	-4 159
Akademiske yrker	10 295	8 151	-2 144	46 462	46 236	-226
Høgskoleyrker	28 882	25 669	-3 213	39 208	40 449	1 241
Kontoryrker	10 134	9 307	-827	29 309	30 481	1 172
Salgs-/serviceyrker	3 291	3 005	-286	29 404	31 114	1 710
Håndverkere	33 700	30 012	-3 688	27 836	29 666	1 830
Operatører og sjåførere	47 220	44 923	-2 297	25 780	28 496	2 716
Yrker uten krav til utdanning	4 781	4 494	-287	21 890	26 099	4 209
Andre yrker	2 070	2 444	374	28 746	31 033	2 287
Nærings- og nytelsesmiddelsindustri						
<b>Totalt</b>	<b>26 994</b>	<b>26 127</b>	<b>-867</b>	<b>27 314</b>	<b>29 852</b>	<b>2 538</b>
Lederyrker	2 323	2 224	-99	49 437	49 668	231
1210 Adm. dir.	155	135	-20	69 551	68 458	-1 093
1300 Ledere i små bedrifter	144	130	-14	40 598	42 913	2 315
Akademiske yrker	854	819	-35	41 757	41 586	-171
Høgskoleyrker	2 967	2 904	-63	33 864	34 533	669
Kontoryrker	2 838	2 771	-67	27 059	29 022	1 963
Salgs-/serviceyrker	1 410	1 326	-84	24 355	26 574	2 219
Håndverkere	3 255	3 123	-132	25 182	27 653	2 471
Operatører og sjåførere	11 601	11 162	-439	23 489	26 612	3 123
Yrker uten krav til utdanning	1 444	1 391	-53	16 483	24 189	7 706
Andre yrker	302	407	105	24 117	27 037	2 920
Tekstil- og bekledningsindustri						
<b>Totalt</b>	<b>2 009</b>	<b>1 937</b>	<b>-72</b>	<b>26 445</b>	<b>29 039</b>	<b>2 594</b>
Lederyrker	197	174	-23	43 786	43 419	-367
1210 Adm. dir.	26	24	-2	54 361	53 074	-1 287
1300 Ledere i små bedrifter	49	41	-8	38 625	42 302	3 677
Akademiske yrker	12	11	-1	32 748	33 680	932
Høgskoleyrker	167	159	-8	33 485	33 894	409
Kontoryrker	233	226	-7	23 342	28 292	4 950
Salgs-/serviceyrker	104	84	-20	33 695	34 547	852

Håndverkere	141	140	-1	24 653	27 913	3 260
Operatører og sjåførere	1 037	974	-63	21 705	25 109	3 404
Yrker uten krav til utdanning	66	62	-4	20 121	24 122	4 001
Andre yrker	52	107	55	32 292	29 563	-2 729
<b>Trelast- og trevareindustri</b>						
<b>Totalt</b>	<b>6 732</b>	<b>6 533</b>	<b>-199</b>	<b>27 878</b>	<b>28 694</b>	<b>816</b>
Lederyrker	447	420	-27	52 353	45 437	-6 916
1210 Adm. dir.	60	54	-6	61 734	61 950	216
1300 Ledere i små bedrifter	65	60	-5	65 842	38 905	-26 937
Akademiske yrker	72	69	-3	44 474	43 250	-1 224
Høgskoleyrker	415	405	-10	40 329	36 358	-3 971
Kontoryrker	394	387	-7	28 865	30 685	1 820
Salgs-/serviceyrker	107	84	-23	40 161	35 398	-4 763
Håndverkere	1 632	1 586	-46	25 888	26 889	1 001
Operatører og sjåførere	3 063	2 981	-82	23 162	25 821	2 659
Yrker uten krav til utdanning	502	498	-4	26 692	27 285	593
Andre yrker	100	103	3	27 481	27 534	53
<b>Treforedlingsindustri</b>						
<b>Totalt</b>	<b>4 675</b>	<b>4 577</b>	<b>-98</b>	<b>31 568</b>	<b>32 295</b>	<b>727</b>
Lederyrker	348	337	-11	59 084	57 209	-1 875
1210 Adm. dir.	23	22	-1	75 172	72 406	-2 766
1300 Ledere i små bedrifter	47	47	0	40 344	40 841	497
Akademiske yrker	185	183	-2	48 601	48 347	-254
Høgskoleyrker	403	395	-8	35 787	36 121	334
Kontoryrker	363	359	-4	28 635	29 802	1 167
Salgs-/serviceyrker	48	45	-3	31 395	32 978	1 583
Håndverkere	562	561	-1	29 488	29 759	271
Operatører og sjåførere	2 590	2 522	-68	27 480	28 584	1 104
Yrker uten krav til utdanning	61	57	-4	22 715	24 967	2 252
Andre yrker	115	118	3	25 673	28 783	3 110
<b>Forlag og grafisk industri</b>						
<b>Totalt</b>	<b>10 916</b>	<b>10 726</b>	<b>-190</b>	<b>36 593</b>	<b>37 006</b>	<b>413</b>
Lederyrker	1 326	1 295	-31	53 397	52 872	-525
1210 Adm. dir.	332	315	-17	67 722	64 263	-3 459
1300 Ledere i små bedrifter	122	120	-2	45 201	45 218	17
Akademiske yrker	862	860	-2	39 392	39 650	258
Høgskoleyrker	3 988	3 947	-41	37 229	37 839	610
Kontoryrker	750	742	-8	30 565	29 817	-748
Salgs-/serviceyrker	732	705	-27	32 527	33 223	696
Håndverkere	1 194	1 180	-14	33 873	34 100	227
Operatører og sjåførere	1 388	1 370	-18	30 801	31 275	474
Yrker uten krav til utdanning	318	293	-25	24 086	27 489	3 403
Andre yrker	358	334	-24	33 434	34 287	853
<b>Kjemisk og mineralisk industri</b>						
<b>Totalt</b>	<b>16 140</b>	<b>15 089</b>	<b>-1 051</b>	<b>30 858</b>	<b>34 555</b>	<b>3 697</b>
Lederyrker	1 420	1 356	-64	56 526	56 142	-384
1210 Adm. dir.	125	117	-8	88 392	78 679	-9 713
1300 Ledere i små bedrifter	138	134	-4	40 932	41 417	485
Akademiske yrker	1 076	931	-145	42 087	48 412	6 325

Høgskoleyrker	2 295	2 138	-157	37 779	39 841	2 062
Kontoryrker	1 167	1 099	-68	28 868	31 219	2 351
Salgs-/serviceyrker	345	276	-69	28 968	35 800	6 832
Håndverkere	2 209	2 128	-81	26 950	30 271	3 321
Operatører og sjåførere	6 768	6 277	-491	23 942	28 890	4 948
Yrker uten krav til utdanning	663	640	-23	22 720	28 227	5 507
Andre yrker	197	244	47	28 675	32 947	4 272
<b>Kjemiske råvarer</b>						
<b>Totalt</b>	<b>6 091</b>	<b>5 700</b>	<b>-391</b>	<b>40 681</b>	<b>39 346</b>	<b>-1 335</b>
Lederyrker	800	616	-184	75 954	68 522	-7 432
1210 Adm. dir.	149	90	-59	129 151	105 864	-23 287
1300 Ledere i små bedrifter	10	9	-1	41 507	39 302	-2 205
Akademiske yrker	479	425	-54	50 085	49 189	-896
Høgskoleyrker	1 080	1 034	-46	43 828	43 942	114
Kontoryrker	335	332	-3	33 442	33 216	-226
Salgs-/serviceyrker	56	52	-4	35 972	36 926	954
Håndverkere	469	450	-19	29 816	30 975	1 159
Operatører og sjåførere	2 756	2 672	-84	31 047	32 107	1 060
Yrker uten krav til utdanning	55	53	-2	18 636	26 194	7 558
Andre yrker	61	66	5	27 020	28 715	1 695
<b>Metallindustri</b>						
<b>Totalt</b>	<b>10 141</b>	<b>9 809</b>	<b>-332</b>	<b>33 209</b>	<b>34 984</b>	<b>1 775</b>
Lederyrker	556	541	-15	64 198	64 426	228
1210 Adm. dir.	73	69	-4	102 423	103 266	843
1300 Ledere i små bedrifter	20	18	-2	43 407	43 449	42
Akademiske yrker	450	445	-5	47 528	47 658	130
Høgskoleyrker	825	820	-5	44 188	44 223	35
Kontoryrker	227	224	-3	29 499	31 771	2 272
Salgs-/serviceyrker	77	75	-2	38 427	39 012	585
Håndverkere	1 089	906	-183	28 477	31 400	2 923
Operatører og sjåførere	6 669	6 545	-124	29 165	31 232	2 067
Yrker uten krav til utdanning	196	186	-10	29 156	30 077	921
Andre yrker	52	67	15	29 605	29 457	-148
<b>Verkstedindustri</b>						
<b>Totalt</b>	<b>40 531</b>	<b>34 727</b>	<b>-5 804</b>	<b>33 473</b>	<b>35 077</b>	<b>1 604</b>
Lederyrker	2 996	2 445	-551	56 939	56 615	-324
1210 Adm. dir.	332	248	-84	76 920	74 511	-2 409
1300 Ledere i små bedrifter	351	332	-19	55 112	47 997	-7 115
Akademiske yrker	3 191	2 687	-504	47 235	49 202	1 967
Høgskoleyrker	10 175	8 259	-1 916	38 590	42 010	3 420
Kontoryrker	2 542	2 114	-428	30 596	31 073	477
Salgs-/serviceyrker	226	198	-28	35 409	35 882	473
Håndverkere	11 068	9 373	-1 695	28 098	29 733	1 635
Operatører og sjåførere	9 137	8 427	-710	26 682	28 233	1 551
Yrker uten krav til utdanning	774	684	-90	22 917	26 473	3 556
Andre yrker	422	540	118	30 596	34 061	3 465
<b>Bygging av skip og oljeplattformer</b>						
<b>Totalt</b>	<b>24 262</b>	<b>19 828</b>	<b>-4 434</b>	<b>35 868</b>	<b>36 029</b>	<b>161</b>
Lederyrker	1 379	1 147	-232	69 808	65 426	-4 382

1210 Adm. dir.	81	65	-16	87 232	77 753	-9 479
1300 Ledere i små bedrifter	27	22	-5	53 965	46 256	-7 709
Akademiske yrker	3 046	1 663	-1 383	49 810	44 899	-4 911
Høgskoleyrker	6 290	5 362	-928	42 927	42 570	-357
Kontoryrker	968	752	-216	31 471	32 141	670
Salgs-/serviceyrker	43	27	-16	25 548	32 746	7 198
Håndverkere	10 934	9 473	-1 461	28 486	30 051	1 565
Operatører og sjåførere	835	681	-154	22 923	29 664	6 741
Yrker uten krav til utdanning	503	447	-56	24 597	25 485	888
Andre yrker	264	276	12	30 438	31 737	1 299
Møbel- og annen industri						
<b>Totalt</b>	<b>4 035</b>	<b>3 846</b>	<b>-189</b>	<b>28 119</b>	<b>30 178</b>	<b>2 059</b>
Lederyrker	361	339	-22	51 776	50 050	-1 726
1210 Adm. dir.	44	42	-2	66 646	58 680	-7 966
1300 Ledere i små bedrifter	39	36	-3	45 566	45 845	279
Akademiske yrker	68	58	-10	41 775	42 149	374
Høgskoleyrker	277	246	-31	36 956	37 082	126
Kontoryrker	317	301	-16	27 411	28 741	1 330
Salgs-/serviceyrker	143	133	-10	36 994	34 434	-2 560
Håndverkere	1 147	1 092	-55	22 325	26 969	4 644
Operatører og sjåførere	1 376	1 312	-64	24 566	26 750	2 184
Yrker uten krav til utdanning	199	183	-16	26 333	26 711	378
Andre yrker	147	182	35	26 651	28 925	2 274

126 906 heltids-observasjoner får ikke endret månedslønn i revisjonsprosessen, noe som vil si at 8,64% av heltidsobservasjonene på produksjonsfila er justert.

Per næringsområde ser vi at det fordeler seg slik:

	antall ikke- endrede obs	antall obs på produksjonsfil	prosent endrede obs
<b>Totalt</b>	<b>126 906</b>	<b>138 899</b>	<b>8,6</b>
Nærings- og nytelsesmiddelindustri	23 513	26 127	10,0
Tekstil- og beklædningsindustri	1 585	1 937	18,2
Trelast- og trevareindustri	5 965	6 533	8,7
Treforedlingsindustri	4 489	4 577	1,9
Forlag og grafisk industri	10 462	10 726	2,4
Kjemisk og mineralsk industri	13 298	15 089	11,9
Kjemiske råvarer	5 512	5 700	3,3
Metallindustri	9 103	9 809	7,2
Verkstedindustri	30 839	34 727	11,2
Bygging av skip og oljeplattformer	18 716	19 828	5,6
Møbel- og annen industri	3 424	3 846	11,0

Og per yrke slik:

	antall ikke- endrede obs	antall obs på produksjonsfil	prosent endrede obs
<b>Totalt</b>	<b>126 906</b>	<b>138 899</b>	<b>8,63</b>
Lederyrker	10 713	10 894	1,66
1210 Adm. dir.	1 170	1 181	0,93
1300 Ledere i små bedrifter	854	949	10,01
Akademiske yrker	7 841	8 151	3,8
Høgskoleyrker	24 543	25 669	4,39
Kontoryrker	8 747	9 307	6,02
Salgs-/serviceyrker	2 810	3 005	6,49
Håndverkere	26 948	30 012	10,21
Operatører og sjåførere	40 157	44 923	10,61
Yrker uten krav til utdanning	3 570	4 494	20,56
Andre yrker	1 577	2 444	35,47

Som en følge av revisjonsrutinene minskes også den samlede spredningen. I tabell 16 er det beregnet ulike spredningsmål for heltidsansatte i industri på ureviderte og reviderte tall for 2008.

**Tabell 16 Effekt av kontroll og revisjon av månedslønn. Spredningsmål pr næringsområde. Antall og kroner.**

Ureviderte tall	Ant obs	Månedslønn			
		Uvektet antall	Gjennomsnitt	Standardfeil	Relativ standardfeil
<b>Totalt</b>	<b>152 526</b>	<b>32 498</b>	<b>58</b>	<b>0,18 %</b>	<b>±114</b>
Nærings- og nytelsesm,	26 994	27 314	109	0,40 %	±214
Tekstil- og bekl,	2 009	26 445	385	1,46 %	±755
Trelast- og trev,	6 732	27 878	446	1,60 %	±874
Treforedling	4 675	31 568	246	0,78 %	±482
Forlag og grafisk ind,	10 916	36 593	182	0,50 %	±357
Kjemisk og min, ind, mv,	16 140	30 858	160	0,52 %	±314
Kjemiske råvarer	6 091	40 681	329	0,81 %	±645
Metallindustri	10 141	33 209	179	0,54 %	±351
Verkstedindustri	40 531	33 473	122	0,37 %	±239
Bygging av skip og oljep,	24 262	35 868	149	0,41 %	±292
Møbelind, og annen ind,	4 035	28 119	320	1,14 %	±627

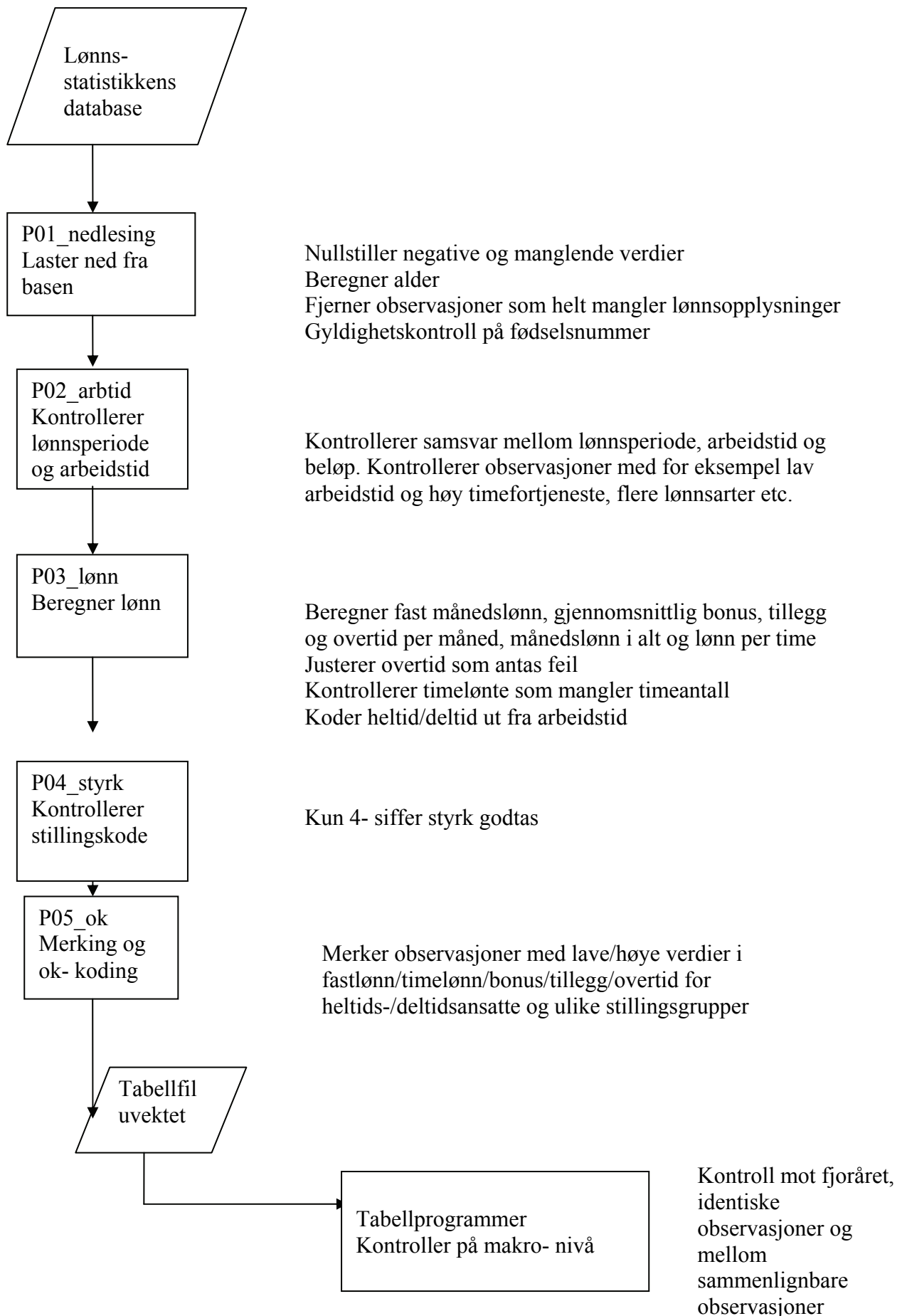
Reviderte tall	Ant obs	Månedslønn			
		Uvektet antall	Gjennomsnitt	Standardfeil	Relativ standardfeil
<b>Totalt</b>	<b>138 899</b>	<b>33 977</b>	<b>41</b>	<b>0,12 %</b>	<b>±80</b>
Nærings- og nytelsesm,	26 127	29 852	82	0,27 %	±160
Tekstil- og bekledningsind,	1 937	29 039	262	0,90 %	±514
Trelast- og trevarehandel,	6 533	28 694	124	0,43 %	±243
Treforedling	4 577	32 295	198	0,61 %	±389
Forlag og grafisk industri,	10 726	37 006	138	0,37 %	±271
Kjemisk og min, ind, mv,	15 089	34 555	121	0,35 %	±237
Kjemiske råvarer	5 700	39 346	260	0,66 %	±510
Metallindustri	9 809	34 984	146	0,42%	±287
Verkstedindustri	34 727	35 077	81	0,23 %	±159
Bygging av skip og oljep,	19 828	36 029	117	0,33 %	±230
Møbelind, og annen industri,	3 846	30 178	224	0,74 %	±438

Vi ser at datarevisjonen minsker spredningen i datamaterialet for alle næringsområder.



## **Vedlegg**

*Vedlegg 1. Prosessdiagram lønnsstatistikk*



## Vedlegg 2. Spredningsmål i lønnsstatistikken

I tabell 16 er det blitt beregnet konfidensintervall, standardfeil og relativ standardfeil (CV) for heltidsansatte i industri i 2008.

Standardavviket benyttes generelt om stokastiske variable mens standardfeilen benyttes når vi snakker om estimater. Standardfeilen for utvalgsgjennomsnittet beregnes ved å ta rota av vekta i foretak a multiplisert med det kvadrerte avviket fra gjennomsnittet i foretak a dividert på rota av n antall observasjoner minus en. Tar da her hensyn til at dette er et stratifisert tilfeldig klyngeutvalg. I vårt tilfelle er populasjonsgjennomsnittet et rent utvalgsgjennomsnitt. Standardavviket og standardfeilen er absolutte mål for spredning av verdier i et datasett eller i en stokastisk variabel. Den relative standardfeilen beregnes ved å dividere standardfeilen på gjennomsnittlig månedslønn. Vi kan derfor si at den relative standardfeilen er et uttrykk for hvor stor andel standardfeilen for utvalgsgjennomsnittet utgjør av gjennomsnittlig månedslønn.

Variasjonskoeffisienten kalles ofte for den relative standardfeilen og er definert som forholdet mellom standardfeilen til estimatoren og forventningsverdien (gjennomsnittet). I vårt tilfelle ønsker vi å beregne variasjonskoeffisienten til den estimerte gjennomsnittlige månedslønna.

Basert på den generelle definisjonen av gjennomsnittet har vi følgende formel:

$$\bar{X}_i = \frac{\sum_{i=1}^N X_i}{N}$$

der

$\bar{X}$  - gjennomsnittlig månedslønn

$i$  - individ (analyseenhet)

$N$  - totalt antall sysselsatte i populasjonen

$X_i$  - månedslønn for individ  $i$ , i populasjonen

Vi kan nå videre definere det estimerte gjennomsnittet for mål populasjonen ved å benytte foretakenes vekter på følgende måte:

$$\hat{\bar{X}} = \frac{\sum_{a=1}^k w_a x_a}{\sum_{a=1}^k w_a}$$

der

$\hat{X}$  - er estimert gjennomsnittlig månedslønn for utvalget

$a$  - foretak (utvalgseenhet)

$x_a$  - er observert gjennomsnittlig månedslønn i foretak  $a$  i utvalget

$w_a$  - er den beregnede vekta for foretak  $a$

Standardfeilen ( $s$ ) til det estimerte gjennomsnittet defineres på følgende måte:

$$s = \sqrt{\frac{\sum_{a=1}^k w_a (x_a - \bar{x})^2}{n-1}}$$

der

$n$  - er antall klynger (som i vårt tilfelle er individer) i utvalget

$(x_a - \bar{x})^2$  - er det kvadrerte avviket fra gjennomsnittet i foretak  $a$  i utvalget

Variasjonskoeffisienten (CV) eller den relative standardfeilen som den også kalles kan da beregnes på følgende måte:

$$CV = \frac{s}{\hat{X}}$$

CV uttrykkes ofte som prosentandel

$$CV = \frac{s}{\hat{X}} 100\%$$

95% konfidensintervallet er i statistikken en måte å angi feilmarginen av en måling eller en beregning på. Således inneholder et 95 -konfidensintervall den sanne verdien med en sannsynlighet på 0,95. Ved normalfordelte variabler kan 95 -konfidensintervallene nokså nøyaktig regnes om fra variabelens gjennomsnitt ( $\hat{X}$ ) og standardfeilen ( $s$ ) som:

$$CI = [\hat{X} - 1,96 \cdot s; \hat{X} + 1,96 \cdot s]$$

## **Litteraturliste**

Beyrer, Kristoffersen, Tyholt (2008): Lønnsstatistikk 2008. Etablering av populasjon og utvalg. Statistisk sentralbyrå, interne notater 2008/27

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