

Kvalitetsdeklaration

Sökområde, 2023:2

Ämnesområde

Arbetsmarknad

Statistikområde

Arbetsförmedlingens arbete med ersättningssökandes yrkesmässiga och geografiska sökområde.

Producerande myndighet

Inspektionen för arbetslöshetsförsäkringen

Referenstid

2019/2021.

Kontaktuppgifter

Niklas Ottosson niklas.ottosson@iaf.se

Patrick Karabona patrick.karabona@iaf.se



Innehåll

1	Relevans	3
1.1	Ändamål och informationsbehov	3
1.2	Statistikens innehåll	3
1.2.1	Objekt och population.....	3
1.2.2	Variabler	4
1.2.3	Statistiska mått	4
1.2.4	Redovisningsgrupper	4
1.2.5	Referenstider	4
2	Tillförlitlighet	4
2.1	Tillförlitlighet totalt	4
2.2	Osäkerhetskällor	5
2.2.1	Precision i den maskinella tolkningen	5
2.2.2	Uppskattning av osäkra tolkningars precision.....	5
2.2.3	Urval	5
2.2.4	Ramtäckning.....	6
2.2.5	Mätning.....	6
2.2.6	Bortfall	6
2.2.7	Bearbetning	7
2.3	Vad är det som inte går att tolka?.....	7
3	Aktualitet	8
3.1	Framställningstid	8
4	Tillgänglighet och tydlighet	8
4.1	Tillgång till statistiken och dokumentation	8
5	Jämförbarhet.....	8
5.1	Jämförbarhet över tid	8
5.2	Jämförbarhet mellan grupper	9
	Allmänna uppgifter	9



1 Relevans

Den här kvalitetsdeklarationen är en bilaga till rapporten *Hur arbetar Arbetsförmedlingen med geografiskt och yrkesmässigt sökområde?*.

Kvalitetsdeklarationen syftar till att beskriva kvaliteten på den statistik som presenteras i rapporten utifrån statistikens ändamål, innehåll, tillförlitlighet, aktualitet, tillgänglighet och jämförbarhet.

1.1 Ändamål och informationsbehov

Statistiken i rapporten beskriver arbetssökandes yrkesmässiga och geografiska sökbeteende. Syftet är att granska hur Arbetsförmedlingen arbetar med geografiskt och yrkesmässigt sökområde för arbetssökande som får arbetslöshetsersättning och aktivt ska söka lämpliga arbeten.

Huvudsaklig användare av statistiken är IAF, Arbetsförmedlingen, arbetslöshetskassorna och Arbetsmarknadsdepartementet.

1.2 Statistikens innehåll

De statistiska målstorheterna utgörs av andelar och antal av olika variabler som mäter arbetssökandes benägenhet att vidga det yrkesmässiga och geografiska sökområdet.

1.2.1 Objekt och population

Villkoren för att tillhöra rampopulationen:

1. Individen ska ha fått arbetslöshetsersättning utbetald någon gång under perioden januari – mars 2019 respektive 2021.
2. Individen tillhör någon av sökandekategorierna 11 Öppet arbetslös; 95 Öppet arbetslös, upphävande av beslut; 96 Öppet arbetslös, felregistrering av beslut, 97 Öppet arbetslös, avbrott/återkallande av beslut, 98 Öppet arbetslös, slutförd beslutsperiod; 21 Deltidsarbetslös och/eller 22 Timanställd.
3. Individen ska ha minst en yrkeskod (SSYK) registrerad i Arbetsförmedlingen ärendehanteringssystem eller ha en X21 eller X33-kod som visar att den arbetssökande inte har tidigare arbetslivserfarenhet.

Rampopulationerna togs fram under oktober 2021 och februari 2022.



Rampopulationen täcker intressepopulationen fullt ut med undantag för eventuella registerbrister.

1.2.2 Variabler

Observationsvariablerna består av textfälten ort och yrke från de aktivitetsrapporter som ersättningssökande lämnat under sin inskrivningsperiod.

Observationsvariablerna speglar intressevariablerna fullt ut.

1.2.3 Statistiska mått

De statistiska mått som används är främst antal och andelar för variabler som mäter den ersättningssökandes sökbeteende (geografiskt och yrkesmässigt).

1.2.4 Redovisningsgrupper

Statistiken särredovisas på sökandes juridiska kön (kvinnor och män).

Ytterligare nedbrytning görs på åldersgrupper, födelseregion, utbildningsbakgrund, funktionsnedsättning, kommuntyp, boendeform, civilstånd och antal barn i den sökandes hushåll.

1.2.5 Referenstider

Datamaterialet täcker in arbetssökande med arbetslöshetsersättning under första kvartalet 2019 respektive 2021. Hela inskrivningsperioden för samtliga arbetssökandes har ingått.

2 Tillförlitlighet

2.1 Tillförlitlighet totalt

Den totala tillförlitligheten påverkas av ett antal osäkerhetskällor som vi redogör för i nästa avsnitt.



2.2 Osäkerhetskällor

2.2.1 Precision i den maskinella tolkningen

Tabell 1 visar hur många aktiviteter som tolkats i den maskinella tolkningen samt hur stor andel av aktiviteterna som kunnat tolkas med hög precision, så kallade säkra tolkningar.

Tabell 1: Användbart material exklusive osäkra tolkningar och innan efterbearbetningar.

År	Antal aktiviteter som tolkats	Andel orter	Andel yrken
2019	4 335 716	95,8 %	79,0 %
2021	5 915 438	96,1 %	80,2 %

2.2.2 Uppskattning av osäkra tolkningars precision

Utöver det som kan tolkas säkert, kan det ibland vara möjligt för programvaran att gissa vilken ort eller vilket yrke den arbetssökande angett. Detta kan vara nödvändigt om det till exempel förekommer stavfel. Om vi kan gissa relativt korrekt kan vi använda en större del av materialet.

Tabell 2 visar andelen redovisade aktiviteter som kan ingå i granskningen om även de osäkra tolkningarna ingår.

Tabell 2: Antal och andel osäkra tolkningar

År	Antal aktiviteter med osäker tolkning	Andel orter	Andel yrken
2019	4 335 650	98,7 %	89,6 %
2021	5 915 438	98,7 %	90,1 %

2.2.3 Urval

Vi har granskat ett urval av osäkra tolkningar för att kunna bilda oss en uppfattning om hur bra den maskinella tolkningen förmår att hantera olika stavfel och liknande problem.

I datamaterialet för både 2019 och 2021 var ungefär 3 procent av orterna och 10 procent av yrkena osäkra tolkningar. För att verifiera kvaliteten på de osäkra tolkningarna har vi genomfört en manuell granskning av ett urval av de redovisade aktiviteter som inte kunnat tolkas med säkerhet. Urvalet var indelat i tre grupper, så kallade stratum:

- Stratum 1: Osäkra tolkningar där unika kombinationer förekommer 100 gånger eller mer.
- Stratum 2: Osäkra tolkningar där unika kombinationer förekommer mellan 10 och 99 gånger.

- Stratum 3: Osäkra tolkningar där unika kombinationer förekommer 9 gånger eller mer sällan.

I avsnitt 2.2.7 redovisar vi hur hanterat de vanligaste tolkningarna, de ord och meningar som finns i stratum 1.

Tabell 3 och 4 visar en uppskattning av hur stor andel av orter och yrken som var korrekta i den andra respektive tredje gruppen.

Tabell 3: Kontroll av osäkra tolkningar, orter.

Yrken	Antal	Granskade	Standardfel	Nedre gräns	Punktskattning	Övre gräns
Grupp 1	52 623	52 623	-	-	-	-
Grupp 2	41 854	21 346	0,003	0,751	0,756	0,762
Grupp 3	33 479	3 386	0,008	0,71	0,726	0,741

Tabell 4: Kontroll av osäkra tolkningar, yrken.

Yrken	Antal	Granskade	Standardfel	Nedre gräns	Punktskattning	Övre gräns
Grupp 1	203 617	203 617	-	-	-	-
Grupp 2	100 919	45 654	0,002	0,728	0,732	0,736
Grupp 3	155 050	3 793	0,008	0,523	0,539	0,554

Den automatiska tolkningen har lägre precision på de mer ovanliga yrkena i den tredje gruppen. I övrigt visar granskningen av tolkningarna att prestandan är förhållandevis god.

2.2.4 Ramtäckning

Inga täckningsfel har noterats.

2.2.5 Mätning

Varje unik kombination (ordpar) har granskats om den tolkade orten eller yrket verkar stämma överens med det som angetts i aktivitetsrapporten. De har antingen angetts som korrekta eller felaktiga.

2.2.6 Bortfall

Inga bortfall har förekommit i kvalitetssäkringen av tolkningar.



2.2.7 Bearbetning

Samtliga fel som förekom i det första stratumet i vår granskning av tolkningar korrigerades i materialet, eftersom de var vanliga. Ytterligare en efterbehandling av data har gjorts för att hantera vissa engelska yrkestitlar, som orsakar felaktigheter. Till exempel kan general manager misstas för den militära titeln general och *head of office* för *officer*. Dessa har korrigerats manuellt i efterhand, men det har inte varit möjligt att åtgärda samtliga, eftersom de uppvisar väldigt stor variation.

Efter de här korrigeringarna var det 57 000 orter och 216 000 yrken med osäkra gissningar kvar i materialet för 2019. Motsvarande siffra för 2021 var ungefär 81 000 orter och 296 000 yrken.

Tabell 5 visar hur stor andel av materialet som har kunnat användas i granskningen efter genomförda efterbearbetningar. Av samtliga redovisade yrken har 84,5 respektive 85 procent tolkats med tillräckligt hög säkerhet att de kunnat ingå i granskningen. Motsvarande siffra för orter är 97,4 respektive 97,3 procent. Antalet aktiviteter som redovisas i tabellen nedan är något lägre

Tabell 5: Träffsäkerheten i den maskinella tolkningen, andel säkra tolkningar i procent.

År	Antal aktiviteter som granskats	Orter	Yrken
2019	4 335 650	97,4 %	84,5 %
2021	5 915 438	97,3 %	85 %


2.3 Vad är det som inte går att tolka?

I datamaterialet för 2019 kunde vi inte översätta cirka 56 000 orter och 458 000 yrken till SSYK eller kommunkod. Det motsvarar ungefär 1 respektive 11 procent av det totala materialet. Motsvarande siffra för 2021 var 81 000 orter och 590 000 yrken. Det motsvarar ungefär 1 respektive 10 procent av det totala materialet.

Bland orterna som vi inte kunnat tolka automatiskt är det vanligt att den som rapporterat uppgett en mycket liten svensk ort, en utländsk ort, eller att arbetet finns i ett annat land. Vi har kunnat ta hänsyn till vissa av dessa, till exempel huvudstäder och större städer i nordiska grannländer. Det är däremot inte praktiskt möjligt att kunna tolka mer ovanliga, utländska orter. Det förekommer också att den sökande uppgett en gatuadress eller ett företagsnamn i stället för en ort. *Kungsgatan* är ett sådant exempel och eftersom ett sådant gatunamn finns i de flesta svenska städer, går det inte att uttyda någon ort.

Bland de yrken vi inte kunnat tolka rör det sig vanligen om

- allmänna termer som till exempel *alla*, *vilket som*, *valfritt* eller liknande,

- 
- allmänna benämningar, till exempel *personal*, *medhjälpare*, *anställd* och *allt i allo*
 - arbetsrelaterade termer, *annonserat* eller *tjänst*
 - olika typer av anställningsformer, till exempel *heltid*, *deltid*, *timanställning* eller *praktikplats*
 - handlingar av typen *lämnat in intresseanmälan*, *hört av mig om...* eller liknande,
 - mycket omfattande felstavningar.

Vi har också redan vid inläsning av materialet rensat för vissa termer, till exempel semestervikariat, säsongsarbete eller sommarjobb. Om den sökande inte skrivit in någon yrkestitel utöver den typen av termer, kvarstår ingen text som går att tolka och översätta till en SSK.

3 Aktualitet

3.1 Framställningstid

Framställningstiden för statistiken är ungefär 12 månader från att ramen fastställdes och till publicering av rapporten.

4 Tillgänglighet och tydlighet

4.1 Tillgång till statistiken och dokumentation

Statistiken publiceras i en rapport med tabeller och diagram, som finns tillgänglig på IAF:s webbplats www.iaf.se. Rapporten expedieras till Arbetsförmedlingen.

5 Jämförbarhet

5.1 Jämförbarhet över tid

Den här rapporten är i sitt utförande den första av sitt slag. Statistiken kan inte jämföras med statistik från tidigare rapporten från IAF.



5.2 Jämförbarhet mellan grupper

Uppgifterna från granskningen används för att bilda olika redovisningsgrupper. I regel kan statistiken jämföras mellan grupper, men jämförelser mellan vissa grupper med få observationer bör göras med viss försiktighet

Allmänna uppgifter

A Sekretess och personuppgiftsbehandling

Uppgifter som IAF samlar in för sin tillsyn omfattas av samma sekretess som uppgifterna omfattades av hos uppgiftslämnaren. Detta följer av 11 kap 1 § i offentlighets- och sekretesslagen (2009:400) (OSL). Uppgifter om sökande omfattas normalt av sekretess enligt 28 kap. 11 § i OSL hos Arbetsförmedlingen. Motsvarande sekretess blir gällande för uppgifterna hos IAF.

IAF behandlar personuppgifter enligt dataskyddsförordningen (GDPR) samt enligt lagen (2006:469) och förordningen (2008:936) om behandling av personuppgifter vid Inspektionen för arbetslöshetsförsäkringen.

B Bevarande och gallring

Reglerna för gallring av uppgifter som IAF behandlar för sin tillsyn framgår av 15 § lagen (2006:469) om behandling av personuppgifter vid Inspektionen för arbetslöshetsförsäkringen.

Personuppgifter som behandlas för tillsyn ska gallras två år efter utgången av det år när behandlingen av uppgifterna har avslutats i ett ärende hos den handläggande myndigheten. Om personuppgifter enligt första stycket behövs för statistiska eller vetenskapliga ändamål ska de gallras tio år efter det att uppgifterna avskildes för de ändamålen. Gallring av uppgifter i diarieförda ärenden framgår av IAF:s dokumenthanteringsplan

C Uppgiftsskyldighet

Arbetsförmedlingen är enligt 48 g § i lagen (1997:238) om arbetslöshetsförsäkring skyldig enligt att lämna de uppgifter som har betydelse för IAF:s tillsyn över ärenden om arbetslöshetsersättning.