

DET STATISTISKA REGISTRETS FRAMSTÄLLNING OCH KVALITET

Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier (LISA)

Produktkod

AM9901

Referenstid

2021

Innehåll

Det statistiska registrets framställning	4
F1 Det statistiska registrets sammanhang	4
F2 Utformning av framställningen	4
F2.1 Det statistiska registrets innehåll i stora drag	4
F2.2 Datakällor	4
F2.3 Tidsaspekter	5
F2.4 Insamlingsförfarande	5
F2.4.1 Datainsamlingsmetoder och leverantörer	5
F2.4.2 Mätning	6
F2.4.3 Brister i leveranser	6
F2.5 Bearbetningar med granskning	6
F2.5.1 Kodning	6
F2.5.2 Dubblettkontroller	6
F2.5.3 Rimlighetskontroller	6
F2.5.4 Imputeringar	6
F2.5.5 Modellbaserade beräkningar	7
F2.5.6 Jämförelser med andra register och datakällor	7
F3 Genomförande	7
F3.1 Kvantitativ information	7
F3.2 Avvikelser från utformningen	8
Det statistiska registrets kvalitet	8
K1 Relevans	8
K1.1 Ändamål och informationsbehov	8
K1.1.1 Registrets ändamål	8
K1.1.2 Registeranvändares informationsbehov	8
K1.2 Registrets innehåll	8
K1.2.1 Objekt och population	9
K1.2.2 Variabler	9
K1.2.3 Referenstider	10
K2 Tillförlitlighet	11
K2.1 Tillförlitlighet totalt	11
K2.2 Osäkerhetskällor	11
K2.2.1 Täckning	11
K2.2.2 Mätning	12
K2.2.3 Bortfall	13
K2.2.4 Bearbetningar	13
K2.2.5 Modellantaganden	13
K2.3 Preliminärt register jämfört med slutligt	13
K3 Aktualitet och punktlighet	13
K3.1 Framställningstid	13
K3.2 Frekvens	13
K3.3 Punktlighet	14

K4	Tillgänglighet och tydlighet.....	14
K4.1	Tillgång till registret	14
K4.2	Informationsspridning	14
K4.3	Dokumentation	14
K5	Jämförbarhet och sammanvändbarhet	14
K5.1	Jämförbarhet över tid	14
K5.2	Sammanvändbarhet med andra register.....	15
Allmänna uppgifter	15
U1	Sekretess och personuppgiftsbehandling.....	15
U2	Bevarande och gallring	15
U3	Uppgiftsskyldighet.....	15
U4	EU-reglering och internationell rapportering.....	15
U5	Historik	15
U6	Kontaktuppgifter	17

Det statistiska registrets framställning

F1 Det statistiska registrets sammanhang

LISA, Longitudinell integrationsdatabas för Sjukförsäkrings- och Arbetsmarknadsstudier, tillhör individfären i SCB:s registersystem. Registret framställs på individnivå och det är individobjektet som avgränsar populationen men det finns kopplingsmöjligheter till familj, hushåll, arbetsställe och företag.

LISA skapades ursprungligen för att bättre utnyttja befintliga registerdata avseende individers situation i livet, relation till arbetsmarknaden och arbetslivet samt ohälsa.

Idag är LISA en källa till den officiella integrationsstatistiken samtidigt som det är ett verktyg för snabb och effektiv hantering av mikrodatauppdrag.

LISA:s egenskaper skapar även möjlighet till att användas som urvalsram vid urvalsdragning. Uppgifterna från LISA kan också komplettera enkätundersökningar med hjälp- och bakgrundsinformation som kan länkas in i alla relevanta faser av en enkätundersökning.

F2 Utformning av framställningen

F2.1 Det statistiska registrets innehåll i stora drag

LISA består av tre varianter; individer, företag och arbetsställen. Varianten individer belyser individens relation till förvärvsarbete, arbetslöshet och sjukdom kopplade till ett stort antal bakgrundsvariabler. För samtliga personer som någon gång ingått i LISA-populationen finns även information om födelseland, in- och utvandring samt när personer har avlidit.

Varianterna företag och arbetsställen innehåller alla företag och arbetsställen med minst 1 förvärvsarbetande enligt den registerbaserade arbetsmarknadsstatistiken, RAMS.

Samtliga varianter är årliga.

F2.2 Datakällor

LISA baseras på data från både interna källor inom SCB och från andra myndigheter, externa källor.

Interna datakällor är Registret över totalbefolkningen (RTB), Registerbaserad arbetsmarknadsstatistik (RAMS), Yrkesregistret, Befolkningens utbildning (UREG), Registret över befolkningens studiedeltagande (RPU), Universitets- och högskoleregistret, Inkomst- och Taxeringsregistret (IoT) och Företagens ekonomi (FEK).

Externa datakällor är Arbetsförmedlingens sökanderegister och Försäkringskassans två datalager Store (STatistik Och REsultat) och MiDAS (Mikro Data för Analys av Socialförsäkringen).

Arbetsförmedlingen är dataleverantör för objektet individer som är inskrivna som arbetssökande eller deltagare i arbetsmarknadspolitiska program. Försäkringskassan är dataleverantör för objektet individer med vissa förmåner som ingår i socialförsäkringen. Leveranserna sker på elektronisk väg. Uppgifterna är inte insamlade för statistiska ändamål men ligger nära ett för SCB önskvärt läge.

Objekt mängden i individ-varianten är individer 15 år och äldre folkbokförda 31/12, dessa samlas in från SCB-registret Registret över totalbefolkningen (RTB). Det

primära objektet är individ men det finns kopplingsmöjligheter till ytterligare objekt såsom familj, hushåll, arbetsställe och företag.

Objekt mängden i företags- och arbetsställe-varianterna är samtliga företag och arbetsställen med minst 1 förvärvsarbetande enligt den registerbaserade arbetsmarknadsstatistiken, RAMS.

I LISA:s variabellista anges källa för varje variabel, [Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier \(LISA\) \(scb.se\)](#).

F2.3 Tidsaspekter

LISA framställs en gång per år och det finns en slutlig version. Produktionsomgången startar när Registret över totalbefolkningen (RTB) offentliggörs vilket i regel sker i februari. Därefter hämtas variabler in allteftersom de olika källorna blir tillgängliga. De externa källorna tas in senare vilket innebär att LISA då får med alla eventuella upprättningar som sker fram tills att uttaget av dessa källor görs. Den sista källan som individvarianten i LISA tar in uppgifter från är Yrkesregistret som finns tillgängligt 14 månader efter referensårets slut. Den sista källan som företagsvarianten i LISA tar in uppgifter från är Företagens ekonomi (FEK) som finns tillgängligt 16 månader efter referensårets slut.

F2.4 Insamlingsförfarande

F2.4.1 Datainsamlingsmetoder och leverantörer

Data från externa leverantörer sammanställs utifrån administrativa system och uppdateras löpande. Syftet med systemen är inte att framställa statistik och innehållet styrs av lagar och administrativa regler hos myndigheterna.

Arbetsförmedlingen har de uppgifter som framgår av bland annat förordning (2007:1030) med instruktion för Arbetsförmedlingen och förordning (2000:628) om den arbetsmarknadspolitiska verksamheten.

Information från Försäkringskassan styrs genom Socialförsäkringsbalken (2010:110), vilken reglerar de flesta ersättningar som betalas ut. Uppgifterna handläggs och administreras på respektive lokalkontor.

I Försäkringskassans datalager STORE fångas det stora mängder data från det administrativa källsystemet. Det rör sig således om mycket omfattande och komplexa data, då de är just administrativa och speglar handläggningsprocessen, med allt vad det innebär. Vill man använda utbetalningarna för att studera individen och följa vad som har hänt under exempelvis en sjukskrivning, behöver komplexiteten hanteras och tolkas för att härleda vad som gäller en specifik dag. Med god kännedom om handläggningsrutiner samt det regelverk och tekniska förutsättningar som handläggningen styrs av, går det dock att med tillräckligt god precision återskapa vad som har hänt i ett ärende. När detta är gjort kan data sedan också paketeras, så att de är tillrättalagda för analys av individen ur olika perspektiv. Processen att omvandla komplexa administrativa data till färdiga analysdata kallar Försäkringskassan för dataförädling. Resultatet av denna dataförädling har blivit MiDAS – MikroData för Analys av Socialförsäkringen. Sjuk- och rehabiliteringspenning, sjuk- och aktivitetsersättning samt föräldrapenning och skattad föräldraledighet hämtas från MiDAS. Övriga variabler i LISA med Försäkringskassan som datakälla kommer från datalagret STORE.

Insamling av informationen som används från andra statistik- och registerprodukter inom SCB dokumenteras av respektive produkter.

F2.4.2 Mätning

Primärt är det respektive myndighet som samlar in och registrerar uppgifterna i administrativa syften. Eftersom informationen samlas in för administrativa syften är det också myndigheters register och administrativa system som styr vad och hur det samlas in. Det sker ingen direktinsamling för LISA.

F2.4.3 Brister i leveranser

Vid leverans görs olika leveranskontroller. Kontroller görs mot föregående års indata och mot leverantörens egen publicerade statistik, i de fall det finns. I de fall brister upptäcks informeras leverantören för att motverka att det upprepas.

F2.5 Bearbetningar med granskning

F2.5.1 Kodning

Ingen kodning genomförs.

F2.5.2 Dubblettkontroller

Individpopulationen, folkbokförda 15 år och äldre 31/12, hämtas från RTB. Dubletter kan förekomma på ett flertal variabler men i RTB:s slutliga observationsregister redovisas en person endast en gång. Ytterligare dubblettkontroller görs kontinuerligt under framställningen av LISA, detta för att säkerställa att matchningar av olika data går rätt till.

Uppgifter i LISA som hämtas från från andra statistik- och registerprodukter inom SCB dokumenteras av respektive produkter.

F2.5.3 Rimlighetskontroller

Rimlighetskontroller görs under hela bearbetningsprocessen, både i indata och i det slutliga observationsregistret.

Vid granskning av data från Arbetsförmedlingen görs jämförelser mot Arbetsförmedlingens egen publicerade statistik. Efter bearbetning görs även rimlighetskontroller, vilket innebär att man på aggregerad nivå jämför t.ex. frekvenser med tidigare produktionsomgångar.

Vid granskning av data från Försäkringskassan görs jämförelser mot tidigare års indata. Granskning sker för samtliga ersättningar och jämförelse görs med flera tidigare år. Efter bearbetning görs även rimlighetskontroller för kombinationer av flera variabler. Detta innebär att man på aggregerad nivå kontrollerar att variabler som ska korrelera med varandra också gör det.

Jämförelser mot tidigare produktionsomgång görs på både indata såsom det ser ut vid leverans och även efter att bearbetningar av indata har gjorts.

F2.5.4 Imputeringar

Ingen imputering genomförs.

F2.5.5 Modellbaserade beräkningar

Arbetsförmedlingen är dataleverantör för objektet individer som är inskrivna som arbetssökande eller deltagare i arbetsmarknadspolitiska program. För varje person finns uppgift om sökandekategori, startdatum och ett eventuellt slutdatum.

Sökandekategori visar om personen är öppet arbetslös, sökande i program med eller utan stöd osv. En person kan förekomma flera gånger om personen har gått mellan olika sökandekategorier under året. I LISA räknas antal dagar fram per person och respektive sökandekategori utifrån datumen i materialet. De olika sökandekategorierna grupperas sedan i arbetslöshet, arbetsökande med förhinder, deltidsarbetslöshet, ”åtgärdsstudier”, ”åtgärdssysselsättning” och nystartsjobb. I LISA:s Bakgrundsfakta går det att läsa vilka sökandekategorier som ingår i respektive gruppering, [Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier \(LISA\) \(scb.se\)](https://scb.se/Longitudinell-integrationsdatabas-for-sjukforsaekrings-och-arbetsmarknadsstudier)

Antal flyttar under året inom riket och över kommungräns härleds utifrån uppgifter i RTB. Antal hemmavarande barn i familjen och hushållet fördelat efter ålder härleds utifrån uppgifter i RTB.

Terminer som en person har varit inskriven vid högskola/universitet härleds utifrån uppgifter i Universitets- och högskoleregistret.

I LISA görs flera summeringar av olika belopp, t.ex. summeras flera olika inkomster som avser studier till en variabel där summan är *inkomster föranledda av studier*. Ett annat exempel är att flera olika inkomster som avser föräldraledighet summeras till en variabel där summan är *inkomster föranledda av föräldraledighet*. I LISA:s Bakgrundsfakta beskrivs samtliga variabler och för de variabler som summerar flera andra variabler anges vilka inkomster som ingår, [Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier \(LISA\) \(scb.se\)](https://scb.se/Longitudinell-integrationsdatabas-for-sjukforsaekrings-och-arbetsmarknadsstudier)

F2.5.6 Jämförelser med andra register och datakällor

För att avgöra om nivån av utbetalningar och antal personer i uppgifterna som kommer från Försäkringskassan är rimliga granskas detta mot den statistik som Försäkringskassan publicerar. Jämförelser görs även mot motsvarande uppgifter i IoT.

Vid granskning av data från Arbetsförmedlingen görs jämförelser mot Arbetsförmedlingens egen publicerade statistik.

F3 Genomförande

F3.1 Kvantitativ information

Nedanstående tabell visar antalsuppgifter för de olika objekten.

Objekt	Variant	Antal 2021	Antal 2020
Individer	Individ	8 613 223	8 541 497
Familjer	Individ	5 626 379	5 572 108
Hushåll	Individ	4 831 811	4 776 239
Företag	Företag	558 549	557 907
Arbetsställen	Arbetsställen	632 815	632 578

För samtliga personer som någon gång ingått i LISA-populationen finns även information om födelseland, in- och utvandring samt när personer har avlidit.

Nedanstående tabell visar antalsuppgifter för de olika objekten i de varianter som avser individer i de årsberoende tabellerna.

Objekt	Antal 1990-2021
Individer som in- och /eller utvandrat	4 082 173
Samtliga individer - födelseland	17 397 021
Individer som avlidit	5 511 805

F3.2 Avvikelser från utformningen

Inga påtagliga avvikelser har gjorts från utformningen.

Det statistiska registrets kvalitet

K1 Relevans

K1.1 Ändamål och informationsbehov

K1.1.1 Registrets ändamål

LISA ligger till grund för den officiella integrationsstatistiken.

Registret kan användas som urvalsram vid urvalsdragning. Uppgifter från LISA kan också komplettera enkätundersökningar med hjälp- och bakgrundsinformation som kan länkas in i alla relevanta faser av en enkätundersökning.

Registret används inom uppdragsverksamheten på SCB. Uppgifterna används främst för mikrodatauppdrag som efterfrågas av forskare, myndigheter, m.fl.

Regeringskansliet använder mikrodata för att följa upp regeringens politik. LISA används även i uppdragsverksamheten som avser tabeller.

K1.1.2 Registeranvändares informationsbehov

Integrationsstatistiken belyser situationen för olika grupper av utrikes och inrikes födda inom områden som arbetsmarknad, utbildning, ekonomi, boende, hälsa och demokrati.

Mikrodataanvändare använder LISA bl.a. för att se dimensioner på etableringen i arbetslivet, anknytningen till arbetslivet, människors situation och hälsa under och efter arbetslivet, rörligheten på arbetsmarknaden, rörligheten ut och in på arbetsmarknaden samt på utträdet från arbetslivet.

K1.2 Registrets innehåll

Intressepopulationen för individer omfattar i de flesta fall samtliga personer som befinner sig i landet, folkbokförda och även personer utan uppehållsrätt, uppehållstillstånd eller motsvarande. Det kan också omfatta personer som arbetar i landet, och har rätt att befinna sig i landet, men som inte är folkbokförda.

Variabelinnehållet i LISA kan huvudsakligen grupperas in som:

- Demografiska variabler
- Familje- och hushållsvariabler
- Utbildningsvariabler
- Sysselsättningsvariabler
- Inkomstvariabler
- Socialförsäkringsvariabler
- Arbetslöshet- och arbetsmarknadsprogramvariabler
- Ekonomiska nyckeltal och grunddata samt personalstruktur på företag
- Personalstruktur på arbetsställe

K1.2.1 Objekt och population

Individregistrets målpopulation är personer 15 år eller äldre som ska vara folkbokförda i Sverige. Observationsobjekten utgörs av poster i RTB, där varje post motsvarar en person. Skillnaden mellan mål- och observationsobjekt är personer som ska vara folkbokförda i Sverige, men som av någon anledning inte är det (återfinns inte i RTB).

Företags- och arbetsställeregistrets målpopulation är företag och arbetsställen med minst 1 förvärvsarbete enligt RAMS och som finns registrerat i SCB:s företagsdatabas. Skillnaden mellan mål- och observationsobjekt är företag och arbetsställen som inte finns registrerade i SCB:s företagsdatabas och/eller som inte lämnar arbetgivardeklaration för anställda eller deklarerar som egna företagare.

K1.2.2 Variabler

Individregistret innehåller variabler som avser sysselsättning, utbildning, inkomster, socialförsäkring och arbetslöshet men även identiteter och kopplingsvariabler samt bakgrundsvariabler.

Typ av variabel	Exempel på variabel
Identitet, koppling	Personnummer, HushållsID, FamilieID, Organisationsnummer, Arbetsställesnummer
Bakgrund	Ålder, Kön, Civilstånd, Födelseland, Medborgarskap, Familjetyp, Familjeställning, Hushållstyp, Hushållsställning, Bostadskommun m.fl.
Utbildning	Utbildningsnivå, Utbildningsinriktning, Examensår, Examenskommun, Studiedeltagande, Registrerad vid högskola/universitet
Sysselsättning	Sysselsättningsstatus, Yrkeställning, Yrke, Arbetsställe, Institutionell sektor, Bransch m.fl. Samtliga variabler finns för det huvudsakliga jobbet samt största förvärvskälla, näst största förvärvskälla och tredje största förvärvskälla.
Inkomster	Inkomst av förvärvskälla, per månad och summerat för hela året (för det huvudsakliga jobbet samt största förvärvskälla, näst största förvärvskälla och tredje största förvärvskälla), Studiemedel och studiehjälp, Utbildningsbidrag, Lärlingsersättning, Grundläggande militärutbildning, ersättningar från Försäkringskassan, ersättningar från

	arbetslöshetskassa, Pension, Familjerelaterade inkomster, Disponibel inkomst
Socialförsäkring	Antal dagar med ersättning för föräldraledighet, sjukdom, arbetsskada, rehabilitering, sjukersättning, aktivitetsersättning
Arbetslöshet och arbetmarknadsprogram	Antal dagar med ersättning för arbetslöshet

Registret för företag innehåller variabler som avser näringsgren, sektor, utbetald lönesumma samt variabler som anger antal förvärvsarbetande personer på företaget grupperat på utbildningsnivå och kön. Registret innehåller även ekonomiska nyckeltal och ekonomisk grunddata från FEK.

Typ av variabel	Exempel på variabel
Identitet, koppling	Organisationsnummer
Karaktäristikor	Belägenhet, näringsgren, sektor
Personalstruktur	Antal förvärvsarbetande personer på företaget grupperat på utbildningsnivå och kön
Ekonomiska nyckeltal	Avkastning, omsättning, tillgångar
Ekonomiska grunddata	Lönekostnader, rörelseresultat, investeringar

Registret för arbetsställen innehåller variabler som avser näringsgren och utbetald lönesumma samt variabler som anger antal förvärvsarbetande personer på arbetsstället grupperat på utbildningsnivå och kön.

Typ av variabel	Exempel på variabel
Identitet, koppling	Organisationsnummer, CFAR-nummer, Arbetsställesnummer
Karaktäristikor	Belägenhet, näringsgren
Personalstruktur	Antal förvärvsarbetande personer på arbetsstället grupperat på utbildningsnivå och kön

Målvariabler i LISA är ofta variabler som beskriver individens aktivitet ett visst år för att med data från flera år se hur den eventuellt går mellan olika aktiviteter såsom t.ex. utbildning – arbete – föräldraledighet – arbetet - pension.

K1.2.3 Referenstider

LISA är uppbyggd med kalenderåret 2021 som mätperiod eller 31/12 2021 som mättpunkt. Denna mätperiod är giltig för flertalet uppgifter med undantag för:

- Samtliga variabler som är knutna till RAMS vilka är avsedda att mäta status i november månad.
- Samtliga LISA-specifika variabler som kommer från Försäkringskassan. Där registreras ersättningslag samt summeras dagar och belopp för utbetald ersättning den 1/1 året efter aktuellt år.
- Beståndsregistren i RTB som innehåller den folkbokförda befolkningen har referenstidpunkt 31 december år 2021. Den regionala indelningen avser här förhållanden den 1 januari 2022. I de demografiska händelseregistren i RTB (födslar, dödsfall, civilståndändringar etc.) är referensperioden 1 januari till 31 december år 2021 och där avser den regionala indelningen istället 1 januari 2021.

- Yrke (SSYK), för att undvika dubbelinsamling använder yrkesregistret sig av ett antal olika källor varvid insamlingstidpunkterna varierar. Se dokumentation om Yrkesregistret.
- Högsta avslutade utbildning, Utbildningsgrupp och Avslutningsår för högsta utbildning hämtas från Utbildningsregistret och avser avslutade utbildningar till och med vårterminen aktuellt år. Studiedeltagande HT avser höstterminen medan Registrering på högskolans grund-/forskarutbildning innehåller information om både vår- och hösttermin det aktuella året.
- Vissa variabler är inte knutna till en specifik årgång. Dessa variabler är Födelseland (eget, moderns och faderns samt olika grupperingar baserade på uppgift om födelseland), Land vid in-/utvandring, Datum för in-/utvandring och Dödsdatum.

K2 Tillförlitlighet

K2.1 Tillförlitlighet totalt

Registrets kvalitet bestäms huvudsakligen av de register som LISA hämtar uppgifter från samt de administrativa system som registret bygger på. Administrativa data från andra myndigheter anses hålla en god kvalitet.

Populationen hämtas från RTB där den sammantagna bedömningen är att den totala tillförlitligheten är god men där täckningen är den stora osäkerhetskällan.

För åren 1990-2010 saknar registret information om arbete för personer som bor i Sverige och arbetar i annat land. Detta gör att inkomst och sysselsättningsgrad i särskilt gränsregioner kan bli missvisande eftersom inkomst i annat land inte kommer med. Sedan 2011 finns inkomstuppgifter för personer som bor i Sverige och arbetar i ett annat nordiskt land. Information om inkomster för jobb för personer som bor i Sverige och arbetar i ett land utanför Norden, saknas dock.

K2.2 Osäkerhetskällor

Sammantaget bedöms varje register som är en källa till LISA vara mycket tillförlitligt. Viss osäkerhet finns vad gäller övertäckning och undertäckning i populationen.

K2.2.1 Täckning

Personer som borde vara folkbokförda men som inte är det orsakar undertäckning, medan personer som är folkbokförda trots att de inte borde vara det orsakar övertäckning. Täckningsbristerna uppstår till följd av brister i rapporteringen av födselar, dödsfall, invandring och utvandring.

Övertäckning

Brister i rapporteringen av dödsfall och utflyttning resulterar i övertäckning. Den största delen består av personer som utvandrar men där utflyttning inte anmäls. Flera studier tyder på att folkbokföringen och därmed RTB innehåller ett betydande antal personer som inte längre bor i landet. Övertäckningen bedöms bidra mest till osäkerheten. Nedan visas skattningen av antal personer i övertäckningen 2012-2019¹.

¹ För mer information om modellspecifikationen, se Skatteverket (2020), Skatteverkets modell för uppskattning av folkbokföringsfelet: modellspecifikation i årsredovisningen 2019.

År	Antal - övertäckning
2012	65 400
2013	67 100
2014	66 900
2015	61 400
2016	62 100
2017	64 700
2018	68 100
2019	73 400

Undertäckning

Brister i rapporteringen av födslar och inflyttning resulterar i undertäckning. Undertäckningen uppstår också om en person flyttar till Sverige och enligt lagar och förordningar ska folkbokföra sig men inte gör det. Det finns dock starka incitament för en person att folkbokföra sig eftersom många rättigheter som t.ex. möjligheten att få olika typer av bidrag hänger ihop med att man är folkbokförd. Detta talar för att undertäckningen på grund av att personer inte folkbokför sig är liten. Nedan visas skattningen av antal personer i undertäckningen 2012-2019².

År	Antal - undertäckning
2012	5 500
2013	6 000
2014	6 700
2015	7 700
2016	7 900
2017	11 800
2018	12 500
2019	17 000

K2.2.2 Mätning

LISA inhämtar uppgifter från flera olika register. Mätfel kan uppstå i produktionen av varje enskilt register.

Arbetsförmedlingen

Uttaget görs i december för uppgifter som avser föregående år. Uppgifterna utgörs av alla personer som var inskrivna hos Arbetsförmedlingen under året och visar antal dagar en person har varit inskriven inom olika sökandekategorier.

I LISA:s Bakgrundsfakta går det att läsa om bearbetningar av data från Arbetsförmedlingen, [Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier \(LISA\) \(scb.se\)](https://www.scb.se/longitudinell-integrationsdatabas-for-sjukforsakrings-och-arbetsmarknadsstudier).

Försäkringskassan

Uttaget görs i december för uppgifter som avser föregående år. Uppgifterna hämtas främst från Försäkringskassans system MiDAS och avser utbetalade årsbelopp och

² För mer information om modellspecifikationen, se Skatteverket (2020), Skatteverkets modell för uppskattning av folkbokföringsfelet: modellspecifikation i årsredovisningen 2019.

antal dagar för händelser som skett under föregående år. De uppgifter som inte finns i MiDAS-systemet hämtas från Försäkringskassans system STORE.

I LISA:s Bakgrundsfakta går det att läsa mer om MiDAS och STORE, [Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier \(LISA\) \(scb.se\)](#).

Inga formella studier har gjorts för att närmare kvantifiera de osäkerheter som uppstår till följd av mätfel men om de finns så bedöms de vara små.

K2.2.3 Bortfall

Information om bortfall i andra statistik- och registerprodukter inom SCB dokumenteras av respektive produkter.

K2.2.4 Bearbetningar

Dataregistrering sker i de olika myndigheternas administrativa datasystem. Då uppgifterna bereds, dataregistreras och bearbetas genomförs både manuella och maskinella kontroller som kan ge upphov till rättelser och korrigeringar innan det kommer till SCB. På SCB granskas och bearbetas de olika källorna och sedan sker ihopkoppling av flera material med olika populationer vilket i sig innebär en osäkerhet.

K2.2.5 Modellantaganden

Underhållsbidrag

Det finns inga administrativa data om underhållsbidrag, detta uppskattas med en modell i IoT. Modellen beskrivs i dokumentationen som avser det statistiska registret IoT eller i ”Resultatrapport Översyn av den disponibla inkomsten i den totalräknade inkomststatistiken” 2013-12-16 utfärdad av Johan Lindberg.

Föräldrapenning, totalt antal dagar

Tills dess att barnet är 18 månader ha föräldern rätt till ledighet med eller utan föräldrapenning. På så sätt är det möjligt att förlänga föräldraledigheten genom att ta ut ledighet utan ersättning. De obetalda dagarna är skattade dagar, skattningen görs av Försäkringskassan. Skattningen beskrivs i ”MiDAS Föräldrapenning - Dokumentation av föräldrapenningdata och skattad föräldraledighet i MiDAS”, PM 2020:2 hos Försäkringskassan.

K2.3 Preliminärt register jämfört med slutligt

Endast slutligt observationsregister tas fram.

K3 Aktualitet och punktlighet

K3.1 Framställningstid

Individvarianten är helt komplett cirka 15 månader efter mättidpunkt. Arbetsställevarianten är klar ca 13 månader efter mättidpunkt och Företagsvarianten är helt komplett ca 17 månader efter mättidpunkt. Varianterna görs tillgängliga tidigare men där det då saknas vissa uppgifter.

K3.2 Frekvens

Registret byggs ut med en ny årgång varje år.

K3.3 Punktlighet

En första version av individvarianten avseende årgång 2021 var planerad att vara klar under februari 2023, den fanns tillgänglig den 20 februari 2023. Den slutliga versionen av individvarianten inkl. yrkesuppgifter som saknades i den första versionen, fanns tillgängligt den 9 mars 2023.

K4 Tillgänglighet och tydlighet

K4.1 Tillgång till registret

Behörighet ges via BBQ på respektive årgång och på de variabler som behörigheten kräver. Slutliga observationsregister finns på server LISA.PROD.SQL.

Uppgifter ur LISA kan lämnas ut för forskning och statistikändamål. Varje förfrågan till SCB om utlämnande av mikrodata sekretessprövas. SCB lyder under statistiksekretessen som återfinns i 24 kap 8 § Offentlighets- och sekretesslagen (2009:400). Vid automatiserad behandling av personuppgifter gäller även reglerna i dataskyddsförordningen.

Efter godkänd prövning kan forskare och utredare få tillgång till avidentifierade mikrodata för egna bearbetningar i t.ex. MONA.

En skriftlig begäran om de uppgifter som önskas skickas till SCB. Mer information om hur man begär uppgifter samt beställningsblankett finns på SCB:s webbplats, www.scb.se. När uppgifterna ska användas i forskning ser SCB gärna att en projektbeskrivning bifogas. Om forskningsprojektet har prövats av en regional etikprövningsnämnd bör även ansökan och beslut bifogas. I de fall som uppgifterna kan lämnas ut lämnas en offert som innehåller en beskrivning av SCB:s åtagande, leveranstidpunkt och pris.

K4.2 Informationsspridning

Användare internt informeras genom registerportalen, särskilda mailutskick samt vid utbildningstillfällen.

K4.3 Dokumentation

Variabelförteckningar och datamängdsbeskrivningar återges i MetaPlus, som bland annat finns på SCB:s webbplats www.metadata.scb.se.

På LISA:s produktsida finns variabellistor och ett omfattande Bakgrundsfakta som i detalj beskriver samtliga variabler och händelser sedan 1990, [Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier \(LISA\) \(scb.se\)](http://Longitudinell%20integrationsdatabas%20f%C3%B6r%20sjukf%C3%B6rs%C3%A4krings%20och%20arbetsmarknadsstudier%20(LISA)%20(scb.se)).

K5 Jämförbarhet och sammanvändbarhet

K5.1 Jämförbarhet över tid

Nedan är en sammanfattning av förändringar som skett i årgång 2021.

Inkomstpensionstillägg

Personer som har arbetat ett långt arbetsliv i Sverige men med låg lön kan ha rätt till inkomstpensionstillägg. Tillägget kan ges som ett tillägg till allmän pension. Den baseras främst på storleken på den inkomstgrundade pensionen, men även på hur många år med pensionsgrundande inkomst personen har tjänat in i Sverige.

K5.2 Samanvändbarhet med andra register

Det finns sammanvändbarhet mellan LISA och RTB. LISA innehåller alla folkbokförda från RTB 31 december, 15 år och äldre.

Det finns sammanvändbarhet mellan LISA och Inkomst- och taxeringsregistret (IoT). IoT innehåller alla personer som finns i LISA, skillnaden är att IoT innehåller alla skattskyldiga personer samt resterande folkbokförda personer och dödsbon. Olika definitioner gäller för inkomster som hämtas från IoT och inkomster från Försäkringskassan gällande sjukförsäkring. Inkomsterna i LISA från Försäkringskassan avser händelseperioden, medan variabler som hämtas från IoT avser utbetalningsperioden, dvs. inkomståret.

Det finns sammanvändbarhet mellan LISA och STATIV. STATIV innehåller alla personer som finns i LISA, skillnaden är att STATIV innehåller samtliga folkbokförda 31 december. Olika definitioner gäller för uppgifter om arbetslöshet i LISA och STATIV. Antal dagar i arbetslöshet och/eller arbetmarknadsprogram i LISA avser hela året och en person kan ha dagar i olika sökandekategorier. I STATIV är det fokus på personens senaste sökandeperioden under året.

Allmänna uppgifter

U1 Sekretess och personuppgiftsbehandling

I myndigheternas särskilda verksamhet för framställning av statistik gäller sekretess enligt 24 kap. 8 § offentlighets- och sekretesslagen ([2009:400](#)).

Vid behandling av personuppgifter, dvs. information som direkt eller indirekt kan hänföras till en person som är i livet, gäller lagen ([2001:99](#)) och förordningen ([2001:100](#)) om den officiella statistiken samt EU:s dataskyddsförordning ([2016/679](#)).

Aidentifierade mikrodata från LISA kan lämnas ut till forskning och statistikändamål efter sekretessprövning om SCB bedömer att det finns stöd för att behandla uppgifterna.

U2 Bevarande och gallring

Bevarandebehov är under utredning.

U3 Uppgiftsskyldighet

Uppgiftsskyldighet föreligger enligt förordning om den officiella statistiken ([2001:100](#)), vilket innebär att myndigheter ska lämna de uppgifter som behövs för framställning av statistiken.

U4 EU-reglering och internationell rapportering

Ingen EU-reglering finns för LISA.

U5 Historik

I mitten av nittioåret byggdes LOUISE-registret på SCB där fokus var på utbildning, inkomst och sysselsättning för forskning rörande arbetsmarknaden. LOUISE byggdes för att göra data mer tillgänglig och flexibel för användaren.

I början av tjugohundratalet kom forskning kring ohälsotalen att efterfrågas i allt större omfattning då sjukskrivningstalen nådde sina högsta nivåer under 2001. Efter ett slutbetänkande från "Utredning om Analys av Hälsa i Arbete" togs hösten 2003 ett

beslut om att SCB skall upprätta LISA-registret. På SCB startades ett uppbyggnadsarbete av ett nytt register (LOUISE var grunden) där syftet var att möjliggöra analyser och utvärderingar inom ohälsområdet. LISA byggdes upp tillsammans med Riksförsäkringsverket (numera Försäkringskassan) och Verket för innovationssystem (VINNOVA).

I LISA:s bakgrundsfakta beskrivs **förändringar utförligt på variabelnivå**, [Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier \(LISA\) \(scb.se\)](#). Nedan är en kortfattad beskrivning av några större förändringar.

Demografiska variabler

Från och med 1 januari 2016 sker inte längre folkbokföring på församling, utan på kommun. Av såväl kulturella och historiska skäl som för statistik- och forskningsändamål är det ändå viktigt att bevara församlingsindelningen från den tid indelningen var stabil. Därför har distrikt införts. Distrikten ska motsvara församlingsindelningen den 31 december 1999. Folkbokföring sker inte på distrikt, men distriktstillhörighet registreras på individnivå i folkbokföringsdatabasen.

Familje- och hushållsvariabler

Familj identifieras med hjälp av familjeidentitet. Uppgifterna för perioden 1990–1997 skapades år 2000 med hjälp av register. För perioden efter 1997 tas familjeuppgifterna fram i samband med att man bildar RTB-årgången. Det finns alltså en konsistens mellan innehållet i RTB för ett visst år och familjeuppgifterna efter 1997. För perioden 1990–1997 baseras uppgifterna om familj på det uppdaterade flergenerationsregistret som färdigställdes till sommaren år 2000.

Sedan 2010 folkbokför Skatteverket personer boende i flerbostadshus på lägenhet och från och med 2011 får RTB in identiteter på lägenhet, adress och fastighet via aviseringarna. Genom att detta sker kan personer och bostäder kopplas ihop och bilda hushåll.

Utbildningsvariabler

Fr.o.m. version 2008 av UREG innehåller registret även uppgift om högsta utbildning för personer äldre än 74 år. I LISA har uppgift om högsta utbildning för personer äldre än 74 år kompletteras även för åren 1990-2007.

Våren 2019 fastställde SCB den nya versionen av SUN, benämnd SUN 2020. I Utbildningsregistret implementeras SUN 2020 för läsåret 2018/19 med publicering våren 2020. Utöver en uppdaterad struktur för inriktningar har innehållsbeskrivningar för såväl nivåer som inriktningar förbättrats och förtydligats med fler exempel på utbildningar som ingår eller inte ingår i respektive kod.

Sysselsättningsvariabler

Den officiella avgränsningen av förvärvsarbetande i RAMS redovisas som variabeln Sysselsättningsstatus i Sysselsättningsregistret. En ny metod för att avgränsa förvärvsarbetande infördes 1993, 2004, 2011 och 2019.

Standarden för svensk näringsgrensindelning, SNI, används för att klassificera företag och arbetsställen efter vilken verksamhet de bedriver. SNI 2007 ersatte den tidigare versionen kallad SNI 2002 som var resultatet av en mindre revision av SNI 92. Gällande standard före SNI 92 benämndes SNI 69. En ny version av SNI arbetas fram i samband med en internationell revidering av näringsgrensindelningar. Någon revision finns för närvarande inte inplanerad.

Standard för svensk yrkesklassificering, SSYK, är ett system för att gruppera individens yrken eller arbetsuppgifter. Det finns två versioner av SSYK, SSYK 96 och SSYK 2012. Från och med 2014 används SSYK 2012 inom den officiella statistiken. Den är en uppdatering av SSYK 96, som den också ersätter.

I enlighet med gällande säkerhetsskyddslagstiftning har SCB genomfört en säkerhetsskyddsanalys som kartlagt om det vid myndigheten finns information som måste skyddas i enlighet med lagstiftningen. Analysen har bland annat slagit fast att tillgången till information gällande vilka personer som arbetar som poliser och militärer enligt standarden för yrkesklassificering (SSYK 2012) måste begränsas. Det innebär att för personer med yrkeskod som avser militär och polis har en omkodning gjorts så att dessa personer istället ingår i gruppen med okänt yrke. Detta gäller från och med årgång 2020.

Inkomstvariabler

Från och med referensår 2019 finns månatliga arbetsgivardeklarationerna på individnivå (AGI), dessa ersätter de årliga kontrolluppgifterna (KU).

Det reformerade pensionssystemet, den allmänna pensionen, trädde i kraft 1999. Reformeringen av ålderspensionen medförde också ett behov av att reformera efterlevandepensionen och den dåvarande förtidspensionen. Detta arbete genomfördes under åren därpå och 2003 var samtliga delar av pensionsreformen på plats.

Från 2020 har registret kompletterats med inkomster intjänade i annat nordiskt land, registret har kompletterats från 2011.

Socialförsäkringsvariabler

Inför framställningen av LISA 2015 beslutades, i samråd med Försäkringskassan, att gå över från STORE till MiDAS som källa för viss ohälsodata. För att undvika tidsseriebrott lades variabler med MiDAS som källa även till för samtliga år bakåt i tiden.

Under 2020 införde regeringen åtgärder på socialförsäkringsområdet till följd av coronapandemin. Åtgärderna syftar till att minska smittspridningen, minska belastningen på sjukvården, stärka den ekonomiska tryggheten samt lindra konsekvenserna för arbetsgivare.

U6 Kontaktuppgifter

Information om det statistiska registret	
Kontaktperson	Sara Ekmark
E-post	sara.ekmark@scb.se
Telefon	010-479 64 34

Mikrodatautlämnande	
E-post	mikrodata@scb.se
Telefon	010-479 50 00 (Statistikservice)

Statistikuppdrag	
E-post	uppdrag@scb.se
Telefon	010-479 50 00 (Statistikservice)