

Kartläggning av företagspopulationen inom life science

- Metodbeskrivning

04.05.23

For information on obtaining additional copies, permission to reprint or translate this work, and all other correspondence. Please contact Amsterdam Data Collective, Stockholm office.



Innehållsförteckning

1	Inledning	4
2	Arbetsprocessen	5
3	Datakällor	6
3.1	Identifiering av företagspopulation	6
3.2	Beräkning av företagsdata	8
4	Modellbeskrivning	11
4.1	Identifikation av företag	11
4.2	Selektering av företag	12
4.3	Klassificering av företag	14
4.4	Beräkning av data	16
5	Känslighetsanalys	18
5.1	Om statistik för life science-sektorn	18
5.2	Avväganden i utvecklingsarbetet	18
5.3	Datakällor och -variabler	20
6	Rutiner vid uppdatering av statistiken	22
6.1	Insamling av datakällor	22
6.2	Framtagning av företagslistor via öppna data	23
6.3	Kvalitetskontroll	23
6.4	Framtagning av statistik i MONA	24
6.5	Rapportsammanställning	24
7	Rekommendation för underhåll och vidare utveckling av metoden	25
7.1	Löpande kvalitetssäkring	25
7.2	Klassificering av marknadsaktiviteter	25
7.3	Tillägg av ytterligare delsektorer och breddning av sektorn	26
8	Appendix	28
8.1	Exkluderande SNI-koder	28
8.2	SNI-koder för att identifiera försäljare/grossister och konsulter	30
8.3	Nyckelord för identifiering	31
8.4	Nyckelord för klassificering	35
8.5	Exkluderade nyckelord	38
8.6	Variabellista från SCB-beställning	39



1 Inledning

Amsterdam Data Collective¹ (ADC) har på uppdrag av och tillsammans med Vinnova utvecklat en modell för att identifiera företagspopulationen inom life science-sektorn i Sverige. Arbetet har tagit sin utgångspunkt i tidigare utförda kartläggningar av företagspopulationen genomförda av Tillväxtanalys och Vinnova. Modellutvecklingen har vidare tagit hänsyn till de rekommendationer som Tillväxtanalys lämnat inom ramen för sitt regeringsuppdrag att utreda metoden för att följa Life Science företag (Regeringsuppdrag N2020/02707).

Målsättningen i arbetet har varit att utveckla en modell som med en hög grad av automatisering kan identifiera företagspopulationen i life science-sektorn. Detta för att säkerställa långsiktighet i projektet och i förlängningen möjliggöra för återkommande kartläggningar och statistik för sektorn.

Arbetet har genomförts i nära samarbete mellan ADC och Vinnova. Vinnova har varit direkt involverad i samtliga delar av projektet och olika inriktningsbeslut har tagits gemensamt och löpande. Detta dokument syftar till att beskriva den modell som utvecklats för att understödja överlämning av resultaten till Vinnova, samt fungera som ett underlag för att drifva metoden i framtiden. Utöver detta dokument omfattar leveransen av arbetet även bilagor i form av Excel-listor samt ett datapaket i R som innehåller koder för databearbetning.

¹ Tidigare Damvad Analytics



2 Arbetsprocessen

Arbetet har genomförts i tre övergripande faser. I den första utforskande fasen sammanställdes existerande kunskap på området och olika möjliga datakällor identifierades. Genom workshops och dialogmöten med en av Vinnova sammansatt referensgrupp samlades olika perspektiv in och avvägande gällande metodval och önskade resultat diskuterades. Arbetet resulterade i en av ADC framtagen genomförandeplan som utgjorde grunden för det fortsatta utvecklingsarbetet.

Ett viktigt inriktningsbeslut som togs under arbetets första fas och på begäran av Vinnova var att arbetet med att kartlägga företagspopulationen, så väl som segmentering i underkategorier, helt skulle baseras på öppna datakällor. Detta innebar i praktiken att information från SCB:s registerdata inte användes i arbetet med att kartlägga företagspopulationen. Huvudanledningen till varför detta beslut togs var för att möjliggöra identifiering och således kvalitetssäkring av företagen. Av sekretesskäl är företagen pseudonymiserade i SCB:s registerdata vilket innebär att den identifierade företagspopulationen inte hade kunnat kvalitetssäkras.

I projektets andra fas utvecklades själva modellen i nära samverkan mellan Vinnova och ADC. Arbetet utgick från den genomförandeplan som tagits fram under arbetets första fas. Arbetet strukturerades i sprintar som genomfördes gemensamt av Vinnova och ADC. Mellan sprintarna genomfördes utvecklingsarbete av både ADC och Vinnova. Arbetet genomfördes iterativt och inriktningsbeslut togs löpande och gemensamt.

Arbetet i utvecklingsfasen genomfördes i tre övergripande steg vilka syftade till att 1) identifiera, 2) selektera och avslutningsvis 3) klassificera företag. Respektive steg beskrivs kortfattat nedan:

- Det första steget syftade till att identifiera relevanta företag: Detta inbegrep att identifiera och utvärdera möjliga datakällor som kunde bidra till att identifiera företag i life science-sektorn. Steget resulterade i en bruttolista på cirka 11 000 företag som bedömdes som relevanta för vidare analys. De datakällor som användes presenteras närmare senare i denna rapport.
- Det andra steget syftade till att utifrån den framtagna bruttolistan selektera de företag som bedömdes tillhöra life science-sektorn. Arbetet genomfördes iterativt och baserades på olika metadata som härstammade från de identifierade datakällorna. Arbetet inkluderade både inkluderingskriterier så väl som exkluderingskriterier. Resultaten från arbetet kvalitetssäkrades löpande. Resultatet från detta steg var en lista med drygt 4 000 företag som bedömdes tillhöra life science-sektorn givet de definitioner som satts upp.
- I det avslutande steget kategoriserades företagspopulationen som selekterats i steg 2. Arbetet utgick från en kategorisering som utvecklats tillsammans med Vinnova där branschföreträdare fått möjlighet att komma med inspel. Klassificeringen genomfördes utifrån metadata kopplad till identifierade och selekterade företag som stammade från de identifierade datakällorna. Arbetet resulterade i en metod för klassificering av företagen i olika delsektorer.

Efter att företagspopulationen selekterats och klassificerats kunde arbetets avslutande fas inledas. Denna del av arbetet har syftat till att berika den identifierade företagspopulationen med officiella företags- och individdata från SCB i linje med den data som tidigare har sammanställts av Tillväxtanalys samt tydligt dokumentera hur dessa data tas fram genom SCB:s plattform för tillgängliggörande av mikrodata MONA (Microdata Online Access). Berikning av data innebär således att vi tillfört information om företagen för att kunna beskriva företagen och dess utveckling över tid. Slutligen har ADC tagit fram rekommendationer för underhåll och vidare utveckling av metod.



3 Datakällor

I detta kapitel beskrivs de öppna datakällor som används till identifiering av företagspopulationen samt officiella data från SCB som används till att berika företagsdata.

3.1 Identifiering av företagspopulation

Ett stort antal öppna datakällor har utvärderats för att kartlägga företagspopulationen inom life science. I denna identifieringsfas har det för arbetsprocessen varit kritiskt att datakällorna har varit öppna. Detta för att genomgående under arbetet kunna kvalitetssäkra företagen som identifieras och kunna jämföra utfallet av olika identifierings-, selekterings-, och klassificeringskriterier. Detta hade inte i samma utsträckning varit möjligt om denna fas genomförts med pseudonymiserade data. Datakällorna som används har således delvis valts ut på grund av att de är öppna. Utöver detta har de valts ut baserat på deras komplementära funktion. Medlemslistor från branschföreningar, medicintekniska- och läkemedelsregister samt start-updatabaser bidrar med hög träffsäkerhet medan näringslivsregister och SNI-koder bidrar med hög täckningsgrad. Dessa datakällor har sedan kombinerats och används i en iterativ kvalitetssäkringsprocess.

Tabellen nedan sammanfattar de datakällor som har använts i arbetet med att identifiera, selektera och klassificera företagen inom life science-sektorn. Respektive datakälla beskrivs därefter mer detaljerat i löptext.

TABELL 3.1

Datakälla	Beskrivning	Ursprung
Medlemslistor från branschföreningar	Listor över medlemmar i branschföreningarna LIF, Swedish Medtech, Labtech, Sweden BIO.	Inrapporterad information från branschföreningar.
Företagsinformation	Utdrag ur näringslivsregistret omfattande SNI-koder, verksamhetsbeskrivningar, omsättning och antal anställda.	Årsredovisningar hos Bolagsverket via tjänsten Proff.
EUDAMED	Utdrag av svenska aktörer registrerade i EU-databasen EUDAMED över registrerade medicintekniska produkter.	Hämtad från EUDAMED:s webbplats genom sökning efter ekonomiska operatörer med säte i Sverige. (länk)
Nationellt produktregister för läkemedel (NPL)	Utdrag av aktörer registrerade i Nationellt produktregister för läkemedel (NPL).	Data hämtas från Läkemedelverkets webbsida där senaste versionen av NPL kan hämtas vilken innehåller en sammanställning av organisationer. (länk)
Dealroom	Utdrag över företag med taggade med "Health" som industri från startupdatabasen Dealroom.	Dealroom via Vinnovas access.



Ansökningsdata

Data över aktörer som ansökt om stöd från Vinnova och andra aktörer i det svenska näringslivsstödjande systemet.

Vinnova och EU horisont.

Ytterligare en datakälla som har används är listan över företag som identifierades som life science-företag av Tillväxtanalys 2018.² Totalt återfinns 3356 företag på denna lista som togs fram under 2017. I detta projekt används företagslistan från Tillväxtanalys enbart i kvalitetssäkrande syfte och således inte till att identifiera företag.

3.1.1 Medlemslistor branschföreningar

Branschföreningarna inom life science (LIF, Swedish Medtech, Labtech, Sweden BIO) är en viktig informationsmängd för att kartlägga företagspopulationen. Listorna samlas in direkt från branschföreningarna och innehåller medlemskapsinformation för år 2022. Listorna innehåller företag som är medlemmar i respektive branschförening. Listorna har hög träffsäkerhet då en stor andel av medlemmarna kan definieras som tillhörande life science-sektorn. Underlaget utgör även en viktig datamängd i klassificeringen av företag tillhörande olika delsektorer. Samtidigt som medlemslistorna har hög träffsäkerhet är täckningsgraden begränsad då endast en mindre del av den samlade företagspopulationen inom life science är medlemmar i en branschförening. Att medlemskapsinformation endast använts för innevarande år innebär en osäkerhet i data längre bak i tiden i de statistiska sammanställningarna. Med andra ord saknas medlemskapsdata för tidigare år. Detta innebär exempelvis att företag som tidigare varit medlemmar men gått i konkurs eller blivit uppköpta innan 2022 inte kommer att identifieras genom medlemslistor från branschföreningar.

3.1.2 Företagsinformation

Företagsinformation för bokslutsåret 2021 hämtas från årsredovisningar som registrerats hos Bolagsverkets via tjänsten Proff. Årsredovisningar sammanställer information om samtliga svenska organisationers räkenskaper och förvaltning för ett räkenskapsår. Variabler som används i arbetsprocessen omfattar bland annat SNI-koder, verksamhetsbeskrivningar, antal anställda och omsättning. Från dessa data identifieras företag genom sökningar efter företag med utvalda SNI-koder. Vidare används SNI-koder i kombination med verksamhetsbeskrivningar för att selektera och klassificera bolag. Se kapitel 4 för beskrivning av urval av SNI-koder.

3.1.3 EUDAMED

EU-databasen EUDAMED är innehåller information över registrerade medicintekniska produkter. Produkter som omfattas av såväl MDR (EU-förordningen om medicintekniska produkter, trädde i kraft 2021) som IVDR (EU-förordningen om medicintekniska produkter för in vitro-diagnostik, trädde i kraft i maj 2022) ska registreras i denna databas. Registret har uppgifter om de aktörer som har registrerade produkter i databasen och det är denna information som används i arbetet. Information från EUDAMED används i samtliga steg (identifiera, selektera och klassificera) av processen. Databasen har hög träffsäkerhet gällande företag som tar fram

²

https://www.tillvaxtanalys.se/download/18.62dd45451715a00666f194a2/1586366154910/pm_2018_06_Den%20svenska%20life%20science-industrins%20utveckling.pdf



medicintekniska produkter samtidigt som den är begränsad genom att den endast innehåller information om medicintekniska produkter som registrerats sedan databasen upprättades från och med 2021.

3.1.4 Nationellt produktregister för läkemedel (NPL)

Nationellt produktregister för läkemedel, NPL, är ett register med information om samtliga läkemedel som är godkända i Sverige. Registret upprätthålls av Läkemedelsverket. Registret har uppgifter om de aktörer som har registrerade produkter i databasen och det är denna information som används i arbetet. NPL innehåller inte uppgifter om pris från Tandvårds- och läkemedelsförmånsverkets (TLV) eller uppgifter från läkemedelsindustrin som till exempel status för försäljning och varunummer. Dessa uppgifter finns i VARA som är E-hälsomyndighetens system. Information från NPL används i samtliga steg (identifiera, selektera och klassificera) av processen. NPL används för att identifiera, selektera och klassificera läkemedelsföretag.

3.1.5 Dealroom

Dealroom.co är främst en dataleverantör med information om startups, tillväxtföretag och tekniska ekosystem. Dealrooms databas innehåller företag från hela världen, men främst tillväxtintensiva och innovativa företag. Data hämtas via en AI-lösning och representanter för företag och investerare kan själva lägga upp information. Databasen är inte att se som heltäckande men är användbart som källa för företag med intresse att ta in kapital, samt för nystartade företag. Dealroom används därför inte ensamt, utan i kombination med andra datakällor i kartläggningen. Dealroom arbetar också med många av världens mest framstående investerare, entreprenörer och statliga organisationer för att ge transparens, analys och insikter om riskkapitalaktivitet.

Med användning av Dealroom som datakälla kunde vi fånga in ytterligare cirka 70 företag som inte återfunnits i bruttolistan vid användning av de övriga datakällorna.

3.1.1 Vinnova & EU horisont ansökningar

Data från ansökningar om projektbidrag till Vinnova och EU horisont har initialt undersökts som källa för att identifiera företag i life science-sektorn. Här är företagen hämtade direkt från Vinnovas datalager.

För företag som identifierats genom ansökningar till Vinnova var tanken att inkludera samtliga aktiva svenska aktiebolag inom de primära och sekundära SNI-koderna som sökte finansiering hos Vinnova för minst ett life science-projekt. Ett life science-projekt definieras som ett projekt inom Vinnovas område Hållbar precisionshälsa med minst tre life science-relaterade nyckelord i projektbeskrivningen (samma nyckelordlista som använts identifiering av företagspopulation).

För ansökningar till EU horisont inkluderas alla aktiva svenska aktiebolag som sökte finansiering från programmet EU horisont 2020 inom de Health delprogram. Den ursprungliga datakällan är här eCORDA databasen.

I selekteringsarbetet konstaterades dock att inga ytterligare företag selekterades baserat på att de kom från dessa datakällor.

3.2 Berikning av företagsdata

I tabellen nedan presenteras och beskrivs de datakällor från SCB som har används för att berika företagsdata. Variablerna som används för att beskriva företagen och dess utveckling över tid har valts baserat på den statistik



som efterfrågats inom ramen för regeringsuppdraget för att beskriva företagspopulationen inom life science. I appendix finns variabellistan över samtliga variabler som används för att berika data.

TABELL 3.2

Källa	Beskrivning
Arbetsställen	Statistiken innehåller företagsdata på arbetsställe-nivå.
Bolagsverket	Data innehåller företagens verksamhetsbeskrivningar.
FAD	Registren för Företagens och arbetsställets dynamik innehåller data från 1986 och framåt. De kan användas till att kartlägga strukturella förändringar av företag och arbetsställen.
FDB	Företagsdatabasen (FDB) är ett register över samtliga företag, myndigheter, organisationer och deras arbetsställen. FDB har en central roll som urvalsram och samordningsinstrument för statistikproduktionen inom SCB. Detta gäller framför allt för den ekonomiska statistiken. FDB är också ett basregister i SCB:s registersystem.
FEK	Statistiken innehåller företagsekonomiska uppgifter såsom antal företag och antal anställda, nettoomsättning, produktions- och förädlingsvärde, tillgångar, brutto- och nettoinvesteringar, intäcks- och kostnadsuppgifter och rörelseresultat.
Födelseuppgift	Statistiken innehåller information om individers bakgrund, utländsk/svensk.
LISA	SCB:s longitudinella databas LISA innehåller detaljerade data om sjuk-, föräldra- och arbetslöshetsförsäkringen på individnivå. Den möjliggör att över tiden följa en persons övergångar mellan exempelvis förvärsarbete, arbetslöshet och sjukdom.
Momsdeklarationsregistret	Statistiken innehåller information om företagens momsdeklarationer.
Population	Statistiken innehåller befolkningsstatistik.
RAMS	Statistiken visar sysselsättning, pendling och personal- och näringsstruktur. Den visar också händelser och flöden på arbetsmarknaden. Statistiken är totalräknad och kan brytas ner på låg regional nivå eller utifrån de anställdas egenskaper; till exempel kön, utbildning och ålder.
RTB	Sedan 1968 finns Registret över totalbefolkningen (RTB) hos SCB. Registret är ett utdrag ur folkbokföringsregistret som Skatteverket ansvarar för.
SyssReg	Statistiken visar förvärsarbetande kvinnor och män med okänd utbildning.
UHV	Statistiken visar Sveriges utrikeshandel med varor fördelat efter länder och olika typer av varuindelningar. Kombinerade nomenklaturen (KN) ger mest detaljerad



	nivå med cirka 9 400 varugrupper. SITC och SPIN är mer översiktliga varuindelningar.
UReg	Statistiken innehåller uppgifter om befolkningens utbildning.
YReg	Statistiken innehåller uppgifter om befolkningens yrken.



4 Modellbeskrivning

Modellen utgår från de identifierade datakällorna och företagspopulationen identifieras i tre steg:

- Identifiera – fokuserar på att säkra analysens täckningsgrad genom att säkra att alla företag i branschen identifieras genom datakällorna.
- Selektera – fokuserar på att säkra exakthet i analysen genom att endast företag som är verksamma inom life science inkluderas.
- Klassificera – fokuserar på att klassificera utvalda life science-företag utifrån följande huvud- och underkategorier: läkemedel, medicinteknik, bioteknik, labtech, diagnostik, konsulter, specialiserade konsulter och försäljare/grossister.

Efter att företagen har identifierats, selekterats och klassificerats har de berikats med officiella registerdata från SCB för att kunna sammanställa statistiken. Denna process beskrivs närmare i avsnitten nedan.

4.1 Identifikation av företag

I arbetets första fas sammanställs en bruttolista över företag som bedöms täcka majoriteten av den svenska life science-sektorn år 2022. Sökningar görs i tidigare beskrivna databaser och resultaten sammanställs i en samlad lista. Sammanställningen görs mot organisationsnummer och datakällor där organisationsnummer saknas matchas på namn mot Näringslivsregistret och berikas med organisationsnummer genom tjänsten Proff.

Följande kriterier används för sammanställning av bruttolistan:

- Samtliga företag som återfinns i datamängderna EUDAMED, NPL, och medlemslistor från branschföreningar inkluderas på listan.
- Företag som är taggade som "Health" i kategorin "Industry" i databasen Dealroom inkluderas på listan.
- Företag som återfinns bland Vinnova (område: Hållbar precisionshälsa) och EU horisont-ansökningar.³
- Företag som har någon av nedanstående SNI-koder inkluderas på listan.

TABELL 4.1

SNI-kod	Branschbeteckning
26.600	Industri för strålningsutrustning samt elektromedicinsk och elektroterapeutisk utrustning
21.100	Industri för farmaceutiska basprodukter
21.200	Läkemedelsindustri
32.501	Industri för medicinska och dentala instrument och tillbehör
32.502	Tandteknikerlaboratorier
46.460	Partihandel med medicinsk utrustning och apoteksvaror
72.110	Biotekniska FoU-institutioner
72.190	Andra naturvetenskapliga och tekniska FoU-institutioner
86.901	Medicinska laboratorier o.d.

³ Se 3.1.1: I selekteringsarbetet konstaterades dock att inga ytterligare företag selekterades baserat på att de kom från dessa datakällor.



Urvalsprocessen för att ta fram dessa SNI-koder har skett genom en iterativ process där ADC och Vinnova har tagit ut fullständiga företagslistor via näringslivsregistret Proff för ett stort urval av SNI-koder. Därefter har företagen i respektive SNI-kod-lista i ett första steg matchats mot företagslistan från Tillväxtanalys rapport i syfte att bedöma SNI-kodens relevans ur ett life science-perspektiv. Om företagen inom en SNI-kod haft hög träffsäkerhet gentemot Tillväxtanalys företagslista har denna SNI-kod direkt valts ut som relevant. Om företagen inom en SNI-kod haft låg träffsäkerhet gentemot Tillväxtanalys företagslista men ändå bedömts som relevant för life science-sektorn, exempelvis 62.010 (dataprogrammering) och 71.122 (teknisk konsultverksamhet inom industriteknik), har företagen med dessa SNI-koder granskats manuellt av ADC. Om det under denna granskning skulle framkommit att företagen inom dessa SNI-koder i relativt stor utsträckning är relevanta för life science-sektorn skulle dessa SNI-koder ha inkluderats. Denna manuella granskning har dock inte lett till att SNI-koder med låg träffsäkerhet gentemot Tillväxtanalys företagslista har lagts till.

Sammantaget identifieras cirka 11 000 företag i detta steg vilket utgör resultatet från steg 1.

4.2 Selektion av företag

Metodens andra steg syftar till att selektera de företag som har verksamhet som faller inom den definition av life science som har tagits fram för projektet och presenteras nedan i form av inkluderings- och exkluderingskriterier. Detta arbete baseras på samma datakällor som tidigare presenterats. Sammantaget används fem inkluderingskriterier och tre exkluderingskriterier.

4.2.1 Inkluderingskriterier

Totalt ställs fem inkluderingskriterier upp vilka kopplar till de ursprungliga datakällorna och metadata som återfinns på dessa. Följande inkluderingskriterier ställs upp:

- Inkluderingskriterium 1 - Företaget har första eller andra SNI-kod som återfinns i kategori 1⁴.
- Inkluderingskriterium 2 - Företaget återfinns på EUDAMED- eller NPL-listan och har första eller andra SNI-kod som återfinns i kategori 1 eller 2.
- Inkluderingskriterium 3 - Företaget återfinns på någon av medlemslistorna år 2022 från branschorganisationerna LIF, Swedish Medtech, Labtech eller Sweden BIO och har första eller andra SNI-kod som återfinns i kategori 1 eller 2.
- Inkluderingskriterium 4 - Företaget återfinns på Dealroom-listan och har första eller andra SNI-kod som återfinns i kategori 1 eller 2.
- Inkluderingskriterium 5 - Minst ett ord i företagets verksamhetsbeskrivning matchar mot nyckelordlista.

Användning av SNI-koder:

SNI-koder grupperas i två kategorier: kategori 1 för primära life science-SNI-koder och kategori 2 för sekundära life science-SNI-koder:

⁴ Se hur SNI-koderna fördelas mellan kategorierna under rubriken "Användning av SNI-koder" längre ned på sidan.



TABELL 4.2

Kategori	SNI-koder
1	21.100, 21.200, 26.600
2	32.501, 32.502, 46.460, 72.110, 72.190, 86.901

Företag kan ha flera SNI-koder kopplade till sig. Ordningen av SNI-koderna bestäms av företagen själva men ska i regel spegla verksamhetens inriktning. Således ses den första SNI-koden som den primära. I selektionskriterierna i denna arbetsprocess används kriterier kopplade till SNI-kodernas ordning där enbart företag med relevanta (för life science-sektorn) första eller andra SNI-kod inkluderas.

Användning av nyckelord:

ADC och Vinnova har för uppdraget utvecklat en nyckelordlista som innehåller nyckelord med bäring mot life science. Nyckelordlistan tar utgångspunkt i ett tidigare arbete som är genomfört av SCB där nyckelord har identifierats baserat på ord som återfinns i företags verksamhetsbeskrivningar. Inom projektet har den ursprungliga listan bearbetats vidare för att matcha den definition av life science som används inom projektet. Nyckelordlistan används i inkluderingskriterium 5 där företag från identifieringsfasen som i sin verksamhetsbeskrivning nämner ord från nyckelordlistan inkluderas.

Nyckelordlistan som används för identifiering av life science-företag innehåller 520 nyckelord och återfinns som appendix till denna rapport.

4.2.2 Exkluderingskriterier

Ett ytterligare lager av selektering görs efter att företagen selekterats genom inkluderingskriterium där företag sorteras bort baserat på tre exkluderingskriterier. Exkluderingskriterier syftar till att ta bort företag som bedöms falla utanför den definition som används av life science. Tre exkluderingskriterier ställs upp:

- Exkluderingskriterium 1 - Företag exkluderas baserat på en lista av SNI-koder som har bedömts som irrelevanta. Lista över dessa SNI-koder återfinns i appendix.
- Exkluderingskriterium 2 - Företag exkluderas baserat på en lista av nyckelord som har bedömts som irrelevanta. Dessutom har företag som bedöms vara verksamma som veterinärer eller ortopedier och således inte anses tillhöra målpopulationen exkluderats på basen av en kombination av ord: till exempel företag som nämner "veterinär" i sin verksamhetsbeskrivning och samtidigt inte nämner "human" eller "mänsklig". Denna exkludering har beslutats om efter manuell granskning av företagen med dessa nyckelord i verksamhetsbeskrivningarna, där dessa företag visats vara irrelevanta i förhållande till definitionen av life science-företag. Lista över exkluderingsnyckelord återfinns i appendix.
- Exkluderingskriterium 3 - Företag som inte är ett aktiebolag exkluderas.

Utöver dessa exkluderingskriterier har ADC och Vinnova löpande under projektet samt i den avslutande delen av kvalitetssäkringsprocessen manuellt gått igenom majoriteten av den identifierade företagspopulationen och tagit beslut om att exkludera företag utifrån de regler som skapats. Trots att majoriteten av de identifierade företagen har granskats är det således möjligt att företag som borde inkluderas har exkluderats samt att



inkluderade företag borde exkluderas. Granskningen har genomförts genom att läsa genom den identifierade företagspopulationens verksamhetsbeskrivningar samt undersöka dessa företags hemsidor och därefter göra en bedömning huruvida företagen tillhör life science-sektorn eller ej.

Användning av SNI-koder:

En lista över irrelevanta SNI-koder har upprättats baserat på analys av företag som inkluderats genom inkluderingskriterier. Listan omfattar ett stort antal SNI-koder som bedöms indikera att ett företag faller utanför den definition av life science-sektorn som används inom ramen för analysen. Ett exempel på företag som exkluderas är praktiserande läkare. SNI-kodslistan som används för exkludering återfinns som appendix till denna rapport.

Användning av nyckelord:

En lista som innehåller nyckelord som används för att exkludera företag från att tillhöra företagspopulationen inom life science har under projektet utvecklats tillsammans av ADC och Vinnova. Nyckelordlistan för exkludering har utvecklats dynamiskt under projektet med huvudsyftet att säkerställa att företag som inte arbetar inom life science men som använder SNI-koder som i detta projekt klassificeras som life science-SNI-koder exkluderas. Exempel på nyckelord som exkluderas är *skidspårsutrustning* och *ölbryggeri*.

Nyckelordlistan som används för exkludering av företag innehåller 34 nyckelord och återfinns som appendix till denna rapport.

4.3 Klassificering av företag

Metodens tredje steg fokuserar på klassificering av de identifierade och selekterade företagen i olika delsektorer. Precis som de två föregående stegen baseras data i detta steg på senast tillgängliga data, det vill säga 2021 eller 2022. Företagen kategoriseras i åtta grupper:

- Läkemedel
- Bioteknik
- Medicinteknik
- Labtech
- Diagnostik
- Specialiserade konsulter⁵
- Konsulter
- Försäljare/grossister

De tre sista kategorierna, specialiserade konsulter, konsulter och försäljare/grossister, är annorlunda jämfört med övriga kategorier genom att företag inte enbart kan bli klassificerade som någon av dessa tre kategorier.

⁵ Specialiserade konsulter är en underkategori till kategorin konsulter. Detta beskrivs i detalj i avsnitt 4.3.6.



Företag som är klassificerade som någon av dessa tre kategorier är således också minst klassificerade som tillhörande någon av de andra fem kategorierna.

Ytterligare en viktig aspekt att notera med klassificeringen är att företag kan klassificeras som tillhörande flera kategorier. Exempelvis kan ett företag både klassificeras som tillhörande läkemedel och bioteknik om verksamhetsbeskrivningen innehåller olika nyckelord som klassificerats som nyckelord för läkemedel och bioteknik.

Klassificeringen är baserad på fyra olika faktorer som presenteras nedan:

- Medlemskap i branschföreningar
- Återfinns på NPL och EUDAMED
- Nyckelord i verksamhetsbeskrivningar
- SNI-koder

Notera att samtliga datakällor som presenteras ovan förutom NPL-data och SNI-koder enbart är baserad på data från innevarande år. Detta innebär större osäkerhet i statistiken som sammanställs från längre bak i tiden eftersom medlemskap och verksamhetsbeskrivningar kan ha sett annorlunda ut än vad de gör idag. Om ett företag exempelvis var medlem i någon av branschorganisationerna under perioden 2009-2020 men sedan upphört att vara medlem kommer det således inte kunna klassificeras baserat på branschorganisationsmedlemskap eftersom historiska medlemslistor inte har använts i detta projekt. Dessa företag kan dock klassificeras baserat på någon av de andra ovannämnda faktorerna.

Metoden för att klassificera företag har skett iterativt där ett stort antal klassificeringskriterier har testats varefter resultaten har verifierats av både ADC och Vinnova. Branschaktörer har även fått möjlighet att komma med inspel kopplat till klassificering av företag genom ett referensgruppmöte samt återkoppling till Vinnova.

Nyckelordlistan för klassificering av företagen innehåller 197 nyckelord och återfinns i appendix.

4.3.1 Klassificering av läkemedelsföretag

Företag klassificeras som läkemedel om det återfinns på NPL-listan, om företaget är medlem i LIF, eller om företaget har ett nyckelord som kategoriserats som läkemedel. Listan över nyckelord som används för att klassificera företag som läkemedel återfinns i appendix.

4.3.2 Klassificering av bioteknikföretag

Företag klassificeras som bioteknik om företaget har ett nyckelord som kategoriserats som bioteknik. Listan över nyckelord som används för att klassificera företag som bioteknik återfinns i appendix.

4.3.3 Klassificering av medicinteknikföretag

Företag klassificeras som medicinteknik om det återfinns på EUDAMED-listan, om företaget är medlem i Swedish Medtech, eller om företaget har ett nyckelord som kategoriserats som medicinteknik. Listan över nyckelord som används för att klassificera företag som medicinteknik återfinns i appendix.



4.3.4 Klassificering av labtech

Företag klassificeras som labtech om företaget har ett nyckelord som kategoriserats som labtech. Listan över nyckelord som används för att klassificera företag som labtech återfinns i appendix.

4.3.5 Klassificering av diagnostik

Företag klassificeras som diagnostik om det har ett nyckelord i sin verksamhetsbeskrivning som kategoriserats som diagnostik. Listan över nyckelord som används för att klassificera företag som diagnostik återfinns i appendix.

4.3.6 Klassificering av specialiserade konsulter

Företag klassificeras som specialiserade konsulter om de har ett nyckelord i sin verksamhetsbeskrivning som kategoriserats som specialiserade konsulter. Dessa nyckelord innehåller ord som förknippas med kontraktstillverkare, kontraktsforskning och regulatoriska konsulter. Denna kategori är en underkategori till kategorin *konsulter*. Listan över nyckelord som används för att klassificera företag som specialiserade konsulter återfinns i appendix.

4.3.7 Klassificering av konsult

Företag klassificeras som konsulter om det har ett nyckelord i sin verksamhetsbeskrivning som kategoriserats som specialiserade konsulter (se ovan) eller om de har en SNI-kod som är klassificerat som "konsult". Denna kategori är således en överkategori till kategorin *specialiserade konsulter*. Listan över nyckelord/SNI-kod som används för att klassificera företag som konsult återfinns i appendix.

4.3.8 Klassificering av försäljare/grossister

Företag klassificeras som försäljare/grossister om de har ett nyckelord i sin verksamhetsbeskrivning som kategoriserats som försäljare/grossister eller om de har en SNI-kod som klassificerats som "försäljare/grossister". Listan över nyckelord/SNI-kod som används för att klassificera företag som konsult återfinns i appendix.

4.3.9 Hantering av residual

En övervägande majoritet av företagen har kunnat kategoriseras baserat på ovan nämnda kriterier. En mindre grupp företag återstår dock vilka kommer att klassificeras manuellt. I ett första skede omfattar gruppen cirka 600 företag. Fortsättningsvis kommer denna manuella hantering av residualen omfatta färre företag eftersom hanteringen endast kommer behöva ske av nyregistrerade företag som inte kan klassificeras baserat på ovan nämnda kriterier. Förslagsvis görs denna manuella hantering kommande år på samma sätt som för detta arbete, det vill säga genom att manuellt granska företagens verksamhetsbeskrivningar och hemsidor och därefter göra en bedömning av företagens verksamhetsfokus.

4.4 Berikning av data

Arbetets sista steg handlar om att berika företagspopulationen inom life science med officiella mikrodata från SCB. Detta har gjorts genom att ansöka och beviljas tillgång till SCB:s plattform för tillgängliggörande av mikrodata: MONA. Efter att tillgång till MONA beviljades skickades listan med selekterade företag in till MONA



vilket möjliggjort att selektera samtliga dessa företag för hela perioden 2009–2020. Listan över selekterade företag innehåller drygt 4 000 unika företag för år 2022 och har utöver att ha förts över till MONA överlämnats till Vinnova.

Selekteringen av företagspopulationen för tidsperioden 2009–2020 följer de inkluderings- och exkluderingskriterier som beskrivs i avsnitt 4.2. Notera dock att branschorganisationernas medlemslistor samt EUDAMED och NPL utgår från år 2022. Företag som fanns med på dessa listor och register år 2022 inkluderas således i selekteringen av företagspopulationen för tidsperioden 2009–2020. Detta innebär exempelvis att företag som var medlem i någon av branschorganisationerna under perioden 2009–2020 men sedan upphört att vara medlem inte inkluderas i företagspopulationen på detta kriterium eftersom historiska medlemslistor inte har använts för att identifiera företag i detta projekt. Dessa företag kan dock selekteras till företagspopulationen på något av de andra inkluderingskriterierna.

Variablerna för att kunna beskriva företagen och dess utveckling över tid grundar sig i Tillväxtanalys och Vinnovas tidigare arbete med life science-statistiken.⁶ Ytterligare variabler med relevans för att beskriva statistiken har valts av Vinnova. Samtliga variabler sammanställs för tidsperioden 2009–2020. Fullständig variabellista återfinns i appendix.

Efter att tillgång till MONA beviljades har ADC genom den integrerade utvecklingsmiljön RStudio och programspråket R tagit fram ett script som sammanställer data i enlighet med den data som efterfrågats inom regeringsuppdraget över den svenska life science-sektorns utveckling från 2018 samt Vinnovas önskingar om ytterligare skärningar. Huvudrubrikerna för den statistiska sammanställningen som görs i MONA är:

- Övergripande utveckling
- Företagens ekonomi
- Geografisk fördelning och utveckling
- Företagens personalsammansättning

Även data om kompetensförsörjning utgör en del av den statistik som rapporteras inom regeringsuppdraget. Kompetensförsörjningsstatistiken baseras inte på den företagspopulation som tagits fram i samarbete mellan Vinnova och ADC som beskrivs i den här rapporten. För att sammanställa statistik om kompetensförsörjning jobbar Vinnova parallellt med data från Universitetskanslersämbetet (UKÄ).

6

https://www.tillvaxtanalys.se/download/18.62dd45451715a00666f194a2/1586366154910/pm_2018_06_Den%20svenska%20life%20science-industrins%20utveckling.pdf



5 Känslighetsanalys

5.1 Om statistik för life science-sektorn

Utveckling av statistik för en sektor handlar om att definiera vilka företag som tillhör utvald sektor. Detta är ett komplext arbete och innebär ett stort antal avväganden. En genomgång av hur olika regioner i Sverige arbetar med statistik för life science-sektorn ger vid handen att många olika definitioner används och ingen allmänt accepterad definition existerar. Detta kan till stor del förklaras med att gränserna för en sektor sällan är skarp. Ytterligare utmaningar kan härledas till att företag kan vara verksamma inom flera olika delsektorer.

Slutligen bör poängteras att det varken globalt eller i Sverige finns någon vedertagen definition av life science-sektorn på den detaljnivå som krävs för att definiera sektorn med hjälp av datakällor. ADC och Vinnova har i detta arbete utgått från den svenska life science-strategin och gemensamt beslutat om hur den kan operationaliseras på bästa sätt.⁷

5.2 Avväganden i utvecklingsarbetet

5.2.1 Användning av öppna datakällor

I arbetet med att identifiera företagspopulationen inom life science-sektorn har flera olika datamängder utvärderats. Efter önskemål från Vinnova togs ett inriktningsbeslut om att i så hög grad som möjligt använda öppna data om företag. Det primära syftet med detta inriktningsbeslut var att kunna generera en lista över företag i sektorn där enskilda företag kan identifieras och således kvalitetssäkras. Samtidigt innebar inriktningsbeslutet att ett antal möjliga datakällor valdes bort. Primärt rörde sig detta om information om företagen som återfinns i SCB:s registerdata, bland annat kopplat till företagens exportverksamhet.

Det är vår övergripande uppfattning att inriktningsbeslutet inte har haft någon negativ påverkan på täckningsgrad och exakthet i den slutgiltiga företagslistan. Arbetet har inneburit en viss begränsning av de datakällor som kunnat användas. Samtidigt har inriktningen medfört en stor fördel genom att det varit möjligt att löpande granska vilka företag som inkluderats på listan.

5.2.2 Användning av registerdata

I arbetets slutskede används registerdata för att ta fram statistik genom att de via öppna data identifierade företagen berikas med registerdata från SCB. Det bör noteras att företagen som identifierats via öppna data utgår från år 2022, medan den officiella statistiken via SCB:s registerdataserver MONA i stor utsträckning publiceras med två års fördröjning. Senast tillgängliga data som används för att sammanställa statistiken över life science-sektorn är således från år 2020.

Vidare är företagspopulationen i dessa registerdata pseudonymiserad vilket betyder att det inte är möjligt att urskilja enskilda företag – statistiken presenteras i stället på aggregerad nivå. Detta innebär att det inte är möjligt att ta fram och således kvalitetssäkra företagslistorna för respektive år som statistiken baseras på vilket bidrar

⁷ Regeringskansliet, (2019). En nationell strategi för life science, N2019.06



till ökad risk att felaktigt inkludera och exkludera företag. I praktiken är dessa företag som inte går att kvalitetssäkra de företag som utifrån den framtagna definitionen identifieras, selekteras och klassificeras som tillhörande life science-sektorn, men som inte existerade år 2022. I statistiksammanställningarna har vi inte noterat några större avvikelser gällande de variabler vi årligen följer varför denna risk bedöms påverka statistiken i låg grad.

5.2.3 Förutsättningar för automation

En uttalad målsättning med projektet har varit att utveckla en metod som i möjligaste mån kan uppdateras automatiskt. Detta innebär att element som omfattar manuella genomgångar av företag i möjligaste mån begränsas. Detta har stora fördelar då det innebär att uppdateringen av statistiken genomförs mer effektivt och utan den stora arbetsinsats som tidigare krävts vid framtagning av statistiken hos både Tillväxtanalys och Vinnova, vilket identifierats som ett hinder för att upprätthålla en uppdaterad statistik.

Samtidigt som det finns stora fördelar med en automatiserad metod för att identifiera och klassificera företagspopulationen medför metoden risker. Framför allt föreligger det risker att företag som inte tillhör branschen inkluderas i populationen, att företag som tillhör branschen inte inkluderas, samt att stora företag men med en mycket begränsad life science-verksamhet inkluderas. Dessa utmaningar har stått i centrum i utvecklingsarbetet.

För att minimera risken att företag inkluderas felaktigt har företagspopulationen kontinuerligt under arbetets gång granskats och kvalitetssäkrats av ADC och Vinnova, samt även i viss mån med hjälp av branschföreningarna i referensgruppen. Genom en iterativ process har regelverk utvecklats för att minimera antalet företag som identifieras och klassificeras felaktigt. Riskerna föreligger dock fortsatt och det är inte möjligt att garantera att samtliga företag är korrekt identifierade och klassificerade, nu och i framtiden. Vår bedömning är dock, baserat på det omfattande arbete som genomförts med att manuellt granska listorna som modellen genererat att antalet felklassificerade företag är få och att deras påverkan på den övergripande statistiken är begränsad.

I Vinnovas fortsatta arbete bör ytterligare fokus läggas på att kvalitetssäkra inkludering av större företag, exempelvis företag med fler än 50 anställda. Under detta projekt har inget särskilt fokus lagts (ex granskning av årsredovisningar) på att kvalitetssäkra att dessa större företags verksamheter i huvudsak sker inom life science, utan enbart att de över huvud taget är verksamma inom life science. Vinnova bör sätta upp rutiner och regelverk för hur hantering av dessa företag ska ske framöver.

För att minimera risken att företag som tillhör branschen felaktigt inte identifieras har flera aktiviteter genomförts. Framför allt har den identifierade företagspopulationen jämförts med företag som identifierats i den senaste utgåvan av life science-statistiken⁸. Sammantaget är bilden att de företag som omfattas av definitionen som används i arbetet i hög grad identifieras genom modellen. Samtidigt finns det utmaningar med att garantera en komplett täckningsgrad, inte minst eftersom en given definition av vilka företag som tillhör den svenska life science-sektorn år 2022 saknas.

8

https://www.tillvaxtanalys.se/download/18.62dd45451715a00666f194a2/1586366154910/pm_2018_06_Den%20svenska%20life%20science-industrins%20utveckling.pdf



Ovan nämnda risker att felaktigt inkludera och exkludera företag ökar dessutom per definition vid en automatisk uppdatering av statistiken om sektorn eller de datakällor som används för identifiering och klassificering förändras. Vidare föreligger en risk att resurser avsatta för granskning av den identifierade populationen är begränsad och att en lika gedigen granskning som genomförts under utvecklingsarbetet inte görs. Detta bidrar sammantaget till en ökad risk att automatiseringen leder till en över tid försämrad statistik. Rutiner för hur detta kvalitetsarbete bör upprätthållas över tid beskrivs i avsnitt 6.3.

5.2.4 Förutsättningar för klassificering

Klassificeringen av företagen är det arbetsmoment som varit mest utmanande. Detta beror bland annat på att det finns överlapp mellan olika delsektorer samt att företag kan vara verksamma i flera delsektorer. I jämförelse med referensarbetet som genomförts i Danmark (se avsnitt 5.2.2) kan exportverksamhet inte användas för klassificering av verksamheter på grund av den initiala avvägningen att använda öppna datakällor för att identifierat, selektera och klassificera företagen. Detta sammantaget gör detta moment till det mest utmanande i arbetsprocessen.

I den initiala diskussionen mellan ADC och Vinnova, samt i dialogen med branschföreträdare har fyra potentiella delsektorer lyfts fram: läkemedel, bioteknik, labtech och medicinteknik. Under projektet har ytterligare delsektorer tillförts klassificeringsarbetet: diagnostik, specialiserade konsulter, konsulter och försäljare/grossister.

Diagnostik har identifierats utifrån att det fanns en relativt stor grupp företag som klassificerades som både medicinteknik och labtech. Efter granskning av dessa företag skapades den nya delsektorn *diagnostik*. Specialiserade konsulter, konsulter och försäljare tillkom i ett sent skede av klassificeringsarbetet baserat på intresse och önskemål från Vinnova och referensgruppen. Dessa tre sist tillkomna delsektorer är de som ägnats minst tid att klassificera och kvalitetssäkra varför de till nästa uppdatering av statistiken bör ses över igen. Ett sådant arbete ligger i linje med det förslag om vidareutveckling av en marknadsdimension i klassificeringen av företag som diskuteras i kapitel 7 i denna rapport. En vidare kvalitetssäkring av specialiserade konsulter, konsulter och försäljare/grossister skulle kunna genomföras parallellt med att ytterligare klassificering för forskning och tillverkning utvecklas.

5.3 Datakällor och -variabler

Nedan presenteras en känslighetsanalys av datakällor och -variabler vi förknippar med störst risk i förhållande till beroende av extern part eller som kan förändras över tid.

5.3.1 SNI-koder

SNI-koder används för att identifiera, selektera och klassificera företag. Vi har ingen kontroll över vilka SNI-koder företagen väljer att förknippa sig med vilket innebär en risk sett till att dessa kan vara både irrelevanta och inaktuella eftersom de inte bestäms centralt av exempelvis Bolagsverket. Detta kan påverka statistiken över sikt om en viss typ av företag är bättre eller sämre på att klassificera sig enligt SNI-koderna.



5.3.2 Verksamhetsbeskrivningar

Vi använder oss i huvudsak av verksamhetsbeskrivningar för klassificering av bolag. Detta möjliggör en mer granulär segmentering än vad enbart SNI-koder tillåter, även om detaljnivån i verksamhetsbeskrivningarna varierar. Samtidigt är syftet med företagens verksamhetsbeskrivningar inte att de ska användas för att klassificera företag och företagen har således inte utformat dessa med hänsyn till att de ska användas för klassificering.

5.3.3 EUDAMED och Nationellt produktregister för läkemedel (NPL)

Genom att använda EUDAMED och NPL får vi en insikt i vilken typ av marknad som företagen verkar på och vilka produkter de erbjuder. Samtidigt har vi ingen kontroll över hur registren används i praktiken. Vi gör oss också beroende av register som kan förändras över tid. Detta kan påverka statistiken på sikt.

5.3.4 Medlemslistor från branschföreningar

Branschlistorna ger oss insikt i företagens perspektiv utifrån vilken delsektor de anser sig tillhöra. Vi har dock ingen kontroll över kriterierna företagen baserar sina val av branschorganisationsmedlemskap på, huruvida branschorganisationerna är villiga att dela med sig av medlemslistorna, eller vilken kvalitet medlemslistorna håller. Således gör vi oss även här beroende av register från extern part.

5.3.5 Dealroom

Data från Dealroom används för att identifiera företag. Vi har ingen kontroll över hur registren underhålls eller uppdateras över tid och är dessutom beroende av att Vinnovas åtkomst förlängs. Eftersom denna datakälla har tillfört relativt få företag i den initiala identifieringen av företag bedöms vikten av denna datakälla vara låg.

5.3.6 Ansökningsdata från Vinnova och EU horisont

Dessa data används för att identifiera företag. Data kommer från Vinnova och baseras på vilka företag som söker om stöd från Vinnova inom området *Hållbar Precisionshälsa* eller andra aktörer i det svenska näringslivsfrämjande systemet och kan således variera beroende på vilka företag som söker stöd. I likhet med data från Dealroom är användningen av denna data begränsad varför vikten och således risken med denna datakälla bedöms som låg.



6 Rutiner vid uppdatering av statistiken

6.1 Insamling av datakällor

Vid uppdatering av statistiken ska information från externa datakällor inhämtas. Dessa beskrivs nedan.

TABELL 6.1

Datakälla/information	Beskrivning	Ursprung
Branschföreningarna LIF, Swedish Medtech, Labtech, Sweden BIO	Listor över medlemmar.	Inrapporterad information från branschföreningar.
Företagsinformation	Utdrag ur näringslivsregistret omfattande SNI-koder, verksamhetsbeskrivningar, omsättning och antal anställda.	Årsredovisningar hos Bolagsverket som kan hämtas via exempelvis tjänsterna Proff, Bisnode, eller direkt via Bolagsverket.
EUDAMED	Utdrag av svenska aktörer registrerade i EU-databasen EUDAMED över registrerade medicintekniska produkter.	Hämtad från EUDAMED:s webbplats genom sökning efter ekonomiska operatörer med säte i Sverige. (länk)
Nationellt produktregister för läkemedel (NPL)	Utdrag av aktörer registrerade i Nationellt produktregister för läkemedel (NPL).	Data hämtas från Läkemedelverkets webbsida där senaste versionen av NPL kan hämtas vilken innehåller en sammanställning av organisationer. (länk)
Dealroom	Utdrag över företag med taggade med "Health" som industri från startupdatabasen Dealroom.	Dealroom via Vinnovas access
Vinnova och EU horisont eCORDA	Data över aktörer som ansökt om stöd från Vinnova och andra aktörer i det svenska näringslivsstödjande systemet.	Vinnova Datalager och EU horisont eCORDA Databas.

Identifiering, selektering och klassificering av företagen har genomförts baserat på data från 2021 eller 2022. Företagsinformationen från Bolagsverket utgår i huvudsak från data från år 2021 eftersom dessa baseras på företagets årsredovisningar. Resterande data uppdateras när den hämtas hem vilket för detta projekt har gjorts mellan september-december 2022.

För att få fram en ny nettolista över företag inom den svenska life science-sektorn för år 2023 behöver således data från dessa datakällor hämtas för senast möjliga år. Detta innebär att Vinnova måste begära ut nya medlemslistor från branschföreningarna, hämta nya data från EUDAMED, NPL och Dealroom samt sammanställa uppdaterade data via Dealroom. Dessutom behöver Vinnova göra nya utdrag från näringslivsregistret över



företagsinformation och utifrån dessa uppdaterade datasammanställningar göra nya körningar av de R-script som utvecklats inom ramen för detta projekt för framtagning av uppdaterade företagslistor.⁹

Det bör dock noteras att datakällorna som används för att berika företagslistorna med officiell statistik i regel ligger två år efter via SCB:s registerdataserver MONA. Senast tillgängliga data som används för att sammanställa statistiken över life science-sektorn i detta projekt är således från år 2020. Givet att inga definitioner förändras eller nya skärningar önskas innebär detta i praktiken att Vinnova fram till år 2025 kan förlita sig på den nettolista av företag som tagits fram inom ramen för detta projekt när statistiken ska uppdateras år 2023 och 2024. Detta eftersom statistiken via SCB:s registerdataserver MONA dessa år kommer innehålla data som senast uppdaterats för år 2021 och 2022, vilket är de år detta projekt har utgått ifrån för att identifiera, selektera och klassificera företag.¹⁰

6.2 Framtagning av företagslistor via öppna data

Avsnitt 6.1 beskriver de datakällor som används för att samla in underlag till att ta fram en bruttolista över företag som enligt de definitioner som tagits fram inom ramen för projektet kan tänkas identifieras som life science-företag. För att ta fram bruttolistan behöver dessa olika datakällor slås samman i en och samma bearbetningsbar fil. I detta projekt har vi valt att sammanställa data i en Excel-fil. Excel-filen är uppbyggd så att raderna innehåller unika företag medan företagsinformationen samt från vilka datakällor företaget inkluderats via visas i kolumnerna.

Det är på denna sammanställning av företagsdata (bruttolistan) Vinnova sedan behöver köra det R-script som utvecklats inom ramen för detta projekt för att identifiera, selektera och klassificera företagen. Detta R-script har av ADC överlämnats till Vinnova. I praktiken genomför R-scriptet samtliga kriterier kopplade till identifiering, selektering och klassificering som arbetats fram inom ramen för detta projekt. Slutprodukten av detta R-script är en nettolista över den svenska företagspopulationen inom life science för senast tillgängliga år (givet att datakällorna har uppdaterats med senast tillgängliga års data enligt vad som beskrivs i avsnitt 6.1).

6.3 Kvalitetskontroll

De listor som genereras från modellen bör kontrolleras inför varje uppdatering av statistiken för att säkerställa att företagen klassificerats enligt gällande definition. Utgångspunkten i utformningen av modellen är att företagen ska kunna klassificerats automatisk, men en årlig manuell genomgång kommer bidra med att säkerställa kvaliteten.

De nytillkomna företag som ej kan klassificeras automatiskt enligt framtagna metod och R-script behöver årligen granskas och klassificeras manuellt av Vinnova på motsvarande vis som gjorts inom ramen för detta projekt (se avsnitt 4.3.9 för beskrivning av hur residualen hanterats).

⁹ Se avsnitt 6.2 för beskrivning av processen för framtagning av företagslistor.

¹⁰ Detta beskrivs närmare i avsnitt 6.4



6.4 Framtagning av statistik i MONA

De R-script som används för framtagande av statistik i MONA har av ADC överlämnats till Vinnova. Utöver nya djupdykningar Vinnova väljer att göra framöver behöver scriptet endast uppdateras för att inkludera senast tillgängliga år i data. I detta arbete finns fullständiga data endast tillgängligt för år 2020 i SCB:s registerdataserver MONA. För nästa års rapport behöver år 2021 läggas till i databeställningen till SCB samt läggas till i R-scripten för samtliga datakällor som används för att år 2021 ska inkluderas.

I de R-script som tagits fram och överlämnats till Vinnova slås först de olika dataregistren som behövs för att ta fram önskad statistik ihop på MONA. Därefter tas önskad aggregerad statistik som beskriver life science-sektorn och delsektorerna fram i form av Excel-filer som lagras på MONA. Slutligen skapas genom R-scriptet önskade figurer och tabeller som sparas i Microsoft Office-kompatibla format (Excel, Powerpoint, Word) på MONA. Dessa tabeller och figurer kan sedan tas ut från MONA och användas av Vinnova i kommande rapportsammanställningar.

6.5 Rapportsammanställning

De R-script som används för att ta fram figurer och tabeller som beskriver och jämför life science-sektorn genererar både Excel-filer samt Microsoft Office-figurer och -tabeller som Vinnova direkt kan använda i sammanställningen av kommande års rapporter över den svenska life science-sektorn.



7 Rekommendation för underhåll och vidare utveckling av metoden

7.1 Löpande kvalitetssäkring

Modellen som tagits fram för att kartlägga företagspopulationen inom life science är utformad för att med hög grad av automation identifiera och klassificera företag i sektorn. Som en del i utvecklingen av modellen har resultaten som genereras i respektive steg granskats ingående. Samtidigt kan förändringar i sektorn eller i de underliggande datakällorna påverka statistikens kvalitet. Av denna anledning är det viktigt att resultaten från modellen löpande kvalitetssäkras så att eventuella felaktigheter upptäcks och kan korrigeras. Vid uppdatering av modellen bör därför listorna över identifierade och klassificerade företag granskas på liknande sätt som beskrivs i avsnitt 6.3.

Den löpande kvalitetssäkringen bidrar till att identifiera underhållsbehov. Eventuella fel kan delvis korrigeras manuellt, men kan även hanteras genom framtida utvecklingsarbete. Ett löpande kvalitetsarbete bör därför genomföras och omfatta översiktlig analys av den deskriptiva statistik som modeller levererar. Vidare bör de ingående listorna översiktligt granskas och feedback från aktörer i branschen systematiskt och löpande tillvaratas.

7.2 Klassificering av marknadsaktiviteter

Segmenteringen av de identifierade företagen har primärt fokuserat på att dela in företagen i olika delsektorer. Modellen delar i dagsläget in företagen i fem delsektorer: läkemedel, bioteknik, medicinteknik, labtech och diagnostik. Därutöver klassificeras företagen delvis baserat på vilken typ av marknadsaktiviteter de är delaktiga igenom att företag kan klassificeras som konsulter samt försäljning och grossister. Vidare utveckling kan omfatta ytterligare delsektorer samt klassificering av fler marknadsaktiviteter hos bolagen.

I diskussioner som förts mellan ADC och Vinnova har primärt behovet av att klassificera företagens marknadsaktiviteter inom forskning och tillverkning identifieras. En sådan klassificering bedöms som möjlig att genomföra baserad på samma metodik som används för övrig klassificering. Med tillägg av dessa två marknadsaktiviteter skulle en tvådimensionell skärning av populationen uppnås vilket skulle kunna användas för fördjupade analyser av olika delpopulationer i sektorn.



FIGUR 7.1

		Marknadsaktiviteter			
		Forskning	Tillverkning	Försäljning och grossister	Konsulter
Delsektorer	Läkemedel				
	Bioteknik				
	Medicinteknik				
	Labtech				
	Diagnostik				

7.3 Tillägg av ytterligare delsektorer och breddning av sektorn

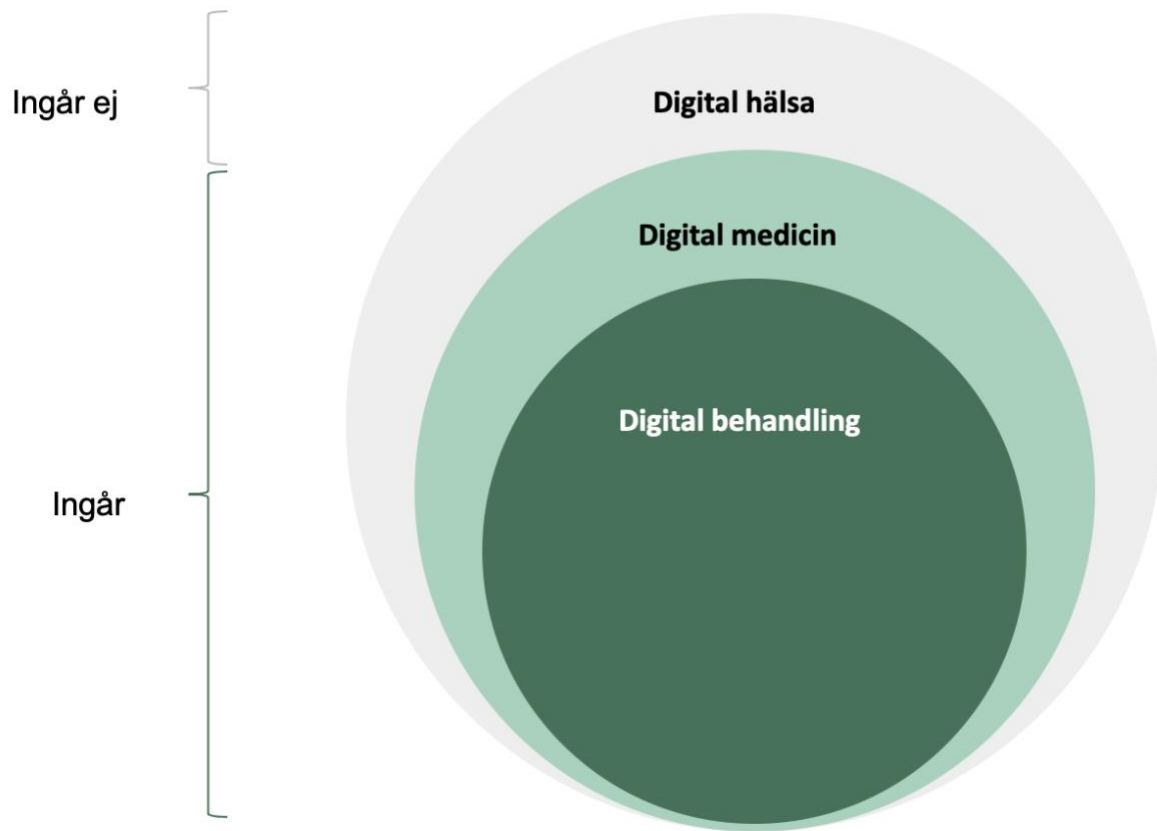
Utöver ovan nämnda utvecklingsbehov tillåter modellen att nya delsektorer inkluderas. I arbetet under 2022 har fokus legat på att kartlägga kärnan av life science-sektorn. Gränser har dragits vid produkter som klassas som medicintekniska, vilket bland annat inneburit att företag som tillverkar och säljer hygienartiklar och andra artiklar som inte klassificeras som medicinska produkter inte varit i fokus för analysen.

Ett resultat av den valda definitionen av life science-sektorn påverkar inkludering av digitala lösningar inom life science. Även här används samma gränsdragning som ovan, där endast företag som säljer tjänster som klassificeras som medicintekniska eftersöks. I praktiken betyder detta att för digitala tjänster inom life science ingår vad som klassificeras som digitala behandlingar och digital medicin (Figur 7.2), medan andra typer av tjänster inom den bredare gruppen digital hälsa inte ingår. I praktiken är gränsdragningen för vilka digitala tjänster som ska omfattas, respektive inte omfattas, komplex. Detta innebär att företag som utvecklar lösningar riktade mot life science-sektorn, men som inte kan klassificeras som medicintekniska inte eftersöks och därmed inte inkluderas i underlaget.

I dialog med branschen har en tydlig trend mot ökad digitalisering identifierats där fler och fler företag utvecklar digitala tjänster med bäring mot life science-sektorn. Det kan av denna anledning finnas skäl att framöver även inkludera företag som utvecklar digitala tjänster mot life science-sektorn i statistiken, även om dessa inte klassificeras som medicintekniska. Det kan bland annat handla om hälso- och vårdappar så väl som andra mjukvarulösningar som inte klassificeras som medicinteknik eller enligt några av de inkluderingskriterier som presenterades i avsnitt 4.2.1. Det finns goda förutsättningar att över tid lägga till fler delsektorer i statistiken. Ett fortsatt utvecklingsarbete skulle kunna ta ett helhetsgrepp om det digitala hälsoområdet, vilket idag endast delvis omfattas av modellen. Ytterligare delsektorer som diskuterats under uppdraget är bland annat foodtech.



FIGUR 7.2



Figur: Inkluderade respektive exkluderade delar av det digitala vårdekosystemet i nuvarande modell



8 Appendix

8.1 Exkluderande SNI-koder

Företag som har någon av de SNI-koder som presenteras nedan som huvudsaklig SNI-kod (första kod) utesluts från företagspopulationen för life science.

SNI-kod	Benämning
00.009	Huvudnäring okänd
01.280	Odling av kryddväxter, drog- och medicinalväxter
01.430	Uppfödning av hästar och andra hästdjur
01.500	Blandat jordbruk
01.620	Service till husdjursskötsel
01.700	Jakt och service i anslutning härtill
02.101	Skogsförvaltning
02.102	Skogsskötsel
03.210	Fiskodling i saltvatten
03.220	Fiskodling i sötvatten
10.200	Beredning och hållbarhetsbehandling av fisk samt skal- och blötdjur
10.519	Annan mejerivarutillverkning
10.612	Tillverkning av frukostflingor, mixer och andra livsmedelsberedningar av kvarnprodukter
10.830	Framställning av te och kaffe
10.910	Framställning av beredda fodermedel
13.921	Sömnad av gardiner, sängkläder och linnevaror
13.922	Tillverkning av presenningar, tält, segel o.d.
16.102	Hyvling av trä
17.220	Tillverkning av hushålls- och hygienartiklar av papper
19.200	Petroleumraffinering
20.200	Tillverkning av bekämpningsmedel och andra lantbrukskemiska produkter
20.300	Tillverkning av färg, lack, tryckfärg m.m.
20.420	Tillverkning av parfymer och toalettartiklar
20.510	Sprängämnestillverkning
32.400	Tillverkning av spel och leksaker
32.910	Tillverkning av borstbinderiarbeten
43.120	Mark- och grundarbeten
43.222	Ventilationsarbeten
43.320	Byggnadssnickeriarbeten
46.160	Provisionshandel med textilier, kläder, skodon och lädervaror
46.170	Provisionshandel med livsmedel, drycker och tobak
46.340	Partihandel med drycker
46.380	Partihandel med andra livsmedel, bl.a. fisk samt skal- och blötdjur
46.390	Icke specialiserad partihandel med livsmedel, drycker och tobak
46.410	Partihandel med textilier
46.420	Partihandel med kläder och skodon



46.431	Partihandel med elektriska hushållsmaskiner och -apparater
46.450	Partihandel med parfym och kosmetika
46.731	Partihandel med virke och andra byggmaterial
47.241	Specialiserad butikshandel med bröd och konditorivaror
47.291	Specialiserad butikshandel med hälsokost
47.711	Specialiserad butikshandel med herr-, dam- och barnkläder, blandat
47.713	Specialiserad butikshandel med damkläder
47.721	Specialiserad butikshandel med skodon
47.730	Apotekshandel
47.750	Specialiserad butikshandel med kosmetika och hygienartiklar
47.919	Postorderhandel och detaljhandel på Internet med övriga varor
49.320	Taxitrafik
50.102	Icke reguljär sjötrafik över hav och kust av passagerare
64.993	Förvaltning av och handel med värdepapper för en begränsad och sluten krets av ägare
68.310	Fastighetsförmedling
69.201	Redovisning och bokföring
73.111	Reklambyråverksamhet
75.000	Veterinärverksamhet
81.210	Lokalvård
81.300	Skötsel och underhåll av grönytor
85.510	Sport- och fritidsutbildning
85.594	Personalutbildning
86.102	Specialiserad sluten somatisk hälso- och sjukvård på sjukhus
86.103	Specialiserad sluten psykiatrisk hälso- och sjukvård på sjukhus
86.211	Primärvårdsmottagningar med läkare m.m.
86.212	Annan allmän öppen hälso- och sjukvård, ej primärvård
86.221	Specialistläkarverksamhet inom öppenvård, på sjukhus
86.222	Specialistläkarverksamhet inom öppenvård, ej på sjukhus
86.230	Tandläkarverksamhet
86.905	Fysioterapeutisk verksamhet o.d.
86.909	Annan öppen hälso- och sjukvård, utan läkare
87.301	Vård och omsorg i särskilda boendeformer för äldre personer
88.102	Öppna sociala insatser för personer med funktionshinder
90.010	Artistisk verksamhet
90.030	Litterärt och konstnärligt skapande
91.040	Drift av botaniska trädgårdar, djurparker och naturreservat
96.022	Skönhetsvård
96.040	Kroppsvård



8.2 SNI-koder för att identifiera försäljare/grossister och konsulter

SNI-kod försäljare/grossist	Benämning
46.XXX	Samtliga SNI-koder inom huvudkategori 46: Parti- och provisionshandel utom med motorfordon.

SNI-kod konsult	Benämning
62.020	Datakonsultverksamhet
70.220	Konsultverksamhet avseende företags organisation
71.122	Teknisk konsultverksamhet inom industrideknik
71.129	Övrig teknisk konsultverksamhet



8.3 Nyckelord för identifiering

Följande 520 nyckelord används för identifiering av företag som tillhör life science-sektorn.

absorbenter	diagnostiska-produkter	kromatografi	Ortoped
acceleratorer	dialys	kromatografisk	ortopedi
adhesiv	digital uppkopplad vård	kroniska	ortopedisk
adjuvant	disk-implantat	kroppspoteser	ortopedteknisk
aerosolprodukter	DMPK	kroppstemperatur	ortoser
affinering	dosdispensering	kroppsvätskebaserade	osmometrar
agonister	dosering	kroppsvätskor	osteoporosdiagnostik
AIDS	dosimetri	kylväst	osterila
allgener	dragering	kärlintervention	OTC-läkemedel
allergi	drug	kärlsjukdomar	otoskop
allergiska	eHäls	labora- torier	oxalatteknologi
anaerob	e-häls	laborativa	oximeter
analys av biomolekyler	elektrokirurgiska	laboratorie	pacemaker
analysinstrument	endopoteser	labratoriearbete	patient
analysreagensprodukter	endoskop	laparoskopiska	patologi
analysutrustning	enzym	larynx	patologiska
analytisk kemi	enzymatisk katalys	laserutrustning	peptider
anatomisk	ergonomi	life science	pharma
andnings	ergonomiska	lifescience	plasma
ansiktspoteser	exosome	life-science	plåster
ansiktsprotetik	Exosomer	lifesciences	prebiotika
antibakteriell	extraktion	lipidkoncept	precisionsinstrument
antibio	farma	livsvetenskaper	preklinisk
antidrog	farmace	ljusbehandling	pre-klinisk
antigena	farmaceutisk	ljuskänsliga adhesiv	prekliniska
antikropp	farmaceutiska	luftvägs	pre-kliniska
antikroppar	farmaci	lungsjuka	prepa
anti-oxidanter	farmako	lustgas	preventivmedel
antitoxiner	farmakologi	läke	probiotika
antivirala	fermentorer	läkemedel	profylaxutrustning
artros	fertilitet	läkemedels	protaminsulfat
aseptiskt	fertilitetsmonitorer	läkemedelsbolag	proteiner
astma	fetma	läkemedelsbranschen	proteinkemiområdet
audiologi	fibros	läkemedelsforskning	proteinproduktion
audiologisk	fingerpoteser	läkemedelsföretag	proteomics
autoimmuna	flerpatiens	läkemedelsindustri	proteomik
bakterie	fluoriderade	läkemedelsindustrin	poteser
bakteriekultur	flödesrelaterade	läkemedelsområdet	psykiatriska
balansrubbingar	forskning inom området viktminskning	läkemedelsprodukter	Psykolog
bandage	fysikaliska	läkemedelsprovning	psykosociala



basläkemedel	fysiologi	läkemedelsprövningar	pulsoximeter
behandlingsmetoder	fysiologiska	läkemedelsutveckling	radiologi
bendensiometri	födoämnesöverkänslighet	makromolekyler	radiologiska
benproteser	förbättring av hjärnans funktioner	maligna	radionuklider
benskador	förlossning	masspektrometrar	radiopharmaka
beräkningskemi	förstahjälpenprodukter	med.tekniska	radioterapi
bildanalys- instrument	galektininhibitorer	medcintekniska	reagen
bindarmolekyler	gaskromatografer	medfödda	reagens
bio- och nanoteknik	gastroenterologi	medi- cinsktekniska	recept
bioadditiv	geléartad	medical	regenerativ
bioaktiva	gemmologi	medici	regulatorisk
biobank	geneti	medicin	rehabiliteringsprodukter
biobaserade	genomik	medicin teknik	repositionering
biocid	genprodukter	medicinleveranssystem	reumatism
bioelektronik	gensekvensering	medicinsk	RNA
bioenergiforskning	genspåring	medicinsk forsk ning	ryggradsimplantat
biofeedback	genteknologi	medicinsk forskning	röntgen
biofysikalisk	genterapi	medicinsk teknik	scintillatorer
biofysikaliska	glasapparater	medicinska	screeningteknologier
bioimpedansbaserade	glaukom	medicinska produkter	serum
bioinformatik	glykosaminoglykaner	medicinska prövningar	sjukdomstillstånd
bioinformatisk	GMP	medicinskt	Sjukvård
biokatalysatorer	gynekologi	medicinskteknisk	sjukvårdsartiklar
biokatalytiska	habilitering	medicinsktekniska	sjukvårdsprodukter
biokemi	handikapp	medicinstekniska	skleros
biokemikalier	handinstrument	medicinteknik	skruvfixation
biokemisk	handledsfrakturer	medicinteknisk	smärt
biokemiska	haptiska	medicintekniska	snark
biologi	hematologiområdet	medicintekniskt	spinala
biologiskt	heparinnatrium	medinsk	sprutor
biologiskt material	herpesvirus	medtech	spårmolekyler
biomarkörer	histologisk	membranvesiklar	stamceller
biomaterial	histopatologi	mentala	steril
biomedicin	histoteknik	metabol	stetoskop
biomedicinsk	HIV	microarray	stroboskopi
biomedicinska	hjälpmedel för funktionshindrade	microdialys	strukturbologi
biomimetisk	hjälpmedel för rörelsehindrade	microdialysis	strålbehandling
Biomimetiskt	hjälpmedel inom hälsa	mikrobiell	strålterapi
biomolekyler	hjälpmedelsprodukter	mikrobio	stödbandage
biooptiska	Hjärtstartare	mikrobiologi	substrat
biopharmaceutisk	HLR	mikrodia	svampsjukdomar
biopolymerer	hormon	mikroorganismer	synförbättrande
biopolymerernas	hornhinne	mikroskop	synhjälpmedel
bioprinting	hudparametrar	mikrotom	synkorrektion
bioreaktor	hudplåster	mikrotomer	synskadade



bioreaktorer	hudvävnadsprover	minispirometer	sårläggning
biosciences	human	molekyl analys	sårområdet
biosensorbaserad	huvudvikt	molekylär	sömnapne
biosensorer	hydrofoba	molekylärbiologi	sömnrelaterade
biosensorsystem	Hygien	muskel	tabletttdispensering
biosyntetisk	hyperhidros	mykologi	tandimplantat
biotech	hypotermi	nanofiber	tandläkarutrustning
bioteknik	Hälsovård	nanostrukturer	tandmaterial
bioteknikbranschen	högpolymera	nanotrådstrukturer	tandprotetik
bioteknikområdet	hörhjälpmedel	naturpreparat	Tandteknisk
bioteknisk forskning	hörselhjälpmedel	nerv	tandtekniskt
bioteknisk industri	immobilisering	neurala	tarmfloras
biotekniska	immun	neuro	tarmsjukdomar
bioteknologi	immun terapi	neurodegenerativa	terapeutisk
bioteknologisk	immunologi	neurofysiologi	terapeutiska
bioteknologiska	immunreceptorer	neurologi	terapi
biotransformation	immuntoxikologi	neurologiska	tobaksavvänjning
biovetenskap	implantat	neurovetenskap	toxikologi
biovetenskaplig	infektionssjukdomar	nikotin	toxikologiska
biverkningar	inflammat	nukleinsy	tranplantaionsområdet
blindhet	inhiber	nålar	transplantationer
blodserum	injektions	näringsprodukter	transplantation
blodsockertestare	inkontinens	näringsstillförsel	tumör
blodtryck	inplantat	obstetrik	ultraljudsteknik
blodtrycksmätning	intensivvård	obstret	uppdragsforskning
bröstproteser	interdentala	obstruktiv	urin
cancer	invitrodiagnosiska	odontologi	urologi
cellbiologi	kanyler	odontologisk	urologiska
celler	karbonater	oftalmologiutrustning	vaccin
cellprodukter	kardio	okularistverksamhet	vaskulära
cellproduktion	katetrar	oligomerer	virus
cellprover	kemiteknik	oligonukleotid	virusproduktion
cellterapi	kirurgi	oligonukleotider	vitro
cro	kirurgiska	oligosackarider	vivo
cytopatologisk	klirik	onkologi	vävnad
datortomografi	klirik	operation	öroninsatser
defibrillatorer	klirik forskning	oral	
demenssjuka	klirik prövning	orphan	
dental	klirik forskning	ortoped	
dentalinstrument	klirik		
dentalutrustning	klirik		
dermato	koagulationssjukdomar		
desinfektionsutrustning	koenzym		
detektering	kofaktorer		
detektion	kognitiv		



diabetes	konstgjorda ögon
diabetshanteringen	kontaktlinser
diagnos	kontrastmedel
diagnostik	kraniofa
diagnosticering	kraniofaciala
diagnostik	
diagnostisera	
diagnostisk	
diagnostiska	
diagnostiska tester	



8.4 Nyckelord för klassificering

Följande 197 nyckelord används för klassificering av företagen i life science-sektorn:

Bioteknik	Labtech	Läkemedel	Medicinteknik
antibakteriell	analysutrustning	agonister	ansiktsproteser
antikroppar	laboratio	antitoxiner	ansiktsprotetik
bindarmolekyler	laboratorie	basläkemedel	audiologi
biofysik	labratorie	biopharmaceutisk	audiologisk
bioinformatik	vitro	cellterapi	behandlingsmetoder
bioinformatisk		drug	benproteser
biokemi		farma	bildanalys- instrument
biokemiska		farmace	biooptiska
biologi		farmaceutisk	biosensorer
biologisk		farmaceutiska	bröstproteser
biologiska		farmako	datortomografi
biomarkörer		farmakologi	defibrillatorer
biomedicin		galektininhibitorer	dental
biopharma		kemisk-tekniska	disk-implantat
biosyntetisk		kemiteknik	elektromedicin
biotech		kromatografi	endoproteser
biotekn		läke	endoskop
bioteknikbranschen		läkemedels	ergonomi
bioteknikområdet		läkemedelsbolag	fertilitetsmonitorer
bioteknik-produkter		läkemedelsbranschen	fingerproteser
bioteknisk forskning		läkemedelsforskning	förstahjälpenprodukter
biotekniska		läkemedelsindustri	gaskromatografer
bioteknologi		läkemedelsindustrin	handinstrument
bioteknologisk		läkemedelsområdet	hjälpmedel
bioteknologiska		läkemedelsprodukter	hjälpmedel för funktionshindrade
biovetenskap		orphan	hjälpmedel för rörelsehindrade
cellbiologi		otc-läkemedel	hjälpmedel inom hälsa
cellterapi		pharma	hjälpmedelsprodukter
exosome		toxikologi	hudplåster
exosomer		vaccin	hörselhjälpmedel
genmodifierade			immobilisering
genomik			implantat
gensekvensering			inplantat
genspårning			instrument
hydrofoba			interdentala
högpolymera			kanyler
kemiteknik			katetrar
koenzym			kirurgi
kofaktorer			kirurgiska
mikrobiell			konstgjorda ögon
mikrobio			kroppsproteser



Bioteknik

mikrobiologi
oligomerer
oligosackarider
proteinkemiområdet
spårmolekyler
stamceller

Medicinteknik

ljuskänsliga adhesiv
medicintekniska
medical Device
medicin tekniska
medicinsk teknik
medicinsk teknisk
medicinsk utrustning
medicinskteknisk
medicinsk-teknisk
medicinsktekniska
medicinteknisk
medicin-teknisk
medicintekniska
medicintekniskt
medtech
odontologisk
oftalmologiutrustning
okularistverksamhet
ortoped
ortopedi
ortopedisk
ortopedteknisk
ortoser
otoskop
oximeter
pacemaker
profylaxutrustning
proteser
pulsoximeter
radiologi
radiologiska
radioterapi
rehabiliteringsprodukter
ryggradsimplantat
röntgen
sensor
sjukvårdsartiklar
sjukvårdsprodukter
snark
sprutor
stetoskop
stroboskopi
strålbehandling
strålterapi



Medicinteknik

stödbandage
synhjälpmedel
synkorrektion
tandtekniskt
utrustning
öroninsatser

Specialiserade konsulter	Diagnostik	Försäljare/grossist
CMO	diagnostik	Distribution
CRO	diagnosticering	Försälja
Kontrakt	diagnostik	Försäljare
Kontraktsanalys	diagnostisera	Handel
Kontraktsforskning	diagnostisk	Grossist
Kontraktstillverkning	diagnostiska	
legotillverkning	diagnostiska tester	
regulatorisk	diagnostiska-produkter	
Regulatoriska		
Regulatoriska konsulter		
uppdragsforskning		



8.5 Exkluderade nyckelord

Företag som nämner något av följande 34 ord i sin verksamhetsbeskrivning utesluts från företagspopulationen för life science:

akupressur
biodrivmedel
Brandskydd
dataspel
fordonsindustrin
energilagring
helikopterdelar
honung
hyrläkar
hårdmetall
hästavel
jordförbättring
kennelverksamhet
medicinsk praktik
miljöteknik
miljötillstånd
naturpreparat
parfym
reseintyg
sexologimottagning
skidspårsutrustning
skogsbruk
skogsresurser
smycken
solceller
tandvårdskliniker
tandvårdsklinik
tatuering borttagning
travhäst
turism
viktminskning
vitaminpreparat
ölbryggeri
tandläkarverksamhet



8.6 Variabellista från SCB-beställning

Följande tabell visar samtliga variabler som beställts från SCB för att berika företagsdata.

Variabelnamn	Beskrivning
annankvant	Annan kvantitet. Redovisas i den enhet som specificeras i den Kombinerade Nomenklaturen.
AnstS	Aktuellt antal anställda
ar	År
Ast_AntalPers	Totalt antal personer på arbetsstället
Ast_AntalSys	Antal förvärvsarbetande på arbetsstället
Ast_Kommun	Arbetsställets belägenhetskommun
Ast_Kv	Antal förvärvsarbetande kvinnor totalt
Ast_KvEGymnK	Antal förvärvsarbetande kvinnor med kort eftergymnasial utbildning
Ast_KvEGymnL	Antal förvärvsarbetande kvinnor med lång eftergymnasial utbildning
Ast_KvForGymn	Antal förvärvsarbetande kvinnor med förgymnasial utbildning
Ast_KvGymn	Antal förvärvsarbetande kvinnor med gymnasial utbildning
Ast_KvOkand	Antal förvärvsarbetande kvinnor med okänd utbildning
Ast_Men	Antal förvärvsarbetande män totalt
Ast_MenEGymnK	Antal förvärvsarbetande män med kort eftergymnasial utbildning
Ast_MenEGymnL	Antal förvärvsarbetande män med lång eftergymnasial utbildning
Ast_MenForGymn	Antal förvärvsarbetande män med förgymnasial utbildning
Ast_MenGymn	Antal förvärvsarbetande män med gymnasial utbildning
Ast_MenOkand	Antal förvärvsarbetande män med okänd utbildning
Ast_Sni2002	Näringsgren, SNI2002
Ast_Sni2007	Näringsgren, SNI2007
AstNr	Arbetsställenummer enligt RAMS
AstNr_LISA	Arbetsställenummer
AterAnv	Återanvänt personnummer
AvregDatKtid	Avregistreringsår
Export_Oms	Total omsättning av varor och tjänster till utlandet både inom och utanför EU.
FelPersonNr	Markering för misstänkt felaktigt personnummer
FodelseAr	Födelseår
Foradlingsvarde	Foradlingsvarde
Ftg_Foradlingsvarde	Förädlingsvärde
Ftg_G7	Nettoomsättning/anställda
Ftg_Nettoomsattning	Nettoomsättning
Ftg_T4	Förädlingsvärde/anställda
infutf	I = import från avsändningsland, U = Export till bestämmelse land
Kommun	Kommun
Kon	Kön
landskod	Landkod
LopNr_Arbstid	Kopplingsvariabel till arbetsställetabell (endast förvärvsarbetande vid faktiska arbetsställen, ej s.k. rörlig personal) (ersätts med löpnummer vid leverans)
LopNr_Arbstid	Arbetsställets identitetsnummer (används för koppling med individtabellen, ersätts med löpnummer vid leverans)



LopNr_CfarNr	Arbetsställets CFAR-nummer (<i>ersätts med löpnummer vid leverans</i>)
LopNr_CfarNr	Löpnummer som ersätter arbetsställets CFAR-nummer
LopNr_CfarNr_LISA	Arbetsställets CFAR-nummer (<i>ersätts med löpnummer vid leverans</i>)
LopNr_PeOrgNr	Företagets organisationsnummer (<i>ersätts med löpnummer vid leverans</i>)
LopNr_PeOrgNr	Företagets organisationsnummer (<i>används för koppling till företagstabellen, ersätts med löpnummer vid leverans</i>)
LopNr_PeOrgNr	Företagets organisationsnummer (<i>ersätts med löpnummer vid leverans</i>)
LopNr_PeOrgNr_LISA	Företagets organisationsnummer (<i>ersätts med löpnummer vid leverans</i>)
LopNr_PeOrgNrHe	Företagets organisationsnummer (<i>ersätts med löpnummer vid leverans</i>)
LopNr_PersonNr	Löpnummer i populationen
M20_25_InkopVarEG	Import av varor från annat EG-land
M21_InkopTjaEG	Inköp av tjänster från annat EG-land
M22_InkopTjaUtEG	Inköp av tjänster från land utanför EG
M35_ForsVarEG	Export av varor till annat EG-land (länder inom EU)
M36_21_ForsVarUtEG	Export av varor till länder utanför EG
M39_ForsTjaEG	Omsättning av tjänster till EG-länder Export av tjänster
M40_ForsTjaUtEg	Omsättning av tjänster till land utanför EU. Export av tjänster
M50_MomsUlagImport	M50 Beskattningsunderlag vid import.
Medb_Sv	Svenskt medborgarskap (Ja/Nej)
MedelantalAnstallda	MedelantalAnstallda
MedeltalAnstallda	MedelantalAnstallda
Nettoomsättning	Nettoomsättning
Ng	Näringsgren SNI2007
Ng2007	Näringsgren SNI2007
NgLve	Näringsgren SNI2007
obtyp	Objekttyp
Org_AntalPers	Totalt antal personer på företaget
Org_AntalSys	Antal förvävsarbetande på företaget
Org_BelKommun	Belägenhetskommun för största arbetsstället
Org_Kv	Antal förvävsarbetande kvinnor totalt
Org_KvEGymnK	Antal förvävsarbetande kvinnor med kort eftergymnasial utbildning
Org_KvEGymnL	Antal förvävsarbetande kvinnor med lång eftergymnasial utbildning
Org_KvForGymn	Antal förvävsarbetande kvinnor med förgymnasial utbildning
Org_KvGymn	Antal förvävsarbetande kvinnor med gymnasial utbildning
Org_KvOkand	Antal förvävsarbetande kvinnor med okänd utbildning
Org_Men	Antal förvävsarbetande män totalt
Org_MenEGymnK	Antal förvävsarbetande män med kort eftergymnasial utbildning
Org_MenEGymnL	Antal förvävsarbetande män med lång eftergymnasial utbildning
Org_MenForGymn	Antal förvävsarbetande män med förgymnasial utbildning
Org_MenGymn	Antal förvävsarbetande män med gymnasial utbildning
Org_MenOkand	Antal förvävsarbetande män med okänd utbildning
Org_SateKommun	Säteskommunen
Org_Sni2002	Näringsgren, SNI2002, fullständig



Org_Sni2007	Näringsgren, SNI2007, fullständig
Produktionsvarde	Produktionsvärde
RadNr	Löpnummer för versioner av omformulerad verksamhetsbeskrivning
RegDatKtid	Registreringsår
Ssyk3	Yrke enligt SSYK, 3-siffernivå
Ssyk3_2012	Yrke enligt SSYK2012, 3-siffernivå
Ssyk3_2012_J16	Yrke enligt SSYK2012, 3-siffernivå, inkl. härledda yrken för företagare
Ssyk3_J16	Yrke enligt SSYK, 3-siffernivå, inkl. härledda yrken för företagare
Ssyk4	Yrke enligt SSYK, 4-siffernivå
Ssyk4_2012	Yrke enligt SSYK2012, 4-siffernivå
Ssyk4_2012_J16	Yrke enligt SSYK2012, 4-siffernivå, inkl. härledda yrken för företagare
Ssyk4_J16	Yrke enligt SSYK, 4-siffernivå, inkl. härledda yrken för företagare
SsykAr	Årtal för yrkesuppgift
SsykAr_J16	Årtal för yrkesuppgift, inkl. härledda yrken för företagare
SsykStatus	Yrkets överensstämmelse med novemberjobbet
SsykStatus_J16	Yrkets överensstämmelse med novemberjobbet, inkl. härledda yrken för företagare
Sun2000Inr	Utbildningsinriktning, högsta avslutade, SUN2000
Sun2000Niva	Utbildningsnivå, högsta avslutade, SUN2000
Sun2020Inr	Utbildningsinriktning, högsta avslutade, SUN2020
Sun2020Niva	Utbildningsnivå, högsta avslutade, SUN2020
SyssStat11	Sysselsättningsstatus enligt justerad metod 2011
SyssStat19	Sysselsättningsstatus enligt justerad metod 2019
SyssStatJ	Sysselsättningsstatus enligt justerad metod 2004
UtlSvBakg	Utländsk/svensk bakgrund
UtlSvBakgAlt	Utländsk/svensk bakgrund, alternativ indelning
varde	Statistiskt värde i SEK
varukod	Varans varukod enligt den Kombinerade Nomenklaturen (militära varukoder lämnas ej ut)
Verksamhetsbeskr	Verksamhetsbeskrivning
vikt	Statistiskt värde i Kilo



We are Amsterdam Data Collective. Together we realise the potential of data for a better future.



AMSTERDAM
DATA
COLLECTIVE

Copyright 2023, Amsterdam Data Collective