

## **Verslag voor de gebruikersgroep**

**December 2020**

## **Inhoud**

<b>1. Algemene stand van zaken .....</b>	<b>2</b>
1.1. Bronnen	2
1.2. Webtoepassingen	4
1.3. Documentatie en het CMS :	5
<b>2. Stand van zaken projecten .....</b>	<b>6</b>
2.1. Onderzoeksproject: indicatoren ontwikkelen die gewoonlijk berekend worden a.d.h.v. enquêtegegevens	6
2.1.1. Het beschikbaar inkomen op huishoudniveau	6
2.1.2. Tewerkstellingsindicator	7
2.1.3. NEET-jongeren	7
2.2. Onderzoeksproject: een meet- en opvolgingsinstrument voor regionale tewerkstellingsmaatregelen	8
2.3. Onderzoeksproject: informatie over het scholingsniveau van de bevolking	9
2.4. Onderzoeksproject: indicatoren voor een boordtabel sociale zekerheid	10
2.5. Onderzoeksproject: variabelen m.b.t. de socio-economische achtergrond	11

## Inleiding

In dit verslag geven we een algemene stand van zaken met betrekking tot het DWH AM&SB en behandelen we de voortgang van de onderzoeksprojecten die werden bepaald in de stuurgroep met de financierders. Deze onderzoeksprojecten worden uitgevoerd door het consortium CeSO, METICES, Steunpunt Werk. Het betreft volgende onderzoeksprojecten:

1. De ontwikkeling van indicatoren m.b.t. de arbeidsmarkt en sociale bescherming op basis van het DWH AM&SB, die, indien mogelijk, de indicatoren reproduceren die gewoonlijk berekend worden a.d.h.v. enquêtegegevens
2. De ontwikkeling van een meet- en opvolgingsinstrument voor regionale tewerkstellingsmaatregelen
3. De verbetering van de informatie over het scholingsniveau van de bevolking
4. De ontwikkeling van indicatoren voor een boordtabel sociale bescherming
5. De ontwikkeling van variabelen m.b.t. de socio-economische achtergrond

We bespreken eerst de algemene stand van zaken en overlopen nadien achtereenvolgens bovenstaande onderzoeksprojecten.

# 1. Algemene stand van zaken

## 1.1. Bronnen

De data van de instellingen m.b.t. 2017 zijn ingeladen, alsook die van 2018, op enkele bronnen na :

- FWB-CREF;
- RSVZ-eClipz.

Voor het bestand RSVZ-Eclipz, dat gebruikt wordt voor de inkomensnotie, is er de afspraak dat eerst 80% van de definitieve inkomens beschikbaar moeten zijn vooraleer het bestand door het RSVZ wordt overgemaakt.

Er zijn een aantal nieuwe bronnen beschikbaar, die een belangrijke aanvulling vormen voor het aanbod aan arbeidsmarktgegevens in het DWH AM&SB :

- FOD Buitenlandse Zaken : gegevens m.b.t. tewerkstelling bij internationale en Europese instellingen. De inlading is aan de gang.
- LIMOSA : dit betreft gegevens over gedetacheerden. Deze gegevens zijn beschikbaar.
- SIGEDIS : de herinlading van de gegevens is aan de gang. De gegevens CAPELO m.b.t. de ambtenaren worden toegevoegd.
- RSVZ-eclipz : de gegevens werden ingeladen teruggaand tot het jaar 1956.
- KBO : de inlading van het bestand wordt afgerond.
- RSZ : het bestand bevat vanaf 2018 ook de data over de zeevarenden.

De combinatie van de gegevens van SIGEDIS en RSVZ geeft de mogelijkheid om veel verder terug te gaan in de tijd dan nu mogelijk is.

Er werden nog een aantal andere nieuwe bronnen ingeladen :

- FPD/RSVZ : gegevens over de minimumpensioenen en -rechten.
- RSVZ : een bestand met pensioenbeslissingsgegevens. Het betreft een hervorming van een bestaand bestand.
- STATBEL : opleidingsgegevens Census 2011.

Met de FWB is de bespreking aan de gang over de integratie van de gegevens van het Franstalig verplicht onderwijs.

Er werden ook een aantal bronnen met data 2020 ingeladen naar aanleiding van de COVID-19-crisis. Een overzicht is terug te vinden op de website van het DWH AM&SB ([https://www.bcsc.fgov.be/nl/dwh/dwh\\_page/content/websites/datawarehouse/data/covid-19.html](https://www.bcsc.fgov.be/nl/dwh/dwh_page/content/websites/datawarehouse/data/covid-19.html)). Deze bronnen zijn voorlopig en worden vervangen door de structurele bestanden die ingeladen worden in het DWH AM&SB, van zodra ze aanwezig zijn.

De bestanden m.b.t. de arbeidsongevallen werden omgevormd naar nieuwe bestanden die meer bevattelijk waren voor de gebruikers. Dit gaf echter geen goed resultaat. FEDRIS heeft intussen nieuwe bestanden aangeleverd in een meer toegankelijke vorm. Deze worden momenteel getest.

De impact van de staats hervorming : op 3 gebieden is er een impact, nl. de kinderbeijlagen, het beleid voor personen met een handicap en de tewerkstellingsmaatregelen.

- Voor de kinderbijslagen is er verscheidene malen overleg geweest met de regio's (Opgroeien/AVIQ/Ministère de la Communauté germanophone/Iriscare). Er werd een gegevenslijst opgesteld met de gewenste gegevens. Momenteel werd het oude bestand van FAMIFED ingeladen, beheerd door ORINT. Dit is een voorlopige oplossing in afwachting van de definitieve gegevens. Vlaanderen zal meteen de nieuwe gegevens aanleveren. De Duitstalige Gemeenschap heeft ook al de nieuwe data aangeleverd.
- 
- Voor het deel personen met een handicap ontbreken de gegevens THAB vanaf september voor Vlaanderen. In Vlaanderen wordt het THAB vervangen door het Zorgbudget voor Ouderen. Deze informatie is beschikbaar bij het Agentschap Vlaamse Sociale Bescherming. Het Agentschap Vlaamse Sociale Bescherming is intussen dataleverancier geworden en heeft de gegevens m.b.t. het Zorgbudget voor Ouderen overgemaakt.
- Voor het deel tewerkstellingsmaatregelen is er binnen het kader van de projecten die de onderzoeksteams uitvoeren ten behoeve van de financierders een werkgroep opgericht.

Er wordt een uitbreiding van de nomenclatuur met 3 populaties voorbereid :

- personen werkend bij Europese en internationale instellingen;
- uitgaande grensarbeiders;
- permanent arbeidsongeschikten omwille van een arbeidsongeval.

De vraag is of de populatie uit LIMOSA moet opgenomen worden.

Nieuwe ontwikkelingen die moeten gevolgd worden :

- De nieuwe gezinsclassificatie die STATBEL heeft voorgesteld.
- De invoering van geo-coördinaten.

## 1.2. Webtoepassingen

De webtoepassingen werden geactualiseerd naar 2018. Er was echter een probleem met de cijfers m.b.t. de pensioentrekkenden. Dit probleem moet nu eerst worden gecorrigeerd. Momenteel zijn de cijfers voor 2018 niet consulteerbaar.

Er werd een nieuwe webtoepassing in gebruik genomen : socio-economische mobiliteit op korte termijn. De webtoepassing werkt anders dan de overige : de documentatie over de variabelen kan nu worden opgevraagd terwijl men met de toepassing werkt. De andere webtoepassingen werden intussen hieraan aangepast.

Er werden een aantal correcties uitgevoerd :

- Indeling nationaliteiten : er werd geen rekening gehouden met de oude nationaliteitscodes van de landen Congo/Rwanda/Burundi. Dit gaf een onderschatting van de cijfers. Dit had een impact op webtoepassing globale cijfers en gezinssamenstelling voor alle jaartallen.
- Herinladen van de gegevens van het NIC : het bestand bleek onvolledig. Dit had een impact op alle webtoepassingen voor de jaren 2016-2017.
- Harmonisatie van de nacecodes voor de jaren 2006-2007 in de webtoepassing globale cijfers : deze jaren bevatten een mengeling van nacecodes versie 2003 en 2008. Alle codes werden nu omgezet naar de versie 2003.

In de webtoepassing globale cijfers zijn de cijfers over THAB vanaf 2017 niet meer opvraagbaar. De reden hiervan is de staatshervorming (zie supra). De cijfers zullen opnieuw opvraagbaar zijn wanneer de data over de opvolger van THAB voor alle regio's opgenomen is in het DWH AM&SB.

In de webtoepassing globale cijfers is vanaf 2018 de bijkomende socio-economische positie (werkend als) PWA niet meer opvraagbaar voor Vlaanderen en de Duitstalige Gemeenschap.

De plaats van tewerkstelling zal worden opgenomen in de webtoepassing globale cijfers en lokale cijfers. Dit is voorzien voor volgend jaar.

Voor Wallonië zal de indeling in RESOC's vervangen worden door Bassins EFE.

### **1.3. Documentatie en het CMS :**

Er werd een pagina aangemaakt die gebruikers moet helpen om vlugger informatie terug te vinden op de website van het DWH AM&SB. Er werd speciaal aandacht besteed aan het verband tussen instellingen, bronnen, variabelen en codes.

Volgende aanpassingen moeten nog steeds worden uitgevoerd :

- Beknopte beschrijving onder de bronnen en minder cryptische namen.
- Tooltips bij de instelling met een korte beschrijving van de instelling.
- Een link of knop op de pagina's van de documentatie waarmee de gebruikers een probleem (bijvoorbeeld foutieve info, verouderde informatie of niet werkende url's) op de pagina kunnen melden.

Een vernieuwing van de themapagina met zoekmogelijkheden op basis van thema's en subthema's, naar het voorbeeld van de portaalsite van de sociale zekerheid, wordt overwogen.

De nieuwe bronnen zullen worden gedocumenteerd.

## 2. Stand van zaken projecten

### 2.1. Onderzoeksproject: indicatoren ontwikkelen die gewoonlijk berekend worden a.d.h.v. enquêtegegevens

#### 2.1.1. Het beschikbaar inkomen op huishoudniveau

Om een aantal indicatoren te ontwikkelen op basis van administratieve gegevens die gewoonlijk berekend worden op basis van enquêtegegevens (zoals bijvoorbeeld de 'AROP'-indicator of de 'relative median income of elderly'-indicator), zijn variabelen over *het beschikbaar inkomen op huishoudniveau* op basis van administratieve data onontbeerlijk.

We gebruiken de HY020-variabele van de Belgische SILC-enquête als referentiepunt voor het ontwikkelen van variabelen over *het beschikbaar inkomen op huishoudniveau* op basis van data uit de IPCAL-databank en het DWH AM&SB. Concreet ontwikkelen we volgende inkomensvariabelen:

- *Basisvariabele*: het beschikbaar inkomen op huishoudniveau
- *Aanvullende variabelen*: afzonderlijke (netto) inkomensvariabelen per huishouden bijvoorbeeld het beschikbaar inkomen uit ouderdomsuitkeringen per huishouden, het beschikbaar inkomen uit loondienst per huishouden etc.

In de voorbije maanden hebben we verder gewerkt aan twee nota's die onderzoeken op welke wijze we bovenstaande inkomensvariabelen kunnen reconstrueren voor de gehele populatie gekend in het DWH AM&SB. De conceptuele nota neemt de HY020-variabele van de BE-SILC als uitgangspunt en onderzoekt de mogelijkheden voor administratieve reconstructie. De methodologische nota beschrijft vervolgens in detail de methodologische keuzes die we maken voor de administratieve reconstructie. Uit deze nota's kunnen we concluderen dat een exacte reconstructie van de HY020-variabele niet mogelijk is omdat de administratieve databronnen bepaalde beperkingen kennen (o.a. het ontbreken van informatie over studiebeurzen). Bovendien is een exacte reconstructie ook niet wenselijk omdat wij soms doelbewust afwijkende keuzes maken (o.a. het niet opnemen van tweedepijlerkapitalen als recurrent inkomen).

Daarnaast hebben we een voorlopige versie van de basisvariabele ontwikkeld. Deze voorlopige variabele is gebaseerd op een eerste (beperkte) gegevensaanvraag die werd ingediend in de zomer van 2019. Alhoewel de resultaten van deze voorlopige variabele veelbelovend zijn, kunnen we geen definitieve conclusies trekken omdat de voorlopige variabele geconstrueerd werd met onvolledige gegevens. Het opstellen van de voorlopige variabele heeft wel een belangrijke bijdrage geleverd aan het verder uitwerken van de methodologie.

Tot slot hebben we een tweede gegevensaanvraag ingediend. Deze gegevensaanvraag bevat, naast gedetailleerde IPCAL-codes en variabelen uit het DWH AM&SB, ook een koppeling met de SILC-enquête. De doelstelling van de tweede gegevensaanvraag is tweeledig. Deze gegevens worden allereerst gebruikt om de definitieve inkomensvariabelen te ontwikkelen. Ten tweede gebruiken we deze gegevens om de complementariteit tussen enquêtegegevens van de SILC en de administratieve inkomensvariabelen te analyseren met als doel om de gebruikers goed te



informereren over de eventuele beperkingen van de door ons ontwikkelde inkomensvariabelen. Eind oktober 2020 werden de eerste gegevens ontvangen. We verwachten dat alle data zullen worden opgeleverd tegen het einde van het jaar.

### **2.1.2. Tewerkstellingsindicator**

In de voorbije maanden hebben we de resultaten van de tewerkstellingsgraden in het Datawarehouse een eerste keer besproken met de financierders van het wetenschappelijk project voor het Datawarehouse. Op basis van dit overleg wijzigden we de tellingswijze van werkende studenten in de tewerkstellingsgraden. De nieuwe resultaten van de tewerkstellingsgraden (met de nieuwe tellingswijze van werkende studenten) hebben we vervolgens geïntegreerd in de begeleidende nota bij de tewerkstellingsindicatoren. In de resultatensectie van de nota bespreken we de globale tewerkstellingsgraad voor de 20- tot 64-jarigen in België en splitsen deze vervolgens uit voor subgroepen naar leeftijd, regio, geslacht, nationaliteit, geboorteland en het huishoudtype. Daarnaast bespreken we ook de verdeling van de loontrekkenden uit beide Datawarehouse indicatoren naar arbeidsregime. In de bespreking van de resultaten namen we ook telkens de tewerkstellingsgraad op volgens de EAK. Zo zijn we op algemene wijze nagegaan hoe de tewerkstellingsgraden van het Datawarehouse zich verhouden tot de tewerkstellingsgraad van de EAK. De begeleidende nota bij de tewerkstellingsindicatoren zal voorgelegd worden aan de financierders en aan Statbel. Op basis van de feedback die daaruit voortkomt, kunnen de tewerkstellingsgraden verder gefinaliseerd worden.

### **2.1.3. NEET-jongeren**

Om NEET-jongeren af te bakenen op basis van de gegevens uit het Datawarehouse is het aangewezen om nieuwe databronnen over onderwijs- en opleidingsdeelname in het Datawarehouse te integreren. In de voorbije maanden werd geen vooruitgang geboekt in de integratie van deze databronnen in het Datawarehouse. We geven hieronder een stand van zaken mee over dit integratieproces.

Voor de onderwijsdeelname in de Franse Gemeenschap werden eerder dit jaar al de gegevens van ARES (Franstalig niet-universitair hoger onderwijs) en de Cref (Franstalig universitair onderwijs) ingeladen in het Datawarehouse. De gegevens van ARES zijn beschikbaar voor de jaren 2017 en 2018. De data van de Cref gaan van 2018 terug tot in de jaren 1980. Van de Fédération Wallonie-Bruxelles (FWB) ontving het onderzoeksteam eerder dit jaar een lijst van gegevens die de FWB bijhoudt over het leerplichtonderwijs van de Franse Gemeenschap. Het overleg met de FWB om hun gegevens in te laden in het Datawarehouse wordt verdergezet. Voor de onderwijsdeelname in de Vlaamse Gemeenschap werden van het Departement Onderwijs & Vorming nog geen data ontvangen. Ook het overleg met het Departement Onderwijs & Vorming om hun gegevens in te laden in het Datawarehouse wordt verdergezet.

## **2.2. Onderzoeksproject: een meet- en opvolgingsinstrument voor regionale tewerkstellingsmaatregelen**

Dit project betreft de ontwikkeling van een meet- en opvolgingsinstrument voor de regionale tewerkstellingsmaatregelen. Het heeft als doel (1) de regionale tewerkstellingsmaatregelen te inventariseren, (2) die te analyseren met het oog op de relevantie voor het Datawarehouse en (3) na te gaan in welke mate een harmonisatie van de verschillende maatregelen (waarvan een groot deel ten gevolge van de zesde staatshervorming is overgeheveld naar de bevoegdheid van de regionale overheden) haalbaar is.

Ter voorbereiding van de integratie van de gegevens over de regionale tewerkstellingsmaatregelen in het Datawarehouse, hebben we een inventaris opgemaakt waarin de titels van de tewerkstellingsmaatregelen van elke regionale en federale instelling van vóór en na de zesde staatshervorming worden opgenomen.

Gezien de breedte van het scala aan tewerkstellingsmaatregelen in België, hebben we vervolgens beslist om een selectie te maken van de tewerkstellingsmaatregelen die we wensen toe te voegen aan het Datawarehouse. Er werd beslist om daarbij in vijf fasen te werken.

In de eerste fase analyseerden we de maatregelen die zijn opgenomen in de Labour Market Policies (hierna LMP) databank van de Europese Commissie. We vergeleken de maatregelen die in de LMP-databank worden opgenomen met die dat we in onze inventaris hebben toegevoegd. Aangezien de LMP-databank gegevens bevat over minder maatregelen dan onze inventaris, vonden we bijna alle maatregelen hieruit terug in de inventaris. Slechts een klein aantal maatregelen uit deze databank werden bijgevolg toegevoegd aan onze inventaris. Deze fase is reeds afgerond.

In de tweede fase, dat momenteel aan de gang is, nemen we contact op met de instellingen om te vragen voor welke maatregelen (opgenomen in de inventaris en/of de databank van de LMP) ze over individuele data (namelijk alle gegevens die aan een natuurlijke persoon gekoppeld kunnen worden) beschikken. Voor het merendeel van de instellingen hebben we inmiddels een overzicht van de beschikbare data.

In de derde fase ontwikkelen we een “tijdslijn van tewerkstellingsmaatregelen” (gebaseerd op de tijdslijn die de LMP gebruikt om diens maatregelen te organiseren). Hierin worden de maatregelen opgenomen waarvoor instellingen over individuele data beschikken, aangevuld met bijkomende informatie. Dit is reeds gebeurd voor de maatregelen die in de databank van de LMP zijn opgenomen en waarvoor de instellingen over individuele data beschikken.

In de vierde fase zullen de instellingen de tijdslijn nakijken en aanvullen waar nodig.

In de vijfde fase zal de werkgroep bepalen welke maatregelen het wenst op te nemen in het Datawarehouse.

### **2.3. Onderzoeksproject: informatie over het scholingsniveau van de bevolking**

De financierders van het wetenschappelijk project voor het Datawarehouse formuleerden de vraag om na te gaan hoe de informatie over het scholingsniveau van de bevolking in het Datawarehouse kan worden verbeterd. Het initiële opzet voor de verbetering van de informatie over het scholingsniveau was om het scholingsniveau zoals bepaald door de FOD WASO & UNIA (2019) over te nemen in het Datawarehouse en te verfijnen waar mogelijk.<sup>1</sup> In de loop van het project werd evenwel duidelijk dat er sterke gelijkenissen zijn tussen deze opdracht van het project wetenschappelijke ondersteuning Datawarehouse en het project van Statbel over de creatie van een onderwijsniveau op basis van administratieve data. We besloten daarom eerder dit jaar de piste te onderzoeken van een integratie in het Datawarehouse van het onderwijsniveau volgens Statbel.

Hiertoe hebben we in een nota de grote lijnen samengevat van de door Statbel ontwikkelde methodologie om het onderwijsniveau te bepalen. De nota bevat naast de bespreking van de methodologie van Statbel ook een vergelijking van het globale resultaat van het onderwijsniveau volgens de Statbel- en de FOD WASO-methodologie. Deze nota werd bezorgd aan de financierders en aan Statbel. Op basis van de reacties van de verschillende partijen zal een agenda worden opgesteld om een eventuele integratie van de Statbel-gegevens in het Datawarehouse te bespreken.

---

<sup>1</sup> FOD WASO & UNIA (2019). *Socio-economische monitoring. Arbeidsmarkt en origine 2019*. Brussel: FOD WASO & UNIA.

## 2.4. Onderzoeksproject: indicatoren voor een boordtabel sociale zekerheid

De Universiteit Antwerpen heeft een boordtabel ontwikkeld die aangeeft rond welke thema's indicatoren moeten worden ontwikkeld om het sociaal beleid in België te monitoren. De medewerkers van het Datawarehouse hebben de taak om een aantal van deze indicatoren te ontwikkelen op basis van administratieve gegevens uit het DWH AM&SB. Momenteel wordt er gewerkt aan het ontwikkelen van indicatoren over het aantal uitkeringsontvangers. Hiertoe werd een nota opgeleverd over de te ontwikkelen indicatoren over uitkeringsontvangers.

In de nota stellen we voor om in de eerste fase indicatoren te ontwikkelen over het aantal uitkeringsontvangers aan de hand van *een basisindicator* en *enkele subindicatoren*. De *basisindicator* duidt aan hoeveel individuen en huishoudens één of meerdere uitkeringen van de Belgische sociale bescherming ontvangen. De *subindicatoren* maken een opdeling naar het aantal uitkeringsontvangers in de twee systemen van sociale bescherming (cf. sociale zekerheid en sociale bijstand), het aantal uitkeringsontvangers in de verschillende takken van de sociale zekerheid en geven meer inzicht in de situatie van multi-uitkeringsontvangers. Dit laatste betreft de uitkeringsontvangers die meerdere uitkeringen van de Belgische sociale bescherming combineren op éénzelfde dag. Omwille van deze reden geven alle indicatoren het aantal uitkeringsontvangers weer dat recht heeft op één uitkering (of meerdere uitkeringen in geval van multi-uitkeringsontvangers) op de laatste dag van ieder kwartaal in het desbetreffende jaar.

In de tweede fase ontwikkelen we aanvullende indicatoren over de hoogte van de ontvangen uitkeringen van uitkeringsontvangers. Concreet construeren we geaggregeerde statistieken over de gemiddelde hoogte en de mediane hoogte van de ontvangen uitkeringen van uitkeringsontvangers. Twee bijkomende indicatoren worden hiertoe ontwikkeld, met name een gemiddelde en een mediaan voor alle indicatoren over het aantal uitkeringsontvangers (cf. de basisindicator en de subindicatoren). De aanvullende indicatoren met betrekking tot de hoogte van de ontvangen uitkeringen van multi-uitkeringsontvangers vereisen wat meer uitleg. Naast het presenteren van de gemiddelde/mediane hoogte van het geheel aan uitkeringen van multi-uitkeringsontvangers, duiden we aan welke uitkering de hoofdkomensbron is en welke uitkeringen supplementaire inkomensbronnen zijn.

De gegevensaanvraag werd opgesteld in samenwerking met de financierders en werd ingediend in april 2020. Ondertussen hebben we de contracten ondertekend waardoor de gegevensaanlevering van start kan gaan.

## **2.5. Onderzoeksproject: variabelen m.b.t. de socio-economische achtergrond**

Dit project betreft de ontwikkeling van variabelen over de socio-economische context van mensen. Het onderzoekcentrum METICES van de Universit  Libre de Bruxelles (ULB) voert momenteel, in samenwerking met de Federale Overheidsdienst Werkgelegenheid, Arbeid en Sociaal Overleg (FOD WASO), een studie uit dat de ontwikkeling van indicatoren omtrent diversiteit en posities op de arbeidsmarkt als doel heeft (socio-economische monitoring). Gezien het belang van empirische analyses naar de arbeidsmarktposities en -situaties van mensen en de samenhang daarvan met de socio-economische achtergrond en -levensloop voor hedendaags onderzoek, trachten we na te gaan of  n of meerdere variabelen met betrekking tot de ‘socio-economische herkomst’ van een persoon kan worden ontwikkeld op basis van de gegevens van het Datawarehouse.

Om de ontwikkeling van dergelijke variabelen te verwezenlijken, maken we gebruik van de gegevens van twee databanken, namelijk die van het Datawarehouse Arbeidsmarkt & Sociale Bescherming en die van Statbel. Eind 2018 werd een gegevensaanvraag ingediend. De gegevens werden eind 2019 (DWH AM & SB) en in januari 2020 (Statbel) geleverd. Omdat we in eerste instantie verkennend onderzoek uitvoeren op basis van een steekproef van de Belgische bevolking, selecteerden we voor de gegevens van 2015  n procent uit de groep van personen van 18 tot 64 jaar. De steekproef is representatief voor gender, leeftijdsgroep, regio, herkomst en socio-economische status. Omdat we onderzoek doen naar de samenhang tussen de achtergrond of de levensloop van een individu en diens latere arbeidsmarktpositie, hebben we gegevens geselecteerd over gegeven personen voor de ‘referentiejaar’ waarin deze 8, 10, 12, 14 of 16 jaar oud waren.

In een eerste nota werden de beschikbare gegevens en de eerste beschrijvende analyses gepresenteerd. Deze laatste bieden een eerste visualisatie met betrekking tot de samenhang tussen de socio-economische positie van een gegeven persoon in 2015 (gedefinieerd volgens de nomenclatuur van het KSZ) en dat van diens ouders voor een bepaald jaar in de jeugd van die persoon.