# Checklist grondige controle databank

**Check zowel externe als interne themaboom**

## Themaboom

* Is de themaboom logisch opgebouwd?
* Is voor alle onderwerpen een metadatafiche aanwezig?
* Zijn de regels voor de opmaak gerespecteerd: hoofdthema in HOOFDLETTERS, rubriek begint met een Hoofdletter, onderwerp met kleine letter?
* Werd er een percentage berekend waar nodig?
* Staan de percentages in een submap indien er meer dan 12 onderwerpen in totaal zijn?
* Naam van de submap is percentages indien enkel percentages in de map, anders relatieve cijfers
* Dupliceren van onderwerpen indien relevant[[1]](#footnote-1)
* Als de data op statsec beschikbaar is, staat het ook in de map GEBIEDSINDELINGEN > data op statsec? Is de data op statsec ook beschikbaar zonder inloggen?

## Onderwerpnamen

Zie informatie over de belangrijkste velden op [Github](https://github.com/provinciesincijfers/JiveDocumentation/blob/master/03.%20Onderwerpentabel/Informatie%20over%20de%20belangrijkste%20velden.md).

* Zichtbaarheid van mapnamen in presentaties: “Visible in title” uitvinken indien nodig (menu Edit achter rechtermuisklik op map)[[2]](#footnote-2)
* Begint het onderwerp met het kenmerk, bv. vrouwen 18-64 jaar?
* Werd het juiste teken gebruikt? (percentageteken enkel bij een echt percentage)
* Geen eenheid (aantal / aandeel) gebruiken in de onderwerpnaam. Woord gemiddelde enkel gebruiken in onderwerpnaam indien nodig voor leesbaarheid (vb "gemiddelde temperatuur" wél, maar gewoon "E-peil appartementen"
* Percentages, promilles, ratio’s: (t.o.v. …) toevoegen in de onderwerpnaam
* Consequente schrijfwijze:
  + 1.000 (met puntje)
  + 65+ (in naam en achter t.o.v.)
  + 18-64 jaar behalve bij "t.o.v." indien er geen kenmerk bij staat, dan 18-64-jarigen
  + € (niet EUR of euro)
* Bij prognoses: achteraan [projecties] toevoegen
* Kortere versies zijn toegestaan in naam onderwerp in databank, o.w.v. beperkt aantal tekens (in dashboard wordt alles voluit geschreven o.w.v. de leesbaarheid)

## Beschrijvingen

Zie informatie over de belangrijkste velden op [Github](https://github.com/provinciesincijfers/JiveDocumentation/blob/master/03.%20Onderwerpentabel/Informatie%20over%20de%20belangrijkste%20velden.md).

* **Begint** bij een berekend onderwerp met [percentage], [promille], [ratio] en bij een prognose met [projecties]
* Beschrijving formule bij formuleonderwerp, steeds (t.o.v. noemer) vermelden, voluit schrijven: percentage, promille (niet % of ‰), per 10.000, ratio, index
* Schrijfwijze:
  + 1.000 (met puntje)
  + 75-jarigen (met streepje)
  + 65 jaar en ouder (niet 65+ of 65-…))
  + €

## Data

* **Zijn de cijfers juist?** Vergelijk de nieuwe jaargang met vorige jaargangen – zijn de evoluties plausibel? Doe dit zeker ook op subgemeentelijk niveau indien van toepassing. Indien meerdere jaren vernieuwd worden: vergelijk de cijfers in de admin/jive met de reeds gepubliceerde cijfers. Vergelijk de ingelezen cijfers met de ruwe data of eventuele elders gepubliceerde cijfers. Controleer de waarden van door jou gekende regio of gemeenten. Lijken de resultaten geloofwaardig?
* Is het onderscheid tussen de waarde “0” en “niet gekend” juist? Zie afspraken over [Missing values](https://github.com/provinciesincijfers/JiveDocumentation/blob/master/01.%20Algemeen%20databeheer/Missing%20values.md)
* Worden in de legende de eenheid en de naam van de bron correct weergegeven? Oorspronkelijke bron | provincies.incijfers.be of Oorspronkelijke bron via Verwerker | provincies.incijfers.be of Naam bron 1 & Naam bron 2 | provincies.incijfers.be
* Consequent gebruik van eenheid (en combinatie met datatype), bijvoorbeeld bij oppervlakte. Gebruik Mean / Ratio
* Standaard afronding op 0,1, uitgezonderd bij promille: 0,01. Uitzonderingen mogelijk op basis van de indicatoren
* Zijn de data voor alle Vlaamse gemeenten ingeladen? Als er data beschikbaar zijn voor Brussel, zijn deze dan ook ingeladen?
* Als de data subgemeentelijk beschikbaar is: zijn ze ZONDER INLOGGEN beschikbaar? Check ook subgemeentelijk wat je op gemeentelijk niveau zou doen.
* Zijn de data ingeladen voor alle jaren (of maanden) of ontbreken er periodes?
* Controleer in de lijst Indicators of er geen Perioden gedefinieerd zijn waar dat niet hoort? (begin- en eindjaar mogen doorgaans niet ingevuld zijn in formule-onderwerpen)
* Komen de totalen overeen met de subcategorieën (waar relevant) / is de som van aandelen gelijk aan 100%?
* Zijn eventuele ingelezen totalen op bijvoorbeeld provincieniveau[[3]](#footnote-3) plausibel in functie van de data ingelezen op gemeenteniveau.
* Zitten evolutiegrafieken en/of -tabellen logisch in elkaar?
* Kloppen klassenindelingen van kaarten?
* Kloppen de dimensies bij de kubussen (bv. geen dubbels)?
* Zijn de gebruikersvoorwaarden ingevuld bij de bron (link naar uitleg data-licentie oorspronkelijke bron)?
* Kijk bij ‘Informatie over…’ de rubrieken na van het tabblad eigenschappen (zoals eenheid, startperiode, drempelwaarde).

## Metadatafiches (tabblad informatie)

* Werd de fiche opgesteld op basis van en conform aan het sjabloon voor de metadata?
* Omvat de titel van het Informatie-tabblad zowel het thema als de bron van de data?
* Zijn alle teksten begrijpelijk voor mensen die niet vertrouwd zijn met de data?
* Omvat het veld ‘definitie’, naast de inhoudelijke definitie van het onderwerp (bv. wat is het totale netto belastbare inkomen), ook een definitie van de data of de statistiek zelf (bv. opgesteld op basis van de aangiften in de personenbelasting), dus de opbouw van de statistiek, doelstelling, eventuele problemen…?
* Worden de verschillende indicatoren (onderwerpen) beschreven?
* Worden afkortingen 1 keer voluit geschreven?
* Wordt voor aandelen duidelijk aangegeven wat teller(s) en noemer(s) zijn?
* Wordt het referentiemoment aangegeven zoals afgesproken?
* Wordt bij dataleverancier voldoende gedetailleerd vermeld wat de bron van de data is? (bv. niet enkel ‘Statbel’, maar ook ‘landbouwtelling’)
* Wordt indien relevant een veld ingevuld in verband met de accuraatheid van de data (wanneer gegevens niet altijd overeenkomen met de werkelijke situatie bv. omwille van de registratieprocedure) en/of de vergelijkbaarheid in de tijd (wanneer er een knik is in de gegevensreeks bv. omwille van een verandering in de manier van registreren of categorieën die worden meegenomen)?
* Wordt bij verwerking duidelijk weergegeven welke verwerkingen door wie gebeuren?
* Verwijzen alle externe linken naar de juiste pagina’s? Verwijst de link voor meer info en contact naar de juiste werkgroep (D&A, economie, landbouw…)?
* Kijk na op taalfouten.

## 

## Rapporten/steekkaarten

Voor een volledig document, [klik hier](https://github.com/provinciesincijfers/JiveDocumentation/blob/master/07.%20Rapporten/Draaiboek%20rapport%20aanmaken.docx).

## Dashboard

Voor alle afspraken, [klik hier](https://github.com/provinciesincijfers/JiveDocumentation/blob/master/08.%20Dashboard/Draaiboek%20dashboard.docx).

* Is het dashboard logisch opgebouwd?
* Horen alle onderwerpen erin thuis en/of ontbreekt een relevant onderwerp?
* Zijn er geen actuelere data beschikbaar (in de databank of elders)?
* Kloppen de onderwerpnaam en eenheid met de databank? (aangezien deze in het dashboard manueel worden aangepast)
* Kijk de verschillende cijfers na en vergelijk met deze met de cijfers in de databank (check dat er niet per ongeluk een foutief onderwerp werd geselecteerd)
* Werken alle links naar rapporten, databank, metadata…?
* Als je een nieuwe bron hebt toegevoegd, heb je die dan ook aan het Bronnenoverzicht toegevoegd?

## Vragen

* Onderwerpen nog niet gepubliceerd (=ok), maar wel al zichtbaar via gebiedsindeling

1. Cijfers beschikbaar op statistisch sectorniveau moeten in de themaboom ook onder *Gebiedsindelingen > Beschikbaar vanaf statistische* sector terug te vinden zijn. Wanneer een indicator relevant is voor twee thema’s in de themaboom moet deze ook op beide plekken kunnen worden teruggevonden (vb. minderjarigen in gezin met minstens 1 werkloze ouder, moet zowel onder *Arbeidsmarkt > Werkloosheid* als onder *Welvaart en armoede* terug te vinden zijn). [↑](#footnote-ref-1)
2. Hiermee wordt NIET verwezen naar het vakje ‘visible’ in de indicatorenlijst van studio maar naar die van het thema in de CategoryTree  
     
   Uitleg: wanneer in themaboom twee indicatoren geselecteerd worden, krijgt presentatie de titel van de eerst bovenliggende gezamenlijke map. Wanneer dit niet nodig is, moet dit uitgevinkt worden (bv. Bij nietszeggende mapnaam als ‘aantal’ of ‘thema’.) [↑](#footnote-ref-2)
3. Meestal worden cijfers slechts op statistisch sectorniveau (of gemeenteniveau) ingelezen en nadien automatisch geaggregeerd binnen Swing. [↑](#footnote-ref-3)