# Handleiding Schoolrapportages

In deze handleiding wordt beschreven hoe je van het SPSS databestand van de gezondheidsmonitor jeugd komt tot de schoolrapportages in powerpoint.

## Stap 1 – Runnen indicatorensyntaxen

Ten behoeve van het aanmaken van de indicatoren, moeten alle indicatorensyntaxen op het regionale bestand worden gerund.

## Stap 2 – Runnen opschoonsytaxen

Om geschoonde data te gebruiken voor het schoolprofiel, moet de regionale opschoonsyntax worden gerund over het databestand.

Daarnaast is het nodig om de syntax voor het opschonen van de extreme waarden te runnen. Dit is onderdeel van de landelijke opschoonsyntax.

## Stap 3 – Trendbestanden

In het schoolprofiel staan ook trendfiguren. Als een desbetreffende school eerder heeft meegedaan, dan kun je deze figuren dus ook gebruiken. Anders kun je op deze plek andere grafieken invoegen.

Voor het uitdraaien van resultaten van 2021, 2019 en 2015 zijn ook stukjes opgenomen in het R-script. Het R-script werkt zowel op basis van een trendbestand waar alle jaren instaan als de losse databestanden voor ieder jaar. Let op: het is noodzakelijk dat de indicatoren waarvoor trends moeten worden gemaakt in alle drie de spss databestanden (2015, 2019 en 2021) dezelfde naam hebben en hetzelfde zijn opgebouwd. Let hierbij ook dat scholen dezelfde naam heeft over de jaren. Hiervoor kunnen de aangeleverde trendsyntaxen worden gebruikt.

## Stap 4 – Benodigde bestanden

Zet de volgende documenten in één map:

1. Configuratie.xlsx
2. Template.pptx
3. Rapportage.R

## Stap 5 – Configuratie Excel bestand

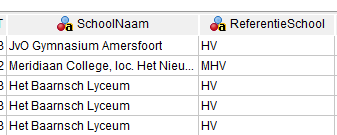
Tip: Wil je wijzigingen aanbrengen in de configuratie? Run eerst het R-script met de standaardconfiguratie en met één school. Mocht je dan tegen problemen aanlopen dan kun je uitsluiten dat de problemen veroorzaakt worden door wijzigingen in de configuratie.

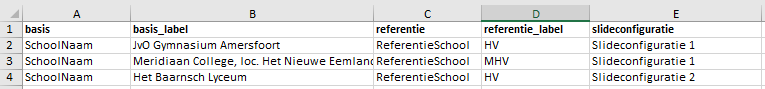
Het Excel bestand ‘1. Configuratie.xlsx’ bestaat uit een tabblad op basis waarvan het R script de PowerPoint rapportage aanmaakt. Je hoeft in dit bestand geen wijzigingen aan te brengen (maar het kan wel). Het configuratiebestand bevat twee tabbladen; ‘Rapportconfiguratie’ en ‘Slideconfiguratie 1’.

In het tabblad 'Rapportconfiguratie' kun je per school aangeven welke slideconfiguratie je wenst te gebruiken. Hier heb je tevens de mogelijkheid om aan te geven met welke referentie je de school wilt vergelijken, zoals bijvoorbeeld alle scholen in de regio of alleen met leerlingen die VMBO-onderwijs volgen. Alle scholen komen hier dus in te staan.

Het tabblad ‘Rapportconfiguratie’ bestaat uit de volgende kolommen:

1. **basis.** Hier geef je het basisniveau aan. In het geval van scholen is dit niveau de school, maar het kan bijvoorbeeld ook de gemeente zijn. Zorg dat de variabelenaam die je hier opgeeft overeen komt met de variabelenaam in het SPSS bestand. In het voorbeeld hieronder ‘SchoolNaam’.
2. **basis\_label.** Hier geef je de naam van de school weer zoals deze in het SPSS bestand is weergegeven onder de hierboven opgegeven variabele. Dit is dan een tekst (string) variabele en niet een numerieke waarde met waardelabels (value labels).
3. **referentie.** Hier geef je het referentieniveau aan. Als een school bijvoorbeeld alleen vmbo-leerlingen heeft, wil je deze mogelijk vergelijken met alle andere vmbo-leerlingen in de regio. Zorg dat de variabelenaam die je hier opgeeft overeen komt met de variabelenaam in het SPSS bestand. De naam die je hier geeft voor de referentie mag je zelf bepalen maar wordt niet getoond in de legenda. Daar staat standaard ‘Regio’ en dit wordt in de toelichting toegelicht. Als je hier iets anders wilt dan kan dit in het R-script aangepast worden.
4. **referentie\_label.** In dit geval geef je de naam op van de referentie zoals je deze onder de referentie variabele hebt gespecificeerd. Zorg ervoor dat de tekst precies overeenkomt met de referentietekst in het SPSS bestand. Ook dit is een tekst variabele.
5. **configuratie.** Hier kun je per rapport aangeven welke slideconfiguratie moet worden gebruikt. Als er voor een bepaalde school een afwijkende slideconfiguratie nodig is, kun je het tabblad 'Slideconfiguratie 1' bijvoorbeeld kopiëren en vervolgens aanpassen naar behoefte. De naam van het tabblad is flexibel, maar het is essentieel dat deze overeenkomt met de naam die wordt ingevoerd onder deze kolom. Dit kun je bijvoorbeeld gebruiken als er voor een paar scholen geen trends aanwezig zijn en je in plaats van trendfiguren een staafgrafiek wilt. Let op: Ondanks dat je verschillende slideconfiguraties kan gebruiken is het wel van belang dat de labels (zie onder punt 12 hieronder) hetzelfde blijven.





Op het tabblad 'Slideconfiguratie 1' (en eventueel andere slideconfiguratie tabbladen) wordt voor elke indicator aangegeven hoe de cijfers worden weergegeven in het rapport, bijvoorbeeld als percentage of in een figuur. Daarnaast wordt per indicator aangegeven op welk niveau het cijfer moet worden gepresenteerd, en of er eventuele uitsplitsingen zijn naar gender of klas.

Het tabblad ‘Slideconfiguratie 1’ bestaat uit de volgende kolommen:

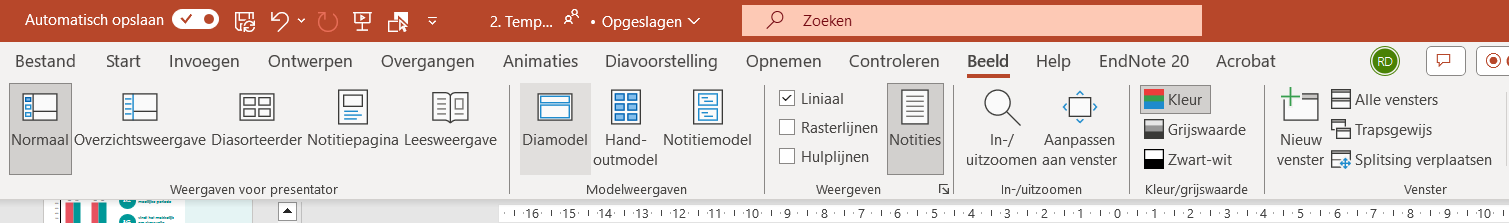
1. **vraag.** Hierin zie je of de indicator uit de landelijke vragenlijst of uit de extra OKO module komt.
2. **omschrijving.** Een korte beschrijving van de indicator. Deze omschrijving wordt bij een aantal grafiektypes als titel gebruikt. Wil je deze grafiektitels aanpassen, pas dan de omschrijving aan.
3. **type.** Deze kolom bepaalt de manier waarop een indicator wordt gepresenteerd. In de eerste rij staat onder type rapportnaam. Op deze manier komt op elk rapport de naam van de school te staand. Verder zijn de volgende types mogelijk:
   * **percentage.** Indicator wordt weergegeven als los percentage. Elk percentage vraagt om een nieuwe regel.
   * **staafgrafiek.** Indicator wordt weergegeven als staafgrafiek.
   * **staafgrafiek gestapeld.** Indicator wordt weergegeven als liggende, gestapelde staafgrafiek.
   * **combi.** Combineert meerdere indicatoren tot een staafgrafiek.
   * **combi liggend.** Combineert meerdere indicatoren tot een liggende staafgrafiek.
   * **trendgrafiek.** Laat een trend over meerdere jaren zien van een indicator.
   * **top 3.** Geeft de top 3 weer van meerdere indicatoren. Als je een top 5 wilt dan kun je hier top 5 neer. Let wel op dat je dan tenminste 5 indicatoren hebt onder indicator.
   * **Percentage in tekst.** In de tekst worden kunnen percentages worden opgenomen. Zorg ervoor dat je onder de omschrijving de desbetreffende tekst plaatst, en waar het percentage moet komen, plaats je <indicator>. De ‘indicator’ is hierbij dezelfde indicator als degene die je onder indicator hebt aangegeven.
4. **indicator.** De SPSS-naam van de indicator. Het is mogelijk om indicatoren (en omschrijvingen) toe te voegen indien je ook voor andere indicatoren figuren wilt aanmaken. Let op dat je hiervoor ook het PowerPoint template aan moet passen. Bij figuren van het type combi vul je in de cel meerdere indicatornamen in. Zorg ervoor dat tussen elke indicator een puntkomma en een spatie staat (; ). Voorbeeld: Indicator1; Indicator2.
5. **waarden.** Hier geef je per indicator aan van welke waarde je het percentage wilt hebben. In de meeste gevallen heb je te maken met 0en en 1en en wil je het percentage weten van de 1en. Als een indicator uit bijvoorbeeld 1en en 2en bestaat en je bijvoorbeeld het percentage van 2 wilt weten dan vul je hier een 2 in. Als een indicator uit 3 of meer categorieën bestaat dan kun je hier de categoriewaarden aangeven. Hiervoor moet per categorie een aparte rij toevoegen. Let op:bij een gestapelde figuur plaats je hier alle categoriewaardes. Zorg ervoor dat tussen elke waarde een puntkomma staat en GEEN spatie.
6. **labels.** Labels die worden gebruikt bij bepaalde (combi) grafieken omdat de standaard SPSS labels te lang of onduidelijk zijn. Deze zijn vrij aanpasbaar, zolang het aantal (en de volgorde van de) labels gelijk is aan de indicatoren in de kolom ‘indicator’. Let op dat de labels in een cel moeten worden gescheiden door een komma en een spatie (; ).
7. **uitsplitsing.** Geeft aan of een indicator in een grafiek uitgesplitst moet worden. Deze uitsplitsing komt op de x-as van de grafiek.
8. **groepering.** Bepaalt de groepering van de data en wordt weergegeven in de legenda van de figuur. Kan worden aangepast naar een andere dichotome variabele, maar deze functionaliteit is experimenteel.
9. **niveau.** In deze kolom geef je de basis- en referentieniveaus aan die je eerder hebt ingesteld in het tabblad 'Rapportconfiguratie'. Voor elke indicator kun je aangeven op welk niveau de cijfers moeten worden weergegeven. Standaard worden de cijfers in de trendgrafieken op zowel school- als regioniveau getoond. Let er bij het maken van trendfiguren op dat de niveaus in een cel moeten worden gescheiden door een puntkomma en een spatie (; ).
10. **jaar.** Hier geef je aan van welk jaar of jaren je de cijfers wilt laten zien. Let op dat bij meerdere jaren elk jaar in een cel moet worden gescheiden door een komma en een spatie (; ).
11. **index.** Nummer van de slide waarop de figuur/het cijfer komt te staan.
12. **label.** Naam van het label van de tijdelijke aanduiding/placeholder. Zorg ervoor dat elke labelnaam uniek is.

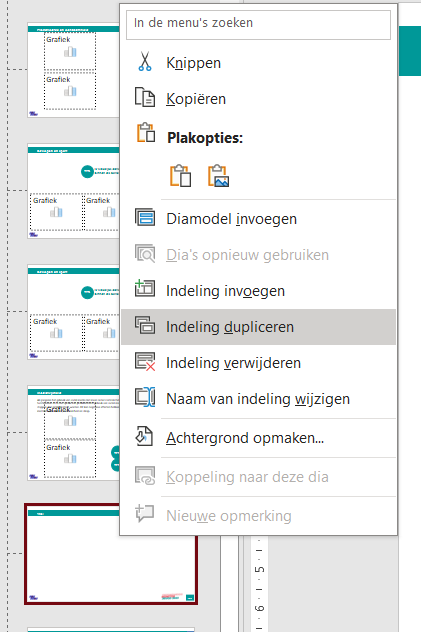
## Stap 6 – Powerpoint template

Het bestand ‘2. Template Rapportage.pptx’ bevat een PowerPoint template. Ook in dit bestand hoef je geen wijzigingen aan te brengen (maar het kan wel). Het is wel aan te raden om in eerste instantie het R script te runnen zonder wijzigingen aan te brengen (om uit te sluiten dat wijzingen aan het template de oorzaak zijn van eventuele errors). In dit template is voor iedere rapportage slide een aparte indeling opgenomen.

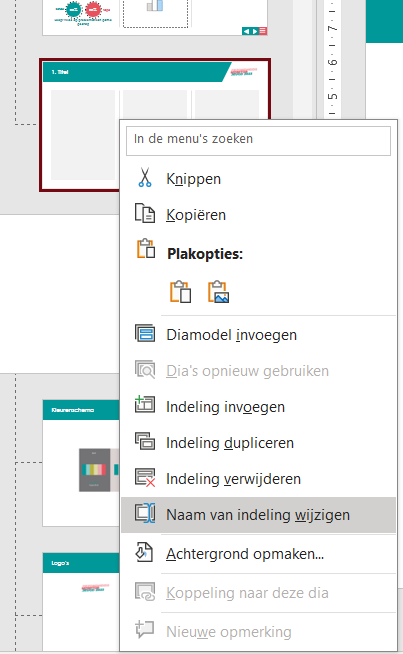
Deze indelingen vind je terug in het diamodel. Alle indelingen zijn volledig aanpasbaar, zowel qua inhoud als layout. De indelingen kunnen bestaan uit twee soorten inhoud. Ten eerste heb je vaste content. Dit zijn titels, tekst, plaatjes en afbeeldingen die op de indelingen zijn geplaatst. Elke slide gebaseerd op een indeling bevat deze vaste content. Ten tweede zijn er tijdelijke aanduidingen (placeholders). Deze tijdelijke aanduidingen zijn in beginsel leeg, maar kunnen worden gevuld vanuit R. De tijdelijke aanduidingen worden vanuit R gevuld met grafieken, cijfers en tekst. In het R-script wordt dan ook de grafiekinstellingen en de kleuren gedefinieerd.

### Nieuwe pagina toevoegen of aanpassen

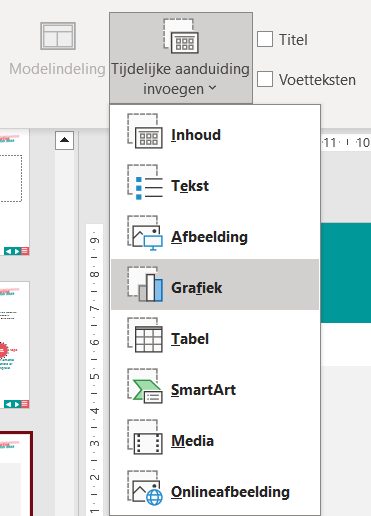
1. Open powerpoint template.
2. Ga via ‘Beeld’ naar het ‘Diamodel’.
3. Onder het diamodel zijn verschillende ‘indelingen’ (vergelijkbaar met slides, maar in het diamodel noemen we het indelingen) te vinden. Scroll naar beneden naar de lege indelingen. Hier vind je ook het kleurenschema en de logo’s.
4. Ga op deze lege indeling staan en dupliceer de indeling door op de rechtermuisknop te klikken en vervolgens ‘Indeling dupliceren’. Verplaats de indeling naar het diamodel erboven.



1. Nu kun je ook deze nieuwe indeling een nieuwe naam geven door op de rechtermuisknop te klikken en vervolgens ‘Naam van indeling wijzigen’.



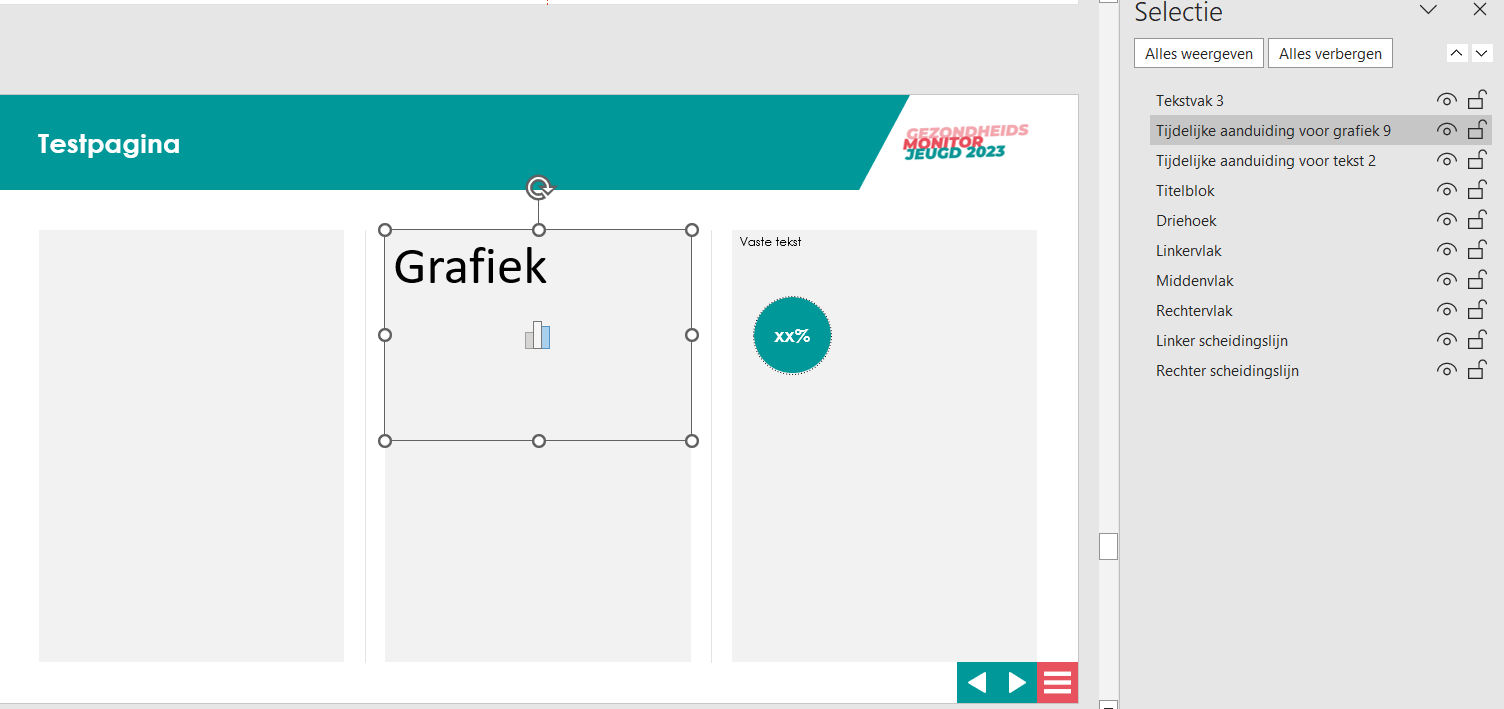
1. Pas de titel aan.
2. Nu kun je inhoud gaan toevoegen. Je kunt nu vaste teksten of pictogrammen invoegen op dezelfde manier als op normale slides. Deze teksten zijn altijd zichtbaar en veranderen niet.
3. Als je grafieken of figuren (tekstbolletjes) met percentages wilt toevoegen dan kun je deze op twee manieren toevoegen. 1) Klik onder modelindeling op ‘Tijdelijke aanduiding invoegen’ (placeholder) en kies vervolgens op grafiek voor figuren en tekst voor percentagebolletjes. 2) Kopieer een tekstbolletje of grafiek van een andere slide en plak deze in de nieuwe slide. In dit geval heb zijn de grootte en vorm van de figuren identiek aan de andere slide.



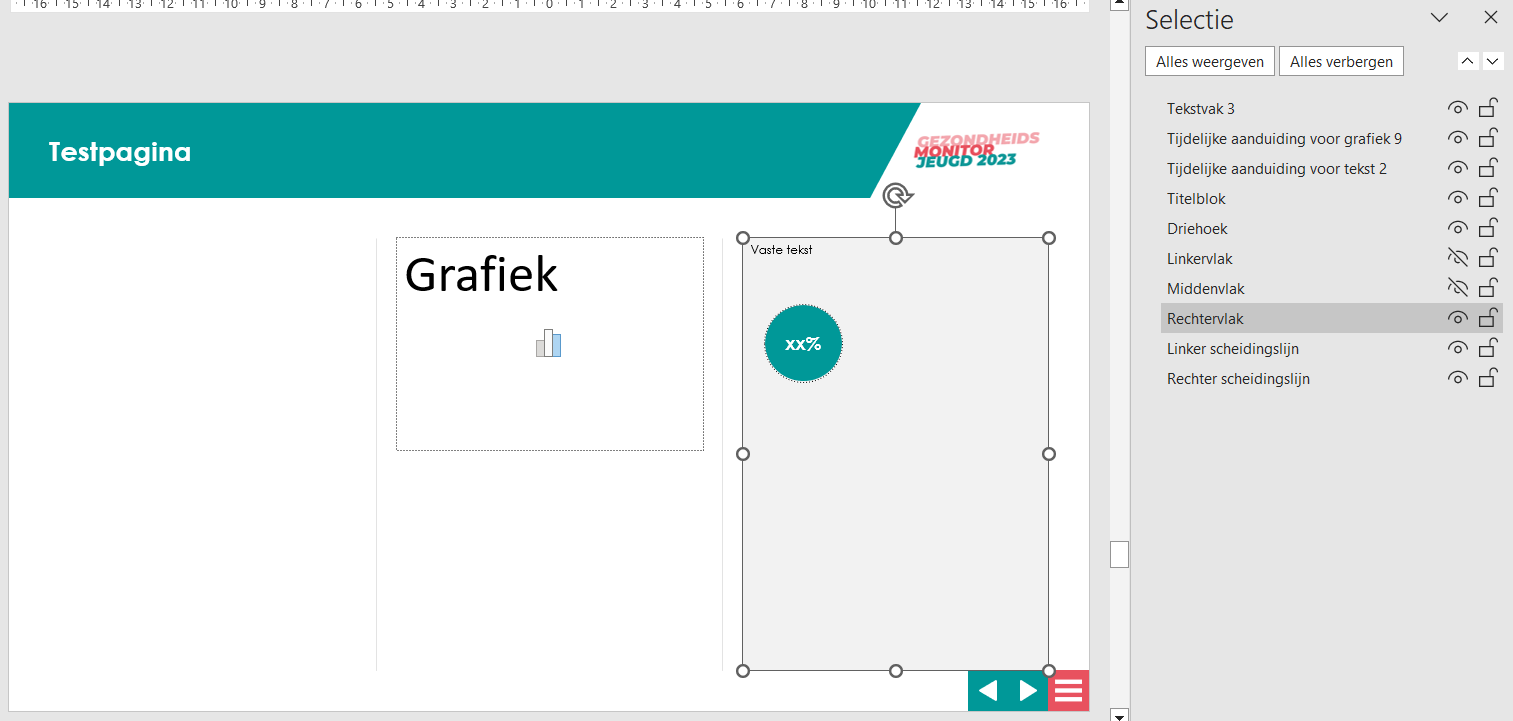
1. Open vervolgens het selectiedoelvenster door ALT+F10 of door boven in het zoekvenster van powerpoint ‘selectievenster’ in te typen en op ‘Het selectiedeelvenster weergeven’ te klikken.

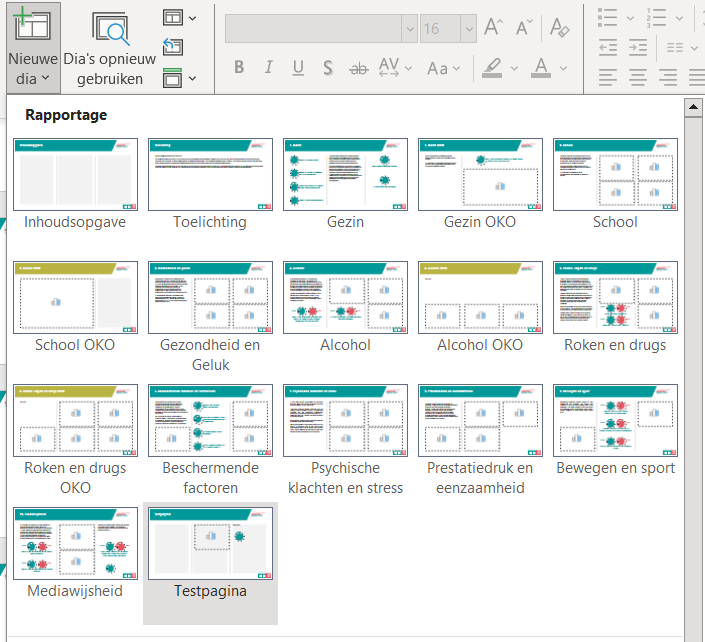
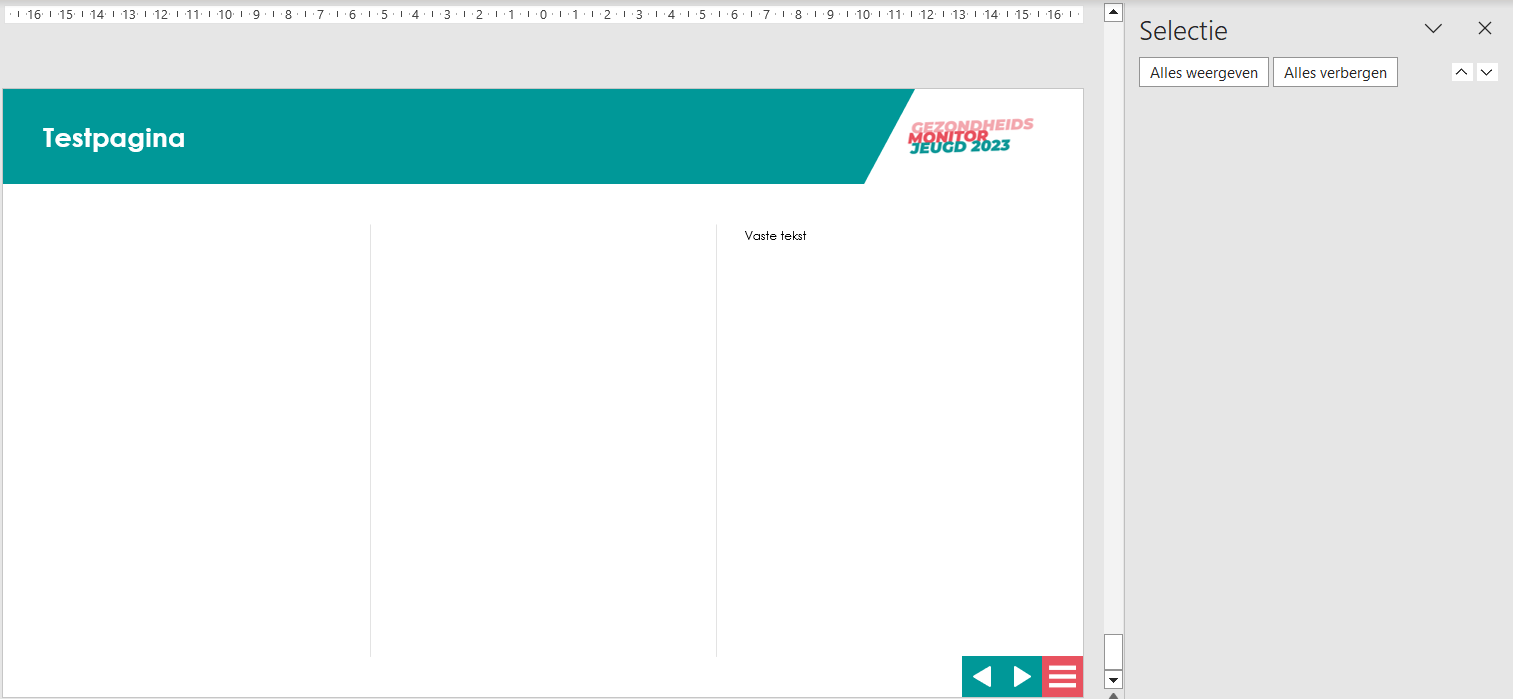


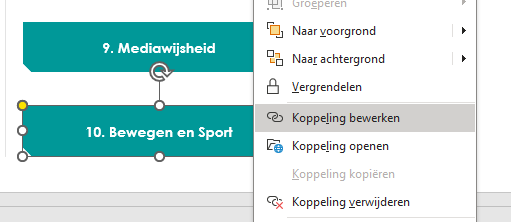
1. Je krijgt het volgende deelvenster te zien.

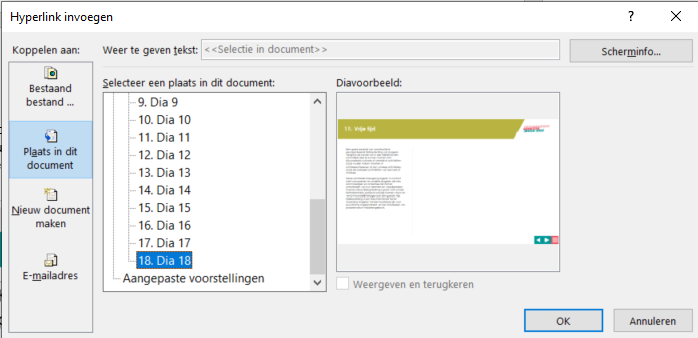


1. In het bovenstaande voorbeeld correspondeert het bolletje met xx% met ‘Tijdelijke aanduiding voor tekst 2’ en de grafiek met ‘Tijdelijke aanduiding voor grafiek 9’. Deze namen kun je in het selectievenster aanpassen. Het is belangrijk dat je hier elk object een unieke naam geeft en dat deze naam ook exact overeenkomt met het label in de kolom ‘label’ in het tabblad ‘Slideconfiguratie 1’ van het configuratiebestand. Als dit niet geval is kan het R-script een foutmelding geven. Op basis van deze naam/dit label bepaald het R-script namelijk waar het cijfer of de figuur op de indeling komt te staan.
2. Naast de twee tijdelijke aanduidingen bevat dit voorbeeld ook vaste content wat in het selectievenster is weergeven als ‘Tekstvak 3’. De naamgeving van de vaste content is niet van belang, deze content wordt altijd toegevoegd.
3. Vul op deze manier de slide met de desbetreffende indicatoren en vaste content en voeg de desbetreffende indicatoren toe aan het configuratiebestand.
4. Elke indeling is opgedeeld in 3 gelijke kolommen weergegeven door de grijze vlakken en in het selectievenster heet dit ‘Linkervlak’, ‘Middenvlak’ en ‘Rechtervlak’. Deze kun je gebruiken bij het uitlijnen van je tekst en figuren en kun je in de uiteindelijk slides onzichtbaar maken door op het oogje te klikken.



1. Sluit vervolgens het diamodel af en ga naar de powerpoint presentatie. Hier kun je de nieuwe dia toevoegen door op ‘Nieuwe dia’ te klikken. Hier zie je alle indelingen die in het diamodel staan. Hieronder zie je bijvoorbeeld ook de hierboven gemaakte indeling van de ‘Testpagina’. Let op: Ook als je een pagina hebt aangepast dan is het belangrijk dat je deze pagina met de aangepaste indeling opnieuw invoegt en de oude pagina verwijderd.
2. Nadat je de slide heb toegevoegd is het belangrijk dat je alle tijdelijke content verwijderd, deze inhoud wordt immers bij het runnen van het R-script gevuld. Dit kun je doen door elke tijdelijk object te selecteren en op delete te drukken. In het selectievenster kun je zien of je alle tijdelijke inhoud hebt verwijderd.
3. Als er nieuwe slides zijn toegevoegd, zijn omgedraaid of zijn verwijderd dan is zal ook de inhoudsopgave aangepast moeten worden. Dit doe je op de slide zelf en in NIET in het diamodel. De blokjes kunnen makkelijk versleept worden en de tekst kan worden aangepast. Als je een nieuwe slide hebt aangemaakt dan kun je het best een bestaand blokje kopiëren en de koppeling aanpassen door met je rechtermuisknop op ‘Koppeling bewerken’ te klikken. Vervolgens kies je bij ‘Koppelen aan:’ uit ‘Plaats in dit document’. In de lijst kun je dan de desbetreffende dia selecteren en op OK drukken.





## Stap 7 – R Script runnen

**Belangrijk: met dit script wordt een PowerPoint bestand aangemaakt. Wil je het script opnieuw runnen zorg er dan voor dat de PowerPoint waarin de output wordt opgeslagen en het Excel configuratiebestand is afgesloten. Doe je dit niet dan kan R de bestand niet uploaden en loopt R vast.**

Om het R-script te runnen zijn de volgende packages nodig:

1. haven
2. labelled
3. readxl
4. tidyverse
5. mschart
6. officer

Deze packages moeten eenmalig voor gebruik worden geïnstalleerd. Voor elke sessie (oftewel wanneer je Rstudio opent) worden deze libraries in het script geladen.

Het R-script bestaat uit verschillende stappen:

1. Dataconfiguratie.
   * In dit stuk wordt het SPSS trendbestand ingeladen of de losse SPSS bestanden per jaar. Deze worden in dit stuk samengevoegd.
2. Excelbestand met configuratie laden
   * Hier wordt het configuratiebestand ingeladen.
3. Grafiekinstellingen en opmaak
   * Definiëren van lettertype.
   * Bepalen van kleuren.
   * Bepalen van grafiekinstellingen met betrekking tot assen en labels.
4. Functie maken om cijfers te berekenen
   * Hier worden alle cijfers berekend op basis van het configuratiebestand. Voor alle cijfers wordt een minimale N van 30 per vraag en 5 per cel gehanteerd.
   * Worden chi-square test uitgevoerd waar relevant (bij tekstvergelijkingen).
5. Functies maken om content te genereren
   * Hier wordt per type een functie aangemaakt om de percentages, teksten en figuren te genereren die in de slides worden weergegeven.
6. Functie aanmaken om content te genereren en te plaatsen

* Deze functie bepaalt op basis van de 'type' kolom uit de configuratie Excel welke functie moet worden uitgevoerd. Daarna wordt de output van de gekozen functie in de PowerPoint geplaatst.

1. Rapportage maken.
2. Rapportages maken op basis van rapportconfiguratie.

In het script zelf staan nog opmerkingen met wat meer uitleg.

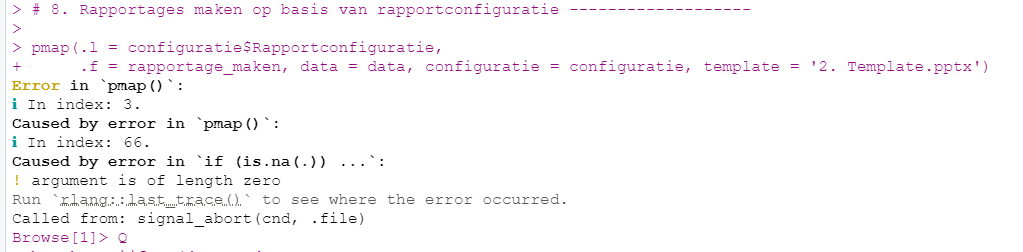
We hopen dat alles in een keer goed gaat. Mocht je toch nog vragen hebben, iets tegenkomen wat niet klopt of een error krijgen dan kun je deze kwijt op [GitHub](https://github.com/ggdatascience/rapportage_monitor_gmj). Hier zie je ook of andere collega’s tegen dezelfde issues aan zijn gelopen en of er updates zijn van het script. Onder stap 8 vind je wat voorbeelden van mogelijke errors en hoe deze op te lossen. Account

## Stap 8 - Foutmeldingen lezen script

Het R-scripts kan waarschuwingen en errors geven. Waarschuwingen kun je negeren. Bij errors stopt het script.

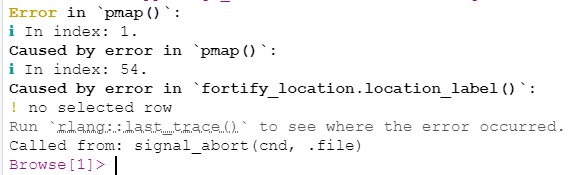
### Errors

Een error melding die voor kan komen is als volgt:

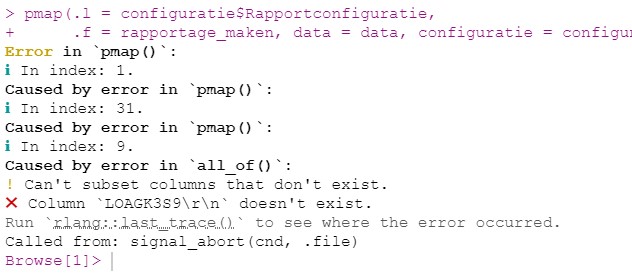


Hier wordt door middel van de twee i’tjes aangegeven waar je naar de fout kan zoeken. Hier geeft de 3 aan dat het gaat om het derde rapport (regel 4 van de rapportconfiguratie) en dat er bij de 66 indicator (regel 67 van desbetreffende slideconfiguratie). In dit voorbeeld is de error veroorzaakt door een lege cel. Dit kan gebeuren als 0 personen ja hebben aangekruist bij een variabele. Deze error is opgelost in het script, maar mocht je deze tegenkomen dan kun je hier een melding van maken op GitHub.

Een andere mogelijke error:



Als deze error wordt weergegeven dan komt of het nummer onder de index niet overeen met de slidenummer of is de tekst onder label niet gelijk aan de tekst in de tijdelijke aanduiding in het diamodel.

Nog een mogelijke error:

In dit geval staan er onder indicator meerdere (9) indicatoren gescheiden door ‘; ‘. Echter hier is een enter teveel geplaatst.

### Foutmelding functie

Onder aan het script is nog een functie ‘Foutmeldingen oplossen’ toegevoegd. Hier kun je bovenstaande indexnummers invoeren en krijg je te zien of alle gegevens die opgeroepen correct zijn.

foutmeldingen(rapportconfiguratie = 'Rapportconfiguratie',  
 rapportregel = 1,  
 slideconfiguratie = 'Slideconfiguratie Testdata',  
 slideregel = 54)

### Percentages controleren

Verder kun je in het R-script zelf controleren of variabele een waarde genereerd door het volgende stukje script te gebruiken:

bereken\_cijfers(data = data, basis = NA, basis\_label = NA, referentie = NA, referentie\_label = NA, omschrijving = NA, indicator = NA , waarden = 1, valuelabel = NA, uitsplitsing = NA, groepering = NA, niveau = NA , jaar = NA, var\_jaar = 'Onderzoeksjaar', Nvar = 30, Ncel = 5, toetsen = F)

Elke NA kun je vervangen door de gegevens zoals je deze gedefinieerd hebt in de slide configuratie. Let wel op de ‘ ‘.

In het voorbeeld hieronder controleren we voor het Ichthus College of er een percentage wordt gegenereerd in het jaar 2023 voor de indicator SBSXK302:

bereken\_cijfers(data = data, basis = 'SchoolNaam', basis\_label = 'Ichthus College', referentie = NA, referentie\_label = NA, omschrijving = NA, indicator = 'SBSXK302', waarden = 1, valuelabel = NA, uitsplitsing = NA, groepering = NA, niveau = 'basis', jaar = 2023, var\_jaar = 'Onderzoeksjaar', Nvar = 30, Ncel = 5, toetsen = F)

## Stap 9 - Overige informatie

Bij grafieken is de limiet van de y-as nu 1 (tot 100%) als de maximumwaarde in de grafiek gelijk is of hoger dan 50%. De limiet van de y-as is 0.5 (tot 50%) als de maximumwaarde in de grafiek lager is dan 50%.

Statistische vergelijkingen kunnen alleen gedaan worden binnen 2 groepen en niet binnen 3 of meer groepen.